

(51) International Patent Classification:

CI2N 15/63 (2006.01) CI2N 15/11 (2006.01)

CI2N 15/113 (2010.01) CI2N 15/86 (2006.01)

A61K 48/00 (2006.01)

(21) International Application Number:

PCT/EP2024/065689

(22) International Filing Date:

07 June 2024 (07.06.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

23178358.0 09 June 2023 (09.06.2023) EP

(71) Applicant (for all designated States except US): F. HOFFMANN-LA ROCHE AG [CH/CH]; Grenzacherstrasse 124, 4070 Basel (CH).

(71) Applicant (for US only): HOFFMANN-LA ROCHE INC. [US/US]; Overlook at Great Notch, 150 Clove Road, 8th Floor, Suite 8 - Legal Department, Little Falls, New Jersey 07424 (US).

(72) Inventors: EBELING, Martin; c/o F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel (CH). JAGASIA, Ravi; c/o F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel (CH). MENDEL, Mateusz Piotr; c/o F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel (CH). ROUDNICKY, Filip; c/o F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel (CH). SCHUKUR, Lina; c/o F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel (CH).

(74) Agent: BUSER, Andres; c/o F. Hoffmann-La Roche AG, Patent Department, Grenzacherstrasse 124, 4070 Basel (CH).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

— in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE

(54) Title: INDUCIBLE GENE EXPRESSION SYSTEM

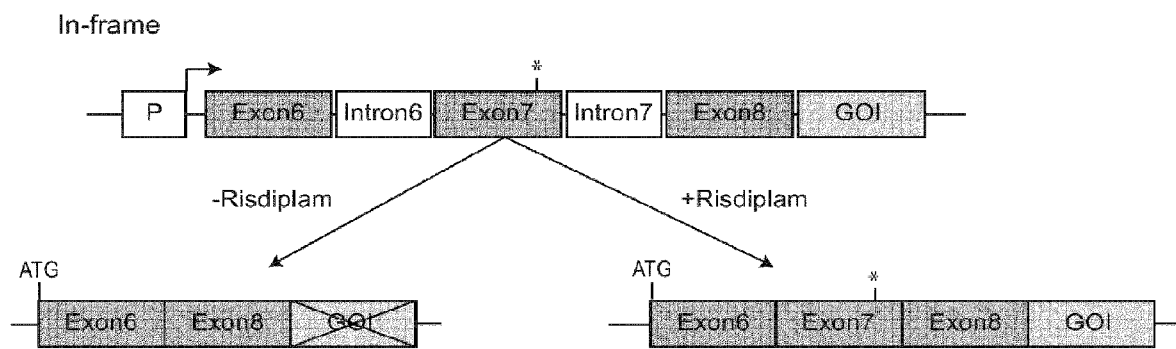


Figure 1

(57) Abstract: The present disclosure relates to the fields of molecular biology and nucleic acid technology. The present disclosure also relates to therapy and prophylaxis of disease.

Inducible Gene Expression System

This application claims priority from EP23178358.0 filed 9 June 2023, the contents and elements of which are herein incorporated by reference for all purposes.

Technical Field

The present disclosure relates to the fields of molecular biology and nucleic acid technology. The present disclosure also relates to therapy and prophylaxis of disease.

5 Background

Gene therapy as a modality aims at ameliorating disease-related phenotypes by using both viral and non-viral delivery systems. Significant efforts have been focusing on optimizing this therapy to be more efficient, non-immunogenic and less toxic while enabling a long-term expression of the gene of interest. One of the key challenges to achieve this goal is to regulate the expression of the transgene to meet the
10 required level and to be expressed upon demand.

For certain targets, where constitutive and unlimited expression leads to adverse events, regulated gene expression is necessary to render gene therapy a viable option as a treatment modality. One example is Rett Syndrome (RTT), an X-linked neurological disorder associated with severe motor abnormalities and
15 reduced lifespan that occurs in 1 in 1,000 females. RTT is characterized by seemingly normal neurological and physical development during early postnatal period followed by a rapid regression with the loss of the purposeful motor skills and the onset of repetitive and autistic behaviors^{1,2}. During the rapid progression stage, the child loses purposeful hand skills and spoken language, experiences motor impairments, and develops breathing abnormalities and may also develop autistic-like features and
20 seizures. At later stages, further motor deterioration begins, typified by severe physical disability, which leads many patients to become dependent on the use of wheelchairs. Current treatment options are limited to symptom control. Mutations in the *MECP2* gene on the X chromosome, encoding for methyl-CpG-binding protein 2 account for 95% of RTT cases³. MeCP2 is highly expressed in neurons and functions as a ubiquitous transcriptional regulator by binding to methylated DNA and recruiting protein
25 partners and regulatory complexes to control transcriptional activity. MeCP2 regulates neuronal physiology and maintenance and landmark studies have demonstrated that restoring levels of the MeCP2 protein dramatically reverses symptoms in mice^{4,5}.

Given that it is a monogenetic disorder driven by the lack of MeCP2, a protein with multiple functions,
30 gene therapy is one potential avenue to treat RTT. Studies in RTT mice have provided encouraging data showing attenuated neurological dysfunctions as well as extended lifespan by intravenous administration of an Adeno-associated virus serotype 9 (AAV9) expressing the wild-type (WT) MeCP2 gene⁵⁻⁸. Despite the promising potential of gene replacement therapy, high doses of MeCP2 by overexpression systems – which rely on the use of ubiquitous promoters resulting in uncontrolled level of the transgene – can lead
35 to toxicity with phenotypes similar to those observed in duplication syndrome, a condition mainly affecting males, characterized by moderate to severe intellectual impairment, and caused by a duplication of the MeCP2 gene on the X-chromosome^{9,10}. The requirement of keeping MeCP2 expression level within a window for normal function is one of the most challenging aspects that has yet to be solved. One solution

would be to develop controlled transgene expression systems in order to overcome this limitation and make gene therapy a viable treatment option.

Inducible gene expression systems allow for reversibility and flexibility and enable the production of
 5 therapeutics upon demand, thereby circumventing overdose-associated side effects. Such regulation can
 be achieved using small molecules, with which transgene expression can be induced (generating ON-
 switches) or repressed (OFF-switch). Among the existing inducible transcriptional gene regulatory
 systems, the tetracycline (Tet)-regulable system is the most widely exploited tool and can be used as an
 ON- or an OFF-switch. Tet-inducible systems, however, are derived from bacteria and require the
 10 expression of the regulator protein, e.g. the ON-switch components include the tetracycline-controlled
 transactivator (tTA), which consists of the bacterial tetR with a C-terminal domain of VP16 (virion protein
 16), derived from the herpes simplex virus. The required co-expression of these components further
 increase the limitations in the size of the transgene for AAV packaging, and can cause both silencing and
 potential immunogenicity effects when introduced into human tissue, which has hampered further
 15 translation into the clinics^{11,12}. Another class of genetic switches originating from bacteria are
 riboswitches, RNA elements that are able to control gene expression in response to ligand binding, have
 a small genomic footprint and do not depend on other proteins for activity, which renders these types of
 regulated gene expression an attractive alternative to protein-based expression control systems¹³.
 However, riboswitches suffer from a poor dynamic range as well as high basal activity¹⁴.

20 A new class of small molecules that regulate splicing of *SMN2* exon 7 was recently identified, via
 screening using a reporter system based on the *SMN2* transcript (described e.g. in WO 2009/151546 A2).
SMN2 reporter systems have since been proposed to be used as gene expression switches, where
 expression of a transgene is regulated through small molecule-controlled splicing of the expression
 25 cassette (Monteys *et al.* Nature (2021) 596: 291-295). Known human *SMN2* exon 6 to exon 8-derived
 transgene expression systems are described e.g. in Zhang, *et al.*, Gene Ther. (2001) 8: 1532-1538, WO
 2022/204471 A1, Monteys *et al.* Nature (2021) 596: 291-295 and WO 2021/163556 A1.

30 However, these known *SMN2*-based switch systems have very long nucleotide sequences, limiting the
 size of the transgene that can be employed under the control of the system. That is, the size of an insert
 comprising the switch and the coding sequence for a therapeutic polypeptide to be delivered as gene
 therapy would very often be much larger than the packaging limit for vectors routinely employed in the
 delivery of gene therapies, e.g. adeno-associated viral (AAV) vectors.

35 Summary

In a first aspect, the present disclosure provides a polynucleotide comprising in 5' to 3' order:

- (i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence
 identity to a nucleotide sequence according to SEQ ID NO:222, wherein the first nucleotide sequence
 does not comprise SEQ ID NO:2;
- 40 (ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80%
 sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least

80% sequence identity to a nucleotide sequence according to SEQ ID NO:378 at its 3' end, and consisting of fewer than 1044 nucleotides;

(iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:226;

5 (iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:15 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;

(v) a fifth nucleotide sequence: (a) consisting of the dinucleotide 'GA', 'TG' or 'TT', or (b) comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:19, wherein the nucleotide sequence comprises 'GA', 'TG' or 'TT' at the positions corresponding to positions 1 and 2 of SEQ ID NO:19, or (c) encoding a polypeptide of interest, and comprising 'GA', 'TG' or 'TT' at positions 1 and 2; and

10 (vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a) or (v)(b), a sixth nucleotide sequence encoding a polypeptide of interest;

15 wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

In a second aspect, the present disclosure provides a polynucleotide comprising in 5' to 3' order:

20 (i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:1, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;

(ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 at its 3' end, and consisting of fewer than 1044 nucleotides;

25 (iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:26, wherein the third nucleotide sequence comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12;

(iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 at its 5' end, and comprising a nucleotide sequence having at least 30 80% sequence identity to SEQ ID NO:16 at its 3' end;

(v) a fifth nucleotide sequence: (a) consisting of the trinucleotide 'GAG' or (b) encoding a polypeptide of interest, and comprising 'GAG' at positions 1 to 3; and

(vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a), a sixth nucleotide sequence encoding a polypeptide of interest;

35 wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

In some embodiments in accordance with the various aspects of the present disclosure, when the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.

In some embodiments, the first nucleotide sequence comprises a nucleotide sequence according to SEQ ID NO:222. In some embodiments, the first nucleotide sequence comprises, or consists of, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:106, SEQ ID NO:219 or SEQ ID NO:220. In some embodiments, the first nucleotide sequence comprises, or consists of, SEQ ID NO:3.

5

In some embodiments, the second nucleotide sequence comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 or SEQ ID NO:377 at its 3' end. In some embodiments, the second nucleotide sequence consists of fewer than 500 nucleotides. In some embodiments, the second nucleotide sequence comprises, or consists of, SEQ ID NO:9, SEQ ID NO:10 or SEQ ID NO:228. In some

10

embodiments, the first nucleotide sequence comprises, or consists of, SEQ ID NO:10.

In some embodiments, the third nucleotide sequence consists of a nucleotide sequence according to SEQ ID NO:226. In some embodiments, the third nucleotide sequence consists of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:223 or SEQ ID NO:224. In some embodiments, the third nucleotide sequence

15

consists of a nucleotide sequence according to SEQ ID NO:26. In some embodiments, the third nucleotide sequence consists of SEQ ID NO:13 or SEQ ID NO:27.

In some embodiments, the fourth nucleotide sequence comprises SEQ ID NO:15 or SEQ ID NO:379 at its 5' end, and comprises SEQ ID NO:16 at its 3' end. In some embodiments, the fourth nucleotide sequence consists of fewer than 500 nucleotides. In some embodiments, the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:227 or SEQ ID NO:340. In some

20

embodiments, the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17.

In some embodiments, the fifth nucleotide sequence: (a) consists of the dinucleotide 'GA' or 'TG', or (b) comprises, or consists of, a nucleotide sequence according to SEQ ID NO:19. In some embodiments, the fifth nucleotide sequence: (a) consists of the dinucleotide 'GA' or 'TG', or (b) comprises, or consists of, SEQ ID NO:20 or SEQ ID NO:21.

25

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence selected from the group consisting of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:111, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:164 and SEQ ID NO:171. In some embodiments, the polynucleotide comprises at least 80% sequence identity to SEQ ID NO:28 or SEQ ID NO:29.

30

35 In some embodiments, the polynucleotide further comprises a promoter sequence 5' to the start codon.

In some embodiments, the polynucleotide further comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.

40 In some embodiments, the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.

The present disclosure also provides a vector comprising a polynucleotide according to the present disclosure. In some embodiments, the vector is an adeno-associated virus (AAV) vector.

5 The present disclosure also provides a pharmaceutical composition comprising a polynucleotide or vector according to the present disclosure, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.

10 The present disclosure also provides a cell comprising a polynucleotide or vector according to the present disclosure. In some embodiments, the cell further comprises a splicing modifier that promotes *SMN2* exon 7 inclusion. In some embodiments, the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

The present disclosure also provides a method for modifying a cell to express a polypeptide of interest, comprising:

- 15 (i) introducing into a cell a polynucleotide or vector according to the present disclosure; and
(ii) subsequently contacting the cell with a splicing modifier that promotes *SMN2* exon 7 inclusion.
In some embodiments, the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

20 The present disclosure also provides a method for expressing a polypeptide of interest in a cell, comprising contacting a cell according to the present disclosure with a splicing modifier that promotes *SMN2* exon 7 inclusion. In some embodiments, the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

25 The present disclosure also provides a method for inhibiting expression of a polypeptide of interest in a cell, comprising contacting a cell according to the present disclosure with a splicing modifier that promotes *SMN2* exon 7 inclusion.

30 The present disclosure also provides a method for modifying a cell to express a polypeptide of interest, comprising introducing into a cell a polynucleotide or vector according to the present disclosure.

The present disclosure also provides a polynucleotide, vector or pharmaceutical composition according to the present disclosure, for use in a method of medical treatment or prophylaxis.

35 The present disclosure also provides a polynucleotide, vector or pharmaceutical composition according to the present disclosure, for use in treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.

40 The present disclosure also provides the use of a polynucleotide, vector or pharmaceutical composition according to the present disclosure, in the manufacture of a medicament for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.

The present disclosure also provides a method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising administering to a subject a polynucleotide, vector or pharmaceutical composition according to the present disclosure.

In some embodiments, treating or preventing the disease or condition further comprises administering to the subject a splicing modifier that promotes *SMN2* exon 7 inclusion. In some embodiments, the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

In some embodiments, the disease or condition is a disease or condition characterised by deficiency of the polypeptide of interest.

The present disclosure also provides a kit, comprising:

- (i) a polynucleotide, vector or pharmaceutical composition according to the present disclosure; and
- (ii) a splicing modifier that promotes *SMN2* exon 7 inclusion. In some embodiments, the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

Description

The present disclosure relates to inducible transgene expression systems using small molecule splicing modifiers (SMSM), and based on the described mechanism of action of the Survival of Motor Neuron 2 (*SMN2*) splicing modifier risdiplam (and variants thereof).

The SMSM binding sites for *SMN2* have previously been described to be two distinct sites within *SMN2* exon 7 pre-mRNA: the ESE2 region and the 5' splice donor site (5'ss) (Figure 2a)¹⁵. Interaction of the small molecules with the mRNA-protein complex is thought to be critical for the high selectivity of the compound. This study suggested that both ESE2 and 5'ss regions are required for full activity of the SMSM compounds in a dose-dependent manner¹⁵. Structural analysis using a compound from the same chemical class as risdiplam (SMN-C5) has demonstrated that the drug selectively promotes the recognition of the weak 5' splice site of *SMN2* exon 7 by U1 snRNP by stabilizing an unpaired adenine at the exon-intron junction in the RNA helix¹⁶.

The present disclosure provides novel *SMN2* exon 6 to exon 8-derived constructs providing for the SMSM-inducible regulation of transgene expression, which are suitable to be employed for chemically-inducible regulation of the expression of gene therapies.

In particular, the present disclosure provides ON-switch constructs: (i) having a size permitting their application for the SMSM-inducible expression of polypeptides to be delivered as gene therapies, (ii) minimising undesirable expression of the polypeptide in the absence of the SMSM, and (iii) which are highly responsive to the SMSM, thus providing for strong induction of expression of the polypeptide in the presence of the SMSM, and (iv) that minimise the number of/completely remove extraneous amino acids

at the N-terminus of the expressed polypeptide. The present disclosure also provides OFF-switch constructs: (i) having a size permitting their application for the SMSM-inducible expression of polypeptides to be delivered as gene therapies, (ii) maximising expression of the polypeptide in the absence of the SMSM, and (iii) which are highly responsive to the SMSM, thus minimising expression of the polypeptide in the presence of the SMSM, and (iv) that minimise the number of/completely remove extraneous amino acids at the N-terminus of the expressed polypeptide.

Polynucleotides

Aspects and embodiments of the present disclosure relate to polynucleotides.

10

A 'polynucleotide' refers to a polymer chain of a plurality of nucleotide monomers linked by bonds between the monomers, typically phosphodiester bonds (*e.g.* in the case of polynucleotides formed by naturally-occurring nucleotide monomers). Polynucleotides include oligonucleotides, which generally comprise ≤ 50 nucleotides. A polynucleotide may be single-stranded, or may be double-stranded (*i.e.* may comprise a duplex formed by hydrogen-bonding between complementary nucleotides). Polynucleotides according to the present disclosure may comprise or consist of: single-stranded DNA, double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single-stranded RNA, double-stranded RNA, RNA that is mixture of single- and double-stranded regions, single-stranded molecules comprising DNA and RNA, double-stranded molecules comprising DNA and RNA, and molecules comprising DNA and RNA having a mixture of single- and double-stranded regions.

20

In some embodiments, a polynucleotide comprises or consists of DNA. In some embodiments, a polynucleotide is a polydeoxyribonucleotide. In some embodiments, a polynucleotide comprises or consists of RNA. In some embodiments, a polynucleotide is a polyribonucleotide.

25

In aspects and embodiments wherein the polynucleotide of the present disclosure is defined by reference to a given nucleotide sequence, and wherein the given nucleotide sequence comprises or consists of RNA and/or is a polyribonucleotide, it will be appreciated that instances of 'T' for thymidine in such sequences are replaced with 'U', for uracil.

30

The present disclosure also contemplates polynucleotides comprising modified nucleotides, *e.g.* in which the phosphonate and/or ribose and/or base of a deoxyribonucleotide or ribonucleotide is/are chemically modified. Nucleotide modifications contemplated in accordance with the present disclosure include those described in Hu *et al.*, Sig. Transduc. Tar. Ther. (2020) 5(101), which is hereby incorporated by reference in its entirety.

35

Phosphonate modifications may be selected from phosphorothioate (*e.g.* Rp isomer, Sp isomer), phosphorodithioate, methylphosphonate, methoxypropylphosphonate, 5'-(*E*)-vinylphosphonate, 5'-methylphosphonate, (S)-5'-C-methyl with phosphate, 5'-phosphorothioate, and peptide nucleic acid modifications. Ribose modifications may be selected from 2'-O-methyl, 2'-O-methoxyethyl, 2'-fluoro, 2'-deoxy-2'-fluoro, 2'-methoxyethyl, 2'-O-alkyl, 2'-O-allyl, 2'-C-allyl, 2'-deoxy, 2'-hydroxyl, 2'-arabino-fluoro, 2'-O-benzyl, 2'-O-methyl-4-pyridine, locked nucleic acid, (S)-cEt-BNA, tricyclo-DNA, PMO, unlocked

40

nucleic acid, hexitol nucleic acid and glycol nucleic acid modifications. Base modifications may be selected from pseudouridine, 2'-thiouridine, N6'-methyladenosine, 5'-methylcytidine, 5'-fluoro-2'-deoxyuridine, N-ethylpiperidine 7'-EAA triazole-modified adenine, N-ethylpiperidine 6'-triazole-modified adenine, 6'-phenylpyrrolo-cytosine, 2',4'-difluorotoluy ribonucleoside and 5'-nitroindole modifications.

5

In some embodiments, a modified nucleotide may be selected from 2'-O-methyluridine-3'-phosphate, 2'-O-methyladenosine-3'-phosphate, 2'-O-methylguanosine-3'-phosphate, 2'-O-methylcytidine-3'-phosphate, 2'-O-methyluridine-3'-phosphorothioate, 2'-O-methyladenosine-3'-phosphorothioate, 2'-O-methylguanosine-3'-phosphorothioate, 2'-O-methylcytidine-3'-phosphorothioate, 2'-fluorouridine-3'-phosphate, 2'-fluoroadenosine-3'-phosphate, 2'-fluoroguanosine-3'-phosphate, 2'-fluorocytidine-3'-phosphate, 2'-fluorocytidine-3'-phosphorothioate, 2'-fluoroguanosine-3'-phosphorothioate, 2'-fluoroadenosine-3'-phosphorothioate, and 2'-fluorouridine-3'-phosphorothioate.

10

Nucleotide sequences of the polynucleotides of the present disclosure

Polynucleotides of the present disclosure are defined herein by reference to constituent nucleotide sequences.

15

It will be appreciated that the constituent nucleotide sequences of polynucleotides according to the present disclosure are provided as subsequences of the complete sequence of the polynucleotide.

20

In some aspects and embodiments, the constituent nucleotide sequences of polynucleotides according to the present disclosure are provided in a particular order in the sequence of the polynucleotide, *e.g.* from 5' to 3'. By way of illustration, the first nucleotide sequence is provided 5' to (*i.e.* upstream of) the second nucleotide sequence, in the context of the sequence of the polynucleotide. Similarly, the second nucleotide sequence is 5' to the third nucleotide sequence, *etc.*

25

The constituent nucleotide sequences of the polynucleotides of the present disclosure are non-overlapping. In some embodiments, constituent nucleotide sequences of the polynucleotides are provided in tandem in the context of the complete sequence of the polynucleotide. In some embodiments, constituent nucleotide sequences of the polynucleotide are immediately adjacent to one another (*i.e.* the 3' nucleotide of a given nucleotide sequence is followed immediately by the 5' nucleotide of another given nucleotide sequence, in the context of the complete sequence of the polynucleotide). By way of illustration, in the polynucleotide of SEQ ID NO:22, positions 1 to 45 form the first nucleotide sequence, positions 46 to 309 form the second nucleotide sequence, positions 310 to 364 form the third nucleotide sequence, positions 356 to 616 form the fourth nucleotide sequence, and positions 617 and 618 form the fifth nucleotide sequence.

30

35

First nucleotide sequence

The first nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence which is a variant of human *SMN2* exon 6. A 'variant' of a given reference nucleotide sequence comprises one or more differences relative to the reference nucleotide sequence. For example, a variant

40

of a given reference nucleotide sequence may comprise insertion, deletion or substitution of one or more nucleotides relative to the reference nucleotide sequence.

Accordingly, in some embodiments, the first nucleotide sequence does not consist of, or does not
 5 comprise, SEQ ID NO:2. In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having <100% sequence identity to SEQ ID NO:2.

In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$,
 10 $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1. It will be appreciated that 'a nucleotide sequence according to SEQ ID NO:1' refers to a nucleotide sequence conforming to the consensus nucleotide sequence of SEQ ID NO:1. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:1 include SEQ ID NOs:3, 4, 5, 6 and 106.

15 In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:3. In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide
 20 sequence according to SEQ ID NO:4. In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:5. In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or
 25 100%) sequence identity to SEQ ID NO:6.

In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:221. Exemplary
 30 sequences conforming to the consensus nucleotide sequence of SEQ ID NO:1 include SEQ ID NOs:219 and 220.

In some embodiments, the first nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$,
 35 $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:222.

In some embodiments, the first nucleotide sequence comprises 'CAG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2 (cf. e.g. positions 109 to 111 of SEQ ID NO:6).

40 As explained herein, position(s) of a nucleotide sequence which 'corresponds to' specified position(s) of a reference nucleotide sequence can be identified by sequence alignment of the subject sequence to the reference sequence, e.g. using sequence alignment software such as ClustalOmega (Söding, J. 2005,

Bioinformatics 21, 951-960). By way of illustration, it will be appreciated that positions 43 to 45 of SEQ ID NO:3 correspond to positions 109 to 111 of SEQ ID NO:2. Similarly, positions 45 to 47 of SEQ ID NO:4 correspond to positions 109 to 111 of SEQ ID NO:2.

- 5 In some embodiments, the first nucleotide sequence comprises 'CTG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2 (*cf. e.g.* positions 109 to 111 of SEQ ID NO:5).

In some embodiments, the first nucleotide sequence comprises 'GGC' at the positions corresponding to positions 64 to 66 of SEQ ID NO:2 (*cf. e.g.* positions 64 to 66 of SEQ ID NO:5, positions 64 to 66 of SEQ ID NO:6).

10

In some embodiments, the first nucleotide sequence comprises 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2 (*cf. e.g.* positions 16 to 18 of SEQ ID NO:3, positions 18 to 20 of SEQ ID NO:4, positions 82 to 84 of SEQ ID NO:5).

15

In some embodiments, the first nucleotide sequence comprises 'GGC' at the positions corresponding to positions 64 to 66 of SEQ ID NO:2, and comprises 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2.

- 20 In some embodiments, the first nucleotide sequence comprises 'GTG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2. In some embodiments, the first nucleotide sequence comprises 'TTG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2. In some embodiments, the first nucleotide sequence comprises 'TAG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2. In some embodiments, the first nucleotide sequence comprises 'CAG' at the positions
- 25 corresponding to positions 39 to 41 of SEQ ID NO:2. In some embodiments, the first nucleotide sequence comprises 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2.

In some embodiments, the first nucleotide sequence consists of fewer than 111 nucleotides. In some embodiments the first nucleotide sequence consists of <100 nucleotides, *e.g.* one of <85 nucleotides, <80 nucleotides, <75 nucleotides, <70 nucleotides, <65 nucleotides, <60 nucleotides, <55 nucleotides, <50 nucleotides or ≤45 nucleotides.

30

In some embodiments, the first nucleotide sequence consists of 45 nucleotides. In some embodiments, the first nucleotide sequence consists of 51 nucleotides. In some embodiments, the first nucleotide sequence consists of 81 nucleotides.

35

- In some embodiments, the first nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (*e.g.* one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and (ii)
- 40 comprises 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iii) consists of ≤45 nucleotides. Examples of such first nucleotide sequences include SEQ ID NOs:3, 4 and 106.

In some embodiments, the first nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and (ii) comprises 'GGC' at the positions corresponding to positions 64 to 66 of SEQ ID NO:2; and (iii) comprises 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iv) comprises 'CTG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2. An example of such a first nucleotide sequence is SEQ ID NO:5.

In some embodiments, the first nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and (ii) comprises 'GGC' at the positions corresponding to positions 64 to 66 of SEQ ID NO:2; and (iii) comprises 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iv) comprises 'CAG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2. An example of such a first nucleotide sequence is SEQ ID NO:6.

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a first nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:3, 5, 6, 219, 220, and 230 to 237.

In some embodiments, a polynucleotide according to the present disclosure does not comprise the nucleotide sequence of SEQ ID NO:2.

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a first nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:3. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a first nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:5. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a first nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:6. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a first nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:219. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a first nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one

of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:220.

Second nucleotide sequence

- 5 The second nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence which is a variant of human *SMN2* intron 6.

Accordingly, in some embodiments, the second nucleotide sequence does not consist of, or does not comprise, SEQ ID NO:30. In some embodiments, the second nucleotide sequence comprises, or consists of, a nucleotide sequence having $<100\%$ sequence identity to SEQ ID NO:30.

In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end.

15 In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:378 at its 3' end. It will be appreciated that 'a nucleotide sequence according to SEQ ID NO:378' refers to a nucleotide sequence conforming to the consensus nucleotide sequence of SEQ ID NO:378. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:378 include SEQ ID NO:8 and SEQ ID NO:377.

In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end. In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:377 at its 3' end.

In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end, and comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end. In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end, and comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:377 at its 3' end.

As used herein, the '5' end' of a given nucleotide sequence refers to the region of the nucleotide sequence formed by nucleotides 5' to (i.e. upstream of, in the context of the complete sequence of the given nucleotide sequence) the midpoint of the given nucleotide sequence, or a subsequence of this region.

In a nucleotide sequence having an even number of nucleotides, the region 5' to the 'midpoint' includes the nucleotide immediately preceding the midpoint. For example, in a nucleotide sequence consisting of 10 nucleotides, the region 5' to the midpoint consists of positions 1 to 5. Similarly, in a nucleotide
 5 sequence having an even number of nucleotides, the region 3' to the 'midpoint' includes the nucleotide immediately after the midpoint. For example, in a nucleotide sequence consisting of 10 nucleotides, the region 3' to the midpoint consists of positions 6 to 10. In a nucleotide sequence having an odd number of nucleotides, the region 5' to the 'midpoint' includes the nucleotides 5' to (*i.e.* upstream of) the nucleotide provided at the midpoint. For example, in a nucleotide sequence consisting of 9 nucleotides, the region 5'
 10 to the midpoint consists of positions 1 to 4. Similarly, in a nucleotide sequence having an odd number of nucleotides, the region 3' to the 'midpoint' includes the nucleotides 3' to (*i.e.* downstream of) the nucleotide provided at the midpoint. For example, in a nucleotide sequence consisting of 9 nucleotides, the region 3' to the midpoint consists of positions 6 to 9.

15 By way of illustration, in SEQ ID NO:9, which consists of 414 nucleotides, a sequence provided at the 5' end of SEQ ID NO:9 refers to a nucleotide sequence formed by positions 1 to 207 of SEQ ID NO:9, or a subsequence thereof.

Conversely, the '3' end' of a given nucleotide sequence refers to the region of the nucleotide sequence
 20 formed by nucleotides 3' to (*i.e.* downstream of, in the context of the complete sequence of the given nucleotide sequence) the midpoint of the given nucleotide sequence, or a subsequence of this region. By way of illustration, in SEQ ID NO:9, a sequence provided at the 3' end of SEQ ID NO:9 refers to a nucleotide sequence formed by positions 208 to 414 of SEQ ID NO:9, or a subsequence thereof.

25 In some embodiments, a subsequence of a given nucleotide sequence provided at the 5' end of the given nucleotide sequence includes at least 5%, *e.g.* one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region of the nucleotide sequence formed by nucleotides 5' to (*i.e.* upstream of, in the context of the complete
 30 sequence of the given nucleotide sequence) the midpoint of the given nucleotide sequence. By way of illustration, a sequence provided at the 5' end of SEQ ID NO:9 may comprise at least 5%, *e.g.* one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region formed by positions 1 to 207 of SEQ ID NO:9.

In some embodiments, a subsequence of a given nucleotide sequence provided at the 3' end of the given
 35 nucleotide sequence includes at least 5%, *e.g.* one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region of the nucleotide sequence formed by nucleotides 3' to (*i.e.* downstream of, in the context of the complete sequence of the given nucleotide sequence) the midpoint of the given nucleotide sequence. By way of
 40 illustration, a sequence provided at the 3' end of SEQ ID NO:9 may comprise at least 5%, *e.g.* one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region formed by positions 208 to 414 of SEQ ID NO:9.

- In some embodiments, a subsequence of a given nucleotide sequence provided at the 5' end of the given nucleotide sequence includes one or more nucleotides provided within 25 nucleotides, e.g. within one of 20, 15, 10 or 5 nucleotides, of the 5' nucleotide of the given nucleotide sequence. In some embodiments, a subsequence of a given nucleotide sequence provided at the 5' end of the given nucleotide sequence includes the 5' nucleotide of the given nucleotide sequence. By way of illustration, a sequence provided at the 5' end of SEQ ID NO:9 may comprise one or more nucleotides provided within positions 1 to 25, e.g. within positions 1 to 20, 1 to 15, 1 to 10, or 1 to 5 of SEQ ID NO:9. By way of illustration, a sequence provided at the 5' end of SEQ ID NO:9 may comprise position 1 of SEQ ID NO:9.
- 10 In some embodiments, a subsequence of a given nucleotide sequence provided at the 3' end of the given nucleotide sequence includes one or more nucleotides provided within 25 nucleotides, e.g. within one of 20, 15, 10 or 5 nucleotides, of the 3' nucleotide of the given nucleotide sequence. In some embodiments, a subsequence of a given nucleotide sequence provided at the 3' end of the given nucleotide sequence includes the 3' nucleotide of the given nucleotide sequence. By way of illustration, a sequence provided at the 3' end of SEQ ID NO:9 may comprise one or more nucleotides provided within positions 389 to 414, e.g. within positions 394 to 414, 399 to 414, 404 to 414, or 409 to 414 of SEQ ID NO:9. By way of illustration, a sequence provided at the 3' end of SEQ ID NO:9 may comprise position 414 of SEQ ID NO:9.
- 15
- 20 In some embodiments the second nucleotide sequence comprises a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 in the region 5' to its midpoint. In some embodiments, the second nucleotide sequence comprises a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 or SEQ ID NO:377 in the region 3' to its midpoint.
- 25 In some embodiments, the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 accounts for at least 5%, e.g. one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region 5' to the midpoint of the second nucleotide sequence. In some embodiments, the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 or SEQ ID NO:377 accounts for at least 5%, e.g. one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region 3' to the midpoint of the second nucleotide sequence.
- 30
- In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 comprises one or more nucleotides provided within 25, e.g. within positions 20, 15, 10, or 5 nucleotides of the 5' nucleotide of the second nucleotide sequence. In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 comprises the 5' nucleotide of the second nucleotide sequence. In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 or SEQ ID NO:377 comprises one or more nucleotides provided within 25, e.g. within positions 20, 15, 10, or 5 nucleotides of the 3' nucleotide of the second nucleotide sequence. In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 or SEQ ID NO:377 comprises the 3' nucleotide of the second nucleotide sequence.
- 35
- 40

An exemplary sequence comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end is SEQ ID NO:9. Positions 1 to 102 of SEQ ID NO:9 correspond to positions 1 to 102 of SEQ ID NO:7.

- 5 An exemplary sequence comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 at its 3' end is SEQ ID NO:9. Positions 253 to 414 of SEQ ID NO:9 correspond to positions 1 to 162 of SEQ ID NO:8.

10 In some embodiments, the second nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:229. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:229 include SEQ ID NOs:10 and 228.

15 In some embodiments, the second nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:9. In some embodiments, the second nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID
20 NO:10.

In some embodiments, the second nucleotide sequence comprises 'C' at the position corresponding to position 5766 of SEQ ID NO:30 (*cf.* e.g. position 411 of SEQ ID NO:9, position 261 of SEQ ID NO:10).

25 In some embodiments, the second nucleotide sequence consists of fewer than 1044 nucleotides, e.g. one of <1000 nucleotides, <900 nucleotides, <800 nucleotides, <750 nucleotides, <700 nucleotides, <650 nucleotides, <600 nucleotides, <550 nucleotides or <500 nucleotides.

30 In some embodiments, the second nucleotide sequence consists of fewer than 500 nucleotides, e.g. one of <450 nucleotides, <400 nucleotides, <350 nucleotides, <300 nucleotides, <250 nucleotides or <200 nucleotides. In some embodiments, the second nucleotide sequence consists of 1044 nucleotides. In some embodiments, the second nucleotide sequence consists of 414 nucleotides. In some embodiments, the second nucleotide sequence consists of 264 nucleotides. In some embodiments, the second nucleotide sequence consists of 189 nucleotides.

35 In some embodiments, the second nucleotide sequence comprises: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%)
40 sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprises 'C' at the position corresponding to position 5766 of SEQ ID NO:30; and (iv) consists of fewer than 1044 nucleotides. Examples of such second nucleotide sequences include SEQ ID NOs:9 and 10.

In some embodiments, the second nucleotide sequence comprises: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) a nucleotide sequence having at least 80%
 5 (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprises 'C' at the position corresponding to position 5766 of SEQ ID NO:30; and (iv) consists of fewer than 500 nucleotides. An example of such a second nucleotide sequence is SEQ ID NO:9.

10 In some embodiments, the second nucleotide sequence comprises: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprises 'C' at the position corresponding to
 15 position 5766 of SEQ ID NO:30; and (iv) consists of fewer than 300 nucleotides. An example of such a second nucleotide sequence is SEQ ID NO:10.

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a second nucleotide sequence) comprising, or consisting of, a nucleotide sequence
 20 having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:9, 10, 228, 238 to 255, 375 and 376.

In some embodiments, a polynucleotide according to the present disclosure does not comprise the
 25 nucleotide sequence of SEQ ID NO:240. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a second nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:9, 10, 228, 238, 239, 241 to 255, 375 and 376.

30 In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a second nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:9. In some embodiments, a polynucleotide according to
 35 the present disclosure comprises a nucleotide sequence (e.g. a second nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a second nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least
 40 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:228.

Third nucleotide sequence

The third nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence which is a variant of human *SMN2* exon 7.

- 5 Accordingly, in some embodiments, the third nucleotide sequence does not consist of, or does not comprise, SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having <100% sequence identity to SEQ ID NO:12.

10 In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11. It will be appreciated that 'a nucleotide sequence according to SEQ ID NO:11' refers to a nucleotide sequence conforming to the consensus nucleotide sequence of SEQ ID NO:11. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:11 include SEQ ID NOs:13 and 14.

15 In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:26. It will be appreciated that 'a nucleotide sequence according to SEQ ID NO:26' refers to a nucleotide sequence conforming to the consensus nucleotide sequence of SEQ ID NO:26. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:26 include SEQ ID NOs:13 and 27.

25 In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:225. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:225 include SEQ ID NOs:223 and 224.

30 In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:226.

35 In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to SEQ ID NO:13. In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to SEQ ID NO:14. In some embodiments, the third nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of ≥85%, ≥90%, ≥91%, ≥92%, ≥93%, ≥94%, ≥95%, ≥96%, ≥97%, ≥98%, ≥99% or 100%) sequence identity to SEQ ID NO:27.

40

In some embodiments, the third nucleotide sequence comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12 (*cf. e.g.* position 2 of SEQ ID NOs:11, 13, 14, 26 and 27).

5 In some embodiments, the third nucleotide sequence comprises insertion of 'GCCACC' after the position corresponding to position 6 of SEQ ID NO:12 (*cf. e.g.* positions 7 to 12 of SEQ ID NO:14).

In some embodiments, the third nucleotide sequence comprises 'TG' at the positions corresponding to positions 8 to 9 of SEQ ID NO:12 (*cf. e.g.* positions 14 and 15 of SEQ ID NO:14).

10 In some embodiments, the third nucleotide sequence comprises insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12 (*cf. e.g.* position 49 of SEQ ID NO:13, position 55 of SEQ ID NO:14).

15 In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 24 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the nucleotide at the position corresponding to position 20 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'T' at the position corresponding to position 27 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'T' at the position corresponding to position 28 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position
20 corresponding to position 21 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'A' at the position corresponding to position 29 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'T' at the position corresponding to position 21 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'A' at the position corresponding to position 31 of SEQ ID NO:12, and comprises 'A' at the position corresponding to position 34 of SEQ ID NO:12. In
25 some embodiments, the third nucleotide sequence comprises deletion of the positions corresponding to positions 9 to 20 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the positions corresponding to positions 33 to 41 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'CAC' at the positions corresponding to positions 34 to 36 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'G' after the position
30 corresponding to position 39 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'GC' at the positions corresponding to positions 47 to 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'CACCATG' after the position corresponding to position 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'C' at the position corresponding to position 48 of SEQ ID NO:12. In some embodiments, the
35 third nucleotide sequence comprises insertion of 'AA' after the position corresponding to position 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'CAT' at the positions corresponding to positions 49 to 51 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'GCC' at the positions corresponding to positions 46 to 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'CCA' at the positions corresponding to
40 positions 49 to 51 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'T' after the position corresponding to position 51 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'GC' at the positions corresponding to positions 39 to 40 of SEQ

ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'CA' after the position corresponding to position 40 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'ATG' at the positions corresponding to positions 43 to 45 of SEQ ID NO:12, and deletion of the position corresponding to position 46 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'GCCACCATG' after the position corresponding to position 9 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'AGCACCATG' after the position corresponding to position 15 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 16 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'CACCATG' after the position corresponding to position 15 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'ATG' after the position corresponding to position 21 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'GG' after the position corresponding to position 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'AT' after the position corresponding to position 30 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 31 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'G' after the position corresponding to position 33 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'A' after the position corresponding to position 43 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'GG' at the positions corresponding to positions 45 and 46 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'AT' after the position corresponding to position 52 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the positions corresponding to positions 49 and 50 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 51 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 40 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'A' at the position corresponding to position 35 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'A' after the position corresponding to position 39 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 41 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'C' at the position corresponding to position 43 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'ATG' after the position corresponding to position 42 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'TG' after the position corresponding to position 45 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'ATGGA' after the position corresponding to position 44 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'A' at the position corresponding to position 44 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 49 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the position corresponding to position 50 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'A' at the position corresponding to position 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding

to position 50 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises 'G' at the position corresponding to position 44 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the position corresponding to position 49 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the positions corresponding to positions 42 to 44 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'G' after the position corresponding to position 39 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises deletion of the position corresponding to position 48 of SEQ ID NO:12. In some embodiments, the third nucleotide sequence comprises insertion of 'GA' after the position corresponding to position 48 of SEQ ID NO:12.

10

In some embodiments, the third nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprises insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12. An example of such a third nucleotide sequence is SEQ ID NO:13.

15

In some embodiments, the third nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprises insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12; and (iv) comprises insertion of 'GCCACC' after the position corresponding to position 6 of SEQ ID NO:12; and (v) comprises 'TG' at the positions corresponding to positions 8 and 9 of SEQ ID NO:12. An example of such a third nucleotide sequence is SEQ ID NO:14.

20

25

In some embodiments, the third nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:26; and (ii) comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprises deletion of the position corresponding to position 20 of SEQ ID NO:12; and (iv) comprises insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12. An example of such a third nucleotide sequence is SEQ ID NO:27.

30

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:13, 14, 27, 223, 224 and 256 to 329.

35

40

In some embodiments, a polynucleotide according to the present disclosure does not comprise the nucleotide sequence of SEQ ID NO:13. In some embodiments, a polynucleotide according to the present

disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:14, 27, 223, 224 and 256 to 329.

5

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:13. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:14. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:27. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:223. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:224.

25 Fourth nucleotide sequence

The fourth nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence which is a variant of human *SMN2* intron 7.

Accordingly, in some embodiments, the fourth nucleotide sequence does not consist of, or does not comprise, SEQ ID NO:18. In some embodiments, the fourth nucleotide sequence comprises, or consists of, a nucleotide sequence having $<100\%$ sequence identity to SEQ ID NO:18.

In some embodiments, the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:380 at its 5' end. It will be appreciated that 'a nucleotide sequence according to SEQ ID NO:380' refers to a nucleotide sequence conforming to the consensus nucleotide sequence of SEQ ID NO:380. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:380 include SEQ ID NO:15 and SEQ ID NO:379.

In some embodiments, the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end. In some embodiments, the fourth nucleotide sequence

comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:379 at its 5' end. In some embodiments, the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end. In some embodiments, the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end, and comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end. In some embodiments, the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:379 at its 5' end, and comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end.

In some embodiments the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 or SEQ ID NO:379 in the region 5' to its midpoint. In some embodiments, the fourth nucleotide sequence comprises a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 in the region 3' to its midpoint.

In some embodiments, the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 or SEQ ID NO:379 accounts for at least 5%, e.g. one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region 5' to the midpoint of the fourth nucleotide sequence. In some embodiments, the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 accounts for at least 5%, e.g. one of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90% or 100% of the region 3' to the midpoint of the fourth nucleotide sequence.

In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 or SEQ ID NO:379 comprises one or more nucleotides provided within 25, e.g. within positions 20, 15, 10, or 5 nucleotides of the 5' nucleotide of the fourth nucleotide sequence. In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 or SEQ ID NO:379 comprises the 5' nucleotide of the fourth nucleotide sequence. In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 comprises one or more nucleotides provided within 25, e.g. within positions 20, 15, 10, or 5 nucleotides of the 3' nucleotide of the fourth nucleotide sequence. In some embodiments the nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 comprises the 3' nucleotide of the fourth nucleotide sequence.

An exemplary sequence comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 at its 5' end is SEQ ID NO:17. Positions 1 to 102 of SEQ ID NO:17 correspond to positions 1 to 102 of SEQ ID NO:15.

An exemplary sequence comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end is SEQ ID NO:18. Positions 295 to 444 of SEQ ID NO:18 correspond to positions 1 to 150 of SEQ ID NO:16.

- 5 In some embodiments, the fourth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16. In some embodiments, the fourth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID
- 10 NO:17. In some embodiments, the fourth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:18.

- In some embodiments, the fourth nucleotide sequence consists of fewer than 444 nucleotides, e.g. one of
- 15 <425 nucleotides, <400 nucleotides, <350 nucleotides, <300 nucleotides, <250 nucleotides or <200 nucleotides. In some embodiments, the fourth nucleotide sequence consists of fewer than 255 nucleotides. In some embodiments, the fourth nucleotide sequence consists of 252 nucleotides. In some embodiments, the fourth nucleotide sequence consists of 177 nucleotides.

- 20 In some embodiments, the fourth nucleotide sequence consists of greater than 444 nucleotides, e.g. one of >450 nucleotides, >475 nucleotides or >500 nucleotides. In some embodiments, the fourth nucleotide sequence consists of 508 nucleotides.

- In some embodiments, the fourth nucleotide sequence comprises 'C' at the position corresponding to
- 25 position 441 of SEQ ID NO:18. In some embodiments, the fourth nucleotide sequence comprises 'T' at the position corresponding to position 441 of SEQ ID NO:18. In some embodiments, the fourth nucleotide sequence comprises 'TT' at the positions corresponding to positions 436 and 437 of SEQ ID NO:18. In some embodiments, the fourth nucleotide sequence comprises 'TCCTC' at the positions corresponding to positions 11 to 15 of SEQ ID NO:18. In some embodiments, the fourth nucleotide sequence comprises
- 30 insertion of 'TTT' after the position corresponding to position 10 of SEQ ID NO:18. In some embodiments, the fourth nucleotide sequence comprises insertion of 'CCC' after the position corresponding to position 10 of SEQ ID NO:18.

- In some embodiments, the fourth nucleotide sequence comprises: (i) a nucleotide sequence having at
- 35 least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consists of fewer than 500 nucleotides.

Examples of such fourth nucleotide sequences include SEQ ID NOs:17 and 18.

40

In some embodiments, the fourth nucleotide sequence comprises: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or

100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consists of fewer than 255 nucleotides. An example of such a fourth nucleotide sequence is SEQ ID NO:17.

5

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fourth nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:18, 227, 330 to 370 and

10

381.

In some embodiments, a polynucleotide according to the present disclosure does not comprise the nucleotide sequence of SEQ ID NO:337. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a third nucleotide sequence) comprising, or

15

consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:18, 227, 330 to 336, 338 to 370 and 381.

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fourth nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:18. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fourth nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$,

20

25

$\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:227.

Fifth nucleotide sequence

The fifth nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence which is a variant of human *SMN2* exon 8.

30

Accordingly, in some embodiments, the fifth nucleotide sequence does not consist of, or does not comprise, SEQ ID NO:32. In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having $<100\%$ sequence identity to SEQ ID NO:32.

35

In some embodiments, the fifth nucleotide sequence consists of the dinucleotide 'GA', 'TG' or 'TT'. In some embodiments, the fifth nucleotide sequence consists of the dinucleotide 'GA'. In some embodiments, the fifth nucleotide sequence consists of the dinucleotide 'TG'. In some embodiments, the fifth nucleotide sequence consists of the dinucleotide 'TT'.

40

In some embodiments, the fifth nucleotide sequence consists of the trinucleotide 'GAG'.

In some embodiments, the fifth nucleotide sequence consists of the nucleotide 'A'.

In some embodiments, the fifth nucleotide sequence encodes a polypeptide of interest (e.g. as described hereinbelow). In some embodiments, the fifth nucleotide sequence encodes a polypeptide of interest, and comprises 'GA', 'TG' or 'TT' at positions 1 and 2. In some embodiments, the fifth nucleotide sequence
 5 encodes a polypeptide of interest, and comprises 'GA' at positions 1 and 2. In some embodiments, the fifth nucleotide sequence encodes a polypeptide of interest, and comprises 'TG' at positions 1 and 2. In some embodiments, the fifth nucleotide sequence encodes a polypeptide of interest, and comprises 'TT' at positions 1 and 2.

10 In some embodiments, the fifth nucleotide sequence encodes a polypeptide of interest, and comprises 'GAG' at positions 1 to 3.

In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$,
 15 $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:19. It will be appreciated that 'a nucleotide sequence according to SEQ ID NO:19' refers to a nucleotide sequence conforming to the consensus nucleotide sequence of SEQ ID NO:19. Exemplary sequences conforming to the consensus nucleotide sequence of SEQ ID NO:19 include SEQ ID NOs:20 and 21.

20 In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:20. In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:21.

25 In some embodiments, the fifth nucleotide sequence comprises 'GA' at the positions corresponding to positions 1 and 2 of SEQ ID NO:32 (cf. e.g. positions 1 and 2 of SEQ ID NO:20).

In some embodiments, the fifth nucleotide sequence comprises 'TG' at the positions corresponding to
 30 positions 1 and 2 of SEQ ID NO:32 (cf. e.g. positions 1 and 2 of SEQ ID NO:21).

In some embodiments, the fifth nucleotide sequence comprises fewer than 577 nucleotides. In some embodiments the fifth nucleotide sequence consists of <500 nucleotides, e.g. one of <400 nucleotides, <300 nucleotides, <200 nucleotides, <100 nucleotides or <50 nucleotides.

35 In some embodiments, the fifth nucleotide sequence comprises fewer than 25 nucleotides, e.g. one of <20 nucleotides, <15 nucleotides, <10 nucleotides or <5 nucleotides.

In some embodiments, the fifth nucleotide sequence comprises 23 nucleotides. In some embodiments,
 40 the fifth nucleotide sequence comprises 2 nucleotides.

In some embodiments, the fifth nucleotide sequence comprises 11 nucleotides. In some embodiments, the fifth nucleotide sequence comprises 8 nucleotides. In some embodiments, the fifth nucleotide sequence comprises 5 nucleotides.

- 5 In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 95\%$ or 100%) sequence identity to positions 1 to 23 of SEQ ID NO:32.

- 10 In some embodiments, the fifth nucleotide sequence comprises insertion of 'G' after the position corresponding to position 2 of SEQ ID NO:32.

- 15 In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. $\geq 90\%$ or 100%) sequence identity to positions 1 to 11 of SEQ ID NO:32. In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. 100%) sequence identity to positions 1 to 8 of SEQ ID NO:32. In some embodiments, the fifth nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 80% (e.g. 100%) sequence identity to positions 1 to 5 of SEQ ID NO:32. In some embodiments, the fifth nucleotide sequence comprises, or consists of, positions 1 and 2 of SEQ ID NO:32.

- 20 In some embodiments, the fifth nucleotide sequence comprises 'A' at the position corresponding to position 1 of SEQ ID NO:32. In some embodiments, the fifth nucleotide sequence comprises insertion of 'T' after the position corresponding to position 1 of SEQ ID NO:32. In some embodiments, the fifth nucleotide sequence comprises 'G' at the position corresponding to position 3 of SEQ ID NO:32.

- 25 In some embodiments, the fifth nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:19; and (ii) comprises 'GA' at the positions corresponding to positions 1 and 2 of SEQ ID NO:32. An example of such a fifth nucleotide sequence is SEQ ID NO:20.

- 30 In some embodiments, the fifth nucleotide sequence comprises, or consists of: (i) a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:19; and (ii) comprises 'TG' at the positions corresponding to positions 1 and 2 of SEQ ID NO:32. An example of such a fifth nucleotide sequence is SEQ ID NO:21.
- 35

- In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of: the dinucleotide 'TG', the dinucleotide 'GA', the dinucleotide 'TT', the trinucleotide 'GAG', the nucleotide 'A', or a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:20, 21 and 371 to 374.
- 40

In some embodiments, a polynucleotide according to the present disclosure does not comprise the nucleotide sequence of SEQ ID NO:20. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of: the dinucleotide 'TG', the dinucleotide 'GA', the dinucleotide 'TT', the trinucleotide 'GAG', the nucleotide 'A', or a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from SEQ ID NOs:21 and 371 to 374.

In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of the dinucleotide 'TG'. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of the dinucleotide 'GA'. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of the dinucleotide 'TT'. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of the trinucleotide 'GAG'. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:20. In some embodiments, a polynucleotide according to the present disclosure comprises a nucleotide sequence (e.g. a fifth nucleotide sequence) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:21.

Nucleotide sequences encoding a polypeptide of interest

In some embodiments, the fifth nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence encoding a polypeptide of interest. Where present, the sixth nucleotide sequence according to the present disclosure comprises, or consists of, a nucleotide sequence encoding a polypeptide of interest.

A polypeptide of interest may be any polypeptide.

In some embodiments, a polypeptide of interest according to the present disclosure may be an antigen-binding polypeptide, an aptamer, an antigen-binding polypeptide complex, an antibody or an antigen-binding fragment or derivative thereof, an Fc fusion protein, an anticoagulant, a blood factor, a bone morphogenetic protein, a decoy receptor for a ligand, a decoy ligand for a receptor, an enzyme, a growth factor, a hormone, an interferon, an interleukin, a thrombolytic, a transcription factor, an epigenetic modifier, a constituent protein of a site-specific nuclease nucleic acid editing system (e.g. a CRISPR/Cas9 system, a CRISPR/Cpf1 system, a CRISPR/C2c1 system, a CRISPR/C2c2 system, a CRISPR/C2c3 system, a ZFN system or a TALEN system), a constituent protein of a ribonucleoprotein, or a viral protein (e.g. a capsid protein or a viral enzyme).

In some embodiments, a polypeptide of interest according to the present disclosure may be an antigen-binding polypeptide or an antigen-binding polypeptide complex. In some embodiments, a polypeptide of interest may be a chimeric antigen-receptor (CAR).

5

In some embodiments, a polypeptide of interest is a polypeptide suitable for use in therapy or prophylaxis of a disease/condition. In some embodiments, a polypeptide of interest is a detectable polypeptide or a polypeptide having detectable activity.

10 A polypeptide suitable for use in therapy or prophylaxis of a disease/condition may be any polypeptide whose administration is useful for the treatment or prevention of a disease/condition. In some embodiments, a polypeptide suitable for use in therapy or prophylaxis of a disease/condition may be a polypeptide for which deficiency thereof is positively associated with, or implicated in the pathology of, a disease or condition. By way of illustration, in some embodiments, the polypeptide of interest may be

15 MeCP2. Deficiency of MeCP2 is associated with Rett syndrome.

In some embodiments, a polypeptide suitable for use in therapy or prophylaxis of a disease/condition may be a polypeptide which inhibits the expression and/or activity of a factor whose expression or activity is positively associated with, or implicated in the pathology of, a disease or condition.

20

A detectable polypeptide may be or comprise a fluorescent polypeptide. Fluorescent polypeptides include green fluorescent protein and variants thereof (e.g. enhanced green fluorescent protein), yellow fluorescent protein (e.g. citrine), red fluorescent protein and variants thereof (e.g. mOrange, mCherry), blue fluorescent protein and variants thereof (e.g. TagBFP), cyan fluorescent protein and variants thereof

25 (e.g. mTurquoise, cerulean), allophycocyanin, phycoerythrin and phycoerythrocyanin.

A detectable polypeptide may be or comprise an epitope tag. Epitope tags include e.g. His, (e.g. 6XHis), FLAG, c-Myc, StrepTag, haemagglutinin, E, calmodulin-binding protein (CBP), glutathione-s-transferase (GST), maltose-binding protein (MBP), thioredoxin, S-peptide, T7 peptide, SH2 domain, avidin,

30 streptavidin, and haptens (e.g. biotin, digoxigenin, dinitrophenol).

A polypeptide having detectable activity may be or comprise an enzymatic moiety. Enzymatic moieties include e.g. luciferases, glucose oxidases, galactosidases (e.g. beta-galactosidase), glucorinidases, phosphatases (e.g. alkaline phosphatase), peroxidases (e.g. horseradish peroxidase) and

35 cholinesterases.

A polypeptide of interest expressed from a polynucleotide of the present disclosure may additionally comprises one or more extraneous amino acids added at the N-terminus of the polypeptide, *i.e.* immediately upstream of the amino acid sequence of the polypeptide of interest. Such extraneous amino

40 acids may be characterised as forming an N-terminal tag on the polypeptide of interest. It may be desirable to minimise the size of, or completely remove, such extraneous amino acids/N-terminal tags on the polypeptide of interest.

In some embodiments, the N-terminal tag consists of fewer than 50 amino acids, e.g. one of ≤ 40 amino acids, ≤ 30 amino acids, ≤ 25 amino acids, ≤ 20 amino acids, ≤ 15 amino acids, ≤ 10 amino acids, ≤ 9 amino acids, ≤ 8 amino acids, ≤ 7 amino acids, ≤ 6 amino acids, ≤ 5 amino acids, ≤ 4 amino acids, ≤ 3 amino acids, ≤ 2 amino acids or 1 amino acid. In some embodiments, the polypeptide of interest lacks an N-terminal tag.

In some embodiments, the polypeptide of the present disclosure comprises one or more cleavage sites. A cleavage site refers to a sequence of amino acids that acts as a substrate for an enzyme capable of cleaving peptide bonds. Many such cleavage sites are known to, and can be employed by, the person skilled in the art of molecular biology. In some embodiments, the cleavage sequence comprises an autocleavage site. Autocleavage sites include the 2A cleavage sequence from Picornavirus 'NPGP', which is cleaved at 'G/P'. Further autocleavage sites are described e.g. in Kim *et al.*, PLoS ONE (2011) 6: e18556 (hereby incorporated by reference in its entirety), and include e.g. T2A, E2A, P2A and F2A cleavage sites. The amino acid sequences of T2A, E2A, P2A and F2A cleavage sites are shown in SEQ ID NOs: 107, 108, 109 and 110, respectively.

A cleavage site may be included in a polypeptide according to the present disclosure to provide for removal of extraneous amino acids added at the N-terminus of the polypeptide, i.e. immediately upstream of the amino acid sequence of the polypeptide of interest. That is, a cleavage site may be included for the removal of an N-terminal tag as described hereinabove. Accordingly, in some embodiments, a polypeptide according to the present disclosure comprises a cleavage site adjacent to (i.e. in the amino acid sequence of the polypeptide, e.g. immediately downstream of) extraneous amino acid(s), e.g. extraneous amino acid(s) forming an N-terminal tag.

In some embodiments, a cleavage site according to the present disclosure is a 2A cleavage site, e.g. selected from a T2A, E2A, P2A and F2A cleavage site. In some embodiments, the cleavage site is a T2A cleavage site. In some embodiments, a polypeptide according to the present disclosure comprises an amino acid sequence having at least 60%, preferably one of $\geq 70\%$, $\geq 75\%$, $\geq 80\%$, $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100% amino acid sequence identity to SEQ ID NO:107, 108, 109 or 110.

In some embodiments, a nucleotide sequence encoding a polypeptide of interest according to the present disclosure comprises, or consists of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:213 or 217.

Further nucleotide sequences of the polynucleotide

Polynucleotides according to the present disclosure may comprise additional nucleotide sequences and/or sequence features in addition to the first, second, third, fourth, fifth and/or sixth nucleotide sequences as described hereinabove.

The polynucleotides of the present disclosure comprise a start codon 5' to (*i.e.* upstream of, in the context of the nucleotide sequence of the polynucleotide) the nucleotide sequence encoding a polypeptide of interest. The start codon is preferably the trinucleotide 'ATG'.

5 In some embodiments, a start codon is provided in the polynucleotide such that following splicing (*e.g.* where the polynucleotide is a polyribonucleotide, or splicing of a polyribonucleotide transcribed from the polynucleotide where the polynucleotide is a polydeoxyribonucleotide), the start codon is provided in the mature mRNA molecule in such a way as to serve as the initiator codon for translation of the polypeptide of interest (encoded by the fifth or sixth nucleotide sequence).

10

In some embodiments, the splicing referred to in the preceding paragraph is splicing performed in the presence of a splicing modifier (*e.g.* as described herein). By way of illustration, splicing of the polyribonucleotide having the nucleotide sequence of SEQ ID NO:22 in the presence of RG7800/RG76196 results in the production of a mature mRNA molecule comprising an initiator codon for translation of a polypeptide of interest. In some embodiments, the splicing referred to in the preceding paragraph is splicing performed in the absence of a splicing modifier (*e.g.* as described herein). By way of illustration, splicing of the polyribonucleotide having the nucleotide sequence of SEQ ID NO:28 in the absence of a splicing modifier results in the production of a mature mRNA molecule comprising an initiator codon for translation of a polypeptide of interest.

15

20

In some embodiments, the polynucleotide further comprises a Kozak sequence. In preferred embodiments, the Kozak sequence is provided immediately upstream of the start codon for initiating translation of the polypeptide of interest. In some embodiments, the Kozak sequence consists of a nucleotide sequence conforming to the consensus of SEQ ID NO:33. In some embodiments, the Kozak sequence consists of SEQ ID NO:34.

25

In some embodiments, the polynucleotide comprises a nucleotide sequence consisting of SEQ ID NO:35 5' to the nucleotide sequence encoding a polypeptide of interest.

30 In some embodiments, SEQ ID NO:35 is provided in the polynucleotide such that following splicing (*e.g.* where the polynucleotide is a polyribonucleotide, or splicing of a polyribonucleotide transcribed from the polynucleotide where the polynucleotide is a polydeoxyribonucleotide), the start codon of SEQ ID NO:35 is provided in the mature mRNA molecule in such a way as to serve as the initiator codon for translation of the polypeptide of interest (encoded by the fifth or sixth nucleotide sequence).

35

In some embodiments, the splicing referred to in the preceding paragraph is splicing performed in the presence of a splicing modifier (*e.g.* as described herein). In some embodiments, the splicing referred to in the preceding paragraph is splicing performed in the absence of a splicing modifier (*e.g.* as described herein).

40

In some embodiments, the polynucleotide further comprises a promoter sequence. The promoter sequence is preferably 5' to the first nucleotide sequence. In some embodiments, the polynucleotide

further comprises one or more enhancer sequences. The one or more enhancer sequences are preferably 5' to the first nucleotide sequence.

5 In some embodiments, the polynucleotide further comprises a stop codon. The stop codon is preferably provided immediately 3' to (*i.e.* downstream of, in the context of the nucleotide sequence of the polynucleotide) the trinucleotide encoding the terminal amino acid of the polypeptide encoded by the fifth or sixth nucleotide sequence.

10 In some embodiments, the polynucleotide further comprises a polyadenylation signal sequence. In preferred embodiments, the polyadenylation signal sequence is provided 3' to (*i.e.* downstream of, in the context of the nucleotide sequence of the polynucleotide) the nucleotide sequence encoding a polypeptide of interest. In some embodiments, the polynucleotide further comprises a terminator sequence. The terminator sequence is preferably 3' to the nucleotide sequence encoding a polypeptide of interest (and 3' to the polyadenylation signal sequence, when present).

15 In preferred embodiments, the constituent nucleotide sequences of the polynucleotides (*i.e.* the first, second, third, fourth, fifth and/or sixth nucleotide sequences as described hereinabove) are provided immediately adjacent to one another. However, in some embodiments, the polynucleotide further comprises one or more linker nucleotide sequences between one or more of the constituent nucleotide sequences of the polynucleotide.

Linker nucleotide sequences may comprise, or consist of, 1-10, *e.g.* one of 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3, or 1-2 nucleotides.

25 Where a polynucleotide according to the present disclosure comprises one more linker nucleotide sequences (*i.e.* provided between the first, second, third, fourth, fifth and/or sixth nucleotide sequences), the linker sequences are preferably selected such that they do not substantially affect post-transcriptional processing of the polynucleotide when the polynucleotide is a polyribonucleotide. In preferred
30 embodiments, where a polynucleotide according to the present disclosure comprises one more linker sequences, splicing of the polynucleotide when the polynucleotide is a polyribonucleotide is substantially the same as splicing of an equivalent polyribonucleotide lacking the linker nucleotide sequence(s).

Similarly, where a polynucleotide according to the present disclosure comprises one more linker nucleotide sequences (*i.e.* provided between the first, second, third, fourth, fifth and/or sixth nucleotide sequences), the linker sequences are preferably selected such that they do not alter the amino acid
35 sequence of a polypeptide encoded by the polynucleotide. In preferred embodiments, where a polynucleotide according to the present disclosure comprises one more linker sequences, the polynucleotide encodes the same polypeptide as the equivalent polyribonucleotide lacking the linker nucleotide sequence(s).

40 In some embodiments, the polynucleotide further comprises inverted terminal repeat (ITR) sequences. In some embodiments, the polynucleotide comprises an ITR 5' to the first nucleotide sequence (and 5' to the

promoter and/or enhancer sequences, when present). In some embodiments, the polynucleotide comprises an ITR 3' to the nucleotide sequence encoding a polypeptide of interest (and 3' to the stop codon, polyadenylation signal sequence and/or terminator sequence, when present).

- 5 In some embodiments, the polynucleotide comprises an ITR sequence at its 5' end, and an ITR sequences at its 3' end. In some embodiments, the first nucleotide of the ITR sequence provided at the 5' end of the polynucleotide is provided within 1 to 25 nucleotides, e.g. within one of 1 to 20, 1 to 15, 1 to 10, or 1 to 5 nucleotides of the first nucleotide of the polynucleotide (*i.e.* position 1 of the nucleotide sequence of the polynucleotide). In some embodiments, the final nucleotide of the ITR sequence provided at the 3'
- 10 end of the polynucleotide is provided within 1 to 25 nucleotides, e.g. within one of 1 to 20, 1 to 15, 1 to 10, or 1 to 5 nucleotides of the final nucleotide of the polynucleotide (*i.e.* the terminal position of the nucleotide sequence of the polynucleotide).

- In some embodiments, the polynucleotide of the present disclosure (*i.e.* comprising the first, second,
- 15 third, fourth, fifth and/or sixth nucleotide sequences described herein) has a size permitting its delivery as a gene therapy, *i.e.* in a suitable vector. In some embodiments, the polynucleotide consists of a nucleotide sequence having a size within the packaging limit of a vector for delivering the polynucleotide.

- In some embodiments, the polynucleotide has a size within the packaging limit of an AAV vector. In some
- 20 embodiments, the polynucleotide has a size within the packaging limit of an AAV vector of one of the following serotypes: AAV9, AAV9.45, AAV-PHP.B, AAV1, AAV2, AAV2i8, AAV5, AAV6, AAV8, AAV10 or AAVrh74. In some embodiments, the polynucleotide has a size within the packaging limit of an AAV vector of one of the following serotypes: AAV9, AAV9.45, AAV-PHP.B, AAV1, AAV2, AAV2.7m8, AAV2i8, AAV5, AAV6, AAV8, AAV10 or AAVrh74.

- 25 In some embodiments, the polynucleotide of the present disclosure (*i.e.* comprising the first, second, third, fourth, fifth and/or sixth nucleotide sequences described herein) consists of fewer than 6,000 nucleotides, e.g. one of $\leq 5,000$, $\leq 4,750$, $\leq 4,500$, $\leq 4,250$, $\leq 4,000$, $\leq 3,750$, $\leq 3,500$, $\leq 3,250$, $\leq 3,000$, $\leq 2,750$, $\leq 2,500$, $\leq 2,250$, $\leq 2,000$, $\leq 1,750$, $\leq 1,500$, $\leq 1,250$ or $\leq 1,000$ nucleotides.

- 30 In preferred embodiments, the polynucleotide of the present disclosure consists of fewer nucleotides than the number of nucleotides of a known *SMN2* exon 6 to exon 8-derived transgene expression system (e.g. an *SMN2* exon 6 to exon 8-derived transgene expression system described in Zhang, *et al.*, Gene Ther. (2001) 8: 1532-1538, WO 2022/204471 A1, Monteys *et al.* Nature (2021) 596: 291-295 or WO
- 35 2021/163556 A1).

- In some embodiments, the polynucleotide of the present disclosure consists of fewer nucleotides than a polynucleotide comprising: (i) SEQ ID NO:36 and (ii) a nucleotide sequence encoding a polypeptide of interest (*i.e.* adjacent and 3' to SEQ ID NO:36, the context of the complete sequence of the
- 40 polynucleotide). Such polynucleotides are described e.g. in WO 2021/163556 A1.

Herein, for conciseness, 'a polynucleotide comprising: (i) SEQ ID NO:36 and (ii) a nucleotide sequence encoding a polypeptide of interest (*i.e.* adjacent and 3' to SEQ ID NO:36, in the context of the complete sequence of the polynucleotide)' is also referred to herein simply as 'a SMN2ind minigene polynucleotide'.

5

In some embodiments, the polynucleotide of the present disclosure consists of a number of nucleotides that is less than 1 times, *e.g.* ≤ 0.99 times, ≤ 0.95 times, ≤ 0.9 times, ≤ 0.85 times, ≤ 0.8 times, ≤ 0.75 times, ≤ 0.7 times, ≤ 0.65 times, ≤ 0.6 times, ≤ 0.55 times, ≤ 0.5 times, ≤ 0.45 times, ≤ 0.4 times, ≤ 0.35 times, ≤ 0.3 times or ≤ 0.25 times the number of nucleotides of a SMN2ind minigene polynucleotide.

10

In some embodiments, the region of a polynucleotide of the present disclosure formed by the first, second, third, fourth and fifth nucleotide sequences consists of fewer nucleotides than the number of nucleotides in SEQ ID NO:36. In some embodiments, the region of a polynucleotide of the present disclosure formed by the first, second, third, fourth and fifth nucleotide sequences consists of a number of nucleotides that is less than 1 times, *e.g.* ≤ 0.99 times, ≤ 0.95 times, ≤ 0.9 times, ≤ 0.85 times, ≤ 0.8 times, ≤ 0.75 times, ≤ 0.7 times, ≤ 0.65 times, ≤ 0.6 times, ≤ 0.55 times, ≤ 0.5 times, ≤ 0.45 times, ≤ 0.4 times, ≤ 0.35 times, ≤ 0.3 times or ≤ 0.25 times the number of nucleotides in SEQ ID NO:36.

15

The polynucleotide of the present disclosure may comprise 5' cap, 5' UTR, 3' UTR and/or PolyA tail nucleotide sequences.

20

In some embodiments, the polynucleotide comprises a 5' UTR 5' to (*i.e.* upstream of, in the context of the nucleotide sequence of the polynucleotide) a start codon.

25

In some embodiments, the polynucleotide comprises a 3' UTR 3' to (*i.e.* downstream of, in the context of the nucleotide sequence of the polynucleotide) a stop codon. In some embodiments, the polynucleotide comprises a 3' UTR 5' to a polyadenylation signal sequence. In some embodiments, the polynucleotide comprises a 3' UTR 3' to a stop codon and 5' to a polyadenylation signal sequence.

30

In some embodiments, the polynucleotide of the present disclosure comprises one or more nucleotide sequences encoding a selectable marker, to facilitate identification and/or selection of cells comprising/expressing the polynucleotide. Selectable markers include proteins that confer resistance to antibiotics or other toxins, *e.g.*, blasticidin, ampicillin, neomycin, methotrexate, or tetracycline, and proteins that complement auxotrophic deficiencies.

35

In some embodiments, the polynucleotide of the present disclosure comprises a nucleotide sequence encoding an internal ribosome entry site (IRES). In some embodiments, the polynucleotide comprises a nucleotide sequence permitting two or more polypeptides to be translated separately from a single polyribonucleotide.

40

The polynucleotides of the present disclosure may be provided in purified or isolated form, *i.e.* from other nucleic acid, or naturally-occurring biological material.

Particular exemplary polynucleotides

In some embodiments, a polynucleotide according to the present disclosure comprises:

- 5 a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from Column A of Table A;
- a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or
- 10 100%) sequence identity to a nucleotide sequence selected from Column B of Table A;
- a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from Column C of Table A;
- a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least
- 15 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from Column D of Table A; and
- a fifth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence selected from Column E of Table A.

20

In accordance with the preceding paragraph, in some embodiments the nucleotide sequence selected from Column A of Table A, and the nucleotide sequence selected from Column B of Table A, and the nucleotide sequence selected from Column C of Table A, and the nucleotide sequence selected from Column B of Table A, and the nucleotide sequence selected from Column D of Table A, and the

25 nucleotide sequence selected from Column B of Table A, and the nucleotide sequence selected from Column E of Table A are all selected from the same row of Table A. By way of illustration, in some embodiments the nucleotide sequences selected from Columns A, B, C, D and E of Table A may be selected from row 3 of Table 1, and therefore may be SEQ ID NOs:5, 9 14 18 and 20, respectively.

30

In some embodiments, the polynucleotide comprises:

- a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and comprising 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iii) consisting of ≤ 45 nucleotides;
- 35 a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position
- 40 5766 of SEQ ID NO:30; and (iv) consisting of fewer than 300 nucleotides;
- a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%)

sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprising 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprising insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12;

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 255 nucleotides; and a fifth nucleotide sequence consisting of the dinucleotide 'GA'.

10

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:106;

15

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:13;

20

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:17; and

a fifth nucleotide sequence consisting of the dinucleotide 'GA'.

25

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:22.

30

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:3;

35

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:13;

40

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:17; and

a fifth nucleotide sequence consisting of the dinucleotide 'GA'.

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:23.

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and comprising 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iii) comprising 'CTG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2;

a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position 5766 of SEQ ID NO:30; and (iv) consisting of fewer than 500 nucleotides;

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprising 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprising insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12; and (iv) comprising insertion of 'GCCACC' after the position corresponding to position 6 of SEQ ID NO:12; and (v) comprising 'TG' at the positions corresponding to positions 8 and 9 of SEQ ID NO:12;

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 500 nucleotides; and

a fifth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:20.

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:5;

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:9;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:14;

5 a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:18; and

a fifth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:20.

10

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:24.

15 In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and comprising 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iii) comprising 'CAG' at the positions corresponding to positions 109 to 111 of SEQ ID NO:2;

20

a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position 5766 of SEQ ID NO:30; and (iv) consisting of fewer than 500 nucleotides;

25

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprising 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprising insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12;

30

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 500 nucleotides; and

35

a fifth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:21.

40

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:6;

5 a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:9;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:13;

10 a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:18; and

a fifth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:20.

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:25.

20

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and comprising 'GGC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iii) consisting of ≤ 45 nucleotides;

25 In some embodiments, the polynucleotide comprises:

a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position 5766 of SEQ ID NO:30; and (iv) consisting of fewer than 300 nucleotides;

30

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprising 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprising insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12;

35

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 255 nucleotides; and

40

a fifth nucleotide sequence consisting of the trinucleotide 'GAG'.

In some embodiments, the polynucleotide comprises:

5 a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:3;

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

10 a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:13;

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:17; and

15 a fifth nucleotide sequence consisting of the trinucleotide 'GAG'.

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:28.

In some embodiments, the polynucleotide comprises:

25 a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:1; and comprising 'GCC' at the positions corresponding to positions 82 to 84 of SEQ ID NO:2; and (iii) consisting of ≤ 45 nucleotides;

In some embodiments, the polynucleotide comprises:

30 a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position 5766 of SEQ ID NO:30; and (iv) consisting of fewer than 300 nucleotides;

35 a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprising 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprising deletion of the position corresponding to position 20 of SEQ ID NO:12; and (iv) comprising insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12;

40 a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80%

(e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 255 nucleotides; and a fifth nucleotide sequence consisting of the trinucleotide 'GAG'.

5 In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:3;

10 a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:27;

15 a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:17; and

a fifth nucleotide sequence consisting of the trinucleotide 'GAG'.

20 In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:29.

In some embodiments, the polynucleotide does not consist of, or does not comprise, SEQ ID NO:36.

25

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to one of SEQ ID NOs:37 to 105.

30 In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:221; and (ii) consisting of ≤ 45 nucleotides;

35 a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position

40 5766 of SEQ ID NO:30; and (iv) consisting of fewer than 300 nucleotides;

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%)

sequence identity to a nucleotide sequence according to SEQ ID NO:11; and (ii) comprising 'A' at the position corresponding to position 2 of SEQ ID NO:12; and (iii) comprising insertion of 'A' after the position corresponding to position 48 of SEQ ID NO:12;

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 255 nucleotides; and a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

10

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:219;

15

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:13;

20

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:17; and

a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

25

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:111.

30

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:221; and (ii) consisting of ≤ 45 nucleotides;

35

a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position

40

5766 of SEQ ID NO:30; and (iv) consisting of fewer than 300 nucleotides;

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:225;

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:15 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:16 at its 3' end; and (iii) consisting of fewer than 255 nucleotides; and a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

10

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:219;

15

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:223;

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:17; and

a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

25

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:116.

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:221; and (ii) consisting of ≤ 45 nucleotides;

35

a second nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:7 at its 5' end; and (ii) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:8 at its 3' end; and (iii) comprising 'C' at the position corresponding to position

5766 of SEQ ID NO:30; and (iv) consisting of fewer than 300 nucleotides;

40

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:225;

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:227; and (iii) consisting of fewer than 280 nucleotides; and

a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:220;

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:10;

a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:224;

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:227; and

a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:164.

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:221; and (ii) consisting of ≤ 45 nucleotides;

a second nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:229;

a third nucleotide sequence (i) comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to a nucleotide sequence according to SEQ ID NO:225;

a fourth nucleotide sequence (i) comprising a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:227; and (iii) consisting of fewer than 280 nucleotides; and

a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

In some embodiments, the polynucleotide comprises:

a first nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%)

5 sequence identity to SEQ ID NO:220;

a second nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:228;

10 a third nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:224;

a fourth nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:227; and

15 a fifth nucleotide sequence consisting of the dinucleotide 'TG'.

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:171.

20

In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to one of SEQ ID NOs:111 to 200. In some embodiments, the polynucleotide comprises a nucleotide sequence having at least 80% (e.g. one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to one of SEQ ID NOs:37 to 105, or 111 to 200.

25

Splicing modifiers

30 Aspects and embodiments of the present disclosure pertain to splicing modifiers. Splicing modifiers are molecules that influence splicing of polyribonucleotides. Such molecules are reviewed e.g. in Tang *et al.*, *Molecules* (2021) 26(8):2263 and Schneider-Poetsch *et al.*, *The Journal of Antibiotics* (2021) 74:603-616, both of which are hereby incorporated by reference in their entirety.

35 Splicing modifiers generally reduce/prevent association of factors required for normal post-transcriptional processing (e.g. components of the spliceosome) with RNA. Splicing modifiers typically bind to a nucleotide sequence of a polyribonucleotide and either inhibit or promote the association of RNA-binding proteins and/or non-coding RNAs which function as splicing activators or repressors. Through competitive inhibition of the recruitment of splicing activators and splicing repressors to the polyribonucleotide, splicing modifiers can alter the equilibrium of the splicing-regulatory RNA structures, and thus promote an

40 increase or decrease in skipping or inclusion of an exon and/or retention or excision of an intron in the mature RNA molecule produced by splicing.

In some embodiments, a splicing modifier is a small molecule or a splice-switching nucleic acid (e.g. a splice-switching oligonucleotide).

5 A 'small molecule' refers to a low molecular weight (<1000 daltons, typically about 300 to about 700 daltons) organic compound.

Splice-switching nucleic acids are reviewed e.g. in Haves and Hastings, *Nucleic Acids Res.* (2016) 44(14): 6549–6563, which is hereby incorporated by reference in its entirety. Splice-switching nucleic acids include e.g. splice-switching oligonucleotides (SSOs). They disrupt the normal splicing of target
10 RNA transcripts by blocking the RNA:RNA base-pairing and/or protein:RNA binding interactions that occur between components of the splicing machinery and pre-mRNA. Splice-switching nucleic acids may be designed to target a specific region of the target transcript, e.g. to promote skipping of exon(s) of interest and/or to promote inclusion of exon(s) of interest and/or to promote retention of introns of interest and/or to promote excision of introns of interest. SSOs often comprise alterations to oligonucleotide
15 sugar-phosphate backbones in order to reduce/prevent RNaseH degradation, such as e.g. phosphorothioate linkages, phosphorodiamidate linkages such as phosphorodiamidate morpholino (PMOs), and may comprise e.g. peptide nucleic acids (PNAs), locked nucleic acids (LNAs), methoxyethyl nucleotide modifications, e.g. 2'-O-methyl (2'OMe) and 2'-O-methoxyethyl (MOE) ribose modifications and/or 5'-methylcytosine modifications.

20 Small molecule splicing modifiers contemplated in accordance with the present disclosure include RG-7800 and RG-7916 (also known as risdiplam), and analogs thereof. RG-7800 and RG-7916 bind to the 5' splice site of intron 7 and exonic splicing enhancer 2 of exon 7 of human *SMN2*, thereby stabilising the transient double-strand RNA structure formed by the *SMN2* pre-mRNA and U1 snRNP complex, and
25 promoting inclusion of exon 7 in the mature RNA molecule obtained following splicing of pre-mRNA transcribed from *SMN2*. RG-7800 and RG-7916 are useful to restore functional *SMN2* protein expression from *SMN2* alleles comprising the spinal muscular atrophy (SMA)-associated polymorphism c.840C>T (which potentiates skipping of exon 7). RG-7916 (DrugBank Acc. No. DB15305) is a close structural analog of RG-7800 having improved potency, pharmacokinetics and safety profile than RG-7800, and is
30 approved by the FDA for the treatment of SMA. Further analogs of RG-7800 and RG-7916 having similar splicing modifier activity include *SMN-C2*, *SMN-C3*, *SMN-C5* and *TEC-1*.

Branaplam (DrugBank Acc. No. DB14918; also known as LMI-070) is another small molecule splicing modifier that promotes inclusion of exon 7 in the mature RNA expressed from *SMN2*. Like RG-7800 and
35 RG-7916, branaplam binds to the 5' splice site of intron 7 of human *SMN2*.

PK4C9 (also known as homocarbonyltopsentin) is another small molecule splicing modifier that promotes inclusion of exon 7 in the mature RNA expressed from *SMN2*. It is thought to bind to the 5' splice site of
40 exon 7 and TSL2, and improve accessibility of the 5' splice site via stabilising a triloop structure of TSL2.

Small molecule splicing modifiers that promote inclusion of exon 7 in the mature RNA expressed from *SMN2* are described e.g. in WO 2015/173181 A1 and WO 2009/151546 A2, which are hereby

incorporated by reference in their entirety. Further small molecule splicing modifiers that promote inclusion of exon 7 in the mature RNA expressed from *SMN2* are described e.g. in WO 2022/204471 A1, which is hereby incorporated by reference in its entirety.

- 5 Nusinersen (DrugBank Acc. No. DB13161) is a splice-switching oligonucleotide that promotes exon 7 retention in mature RNA expressed from human *SMN2*. Nusinersen is an 18-mer 2'-MOe phosphorothioate antisense oligonucleotide that hybridises to intronic splicing silencer site 1 of intron 7, occupying the site and thereby inhibiting the association of the splicing suppressor ribonucleoproteins hnRNPs A1/A2, thus promoting inclusion of exon 7 in the mature RNA molecule.

10

- The splicing modifier according to the present disclosure preferably promotes inclusion of *SMN2* exon 7 in RNA obtained following splicing of pre-mRNA transcribed from human *SMN2*. Such molecules increase the proportion of RNA molecules comprising *SMN2* exon 7 among RNA molecules obtained following splicing of pre-mRNA transcribed from human *SMN2* (i.e. relative to the proportion obtained in the absence of the splicing modifier). For conciseness, 'a splicing modifier that promotes inclusion of *SMN2* exon 7 in RNA obtained following splicing of pre-mRNA transcribed from human *SMN2*' may be referred to herein simply as 'a splicing modifier that promotes *SMN2* exon 7 inclusion'

15

- In some embodiments, the splicing modifier promotes inclusion of exon 7 in RNA obtained following splicing of pre-mRNA transcribed from an allele of human *SMN2* comprising c.840C>T. That is, in some embodiments, the splicing modifier increases the proportion of RNA molecules comprising *SMN2* exon 7 among RNA molecules obtained following splicing of pre-mRNA transcribed from an allele of human *SMN2* comprising c.840C>T (i.e. relative to the proportion obtained in the absence of the splicing modifier).

20

- In some embodiments, the splicing modifier increases the level of human SMN2 protein comprising amino acids encoded by exon 7 of human *SMN2* (i.e. relative to the level detected in the absence of the splicing modifier). In some embodiments, the splicing modifier increases the proportion of polypeptides comprising amino acids encoded by exon 7 of human *SMN2* among polypeptides expressed from human *SMN2* (i.e. relative to the proportion obtained in the absence of the splicing modifier). In some embodiments, the splicing modifier increases the proportion of polypeptides comprising amino acids encoded by exon 7 of human *SMN2* among polypeptides expressed from an allele of human *SMN2* comprising c.840C>T (i.e. relative to the proportion obtained in the absence of the splicing modifier).

25

- In some embodiments, the splicing modifier promotes inclusion of the third nucleotide sequence in the product of splicing of a polynucleotide according to the present disclosure (i.e. where the polynucleotide is a polyribonucleotide). In some embodiments, the splicing modifier increases the proportion of molecules comprising the third nucleotide sequence among molecules obtained following splicing of a polynucleotide according to the present disclosure (i.e. where the polynucleotide is a polyribonucleotide).

30

- In some embodiments, the splicing modifier promotes inclusion of the nucleotide sequence consisting of SEQ ID NO:13 in molecules obtained following splicing of a polyribonucleotide consisting of the sequence

of SEQ ID NO:22. In some embodiments, the splicing modifier increases the proportion of molecules comprising the nucleotide sequence consisting of SEQ ID NO:13 among molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:22.

5 In some embodiments, the splicing modifier promotes inclusion of the nucleotide sequence consisting of SEQ ID NO:13 in molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:23. In some embodiments, the splicing modifier increases the proportion of molecules comprising the nucleotide sequence consisting of SEQ ID NO:13 among molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:23.

10

In some embodiments, the splicing modifier promotes inclusion of the nucleotide sequence consisting of SEQ ID NO:14 in molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:24. In some embodiments, the splicing modifier increases the proportion of molecules comprising the nucleotide sequence consisting of SEQ ID NO:14 among molecules obtained following

15 splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:24.

In some embodiments, the splicing modifier promotes inclusion of the nucleotide sequence consisting of SEQ ID NO:13 in molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:25. In some embodiments, the splicing modifier increases the proportion of molecules comprising the nucleotide sequence consisting of SEQ ID NO:13 among molecules obtained following

20

In some embodiments, the splicing modifier promotes inclusion of the nucleotide sequence consisting of SEQ ID NO:13 in molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:28. In some embodiments, the splicing modifier increases the proportion of molecules comprising the nucleotide sequence consisting of SEQ ID NO:13 among molecules obtained following

25

In some embodiments, the splicing modifier promotes inclusion of the nucleotide sequence consisting of SEQ ID NO:27 in molecules obtained following splicing of a polyribonucleotide consisting of the sequence of SEQ ID NO:29. In some embodiments, the splicing modifier increases the proportion of molecules comprising the nucleotide sequence consisting of SEQ ID NO:27 among molecules obtained following

30

35 Splicing modifiers having such functional properties can be identified by analysis *e.g.* in suitable *in vitro* assays. Such assays may comprise culturing cells *in vitro* in the presence or absence of a candidate splicing modifier, and analysing the RNA and/or protein produced after an appropriate period of time for an effect of the candidate splicing modifier to be observed. By way of illustration, cells can be transfected with a vector comprising DNA having the sequence of SEQ ID NO:22 and cultured in the presence or

40 absence of a candidate splicing modifier for a suitable period of time (*e.g.* 24, 48, 72 hours). RNA can subsequently be isolated from the cells, and analysed (*e.g.* by qRT-PCR) to determine the

level/proportion of mature RNA molecules comprising the nucleotide sequence consisting of SEQ ID NO:13.

In some embodiments, a splicing modifier that promotes *SMN2* exon 7 inclusion according to the present disclosure is a compound of formula (I) of WO 2015/173181 A1. In some embodiments, a splicing modifier that promotes *SMN2* exon 7 inclusion according to the present disclosure is a compound selected from those listed in claim 39 of WO 2015/173181 A1. In some embodiments, a splicing modifier that promotes *SMN2* exon 7 inclusion according to the present disclosure is a compound selected from those listed in claim 40 of WO 2015/173181 A1.

In some embodiments, the splicing modifier according to the present disclosure is selected from: RG-7916, RG-7800, SMN-C2, SMN-C3, SMN-C5, TEC-1, branaplam, PK4C9 and nusinersen. In some embodiments, the splicing modifier is selected from: RG-7916, RG-7800, SMN-C2, SMN-C3, SMN-C5 and TEC-1. In some embodiments, the splicing modifier is selected from RG-7916 and RG-7800. In preferred embodiments, the splicing modifier is risdiplam (RG-7916).

Functional properties of the polynucleotides

In aspects and embodiments of the present disclosure, the polynucleotides of the present disclosure may be characterised by reference to one or more functional properties.

In some embodiments, where the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence. In some embodiments, where the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides comprising the third nucleotide sequence.

In some embodiments, in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion, inclusion of the third nucleotide sequence is promoted in splicing of the polyribonucleotide, favouring the production of mature RNA molecules encoding the polypeptide of interest, and thus expression of the polypeptide of interest at the protein level. Such polynucleotides may be referred to herein as 'ON-switch' polynucleotides.

In some embodiments, cells comprising an ON-switch polynucleotide of the present disclosure substantially do not express the polypeptide of interest in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion. In some embodiments, cells comprising the ON-switch polynucleotide express the polypeptide of interest in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion.

Exemplary ON-switch polynucleotides according to the present disclosure include polynucleotides comprising SEQ ID NO:22, 23, 24 or 25. Further exemplary ON-switch polynucleotides according to the present disclosure include polynucleotides comprising one of SEQ ID NOs:37 to 105.

Inclusion of the third nucleotide sequence as described herein in the products of splicing of a polyribonucleotide according to the present disclosure can be evaluated using methods that are well known to the person skilled in the art. Such methods include qRT-PCR based methods for the detection and/or quantification of RNA molecules comprising nucleotide sequences of interest.

5

For example, cells may be transduced *in vitro* with a vector comprising a polynucleotide according to the present disclosure and subsequently cultured *in vitro* in the presence or absence of a splicing modifier that promotes *SMN2* exon 7 inclusion for a period of time appropriate for an effect on splicing of a polyribonucleotide according to the present disclosure to be observed. After such period of time, total
 10 RNA may be isolated from the cells, cDNA may be prepared from the total RNA, and the number/proportion of mature RNA molecules comprising the third nucleotide sequence may be evaluated by qPCR using oligonucleotides providing for the specific amplification and/or detection of the products of splicing of the polyribonucleotide (*e.g.* oligonucleotides hybridising to nucleotide sequences spanning exon:exon boundaries). The qPCR analysis may employ oligonucleotides enabling distinction between
 15 RNA molecules comprising the third nucleotide sequence, and RNA molecules lacking the third nucleotide sequence. Such qRT-PCR-based methods for analysing the products of splicing of a polyribonucleotide are described in the experimental examples of the present disclosure.

Expression of a polypeptide of interest can be evaluated using any suitable technique for the detection
 20 and/or quantification of the relevant polypeptide. Such techniques include *e.g.* antibody-based methods, (for example flow cytometry, immunocytochemistry, western blot, ELISA), fluorescence microscopy and flow cytometry. In some embodiments, expression of a polypeptide of interest can be evaluated as described in the experimental examples of the present disclosure. In preferred embodiments, expression of a polypeptide of interest may be evaluated by flow cytometry.

25

Herein, within a plurality/population of polyribonucleotides obtained after splicing that 'substantially lack' the third nucleotide sequence, the third nucleotide sequence may be present in fewer than 15%, *e.g.* one of $\leq 10\%$, $\leq 5\%$, $\leq 4\%$, $\leq 3\%$, $\leq 2\%$ or $\leq 1\%$ of the polyribonucleotides. Conversely, within a
 plurality/population of polyribonucleotides obtained after splicing that 'comprise' the third nucleotide
 30 sequence, the third nucleotide sequence may be present in more than 80%, *e.g.* one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100% of the polyribonucleotides.

Herein, within a plurality/population of cells that 'substantially do not express' a polypeptide of interest, the polypeptide may be expressed by fewer than 15%, *e.g.* one of $\leq 10\%$, $\leq 5\%$, $\leq 4\%$, $\leq 3\%$, $\leq 2\%$ or $\leq 1\%$ of the
 35 cells. Conversely, within a plurality/population of cells that 'express' a polypeptide of interest, the polypeptide may be expressed by more than 80%, *e.g.* one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100% of the cells.

In some embodiments, cells that 'substantially do not express' a polypeptide of interest may display a
 40 level of expression of the polypeptide of interest which is less than 0.2 times, *e.g.* one of ≤ 0.1 times, ≤ 0.09 times, ≤ 0.08 times, ≤ 0.07 times, ≤ 0.06 times, ≤ 0.05 times, ≤ 0.04 times, ≤ 0.03 times, ≤ 0.02 times, or ≤ 0.01 times the level of expression by cells that express the polypeptide of interest. In some

embodiments, cells that 'express' a polypeptide of interest may display a level of expression of the polypeptide of interest which is greater than 5 times, e.g. one of ≥ 10 times, ≥ 20 times, ≥ 50 times, ≥ 100 times, ≥ 1000 times, ≥ 5000 times or ≥ 10000 times level of expression by cells that 'substantially do not express' the polypeptide of interest.

5

In some embodiments, the level of RNA comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 in cells comprising a polynucleotide according to the present disclosure following culture in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 10 times, ≥ 20 times, ≥ 50 times, ≥ 100 times, ≥ 1000 times, ≥ 5000 times or ≥ 10000 times the level in equivalent cells cultured in the absence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

10

In some embodiments, the level of RNA comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 in cells comprising a polynucleotide according to the present disclosure following culture in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 100 times, e.g. one of ≤ 50 times, ≤ 20 times, ≤ 10 times, ≤ 5 times, ≤ 4 times, ≤ 3 times, ≤ 2 times or ≤ 1 times the level of RNA comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 in equivalent cells not comprising the polynucleotide.

15

20

In some embodiments, the level of a polypeptide of interest in cells comprising a polynucleotide according to the present disclosure (e.g. an ON-switch polynucleotide) in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 10 times, ≥ 20 times, ≥ 50 times, ≥ 100 times, ≥ 1000 times, ≥ 5000 times or ≥ 10000 times the level in equivalent cells cultured in the absence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

25

In some embodiments, the level of expression of a polypeptide of interest by cells comprising a polynucleotide according to the present disclosure (e.g. an ON-switch polynucleotide) in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 100 times, e.g. one of ≤ 50 times, ≤ 20 times, ≤ 10 times, ≤ 5 times, ≤ 4 times, ≤ 3 times, ≤ 2 times or ≤ 1 times the level of the polypeptide of interest in equivalent cells not comprising the polynucleotide.

30

In some embodiments, the proportion of cells expressing a polypeptide of interest within a population of cells comprising a polynucleotide according to the present disclosure (e.g. an ON-switch polynucleotide) cultured in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 10 times, ≥ 20 times, ≥ 50 times, ≥ 100 times, ≥ 1000 times, ≥ 5000 times or ≥ 10000 times the proportion of such cells expressing the polypeptide of interest cultured in the absence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

35

In some embodiments, the proportion of cells expressing a polypeptide of interest within a population of cells comprising a polynucleotide according to the present disclosure (e.g. an ON-switch polynucleotide) cultured in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 100

40

times, e.g. one of ≤ 50 times, ≤ 20 times, ≤ 10 times, ≤ 5 times, ≤ 4 times, ≤ 3 times, ≤ 2 times or ≤ 1 times the proportion of such cells expressing the polypeptide of interest cultured in the presence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

5 In some embodiments, in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion, inclusion of the third nucleotide sequence is promoted in splicing of the polyribonucleotide, favouring the production of mature RNA molecules encoding a premature stop codon, and thus preventing expression of the polypeptide of interest at the protein level. Such polynucleotides may be referred to herein as 'OFF-switch' polynucleotides.

10

In some embodiments, cells comprising an OFF-switch polynucleotide of the present disclosure express the polypeptide of interest in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion. In some embodiments, cells comprising the OFF-switch polynucleotide do not express the polypeptide of interest in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion.

15

Exemplary OFF-switch polynucleotides according to the present disclosure include polynucleotides comprising SEQ ID NO:28 or 29.

20 In some embodiments, the level of RNA comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:26 in cells comprising a polynucleotide according to the present disclosure following culture in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 10 times, ≥ 20 times, ≥ 50 times, ≥ 100 times, ≥ 1000 times, ≥ 5000 times or ≥ 10000 times the level in equivalent cells cultured in the absence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

25

In some embodiments, the level of RNA comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:26 in cells comprising a polynucleotide according to the present disclosure following culture in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 100 times, e.g. one of ≤ 50 times, ≤ 20 times, ≤ 10 times, ≤ 5 times, ≤ 4 times, ≤ 3 times, ≤ 2 times or ≤ 1 times the level of the polypeptide of interest in equivalent cells not comprising the polynucleotide.

30

35 In some embodiments, the level of a polypeptide of interest in cells comprising a polynucleotide according to the present disclosure (e.g. an OFF-switch polynucleotide) in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 10 times, ≥ 20 times, ≥ 50 times, ≥ 100 times, ≥ 1000 times, ≥ 5000 times or ≥ 10000 times the level in equivalent cells cultured in the presence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

35

40 In some embodiments, the level of expression of a polypeptide of interest by cells comprising a polynucleotide according to the present disclosure (e.g. an OFF-switch polynucleotide) in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 100 times, e.g. one of ≤ 50 times,

≤20 times, ≤10 times, ≤5 times, ≤4 times, ≤3 times, ≤2 times or ≤1 times the level of the polypeptide of interest in equivalent cells not comprising the polynucleotide.

In some embodiments, the proportion of cells expressing a polypeptide of interest within a population of cells comprising a polynucleotide according to the present disclosure (e.g. an OFF-switch polynucleotide) cultured in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥2 times, ≥3 times, ≥4 times, ≥5 times, ≥10 times, ≥20 times, ≥50 times, ≥100 times, ≥1000 times, ≥5000 times or ≥10000 times the proportion of such cells expressing the polypeptide of interest cultured in the presence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

In some embodiments, the proportion of cells expressing a polypeptide of interest within a population of cells comprising a polynucleotide according to the present disclosure (e.g. an OFF-switch polynucleotide) cultured in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 100 times, e.g. one of ≤50 times, ≤20 times, ≤10 times, ≤5 times, ≤4 times, ≤3 times, ≤2 times or ≤1 times the proportion of such cells expressing the polypeptide of interest cultured in the absence of the splicing modifier that promotes *SMN2* exon 7 inclusion.

The polynucleotides of the present disclosure possess novel and/or improved properties relative to known transgene expression systems, e.g. known transgene expression systems comprising a SMSM-mediated switch derived from *SMN2* exons 6 to 8 (e.g. *SMN2* exon 6 to exon 8-derived transgene expression systems described in Zhang, *et al.*, Gene Ther. (2001) 8: 1532-1538, WO 2022/204471 A1, Monteys *et al.* Nature (2021) 596: 291-295 or WO 2021/163556 A1).

In some embodiments, the polynucleotides of the present disclosure possess novel and/or improved properties relative to a *SMN2*ind minigene polynucleotide (which is defined hereinabove).

In some embodiments, cells comprising a polynucleotide of the present disclosure cultured in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion produce a population of RNA molecules having an increased proportion of RNA molecules comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26, compared to the population of RNA molecules produced by cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions. That is, in some embodiments the polynucleotide of the present disclosure is more effective at promoting inclusion of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26 in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to a *SMN2*ind minigene polynucleotide. That is, the polynucleotides of the present disclosure are more responsive to induction of *SMN2* exon 7 variant inclusion in the products of splicing in response to a splicing modifier that promotes *SMN2* exon 7 inclusion than *SMN2*ind minigene polynucleotides.

In some embodiments, the proportion of RNA molecules comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26 among RNA molecules obtained from cells comprising a polynucleotide of the present disclosure cultured in the

presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 1.01 times, ≥ 1.02 times, ≥ 1.03 times, ≥ 1.04 times, ≥ 1.05 times, ≥ 1.1 times, ≥ 1.2 times, ≥ 1.3 times, ≥ 1.4 times, ≥ 1.5 times, ≥ 1.6 times, ≥ 1.7 times, ≥ 1.8 times, ≥ 1.9 times, ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 6 times, ≥ 7 times, ≥ 8 times, ≥ 9 times or ≥ 10 times the proportion of RNA molecules comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26 among RNA molecules obtained from cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions.

It will be appreciated that the functional properties described herein are evaluated employing the same experimental conditions for the evaluation of cells comprising the different polynucleotides (*i.e.* the polynucleotide of the present disclosure, and the *SMN2*ind minigene polynucleotide). For example, the same cell type, the same splicing modifier that promotes *SMN2* exon 7 inclusion, the same concentration of the splicing modifier that promotes *SMN2* exon 7 inclusion, *etc.* are used, the same culture period is provided, and the cells are analysed in order to determine the proportion of RNA molecules comprising the relevant nucleotide sequence and/or the level of the polypeptide of interest in the same way.

In some embodiments, cells comprising a polynucleotide of the present disclosure cultured in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion produce a population of RNA molecules having a decreased proportion of RNA molecules comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26, compared to the population of RNA molecules produced by cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions.

That is, in some embodiments the polynucleotide of the present disclosure is more effective at excluding a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26 in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to a *SMN2*ind minigene polynucleotide. That is, the polynucleotides of the present disclosure are less 'leaky' with respect to *SMN2* exon 7 variant inclusion in the products of splicing in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to *SMN2*ind minigene polynucleotides. Leaky production of mature RNA molecules comprising an *SMN2* exon 7 variant (*i.e.* in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion) may also be referred to herein as 'background' production of such RNA molecules.

In some embodiments, the proportion of RNA molecules comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or 26 among RNA molecules obtained from cells comprising a polynucleotide of the present disclosure cultured in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 1 times, e.g. ≤ 0.99 times, ≤ 0.95 times, ≤ 0.9 times, ≤ 0.85 times, ≤ 0.8 times, ≤ 0.75 times, ≤ 0.7 times, ≤ 0.65 times, ≤ 0.6 times, ≤ 0.55 times, ≤ 0.5 times, ≤ 0.45 times, ≤ 0.4 times, ≤ 0.35 times, ≤ 0.3 times, ≤ 0.25 times, ≤ 0.2 times, ≤ 0.15 times, ≤ 0.1 times, ≤ 0.05 times, or ≤ 0.01 times the proportion of RNA molecules comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11 or

26 among RNA molecules obtained from cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions.

In some embodiments, cells comprising a polynucleotide of the present disclosure (e.g. an ON-switch polynucleotide) cultured in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion display an increased level of protein expression of the polypeptide of interest, compared to the level of protein expression of the polypeptide of interest by cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions. That is, in some embodiments the polynucleotide of the present disclosure is more effective at promoting protein expression of the polypeptide of interest in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to a *SMN2*ind minigene polynucleotide. That is, the polynucleotides of the present disclosure are more responsive to induction of expression of the polypeptide of interest in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to *SMN2*ind minigene polynucleotides.

It will be appreciated that for the purposes of such comparison, the different polynucleotides (i.e. the polynucleotide of the present disclosure, and the *SMN2*ind minigene polynucleotide) preferably encode the same polypeptide of interest.

In some embodiments, the level of protein expression of the polypeptide of interest from cells comprising a polynucleotide of the present disclosure (e.g. an ON-switch polynucleotide) cultured in the presence of a splicing modifier that promotes *SMN2* exon 7 inclusion is greater than 1 times, e.g. one of ≥ 1.01 times, ≥ 1.02 times, ≥ 1.03 times, ≥ 1.04 times, ≥ 1.05 times, ≥ 1.1 times, ≥ 1.2 times, ≥ 1.3 times, ≥ 1.4 times, ≥ 1.5 times, ≥ 1.6 times, ≥ 1.7 times, ≥ 1.8 times, ≥ 1.9 times, ≥ 2 times, ≥ 3 times, ≥ 4 times, ≥ 5 times, ≥ 6 times, ≥ 7 times, ≥ 8 times, ≥ 9 times or ≥ 10 times the level of protein expression of the polypeptide of interest from cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions.

In some embodiments, cells comprising a polynucleotide of the present disclosure (e.g. an ON-switch polynucleotide) cultured in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion display a decreased level of protein expression of the polypeptide of interest, compared to the level of protein expression of the polypeptide of interest by cells comprising a *SMN2*ind minigene polynucleotide and cultured under the same conditions. That is, in some embodiments the polynucleotide of the present disclosure is more effective at preventing protein expression of the polypeptide of interest in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to a *SMN2*ind minigene polynucleotide. That is, the polynucleotides of the present disclosure are less 'leaky' with respect to expression of the polypeptide of interest in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion, compared to *SMN2*ind minigene polynucleotides. Leaky expression of the polypeptide of interest (i.e. in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion) may also be referred to herein as 'background' expression of the polypeptide.

In some embodiments, the level of protein expression of the polypeptide of interest from cells comprising a polynucleotide of the present disclosure (e.g. an ON-switch polynucleotide) cultured in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion is less than 1 times, e.g. ≤ 0.99 times, ≤ 0.95 times,

≤0.9 times, ≤0.85 times, ≤0.8 times, ≤0.75 times, ≤0.7 times, ≤0.65 times, ≤0.6 times, ≤0.55 times, ≤0.5 times, ≤0.45 times, ≤0.4 times, ≤0.35 times, ≤0.3 times, ≤0.25 times, ≤0.2 times, ≤0.15 times, ≤0.1 times, ≤0.05 times, or ≤0.01 times the level of protein expression of the polypeptide of interest from cells comprising a SMN2ind minigene polynucleotide and cultured under the same conditions.

5

In some embodiments, a polypeptide of interest expressed from a polynucleotide of the present disclosure comprises fewer extraneous additional amino acids (*i.e.* amino acids additional to the amino acid sequence of the polypeptide of interest) compared to a polypeptide of interest expressed from SMN2ind minigene polynucleotide. In some embodiments, a polypeptide of interest expressed from a polynucleotide of the present disclosure comprises a smaller N-terminal tag (*i.e.* formed of extraneous, additional amino acids in addition to the amino acid sequence of the polypeptide of interest) than a polypeptide of interest expressed from SMN2ind minigene polynucleotide. That is, in some embodiments, the polynucleotides of the present disclosure provide for the inducible expression of polypeptides of interest comprising fewer extraneous additional amino acids, relative to polypeptides of interest expressed from SMN2ind minigene polynucleotide.

15

Vectors

The present disclosure provides vectors comprising the polynucleotides according to the present disclosure. It will be appreciated that a vector is also a polynucleotide, and so in some embodiments a polynucleotide according to the present disclosure may be a vector.

20

A 'vector' as used herein refers to a polynucleotide used as a vehicle to transfer exogenous nucleic acid into a cell. The vector may be a vector for expression of the nucleic acid in the cell (*i.e.* the vector may be an expression vector). Such vectors may include a promoter sequence operably linked to the nucleotide sequence to be expressed. Vectors may also include a termination codon and expression enhancers. Any suitable vectors, promoters, enhancers and termination codons known in the art may be used in a vector according to the present disclosure.

25

The term 'operably linked' may include the situation where nucleic acid encoding a polypeptide of interest according to the present disclosure and regulatory nucleic acid sequence(s) (*e.g.* a promoter and/or enhancers) are covalently linked in such a way as to place the expression of the nucleic acid encoding a polypeptide under the influence or control of the regulatory nucleic acid sequence(s) (thereby forming an expression cassette). Thus, a regulatory sequence is operably linked to the selected nucleic acid sequence if the regulatory sequence is capable of effecting transcription of the nucleic acid sequence. The resulting transcript(s) may then be translated into the desired polypeptide(s).

35

Vectors contemplated in connection with the present disclosure include DNA vectors, RNA vectors, plasmids (*e.g.* conjugative plasmids (*e.g.* F plasmids), non-conjugative plasmids, R plasmids, col plasmids, episomes), viral vectors (*e.g.* retroviral vectors, *e.g.* gammaretroviral vectors (*e.g.* murine Leukemia virus (MLV)-derived vectors, *e.g.* SFG vector), lentiviral vectors, adenovirus vectors, adeno-associated virus vectors, vaccinia virus vectors, baculoviral vectors and herpesvirus vectors), transposon-based vectors, and artificial chromosomes (*e.g.* yeast artificial chromosomes), *e.g.* as described in Maus

40

et al., Annu Rev Immunol (2014) 32:189-225 and Morgan and Boyerinas, Biomedicines (2016) 4:9, which are both hereby incorporated by reference in their entirety.

In some embodiments, the vector may be a eukaryotic vector, *i.e.* a vector comprising the elements
 5 necessary for expression of protein from the vector in a eukaryotic cell. In some embodiments, the vector may be a mammalian vector, *e.g.* comprising a cytomegalovirus (CMV) or SV40 promoter to drive protein expression. In some embodiments, the vector comprises a CMV (*e.g.* mCMV), SV40, RSV or PGK promoter.

10 In some embodiments, the polynucleotide according to the present disclosure (*e.g.* the vector according to the present disclosure) comprises a CMV promoter, a CAG promoter, a hEF1a promoter, a hUbiC promoter, an RSV promoter, a TK promoter, a PGK promoter, or a CAG minimal promoter.

15 In some embodiments, the polynucleotide according to the present disclosure (*e.g.* the vector according to the present disclosure) comprises a promoter having a nucleotide sequence comprising, or consisting of, a nucleotide sequence having at least 80% (*e.g.* one of $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100%) sequence identity to SEQ ID NO:201, 202, 203, 204, 205, 206, 207 or 208.

20 In some embodiments, a vector is selected based on tropism for a cell type/tissue/organ to which it is desired to deliver the polynucleotide according to the present disclosure. In some embodiments, a vector is selected based on tropism for a cell type/tissue/organ in which it is desired to express the polypeptide of interest. For example, it may be desired to deliver the polynucleotide to, and/or express the polypeptide of interest in, a cell type/tissue/organ affected by a disease/condition to be treated/prevented in
 25 accordance with the present disclosure (*e.g.* a cell type/tissue/organ in which the symptoms of the disease/condition manifest).

For example, it might be desirable to deliver a polynucleotide of the present disclosure encoding a polypeptide of interest to neuronal cells/tissue, and vectors having a tropism for such cells/tissue may be
 30 employed in such instances (*i.e.* neurotropic vectors).

In preferred embodiments, the vector is an adeno-associated virus (AAV) vector. Adeno-associated virus vectors and their use to vector gene therapy is reviewed *e.g.* in Wang *et al.*, Nat. Rev. Drug Discov. (2019) 18: 358-378 and Li and Samulski, Nat. Rev. Genet. (2020) 12: 255-272, both of which are hereby
 35 incorporated by reference in their entirety. In some embodiments, a vector may be an adeno-associated virus vector described in Wang *et al.*, Nat. Rev. Drug Discov. (2019) 18: 358-378. In some embodiments, a vector may be an adeno-associated virus vector described in Li and Samulski, Nat. Rev. Genet. (2020) 12: 255-272.

40 In some embodiments, the vector is a self-complementary adeno-associated virus (scAAV) vector. Self-complementary adeno-associated virus vectors are described *e.g.* in McCarty, Mol Ther. (2008) 16(10):1648-56, which is hereby incorporated by reference in its entirety. Conventional AAV have a

single-stranded DNA genome, and depend on the DNA replication machinery of a transduced cell to synthesise the complementary strand, delaying transgene expression. By contrast, scAAV contain complementary sequences that spontaneously anneal upon infection, eliminating the requirement for DNA synthesis in the transduced host cell. Compared to classical, single-stranded AAV vectors, scAAV
 5 vectors have been shown to provide for accelerated onset of transgene expression, and an increased level of transgene expression.

In some embodiments, a vector may be an adeno-associated viral vector of one of the following serotypes: AAV9 (including AAV9 variants AAV-PHP.B and AAV9.45), AAV1, AAV2 (including AAV2
 10 variant AAV2i8), AAV5, AAV6, AAV8, AAV10 or AAVrh74. In some embodiments, a vector may be an adeno-associated viral vector of one of the following serotypes: AAV9 (including AAV9 variants AAV-PHP.B and AAV9.45), AAV1, AAV2 (including AAV2 variants AAV2.7m8 and AAV2i8), AAV5, AAV6, AAV8, AAV10 or AAVrh74. In some embodiments, the vector is an AAV9 vector.

15 In some embodiments a vector comprises modification to increase binding to and/or transduction of a cell-type of interest (*i.e.* as compared to the level of binding/transduction by the unmodified vector). In some embodiments modification is to a capsid protein.

In some embodiments a vector comprises a capsid protein comprising a cell-targeting peptide. In some
 20 embodiments the cell-targeting peptide is a cell-targeting peptide described in Büning and Srivastava, Molecular Therapy: Methods & Clinical Development (2019) 12: 248-265, which is hereby incorporated by reference in its entirety, *e.g.* a cell-targeting peptide shown in Table 1, 2, 3 or 4 thereof.

In some embodiments a vector comprises a capsid protein comprising substitution to one or more
 25 tyrosine residues, *e.g.* one or more surface-exposed tyrosine residues. In some embodiments, one or more tyrosine residues of the capsid protein are substituted with phenylalanine. In some embodiments a vector comprises a capsid protein in which one or more tyrosine residues are substituted with another amino acid as described in Iida *et al.*, Biomed Res Int. (2013) 2013: 974819, which is hereby incorporated by reference in its entirety.

30 In some embodiments, a vector may be an adeno-associated virus vector described in Büning and Srivastava, *supra*. In some embodiments, a vector may be an adeno-associated virus vector described in Iida *et al.*, *supra*.

35 In some embodiments the vector comprises a control element for inducible expression of the polynucleotide of the disclosure.

A sequence for controlling expression of the polynucleotide may provide for expression of the polynucleotide by cells of a particular type or tissue. For example, expression may be under the control of
 40 a cell type- or tissue-specific promoter.

Promoters for cell type- or tissue-specific expression of a polynucleotide in accordance with the present disclosure can be selected in accordance with a disease/condition to be treated/prevented. For example, the promoter may drive expression in a cell type/tissue/an organ affected by the disease/condition (e.g. a cell type/tissue/an organ in which the symptoms of the disease/condition manifest).

5

In some embodiments, a promoter may provide for expression of the polynucleotide in neuronal cells/tissue. In some embodiments, a promoter may be a neuron-specific promoter (e.g. a CaMKII, NSE or SynI-miniCMV promoter). In some embodiments, a promoter may provide for expression of the polynucleotide in muscle cells/tissue (e.g. cardiac and/or skeletal muscle cells/tissue). In some
 10 embodiments, a promoter may be a cardiac or cardiomyocyte-specific promoter (e.g. a cTNT, α -MHC or MLC2v promoter). In some embodiments, a promoter may be a skeletal muscle/striated muscle cell-specific promoter (e.g. a MCK, MHCK7 or desmin promoter). In some embodiments, a promoter may be a vascular endothelial cell-specific promoter (e.g. a Tie2 promoter). In some embodiments, a promoter may be a vascular smooth muscle cell-specific promoter (e.g. a SM22a promoter). In some embodiments, a
 15 promoter may be a monocyte/macrophage-specific promoter (e.g. a LysM promoter).

A sequence for controlling expression of the polynucleotide may provide for expression of the polynucleotide in response to e.g. a given agent/signal. For example, expression may be under the control of inducible promoter. The agent may provide for inducible expression of the polynucleotide *in vivo*
 20 by administration of the agent to a subject having been administered with a modified cell according to the disclosure, or *ex vivo/in vitro* by administration of the agent to cells in culture *ex vivo* or *in vitro*.

In some embodiments a polynucleotide or vector according to the present disclosure may employ a conditional expression system for controlling expression of the polynucleotide by cells comprising the
 25 polynucleotide/vector. 'Conditional expression' may also be referred to herein as 'inducible expression', and refers to expression contingent on certain conditions, e.g. the presence of a particular agent. Conditional expression systems are well known in the art and are reviewed e.g. in Ryding *et al.* Journal of Endocrinology (2001) 171, 1-14, which is hereby incorporated by reference in its entirety.

30 **Cells**

The present disclosure also provides a cell comprising or expressing a polynucleotide according to the present disclosure. Also provided is a cell comprising or expressing vector according to the present disclosure.

35 A polynucleotide according to the present disclosure (e.g. a polyribonucleotide) may be produced within a cell by transcription from a polynucleotide (e.g. a polydeoxyribonucleotide) encoding the polynucleotide.

The cell may be a eukaryotic cell, e.g. a mammalian cell. The mammal may be a primate (rhesus, cynomolgous, non-human primate or human) or a non-human mammal (e.g. rabbit, guinea pig, rat,
 40 mouse or other rodent (including any animal in the order *Rodentia*), cat, dog, pig, sheep, goat, cattle (including cows, e.g. dairy cows, or any animal in the order *Bos*), horse (including any animal in the order *Equidae*), donkey, and non-human primate). In preferred embodiments, the cell may be a human cell.

The cell may be an immune cell. The cell may be a cell of hematopoietic origin, e.g. a neutrophil, eosinophil, basophil, dendritic cell, lymphocyte, or monocyte. The lymphocyte may be e.g. a T cell, B cell, NK cell, NKT cell or innate lymphoid cell (ILC), or a precursor thereof. The cell may express e.g. CD3 polypeptides (e.g. CD3 γ CD3 ϵ CD3 ζ or CD3 δ), TCR polypeptides (TCR α or TCR β), CD27, CD28, CD4 or CD8. In some embodiments, the cell is a T cell. In some embodiments, the T cell is a CD3+ T cell. In some embodiments, the T cell is a CD3+, CD8+ T cell. In some embodiments, the T cell is a cytotoxic T cell (e.g. a cytotoxic T lymphocyte (CTL)).

10 The present disclosure also provides a method for producing a cell comprising or expressing a polynucleotide/vector according to the present disclosure, the method comprising introducing a polynucleotide/vector of the present disclosure into a cell. In some embodiments, introducing a polynucleotide/vector according to the present disclosure into a cell comprises transformation, transfection, electroporation or transduction (e.g. adeno-associated viral transduction). In some
15 embodiments, the polynucleotide/vector is introduced to the cell *in vivo*, e.g. by administration of a vector according to the present disclosure (e.g. a viral vector, e.g. an adeno-associated viral vector) to a subject. In some embodiments, the polynucleotide/vector is introduced into cells in culture *ex vivo* or *in vitro*.

Any suitable method may be employed to produce a cell according to the present disclosure. Such
20 methods may comprise nucleic acid transfer for permanent (*i.e.* stable) or transient expression of the polynucleotide of the present disclosure. In some embodiments, following introduction into a cell, the polynucleotide may be integrated into or form part of the genomic DNA of the cell. In some embodiments, following introduction into a cell, the polynucleotide may be maintained extrachromosomally.

25 Any suitable genetic engineering platform may be used, and include gammaretroviral vectors, lentiviral vectors, adenovirus vectors, DNA transfection, transposon-based gene delivery and RNA transfection, for example as described in Maus *et al.*, Annu Rev Immunol. (2014) 32:189-225, hereby incorporated by reference in its entirety. Methods also include those described e.g. in Wang and Rivière Mol Ther Oncolytics. (2016) 3:16015, which is hereby incorporated by reference in its entirety. Suitable methods for
30 introducing nucleic acid(s)/vector(s) into cells include transduction, transfection and electroporation.

In some embodiments, the methods additionally comprise maintaining the cell under conditions suitable for expression of the polynucleotide/vector by the cell.

35 The present disclosure also provides cells obtained or obtainable by the methods according to the present disclosure.

Compositions

40 The present disclosure also provides compositions comprising the polynucleotides, vectors and cells described herein. In particular, the present disclosure provides pharmaceutical compositions and medicaments comprising the polynucleotides, vectors and cells of the present disclosure.

Such compositions may comprise the relevant article (*i.e.* the polynucleotide/vector/cell) in a formulation suitable for clinical use. The present disclosure is concerned in particular with pharmaceutical compositions/medicaments comprising polynucleotides and vectors according to the present disclosure.

- 5 The compositions of the present disclosure may comprise one or more pharmaceutically-acceptable carriers (*e.g.* liposomes, micelles, microspheres, nanoparticles), diluents/excipients (*e.g.* starch, cellulose, a cellulose derivative, a polyol, dextrose, maltodextrin, magnesium stearate), adjuvants, fillers, buffers, preservatives (*e.g.* vitamin A, vitamin E, vitamin C, retinyl palmitate, selenium, cysteine, methionine, citric acid, sodium citrate, methyl paraben, propyl paraben), anti-oxidants (*e.g.* vitamin A, vitamin E, vitamin C, retinyl palmitate, selenium), lubricants (*e.g.* magnesium stearate, talc, silica, stearic acid, vegetable stearin), binders (*e.g.* sucrose, lactose, starch, cellulose, gelatin, polyethylene glycol (PEG), polyvinylpyrrolidone (PVP), xylitol, sorbitol, mannitol), stabilisers, solubilisers, surfactants (*e.g.*, wetting agents), masking agents or colouring agents (*e.g.* titanium oxide).
- 15 The term 'pharmaceutically-acceptable' as used herein pertains to compounds, ingredients, materials, compositions, dosage forms, *etc.*, which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of the subject in question (*e.g.* a human subject) without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio. Each carrier, diluent, excipient, adjuvant, filler, buffer, preservative, anti-oxidant, lubricant, binder, stabiliser, solubiliser, surfactant, masking agent, colouring agent, flavouring agent or sweetening agent of a composition according to the present disclosure must also be 'acceptable' in the sense of being compatible with the other ingredients of the formulation. Suitable carriers, diluents, excipients, adjuvants, fillers, buffers, preservatives, anti-oxidants, lubricants, binders, stabilisers, solubilisers, surfactants, masking agents, colouring agents, flavouring agents or sweetening agents can
- 25 be found in standard pharmaceutical texts, for example, Remington's 'The Science and Practice of Pharmacy' (Ed.(A)Adejare), 23rd Edition (2020), Academic Press.

- The pharmaceutical compositions/medicaments according to the present disclosure may be formulated for administration to a subject, *e.g.* administration via a route of administration as appropriate for the nature of the therapeutic agent and the disease to be treated/prevented. In some embodiments, a pharmaceutical composition/medicament may be formulated for parenteral, systemic, topical, intracavitary, intravascular, intravenous, intra-arterial, intramuscular, intrathecal, intraocular, intraconjunctival, intratumoral, subcutaneous, intradermal,, oral or transdermal administration. In some embodiments, a pharmaceutical composition/medicament may be formulated for administration by
- 35 injection or infusion, or administration by ingestion.

- Medicaments and pharmaceutical compositions may be formulated for administration to a blood vessel, or to a tissue/organ of interest (*e.g.* a tissue/organ affected by a disease/condition, *e.g.* a tissue/organ in which symptoms of the disease/condition manifest).

- 40 The pharmaceutical compositions/medicaments may comprise the polynucleotide/vector/cell in a sterile or isotonic medium. The pharmaceutical compositions/medicaments may be provided in fluid, including gel,

form. Fluid formulations may be formulated for administration by injection or infusion (e.g. via cannula) to a blood vessel, or a selected region of the human or animal body. The pharmaceutical compositions/medicaments may be provided in solid form, e.g. in lyophilised form.

5 The present disclosure also provides methods for producing pharmaceutical compositions/medicaments according to the present disclosure. Such methods may comprise mixing a polynucleotide/vector/cell described herein with a pharmaceutically-acceptable carrier, diluent, excipient, adjuvant, filler, buffer, preservative, anti-oxidant, lubricant, binder, stabiliser, solubiliser, surfactant, masking agent, colouring agent, flavouring agent or sweetening agent. Such methods generally include the step of bringing into
10 association the polynucleotide/vector/cell with a carrier which constitutes one or more accessory ingredients. In general, the compositions are prepared by uniformly and intimately bringing into association the active compound with carriers (e.g., liquid carriers, finely divided solid carrier, etc.), and then shaping the product, if necessary.

15 Polynucleotides, vectors, cells and compositions according to the present disclosure may be modified and/or formulated to facilitate delivery to, and/or uptake by, a cell type/tissue/organ of interest (e.g. a cell type/tissue/organ in which symptoms of a disease/condition manifest).

Strategies for targeted delivery of polynucleotides are reviewed e.g. in Li *et al.*, Int. J. Mol. Sci. (2015) 16: 19518-19536 and Fu *et al.*, Bioconjug Chem. (2014) 25(9): 1602-1608, which are hereby incorporated by
20 reference in their entirety.

In some embodiments, articles of the present disclosure may be encapsulated in a nanoparticle or a liposome. In some embodiments, articles of the present disclosure may be (covalently or non-covalently)
25 associated with a cell-penetrating peptide (e.g. a protein transduction domain, trojan peptide, arginine-rich peptide, vectocell peptide), a cationic polymer, a cationic lipid or a viral carrier.

Nanoparticles may be organic, e.g. micelles, liposomes, proteins, solid-lipid particles, solid polymer particles, dendrimers, and polymer therapeutics. Nanoparticles may be inorganic, e.g. such as nanotubes
30 or metal particles, optionally with organic molecules added. In some embodiments, a nanoparticle is a nanoparticle described in Chen *et al.*, Mol Ther Methods Clin Dev. (2016) 3:16023, which is hereby incorporated by reference in its entirety. In some embodiments, a nanoparticle is a PLGA, polypeptide, poly(β -amino ester), DOPE, β -cyclodextrin-containing polycation, linear PEI, PAMAM dendrimer, branched PEI, chitosan or polyphosphoester nanoparticle.

35 In some embodiments, polynucleotides and vectors according to the present disclosure comprise modification to incorporate one or more moieties facilitating delivery to, and/or uptake by, a cell type, organ or tissue of interest (e.g. a cell type/tissue/organ in which symptoms of a disease/condition manifest). In some embodiments, polynucleotides or vectors according to the present disclosure are
40 linked (e.g. chemically conjugated to) one or more moieties facilitating delivery to, and/or uptake by, a cell type, tissue or organ of interest.

Moieties facilitating delivery to, and/or uptake by, cell types, tissues or organs of interest are described e.g. in Benizri *et al.*, Bioconjug Chem. (2019) 30(2): 366–383, which is hereby incorporated by reference in its entirety. Such moieties include e.g. N-acetylgalactosamine (GalNAc), α -tocopherol, cell-penetrating peptides, nucleic acid aptamers, antibodies and antigen-binding fragments/derivatives thereof,
5 cholesterol, squalene, polyethylene glycol (PEG), fatty acids (e.g. palmitic acid) and nucleolipid moieties.

Articles of the present disclosure may be formulated in a sustained release delivery system, in order to release the polynucleotide, vector, cell or composition at a predetermined rate. Sustained release delivery systems may maintain a constant drug/therapeutic/prophylactic concentration for a specified period of
10 time. In some embodiments, articles of the present disclosure are formulated in a liposome, gel, implant, device, or drug-polymer conjugate e.g. hydrogel.

In some embodiments, a composition according to the present disclosure may further comprise a splicing modifier that promotes *SMN2* exon 7 inclusion (e.g. as described herein).

15

Therapeutic/prophylactic applications

The polynucleotides, vectors, cells and compositions of the present disclosure find use in therapy and prophylaxis.

20 Accordingly, the present disclosure provides a polynucleotide, vector, cell or composition described herein for use in a method of medical treatment or prophylaxis. Also provided is a polynucleotide, vector, cell or composition described herein for use in a method of treating or preventing a disease/condition described herein. Also provided is the use of a polynucleotide, vector, cell or composition described
25 herein in the manufacture of a medicament for treating or preventing a disease or condition described herein. Also provided is a method of treating or preventing a disease or condition described herein, comprising administering to a subject a therapeutically- or prophylactically- effective amount of a polynucleotide, vector, cell or composition described herein.

The intervention described in the preceding paragraph may be effective to reduce the development or
30 progression of a disease/condition, alleviate the symptoms of a disease/condition or reduce the pathology of a disease/condition. The intervention may be effective to prevent progression of the disease/condition, e.g. to prevent worsening of, or to slow the rate of development of, the disease/condition. In some embodiments, the intervention may lead to an improvement in the disease/condition, e.g. a reduction in the symptoms of the disease/condition or reduction in some other correlate of the severity/activity of the
35 disease/condition. In some embodiments, the intervention may prevent progression/development of the disease/condition a later stage (e.g. a chronic stage).

It will be appreciated that the polynucleotides, vectors, cells and compositions described herein may be used for the treatment/prevention of any disease/condition that would derive therapeutic or prophylactic
40 benefit from an increase in the level of the polypeptide of interest (i.e. the polypeptide of interest encoded by the polynucleotide).

For example, the disease/condition may be a disease/condition associated with and/or characterised by deficiency/insufficiency of the polypeptide of interest. Deficiency/insufficiency of the polypeptide of interest may be positively associated with the onset, development or progression of the disease/condition, and/or positively associated with the severity of one or more symptoms of the disease/condition.

- 5 Deficiency/insufficiency of the polypeptide of interest may be a risk factor for the onset, development or progression of the disease/condition. The disease/condition may be characterised by a decreased level of expression or activity of the polypeptide of interest, *e.g.* as compared to the level of expression/activity in the absence of the disease/condition. In some embodiments, the disease/condition may be characterised by a decrease in the number/proportion/activity of cells expressing the polypeptide of interest, *e.g.* as compared to the level/number/proportion/activity in the absence of the disease/condition (*e.g.* in a healthy subject, or in equivalent non-diseased tissue).

- 15 By way of illustration, in some embodiments, the polypeptide of interest may be MeCP2, and the disease/condition to be treated/prevented in accordance with the present disclosure may be a disease/condition caused by deficiency/insufficiency of MeCP2, *e.g.* Rett syndrome.

- By way of further example, in embodiments wherein the polypeptide of interest is a polypeptide capable of inhibiting the expression and/or activity of a target antigen of interest, the disease/condition may be a disease/condition in which the target antigen, or cells comprising/expressing the target antigen are pathologically-implicated, *e.g.* a disease/condition in which an increased level/activity of the target antigen, or an increase in the number/proportion/activity of cells comprising/expressing the target antigen is positively associated with the onset, development or progression of the disease/condition, and/or severity of one or more symptoms of the disease/condition. In some embodiments, an increased level/activity of the target antigen, or an increase in the number/proportion/activity of cells comprising/expressing the target antigen may be a risk factor for the onset, development or progression of the disease/condition. The disease/condition may be characterised by an increase in the level of expression or activity of the target antigen, *e.g.* as compared to the level of expression/activity in the absence of the disease/condition. In some embodiments, the disease/condition may be characterised by an increase in the number/proportion/activity of cells expressing the target antigen, *e.g.* as compared to the level/number/proportion/activity in the absence of the disease/condition (*e.g.* in a healthy subject, or in equivalent non-diseased tissue). Therapeutic/prophylactic intervention in accordance with the present disclosure may achieve one or more of the following in a subject (compared to an equivalent untreated subject, or subject treated with an appropriate control): a reduction in the level of the target antigen; a reduction in the activity of the target antigen; and/or a reduction in the number/proportion/activity of cells comprising/expressing the target antigen.

- 40 By way of further example, the disease/condition may be a disease/condition to be treated by nucleic acid editing, and the polypeptide of interest may be a constituent protein of an appropriate site-specific nuclease nucleic acid editing system.

The present disclosure provides the articles of the present disclosure for use, uses of articles of the present disclosure, and methods comprising administering polynucleotides, vectors, cells and

compositions according to the present disclosure to a subject (e.g. a subject in need of treatment). In some embodiments, the methods comprise administering to the subject a splicing modifier that promotes *SMN2* exon 7 inclusion (e.g. a splicing modifier that promotes *SMN2* exon 7 inclusion as described herein).

5

Administration of the articles of the present disclosure is preferably in a 'therapeutically-effective' or 'prophylactically-effective' amount, this being sufficient to show therapeutic or prophylactic benefit to the subject. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of the disease/condition and the particular article administered. Prescription of treatment, e.g. decisions on dosage *etc.*, is within the responsibility of general practitioners and other medical doctors, and typically takes account of the disease/disorder to be treated, the condition of the individual subject, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned above can be found in Remington's 'The Science and Practice of Pharmacy' (ed.(A)Adejare), 23rd Edition (2020), Academic Press.

15

Administration of the articles of the present disclosure may be parenteral, systemic, intravenous, intra-arterial, intramuscular, intracavitary, intrathecal, intraocular, intravitreal, intraconjunctival, subretinal, suprachoroidal, subcutaneous, intradermal, intrathecal, oral, nasal, topical or transdermal. Administration may be by injection or infusion. Administration of the articles of the present disclosure may be intratumoral. In some cases, the articles of the present disclosure may be formulated for targeted delivery to specific cells, a tissue, an organ and/or a tumor.

20

Multiple doses of an article of the present disclosure may be provided. Multiple doses may be separated by a predetermined time interval, which may be selected to be one of 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, or 1, 2, 3, 4, 5, or 6 months.

25

In some embodiments, the present disclosure provides the articles of the present disclosure for use, uses of articles of the present disclosure, and methods comprising administering: (i) a polynucleotide, vector, cell or composition according to the present disclosure, and (ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, to a subject (e.g. a subject in need of treatment).

30

In embodiments in accordance with aspects of the preceding paragraph, provision of (i) and (ii) may be as a combination therapy. In some embodiments, (i) and (ii) may be provided simultaneously or sequentially.

35

Simultaneous administration refers to administration of the two or more agents (e.g. a polynucleotide, vector, cell or composition according to the present disclosure, and a splicing modifier that promotes *SMN2* exon 7 inclusion) together, for example as a pharmaceutical composition containing both agents (*i.e.* as a combined preparation), or immediately after each other (e.g. within 1, 4, 6, 8 or 12 hours), and optionally via the same route of administration, e.g. to the same artery, vein or other blood vessel.

40

Sequential administration refers to administration of one of the agents followed after a given time interval

by separate administration of another agent. It is not required that the agents are administered by the same route, although this is the case in some embodiments. The time interval may be any time interval.

In some embodiments, a splicing modifier that promotes *SMN2* exon 7 inclusion is administered to a subject after administration of a polynucleotide, vector, cell or composition according to the present disclosure. In some embodiments, a polynucleotide, vector, cell or composition according to the present disclosure is administered to a subject, and a splicing modifier that promotes *SMN2* exon 7 inclusion is administered continuously to the subject thereafter.

In some embodiments, the splicing modifier that promotes *SMN2* exon 7 inclusion is administered in a quantity, and/or with a periodicity, selected to achieve a desired level of expression of the polypeptide of interest. That is, in some embodiments the quantity of, and/or periodicity with which, a splicing modifier that promotes *SMN2* exon 7 inclusion administered to a subject in accordance with the present disclosure is selected to achieve a desired level of expression of the polypeptide of interest. In some embodiments, the quantity of, and/or periodicity with which, a splicing modifier that promotes *SMN2* exon 7 inclusion is administered to a subject is adjusted through the course of its administration to change (*i.e.* decrease or increase) the level of expression of the polypeptide of interest in the subject.

In embodiments wherein the polynucleotide of the disclosure is an ON-switch polynucleotide, the quantity and/or periodicity of administration of a splicing modifier that promotes *SMN2* exon 7 inclusion may be increased to increase the level of expression of the polypeptide of interest. Conversely, the quantity and/or periodicity of administration of a splicing modifier that promotes *SMN2* exon 7 inclusion may be decreased to decrease the level of expression of the polypeptide of interest.

In embodiments wherein the polynucleotide of the disclosure is an OFF-switch polynucleotide, the quantity and/or periodicity of administration of a splicing modifier that promotes *SMN2* exon 7 inclusion may be decreased to increase the level of expression of the polypeptide of interest. Conversely, the quantity and/or periodicity of administration of a splicing modifier that promotes *SMN2* exon 7 inclusion may be increased to decrease the level of expression of the polypeptide of interest.

Further methods

The present disclosure also provides a method for modifying a cell to comprise or express a polynucleotide according to the present disclosure, comprising introducing into a cell a polynucleotide or vector according to the present disclosure.

In some embodiments, introducing a polynucleotide or vector according to the present disclosure into a cell comprises transformation, transfection, electroporation or transduction (*e.g.* retroviral transduction).

Transfection relates to the process of introducing nucleic acid into cells using means other than viral infection and is hence a non-viral method. Transfection may be performed by physical/mechanical methods (including electroporation, sonoporation, magnetofection, gene microinjection and laser irradiation) or chemical methods (liposomal-based or non-liposomal based). Liposomal-based transfection

reagents are chemicals which enable the formation of positively charged lipid aggregates, which can then merge with the phospholipid bilayer of the cell to facilitate the entry of foreign genetic material. Examples of liposomal-based transfection reagents include, but are not limited to Oligofectamine®, Lipofectamine® and DharmaFECT®. Non-liposomal transfection reagents include, but are not limited to, calcium phosphate, nanoparticles, polymers, dendrimers and non-liposomal lipids. One example of a non-liposomal transfection reagent is polyethylenimine (PEI).

Electroporation may be performed *e.g.* as described in Koh *et al.*, Molecular Therapy – Nucleic Acids (2013) 2, e114, which is hereby incorporated by reference in its entirety.

10

Transduction is a process by which nucleic acids may be introduced into a cell by a virus or a viral vector. Accordingly, in some embodiments the polynucleotide is, or is comprised in, a viral vector, or the vector is a viral vector. Transduction of immune cells with viral vectors is described *e.g.* in Simmons and Alberola-Illa, Methods Mol Biol. (2016) 1323:99-108, which is hereby incorporated by reference in its entirety.

15

Agents may be employed in the methods of the present disclosure to enhance the efficiency of transduction. Hexadimethrine bromide (polybrene) is a cationic polymer which is commonly used to improve transduction, through neutralising charge repulsion between virions and sialic acid residues expressed on the cell surface. Other agents commonly used to enhance transduction include *e.g.* the poloxamer-based agents such as LentiBOOST (Sirion Biotech), Retronectin (Takara), Vectofusin (Miltenyi Biotech) and also SureENTRY (Qiagen) and ViraDuctin (Cell Biolabs).

20

In some embodiments the methods comprise centrifuging the cells into which it is desired to introduce polynucleotide or vector according to the present disclosure in the presence of cell culture medium comprising viral vector(s) comprising the polynucleotide (referred to in the art as 'spinfection').

25

In some embodiments, the methods comprise culturing the cell under conditions suitable for expression of the polynucleotide or vector by the cell. In some embodiments, the methods comprise culturing the cell under conditions suitable for transcription of a polydeoxyribonucleotide. In some embodiments, the methods comprise culturing the cell under conditions suitable for post-transcriptional processing (*e.g.* splicing) of a polyribonucleotide. In some embodiments, the methods comprise culturing the cell under conditions suitable for translation of a polypeptide from a polyribonucleotide.

30

Methods for culturing (including generating and/or expanding) populations of cells *in vitro/ex vivo* – including suitable culture conditions (*i.e.* cell culture media, additives, stimulations, temperature, gaseous atmosphere), cell numbers, culture periods *etc.* – are well known to the skilled person. Conveniently, cultures of cells according to the present disclosure may be maintained at 37°C in a humidified atmosphere containing 5% CO₂.

35

The present disclosure also provides a method for modifying a cell to express a polypeptide of interest, comprising introducing into a cell a polynucleotide or vector according to the present disclosure. Where the polynucleotide is, or wherein the vector comprises/encodes, an OFF-switch, the cell may express the polypeptide of interest following introduction of the polynucleotide/vector into the cell.

40

Where the polynucleotide is, or wherein the vector comprises/encodes, an ON-switch, the method may further comprise contacting the cell with a splicing modifier that promotes *SMN2* exon 7 inclusion as described herein. Thus, in some embodiments, the method comprises (i) introducing into a cell a polynucleotide or vector according to the present disclosure; and (ii) subsequently contacting the cell with a splicing modifier that promotes *SMN2* exon 7 inclusion.

'Contacting' a cell with a splicing modifier may comprise bringing a cell into contact with a splicing modifier in a cell culture, and may be achieved by applying the splicing modifier to the cells in culture.

The present disclosure also provides a method for inhibiting expression of a polypeptide of interest in a cell. The method comprises contacting a cell comprising a polynucleotide or vector according to the present disclosure comprising/encoding an OFF-switch with a splicing modifier that promotes *SMN2* exon 7 inclusion as described herein.

Any suitable quantity/concentration of a splicing modifier may be employed in the methods of the present disclosure. It will be appreciated that the quantity/concentration of the splicing modifier is preferably selected such as to achieve the desired effect, *i.e.* increased inclusion of the *SMN2* exon 7 variant nucleotide sequence in the products of splicing of the relevant polyribonucleotide.

Subjects

A subject in accordance with the various aspects of the present disclosure may be any animal or human. Therapeutic and prophylactic applications may be in human or animals (veterinary use).

The subject to be administered with an article of the present disclosure (*e.g.* in accordance with therapeutic or prophylactic intervention) may be a subject in need of such intervention. The subject is preferably mammalian, more preferably human. The subject may be a non-human mammal, but is more preferably human. The subject may be male or female. The subject may be a patient.

A subject may have (*e.g.* may have been diagnosed with) a disease or condition described herein, may be suspected of having such a disease/condition, or may be at risk of developing/contracting such a disease/condition. In embodiments according to the present disclosure, a subject may be selected for treatment according to the methods based on characterisation for one or more markers of such a disease/condition.

Kits

The present disclosure also provides kits of parts.

In some aspects and embodiments, a kit of parts according to the present disclosure comprises (i) a polynucleotide, vector, or a pharmaceutical composition according to the present disclosure, and (ii) a splicing modifier that promotes *SMN2* exon 7 inclusion.

Kits of parts according to the present disclosure may comprise a predetermined quantity of articles according to (i) and/or (ii), as described in the preceding paragraph. In some embodiments, articles according to (i) and/or (ii) are provided in containers (e.g. in vials or bottles). The kit may provide articles according to (i) and/or (ii) together with instructions (e.g. a protocol) as to how to employ them in accordance with a therapeutic or prophylactic intervention as described herein.

In some embodiments, the kit of parts may comprise a polynucleotide or vector according to the present disclosure, and optionally materials for introducing the polynucleotide/vector into a cell. In some embodiments, the kit of parts may comprise a system for producing a cell according to the present disclosure. In some embodiments, the kit of parts may comprise a (closed) bag cell incubation system in which a polynucleotide or vector can be introduced into a cell.

In some embodiments, the kit of parts may comprise materials for formulating a polynucleotide or vector according to the present disclosure to a pharmaceutical composition, e.g. a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.

In some embodiments, the kit of parts further comprises reagents, buffers and/or standards required for execution of a method according to the present disclosure. Kits according to the present disclosure may include instructions for use, e.g. in the form of an instruction booklet or leaflet. The instructions may include a protocol for performing any one or more of the methods described herein.

The manufacture of kits of parts according to the present disclosure preferably follows standard procedures which are known to the person skilled in the art.

Sequence identity

The 'sequence identity' between a given nucleotide sequence (e.g. of a polynucleotide) and a reference nucleotide sequence is calculated by determining the percentage of the nucleotides in the given nucleotide sequence that are identical to those of the reference nucleotide sequence, after aligning the sequences and, if necessary, introducing gaps, to achieve the maximum percentage sequence identity between the two sequences. Similarly, 'sequence identity' between a given amino acid sequence (e.g. of a polypeptide) and a reference amino acid sequence is calculated by determining the percentage of the amino acids in the given amino acid sequence that are identical to those of the reference amino acid sequence, after aligning the sequences and, if necessary, introducing gaps, to achieve the maximum percentage sequence identity between the two sequences.

Pairwise and multiple sequence alignment for the purposes of evaluating sequence identity between two or more nucleotide or amino acid sequences can be achieved in various ways known to a person of skill in the art, for instance, using publicly available computer software such as ClustalOmega (Söding, J. 2005, *Bioinformatics* 21, 951-960), T-coffee (Notredame *et al.* 2000, *J. Mol. Biol.* (2000) 302, 205-217), Kalign (Lassmann and Sonnhammer 2005, *BMC Bioinformatics*, 6(298)) and MAFFT (Katoh and Standley 2013, *Molecular Biology and Evolution*, 30(4) 772-780) software. When using such software, the default parameters, e.g. for gap penalty and extension penalty, are preferably used.

Sequences

SEQ ID NO:	DESCRIPTION	SEQUENCE
1	SMN2 exon 6 variant consensus 1	GCX _{1,2} AATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATX _{3,4} G wherein X _{1,2} = absent or TT; X _{3,4} = AT, CT or CA
2	SMN2 exon 6 (wildtype)	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATG
3	45 nucleotide SMN2 exon 6 comprising ATG>GGC at positions 82 to 84 (SMN2 exon 6 variant 1)	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATG
4	SMN2 exon 6 variant consensus 2	GCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCX ₁ G wherein X ₁ = T or A
5	SMN2 exon 6 comprising ATG>GGC at positions 64 to 66, ATG>GGC at positions 82 to 84 and ATG>CTG at positions 109 to 111 (SMN2 exon 6 variant 2)	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTG
6	SMN2 exon 6 comprising ATG>GGC at positions 64 to 66, ATG>GGC at positions 82 to 84 and ATG>CAG at positions 109 to 111 (SMN2 exon 6 variant 3)	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAG
7	Positions 1 to 102 of SMN2 intron 6	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACT
8	3' 162 positions of SMN2 intron 6, comprising A>C 4 positions from the 3' end (position 5766 of intron 6)	TTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTT GTGAAACAAAATGCTTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATCTAT ATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAG
9	414 nucleotide SMN2 intron 6, comprising A>C 4 positions from the 3' end (position 5766 of intron 6) (intron 6 variant 1)	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTTAAAAATAAAAAATAAC AACTGTCTCAGGCTTAAATTACTCTTGCAATTGTCCTAAGTATAATTTTAGTTAATTTTAA AAGCTTTTCATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGT ACATTTAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGT CTTGTGAAACAAAATGCTTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATC TATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAG
410	264 nucleotide SMN2 intron 6, comprising A>C 4 positions from the 3' end (position 5766 of intron 6) (intron 6 variant 2)	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTT CTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTTAA TCCATATAAAGCTATCTATATAGCTATCTATATAGCTATTTTTTTTAACTTCCTTT TATTTTCTTCCAG
11	SMN2 exon 7 variant consensus 1	GATTTTX _{1,6} AX _{7,8} CAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGA wherein X _{1,6} = absent or GCCACC; X _{7,8} = GA or TG
12	SMN2 exon 7 (wildtype)	GGTTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATTAAGGA
13	SMN2 exon 7 comprising G>A at position 2, and insertion of A after position 48 (exon 7 variant 1)	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGA
14	SMN2 exon 7 comprising G>A at position 2, insertion of GCCACC	GATTTTGCCACCATGCAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGG A

	after position 6, GA>TG at positions 8 and 9, and insertion of A after position 48 (exon 7 variant 2)	
15	Positions 1 to 102 of <i>SMN2</i> intron 7	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAA
16	3' 150 positions of <i>SMN2</i> intron 7	CTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCA AGCCTCTGGTTCTAATTTCTCATTTCAG
17	252 nucleotide <i>SMN2</i> intron 7	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTCAG
18	444 nucleotide <i>SMN2</i> intron 7 (wildtype)	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTGAAAGGTTAATGTAAA ACAATCAATATTAAGAATTTTGATGCCAAAACATTAGATAAAAGGTTAATCTACATCCC TACTAGAATTCTCATACTTAACTGGTTGGTGTGTGGAAGAAACATACTTTCACAATAAAG AGCTTTAGGATATGATGCCATTTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTAT TGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAG TCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT TCTAATTTCTCATTTCAG
19	Positions 1 to 23 of <i>SMN2</i> exon 8 consensus	X ₁₋₂ AATGCTGGCATAGAGCAGCAC wherein X ₁₋₂ = GA or TG;
20	Positions 1 to 23 of <i>SMN2</i> exon 8 (exon 8 variant 1)	GAAATGCTGGCATAGAGCAGCAC
21	Positions 1 to 23 of <i>SMN2</i> exon 8, comprising GA>TG at positions 1 to 2 (exon 8 variant 2)	TGAATGCTGGCATAGAGCAGCAC
22	pMM112 ON-switch	GCCACCATGGCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGT AATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTT ATAAAATACTACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGAT CATATTTTGTGAAATAAAATAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCAT ATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTT TCCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAA GGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTG GAAAACAAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTATTGTGATA TGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTT TAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATT TCTCATTTCAGGA
23	pMM130 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAAAACAAA TGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTC CAGGA
24	pMM59 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTTATGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTGGCACCATGCAAAATCAAAAAGA AGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAA TCTTACTTTTGTAACATTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTT CAGATGTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCA

		AAACTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGG TTGTGTGGGAAGAAACATACTTTTCCACAATAAAGAGCTTTAGGATATGATGCCATTTTATATC ACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTG GCATAGAGCAGCAC
25	pMM198 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACTGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAGGAA GTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTTAAAAAGTTTCAGATGT TAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAAATCTCATACTGGTTGGTTGTGTG GAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCCTGTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTGAATGCTGGCATAGA GCAGCAC
26	SMN2 exon 7 variant consensus 2	GATTTTAGACAAAATCAAAX ₁ AGAAGGAAGGTGCTCACATTCCCTAAATATAAGGA wherein X ₁ = absent or A
27	SMN2 exon 7 comprising G>A at position 2, comprising deletion of position 20, and comprising insertion of A after position 48 (exon 7 variant 3)	GATTTTAGACAAAATCAAAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGA
28	pMM193 OFF-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTTAAAAAGTTTCAAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAG
29	pMM194 OFF-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCT GCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATG TTTTTGAACATTTTAAAAAGTTTCAAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCT AGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTC CACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCA GGAG
30	SMN2 intron 6 (wildtype)	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATAC AACTGTCTGAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTAGTTAATTTTAAA AAGCTTTCATGCTATTGTTAGATTATTTTGATTATACACTTTTGAATTGAAATTATACTTTTT CTAAATAATGTTTAAATCTCTGATTTGAAATTGATTGTAGGGAATTGAAAAGATGGGATAA TTTTTCATAAATGAAAAATGAAATCTTTTTTTTTTTTTTTTTTTTTTGTAGACGGAGTCTTGC TCTGTTGCCCAGGCTGGAGTGCAATGGCGTGATCTTGCTCACAGCAAGCTCTGCCTCC TGGATTCACGCCATTCTCCTGCCTCAGCCTCAGAGGTAGCTGGGACTACAGGTGCCTGC CACCACGCCCTGTCTAATTTTTTGTATTTTTTTGTAAAGACAGGGTTTCACTGTGTTAGCCA GGATGGTCTCAATCTCCTGACCCCGTGATCCACCCGCCCTCGGCCCTTCCAAGAGAAATGA AATTTTTTAAATGCACAAAGATCTGGGGTAATGTGTACCACATTGAACCTTGGGGAGTAT GGCTTCAAACCTGTCACTTTATACGTTAGTCTCCTACGGACATGTTCTATTGTATTTAGT

[illegible]

		TTTTTACATTTAAATTTTATTTTTAGAGACCAGGTCTCACTCTATTGCTCAGGCTGGAGT GCAAGGGCACATTACAGCTCACTGCAGCCTTGACCTCCAGGGCTCAAGCAGTCCCTCTC ACCTCAGTTTCCCGAGTAGCTGGGACTACAGTGATAATGCCACTGCACCTGGCTAATTTT TATTTTTATTTATTTATTTTTTTTGGAGACAGAGTCTTGCTCTGTCAACCCAGGCTGGAGTG CAGTGGTGTAATCTCAGCTCACTGCAGCCTCCGCTCCTGGGTTCAAGTGATTCTCCT GCCTCAACCTCCCAAGTAGCTGGGATTAGAGGTCCCAACCCATGCCTGGCTAATTTT TTGTACTTTAGTAGAAACGGGGTTTTGCCATGTTGCCAGGCTGTTCTCGAACTCCTGA GCTCAGGTGATCCAAGTGTCTCGGCCTCCCAAGTGTGGGATTACAGGCTGAGCCA CTGTGCCTAGCCTGAGCCACCACGCCGCCCTAATTTTTAAATTTTTTGTAGAGACAGGT CTCATTATGTTGCCAGGTTGGTGTCAAGTCCAGGTCTCAAGTGATCCCCCTACCTCC GCCTCCCAAGTTGTGGATTGTAGGCATGAGCCACTGCAAGAAAACCTTAAGTGCAGC CTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAGAACTATCAACTTAATTTT TGATCATATTTTGTGAATAAAATAAGTAAATGTCTGTGAAACAAAATGCTTTTTAACAT CCATATAAAGCTATCTATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTT ATTTTCCTTACAG
32	SMN2 exon 8 (wildtype)	GAAATGCTGGCATAGAGCAGCACTAAATGACACCACTAAAGAAACGATCAGACAGATCT GGAATGTGAAGCGTTATAGAAGATAACTGGCCTCATTTCTTCAAAATATCAAGTGTTGGG AAAGAAAAAGGAAGTGAATGGGTAACCTTTCTTGATTAAGTTATGTAATAACCAAT GCAATGTGAAATATTTTACTGGAATCTATTTTGA AAAAACCATCTGTAAAGACTGAGGTG GGGTTGGGAGGCCAGCACGGTGGTGAGGCAGTTGAGAAAATTTGAATGTGGATTAGAT TTTGAATGATATTGGATAATTATTGGTAATTTTATGAGCTGTGAGAAGGGTGTGTAGTTT ATAAAGACTGTCTTAATTTGCATACTTAAGCATTTAGGAATGAAGTGTAGAGTGTCTTA AAATGTTTCAAAATGGTTTAAACAAAATGTATGTGAGGCGTATGTGGCAAAAGTTACAGAAT CTAACTGGTGGACATGGCTGTTCTATTGTACTGTTTTTTTCTATCTTCTATATGTTTAAAGT ATATAATAAAATATTTAATTTTTTTTTTAAATTA
33	Kozak consensus	GCCX ₁ CC wherein X ₁ = A or G
34	Kozak	GCCACC
35	Kozak plus start codon	GCCACCATG
36	SMN2ind minigene of WO 2021/163556 A1 (SEQ ID NO:2 of WO 2021/163556 A1)	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGTTAAGTAATC ACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAA AATACTACTTGTCTCTCTTTTATATTACTAAAAAATAAAAAATAAAAAATACAAGTGTCTG AGGCTTAAATTACTCTTGCATTGTCCCTAAGTATAATTTTGTAAATTTTAAAAAGCTTTCA TGCTATTGTTAGATTATTTTGAATATACACTTTTGAATTGAAATTAATCTTTTCTAAATAAT GTTTTAATCTCTGATTGAAATTGATTGTAGGGAATGGAAAGATGGGATAATTTTTCATA AATGAAAAATGAAATTCCTTTTTTTTTTTTTTTTTTTTTTGGAGACGGAGTCTTGCTCTGTTGC CCAGGCTGGAGTGCAATGGCGTGATCTTGGCTCACAGCAAGCTCTGCCTCCTGGATTCA CGCCATTCTCCTGCCTCAGCCTCAGAGGTAGCTGGGACTACAGGTGCCTGCCACCACG CCTGGCTAGCTGGGATTAGAGTCCCCACCACCATGCCTGGCTAATTTTTTGTACTTTTCA GTAGAAACGGGGTTTTGCCATGTTGGCCAGGCTGTTCTCGAACTCCTGAGCTCAGGTGA TCCAAGTGTCTCGGCCTCCCAAGTGTCTGGGATTACAGGCGTGAGCCACTGTGCCTAGC CTGAGCCACCACGCCGGCCTAATTTTTTAAATTTTTTGTAGAGACAGGCTCTCATTATGTT GCCAGGGTGGTGTCAAGTCCAGGTCTCAAGTGATCCCCCTACCTCGGCCTCCAAA GTTGTGGGATTGTAGGCATGAGCCACTGCAAGAAAACCTTAAGTGCAGCCTAATAATTGT TTTCTTTGGGATAACTTTTAAAGTACATTAAGAACTATCAACTTAATTTCTGATCATATTT TGTTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAG CTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCTTTATTTTCTTCT CAGGATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGAGT AAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACCTTTATGGTTTGTGAAAA CAAATGTTTTGAACATTTAAAGTTTCAGATGTTAGAAAGTTGAAAGGTTAATGTAAAC AATCAATATTAAGAATTTTGTGCCAAAATATTAGATAAAAGGTTAATCTACATCCCTA CTAGAATTCTCATACTTAACGTGTTGGTTGTGTGGAAGAAACATACTTTCAATAAAGAG CTTTAGGATATGATGCCATTTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTTTATTG TGATATGGGATAACCTAGGCATCTGCACTGTACACTCTGACATATGAAGTGTCTAGTC AAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGATCTGACATATGAAGTGTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAGCAGCAC
37	pLS41 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGTTAAGTAATC ACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAA AATACTACTTGTCTCTCTTTTATATTACTAAAAAATAAAAAATAAAAAATACAAGTGTCTG AGGCTTAAATTACTCTTGCATTGTCCCTAAGTATAATTTTGTAAATTTTAAAAAGCTTTCA TGCTATCTTAAGTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAATA GACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAA CAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCT ATTTTTTTAACTTCCTTTATTTTCTTACAGGTTTTAGACAAAATCAAAAAGAAGGAAGG

		TGCTCACATTCTTAAATATAAGGAGAAATGCTGGCATAGAGCAGCACGTAAGTCTGCCA GCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTT GAACATTTAAAAAGTTCAGATGTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTA AAGAATTTTGATGCCAAAACATTAGATAAAAAGGTTAATCTACATCCCTACTAGAAATCTC ATACTTAAGTGGTTGGTTGTGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATAT GATGCCATTTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGAT AACCTAGGCATACTGCCTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTG GTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCAT TTGCAG
38	pLS76 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGTAAGTAATC ACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAA AATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACCTGTCTG AGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTTCA TGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGCGATAACTTTTAAAGTACATTAAGA GACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAA CAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCT ATTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTTAGACAAAAACAAAAAGAAGGAAGG TGCTCACATTCTTAAATATAAGGAGTAAAGTCTGCCAGCATTGAAAGTGAATCTTACTT TTGTAAAACCTTTATGGTTTGTGAAAACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTT AGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAATTTTGATGCCAAAACCTATT AGATAAAAGGTTAATCTACATCCCTACTAGAATTCCTACTTAACCTGGTTGGTTGTGTGG AAGAAACATACTTTTACATAAAGAGCTTTTAGGATATGATGCCATTTTATGATCACTAGTAG GCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCCTGTAC ACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTA ACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAG CAGCAC
39	pLS159 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGTAAGTAATC ACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAA AATACTACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATAT TTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAA GCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCTTTATTTTCTCT CCAGGATTTTAGACAAAATCAAAAAGAAGGAGTGTCTACATTCCTTAAATATAAGGAGTAAG TAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAA ACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAATTTTTTTTATTGTGATATGGG ATAACCTAGGCATACTGCCTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACT TGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTC ATTTGCAGGAAATGCTGGCATAGAGCAGCAC
40	pLS160 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGTAAGTAATC ACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAA AATACGTAATAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATA TAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTTAG ACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAG CATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAACAAATGTTTTTG AACTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAA CTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAGC AGCAC
41	pLS167 ON-switch	GATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTTCATGGTACATGAGTGGC TATCATACTGGCTATTATATGGTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTGT AGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTAA AAAAATAAAAAATAAAAAATACAACCTGTCTGAGGCTTAAATTACTCTTGCAATTGTCCCTAAG TATAATTTTAGTTAATTTTAAAAAGCTTTCATGCTATCTTAACTGCAGCCTAATAATTGTTT TCTTTGGGATAACTTTTAAAGTACATTAAGAGACTATCAACTTAATTTCTGATCATATTTTG TTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTA TCTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCTTTATTTTCTTCCAG GATTTTAGACAAAATCAAAAAGAAGGAGGTGCTCACATTCCTTAAATATAAGGAGTAAG TCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAACAA ATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTGAAAGGTTAATGTAAAACAAT CAATATTAAGAATTTTATGATGCCAAAACATTAGATAAAAAGGTTAATCTACATCCCTACTA GAATTCCTACATACTTAACCTGGTTGGTTGTGTGGAAGAAACATACTTCAACAATAAAGAGCTT TAGGATATGATGCCATTTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGA TATGGGATAACCTAGGCATACTGCCTGTACACTCTGACATATGAAGTGCTCTAGTCAAG TTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAA TTTCTCATTTGCAGGAAATGCTGGCATAGAGCAGCAC
42	pLS168 ON-switch	AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATGGTAAGTAATC ACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAA AATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACCTGTCTG AGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTTCA

		TGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTA GACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAA CAAAATGCTTTTAAACATCCATATAAAGCTATCTATATAGCTATCTATATCTATAGCT ATTTTTTTAACTTCCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAGG TGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTT TTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGATGTT AGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTATGATGCCAAACTATT AGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGG AAGAAACATACTTTCACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGTAG GCAGACCAGCAGACTTTTTTTTATTTGTGATATGGGATAACCTAGGCATACTGCACTGTAC ACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGCCACAGAGGACATGGTTTA ACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCAGGAAATGCTGGCATAGAG CAGCAC
43	pLS179 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTAAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAATTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTC CAGGAAATGCTGGCATAGAGCAGCAC
44	pMM70 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTAAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAATTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTC CAGGAAATGCTGGCATAGAG
45	pMM71 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTAAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAATTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTC CAGGAAATGCTGGC
46	pMM72 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTAAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAATTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTC CAGGAAATGCT
47	pMM73 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTAAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAATTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTC CAGGAAAT
48	pMM143 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA

		GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCTCCTCTTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAAACAA TGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAAATGCTGGCATAGAGCAGCAC
49	pMM144 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCAAGGCTTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAAACAA TGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAAATGCTGGCATAGAGCAGCAC
50	pMM145 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCTAGTCTTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAAACAA TGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAAATGCTGGCATAGAGCAGCAC
51	pMM146 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCTTTTCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAAAC AAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGAT AACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTG GTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCAT TTGCAGGAAATGCTGGCATAGAGCAGCAC
52	pMM147 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCCCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAAAC AAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGAT AACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTG GTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCAT TTGCAGGAAATGCTGGCATAGAGCAGCAC
53	pMM136 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGTAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT TGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAAACAA GTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACC TAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG CCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG AGGAAATGCTGGCATAGAGCAGCAC
54	pMM137 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT

		CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGATGGTGCTCACATTCCTTAAATATAAGGAGTAAGTC TGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAAT GTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACC TAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGT CCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGC AGGAAATGCTGGCATAGAGCAGCAC
55	pMM138 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGT TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAAATGCTGGCATAGAGCAGCAC
56	pMM139 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTC TGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAAT GTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACC TAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGT CCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG AGGAAATGCTGGCATAGAGCAGCAC
57	pMM140 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTC TGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAAT GTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACC TAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGT CCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG AGGAAATGCTGGCATAGAGCAGCAC
58	pMM141 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTC CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGT TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAAATGCTGGCATAGAGCAGCAC
59	pMM142 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGACAAAATCAAAAAGGAAGGAGCAGCATTCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGT TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGGAAATGCTGGCATAGAGCAGCAC
60	pMM151 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTCCAGG ATTTTAGAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATT TGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAAACAAATGTTTTTGAACAT

		TTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTG GCATAGAGCAGCAC
61	pMM152 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGAGAAGGAAGGTGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGA ATCTTACTTTTTGTAAACCTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGT TCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACA CTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAA CTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAGC AGCAC
62	pLS174 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAAATGTCTTGTGAA ACAAAATGCTTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAAACCTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTCAGATGT TAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAATTTTGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCCTACTTAACTGGTTGGTTGTGTG GAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGA GCAGCAC
63	pMM56 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAAATGTCTTGTGAA ACAAAATGCTTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCCACCATGTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAAACCTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCCTACTTAACTGGTTGGTTGTG TGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
64	pMM60 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAAATGTCTTGTGAA ACAAAATGCTTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCCTTAAAGCCACCATGATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGA ATCTTACTTTTTGTAAACCTTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGT TCAGATGTTAGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAATTTTGATGCC AAAACATTTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCCTACTTAACTGGTTG GTTGTGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATAT CACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACT GCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGA CATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCT GCATAGAGCAGCAC
65	pMM61 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA

		AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTGTAAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAACAACATGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTG TGGAAGAAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
66	pMM62 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTGTAAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTACGCCACCATGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTG TGGAAGAAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
67	pMM63 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTGTAAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACAGCCACCATGAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTT ACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGA TGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAAC TATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGT GTGGAAGAAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTA GTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACT GTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGG TTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
68	pLS176 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTGTAAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTGCACCATGCAAAAATCAAAAAGA AGGAAGGTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAA TCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTT CAGATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCA AACTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGG TTGTGTGGAAGAAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATC ACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTG

		GCATAGAGCAGCAC
69	pMM108 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTGCACCATGCAAAATCAAAAAGA AGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAA TCTTACTTTTGTAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTT CAGATGTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCA AAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGG TTGTGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATC ACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTG GCATAGAGCAGCAC
70	pMM110 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATGTGGTAAGTAAT CACTCAGCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTGCACCATGCAAAATCAAAAAGA AGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAA TCTTACTTTTGTAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTT CAGATGTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCA AAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGG TTGTGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATC ACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTG GCATAGAGCAGCAC
71	pMM111 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATTTGGTAAGTAAT CACTCAGCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTGCACCATGCAAAATCAAAAAGA AGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAA TCTTACTTTTGTAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTT CAGATGTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCA AAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGG TTGTGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATC ACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTG GCATAGAGCAGCAC
72	pMM242 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGAGCCACCATGCAAAATCAAAA AGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGT GAATCTTACTTTTGTAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAA GTTCAGATGTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATG CCAAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGT TGTTGTGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTAT ATCACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATA

		CTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAG GACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCGAGGAAATG CTGGCATAGAGCAGCAC
73	pMM243 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATAGCACCATTGCAAAA AGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGT GAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAA GTTGAGATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATG CCAAAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGT TGGTTGTGTGGAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTAT ATCACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATA CTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAG GACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCGAGGAAATG CTGGCATAGAGCAGCAC
74	pMM244 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATAGCACCATTGCAAAA AGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGT GAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAA GTTGAGATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATG CCAAAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGT TGGTTGTGTGGAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTAT ATCACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATA CTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAG GACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCGAGGAAATG CTGGCATAGAGCAGCAC
75	pMM245 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCACCATTGAAAAAG AAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGA ATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGT TCAGATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCC AAAATATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTG GTTGTGTGGAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATAT CACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACT GCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGA CATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCGAGGAAATGCT GGCATAGAGCAGCAC
76	pMM246 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATATGAAAAAGAAGG AAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCT TACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTGAG ATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAA CTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTG

		TGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACT AGTAGGCAGACCAGCAGACTTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC TGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATG GTTTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCA TAGAGCAGCAC
77	pMM247 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAATGGAAGG AAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCT TACTTTTGTAAACTTTATGGTTTGTGGAAAAACAAATGTTTTGAACATTTAAAAAGTTCAG ATGTTAGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAAATTTTGTGCCAAAA CTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTG TGTGGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCAGCATTATGAAAGTGAATCT AGTAGGCAGACCAGCAGACTTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC TGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATG GTTTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCA TAGAGCAGCAC
78	pLS175 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAAACTTTATGGTTTGTGGAAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATG TAGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAAATTTTGTGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGTG GAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGTA GGCAGACCAGCAGACTTTTTTTTTATTGTGATATGGGATAACCTAGGCATAGCTGCACTGT CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGCCATAGA GCAGCAC
79	pMM123 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCCCTTAAATATAATGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTAC TTTTGTAAACTTTATGGTTTGTGGAAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATG TTAGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAAATTTTGTGCCAAAACAT TTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGT GGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGT AGGCAGACCAGCAGACTTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGT ACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTT TAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGCCATAG AGCAGCAC
80	pMM124 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAAAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCCCTTAAATGTAAAGGTAAGTCTGCCAGCATTATGAAAGTGAATCTTCTTA CTTTTGTAAACTTTATGGTTTGTGGAAAAACAAATGTTTTGAACATTTAAAAAGTTCAGAT

		GTTAGAAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACCTGGTTGGTTGTG TGGAAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACCTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
81	pMM125 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GATGGCGTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCT TACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAG ATGTTAGAAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGTATTTGATGCCAAAAC CTATTAGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACCTGGTTGGTTG TGTGGAAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACT AGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACCTGCAC TGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATG GTTTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCA TAGAGCAGCAC
82	pMM126 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACCTGGTTGGTTGTG TGGAAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACCTGCACCTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
83	pMM127 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTATGGATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACCTGGTTGGTTGTG TGGAAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACCTGCACCTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
84	pMM128 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTATGGATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAAC ATTAGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACCTGGTTGGTTGTG TGGAAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACCTGCACCTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGTCAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC

		GTGCTCACATTCCCTTAAATATAAGATGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAACCTTTATGGTTTGTGGAAAAACAAATGTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAAC ATTAGATAAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACTGGTTGGTTGTG TGGAAGAAACATACTTTCACAATAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATCTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
85	pMM129 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACCTGTCT GAGGCTTAAATTACTCTTGCAATTGTCCTAAGTATAATTTTAGTTAATTTAAAAAGCTTTT ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATGAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTCTT TTGTAACCTTTATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAGTTTCAGATGTT AGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAACCTATT AGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACTGGTTGGTTGTGTGG AAGAAACATACTTTCACAATAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTAG GCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATCTGCACTGTAC ACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTA ACTGGAATTCGCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAG CAGCAC
86	pMM201 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACCTGTCT GAGGCTTAAATTACTCTTGCAATTGTCCTAAGTATAATTTTAGTTAATTTAAAAAGCTTTT ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATGGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAACCTTTATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAACCT ATTAGATAAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACTGGTTGGTTGTG TGGAAGAAACATACTTTCACAATAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATCTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
87	pMM202 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACCTGTCT GAGGCTTAAATTACTCTTGCAATTGTCCTAAGTATAATTTTAGTTAATTTAAAAAGCTTTT ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTAACATGGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAACCTTTATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAACCT ATTAGATAAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACTGGTTGGTTGTG TGGAAGAAACATACTTTCACAATAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATCTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
88	pMM203 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACCTGTCT GAGGCTTAAATTACTCTTGCAATTGTCCTAAGTATAATTTTAGTTAATTTAAAAAGCTTTT ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTAACATGGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTA CTTTTGTAACCTTTATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAGTTTCAGAT GTTAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAAGAATTTTGATGCCAAAACCT ATTAGATAAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTTAACTGGTTGGTTGTG TGGAAGAAACATACTTTCACAATAAGAGCTTTAGGATATGATGCCATTTTATATCACTAG TAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATCTGCACTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAACTGGAATTCGCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC

		ACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATATGCCTAAATTAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTAC TTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATG TTAGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAACATA TTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGT GGAAGAAACATACTTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGT AGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACGT ACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTT TAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAG AGCAGCAC
89	pMM204 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCCATGTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCT TACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCA ATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAA CTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTG TGTGGAAGAAACATACTTTCACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACT AGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC TGACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATG GTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCA TAGAGCAGCAC
90	pMM205 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCTTATGAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTT ACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTTCA TGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAAC TATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTG GTGGAAGAAACATACTTTCACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTA GTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC GTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGG TTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATA GAGCAGCAC
91	pMM206 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCTTATGGAATAAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTT CTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTT AGATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAA AACTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGT TGTGTGGAAGAAACATACTTTCACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCA CTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGC ACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACA TGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGG CATAGAGCAGCAC
92	pMM248 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGTCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGGAAG GTGCTCACATTCTTATGGAATAAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTT CTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTT AGATGTTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAA AACTATTAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGT TGTGTGGAAGAAACATACTTTCACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCA CTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGC ACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACA TGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGG CATAGAGCAGCAC

		ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATTAAGATGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTAC TTTTGTAAACCTTTATGGTTTGTGGAAAAAATGTTTTGAACATTTAAAAAGTTCCAGATG TTAGAAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAACATA TTAGATAAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTAACTGGTTGGTTGTGT GGAAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACTAGT AGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGT ACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTT TAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAG AGCAGCAC
93	pMM249 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATGAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTT TTGTAAACCTTTATGGTTTGTGGAAAAACAATGTTTTGAACATTTAAAAAGTTCCAGATGTT AGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAACCTATT AGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTAACTGGTTGGTTGTGTGG AAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTAG GCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTAC ACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTA ACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAG CAGCAC
94	pMM250 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTGAAATGAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTT TTGTAAACCTTTATGGTTTGTGGAAAAACAATGTTTTGAACATTTAAAAAGTTCCAGATGTT AGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAACCTATT AGATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTAACTGGTTGGTTGTGTGG AAGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTAG GCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTAC ACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTA ACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAG CAGCAC
95	pMM251 ON-switch	ATAATTCCTCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATA AAATACTACTTGCTTCTCTCTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTAATCTTGCATTGTCCCTAAGTATAATTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTTGAATAAAATAAGTAAATGTCTTGTGAA ACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCAAAATGAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTT GTAAACCTTTATGGTTTGTGGAAAAACAATGTTTTGAACATTTAAAAAGTTCCAGATGTTA GAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAGAATTTTGATGCCAAAACCTATT GATAAAAGGTTAATCTACATCCCTACTAGAAATTCATACCTAACTGGTTGGTTGTGTGG AGAAACATACTTTTCAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTAGG CAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACA CTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTA CTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGAAATGCTGGCATAGAGC AGCAC
96	pMM252 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT

		TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGGA
97	pMM254 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATGTCCTTAAATATAAGGAGTAAG TCTGCCAGCATTATGAAAGTGAATCTTACTTTTTGTAAAACCTTTATGGTTTGTGGAAAACAA ATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAA CCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAGGA
98	pMM255 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGATGAGTAAG TCTGCCAGCATTATGAAAGTGAATCTTACTTTTTGTAAAACCTTTATGGTTTGTGGAAAACAA ATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAA CCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAGGA
99	pMM256 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATGAGGAGTAAGTCT GCCAGCATTATGAAAGTGAATCTTACTTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATG TTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCT AGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTC CACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCA GGA
100	pMM257 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATGATAAGGAGTAAG TCTGCCAGCATTATGAAAGTGAATCTTACTTTTTGTAAAACCTTTATGGTTTGTGGAAAACAA ATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAA CCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAGGA
101	pMM199 ON-switch	ATAATTCACCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACTGTCT GAGGCTTAAATTAATCTTGCAATTTGCCCTAAGTATAATTTAGTTAATTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAAATGCTTGTGAA ACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATGT TAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTATGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCCTACTTAACTGGTTGTTGTG GAAGAAACATACTTTTCAACAATAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCCAGTGAATGCTGGCATAGA GCAGCAC

102	pMM200 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAACCTTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTTCAGATGT TAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGTG GAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATAGTGCAGTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTTTTTTTCAGTGAATGCTGGCATAG CAGCAC
103	pMM235 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAACCTTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTTCAGATGT TAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGTG GAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATAGTGCAGTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCAGATGAATGCTGGCATAG AGCAGCAC
104	pMM237 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATTAGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAACCTTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTTCAGATGT TAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGTG GAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATAGTGCAGTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCAGATGAATGCTGGCATAG AGCAGCAC
105	pMM238 ON-switch	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATTGGTAAGTAAT CACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATA AAATACTACTTGGCTTCTCTCTTTATATTACTAAAAAATAAAAAATAAAAAATACAACGTCT GAGGCTTAAATTACTCTTGCAATTGCCCTAAGTATAATTTTAGTTAATTTTAAAAAGCTTTC ATGCTATCTTAACTGCAGCCTAATAATTGTTTTCTTTGGGATAACTTTTAAAGTACATTAAA AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAA ACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGC TATTTTTTTTAACTTCCTTTATTTTCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAG GTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACT TTTGTAACCTTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTTCAGATGT TAGAAAGTTGAAAGGTTAATGTAAACAATCAATATTAAGAATTTTGATGCCAAAACAT TAGATAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTGTG GAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATCACTAGTA GGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATAGTGCAGTGTA CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCAGATGAATGCTGGCATAG AGCAGCAC

		AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGATGAATGCTGGCATAGAGCAGCAC
106	45 nucleotide <i>SMN2</i> exon 6 comprising ATG>GGC at positions 82 to 84, additionally comprising 5' 'GCCACCATG'	GCCACCATGGCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATG
107	T2A cleavage site	GSSEGRGSLLTCGDVEENPGP
108	E2A linker	GSQCTNYALLKLAGDVESNPGP
109	P2A cleavage site	GSQATNFSLLKQAGDVEENPGP
110	F2A cleavage site	GSQVKQTLNFDLLKLAGDVESNPGP
111	pMM273 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGTG
112	pMM362 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTG GCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAAATG TTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCT AGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTC CACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCA GTG
113	pMM363 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATGATGAGTAAGAT TGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAAAT GTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCT TAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGC AGTG
114	pMM364 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT GTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGTG
115	pMM358 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGTGTGCCAG CATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTG AACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCA TACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAG AGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCACTG
116	pMM359 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT

		TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTTGTAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAACATTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACCTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
117	pMM360 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAA GTGAATCTTACTTTTTGTAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAACATTTAAAA AAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACCTG TACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGT TTAAGTGGAAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
118	pMM361 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACTCAGATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATC TTACTTTTGTAAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAACATTTAAAAAGTTCAG ATGTTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACCTGTACACTC TGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTG GAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
119	pMM365 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATG AAAGTGAATCTTACTTTTTGTAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAACATTT AAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCA CTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACAT GGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
120	pMM366 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGA AAGTGAATCTTACTTTTGTAAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAACATTTA AAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC TGACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATG GTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
121	pMM367 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGAA AGTGAATCTTACTTTTGTAAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAACATTTAA AAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC TGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGG TTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
122	pMM368 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATT ATGAAAGTGAATCTTACTTTTGTAAACCTTTATGGTTTGTGGAAACAAATGTTTTTGAAC ATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAACCTAGGCATAC TGCACCTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGG ACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
123	pMM369 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC

		TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCA TTATGAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGAA CATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATA CTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAG GACATGGTTTAACTGGAATTTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
124	pMM370 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCAT ACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGA GGACATGGTTTAACTGGAATTTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
125	pMM371 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCC AGCATTATGAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTT TGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGG CATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCAC AGAGGACATGGTTTAACTGGAATTTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
126	pMM372 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAG TGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTTAAAAA AGTTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGT ACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT TAACTGGAATTTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
127	pMM373 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTG AATCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTTAAAAAG TTCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTAC ACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAA CTGGAATTTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
128	pMM374 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAA TCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTTAAAAAGTT CAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACAC TCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACT TGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
129	pMM375 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGGTGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGA ATCTTACTTTTGAAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTTAAAAAGT TCAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACA CTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAA CTGGAATTTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
130	pMM376 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA

		GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGA AAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTTA AAAAGTTCAGATGTTAGAAGTCTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCAC TGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATG GTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
131	pMM377 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAG TGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTAAAAA AGTTCAGATGTTAGAAGTCTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACGT ACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTT TAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
132	pMM378 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGAGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGT GAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTAAAAA GTTTCAGATGTTAGAAGTCTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACGT CACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
133	pMM379 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGAGAAGGAAGGTGCTCACATTCCCTAAATATAAGGAGTAAGTCTGCCAGCATTAT TGAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAACAAAATGTTTTTGAACAT TTAAAAAGTTCAGATGTTAGAAGTCTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTG CACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGAC ATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
134	pMM380 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTTAAATATAAGGAGTAAGTCTGCCAGCA TTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAACAAAATGTTTTTGAA CATTTAAAAAGTTCAGATGTTAGAAGTCTTTTTTTTATTGTGATATGGGATAACCTAGGCATA CTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAG GACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
135	pMM381 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGAGAAGGAAGGTGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGA ATCTTACTTTTGTAAACTTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTAAAAAGT TCAGATGTTAGAAGTCTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACGTACA CTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAA CTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
136	pMM382 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGT AAAACCTTTATGGTTTGTGGAACAAAATGTTTTTGAACATTAAAAAGTTCAGATGTTAGA ACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACGTGTACACTCTGACATATG AAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGT AAGCCTCTGGTTCTAATTTCTCATTTGCAGTG

137	pMM385 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTA GGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCC ACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAG TG
138	pMM387 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGTCT GCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCT AGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCC CACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCA GTG
139	pMM435 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGGCACATTCCCTTAAATATAAGGAGTAAGTCTGCCAGCATT ATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGTTTTGAAC ATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGCATAC TGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGG ACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
140	pMM437 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAACCTCACATTCCCTTAAATATAAGGAGTAAGTCTGCC AGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGTTTT TGAACATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGG CATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
141	pMM438 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAACACATTCCCTTAAATATAAGGAGTAAGTCTGCCA GCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGTTTT GAACATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAACCTAGGC ATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACA GAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
142	pMM477 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTTGTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGG AATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
143	pMM478 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAAGATGTTAGAACCTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTTGTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGG AATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG

		TGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
144	pMM479 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTCTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
145	pMM480 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACAACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
146	pMM481 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGTGAAGTGCTCTAGTCAAGTTTAACTGGTGTCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
147	pMM482 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGCTCTAGTCAAGTTTAACTGGTGTCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
148	pMM483 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCAC AGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTGCGAGTG
149	pMM484 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAAAACAAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGGCATATGAAGTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTG CAGTG
150	pMM485 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCCTTTATTTTCCCTCCAGG

		ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAAACAA TGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTATTGTGATATGGGATAAC CTAGGCATACTTGAATTAGGAGGGGAGGATTCACTTGCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGTG
151	pMM454 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTAAAAAGTACATTAAAAGACTATCAACTTAATTTCTGA TCATATTTTGTGAATAAAATAAGTAAATGTC TTGTGAAACAAAATGCTTTTTAACATCCA TATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCTTTATT TTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATA AGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGT GGAAAAACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTAGCAGACTTTTT TTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGC TCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCT CTGGTTCTAATTTCTCATTTGCAGTG
152	pMM455 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTAAAAATAAATTAAGTACATTAAAAGACTATCAACT TAATTTCTGATCATATTTTGTGTAATAAAATAAGTAAATGCTTGTGAAACAAAATGCTTT TTAACATCCATATAAAGCTATCTATATAGCTATCTATATCTATATAGCTATTTTTTTAACT TTCTTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCT CTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCT TATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTGA AAGAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTA CACTCTGACATATGAAGTGTCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
153	pMM456 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTAAAAATAAATTAAGTACATTAAAAGACTATCAACT ACTATCAACTTAATTTCTGATCATATTTTGTGTAATAAAATAAGTAAATGCTTGTGAAAC AAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTA TTTTTTTTAACTTCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGT GCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTT TGTAAACTTTATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTA GAAAGTTGAAAGGTTAATAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAAC CTAGGCATACTGCACTGTACACTCTGACATATGAAGTGTCTCTAGTCAAGTTTAACTGGTG TCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTG CAGTG
154	pMM457 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTAAAAATAAATAAAGGATAACTTTTAAAGT ACATTAAGACTATCAACTTAATTTCTGATCATATTTTGTGTAATAAAATAAGTAAATGT CTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATC TATATAGCTATTTTTTTAACTTCTTTTAACTTCTTTTAACTTCTTCCAGGATTTAGACAAAATCAAAA GAAGGAAGGTGCTCACATTCTTAAATATAAGGAGTAAGTCTGCCAGCATTATGAAAGTG AATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAAACAAATGTTTTTGAACATTTAAAAAG TTCAGATGTTAGAAAGTTGAAAGGTTAATGTAACTAGTAGGCAGACCAGCAGACTTTTTT TTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGC TCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCT CTGGTTCTAATTTCTCATTTGCAGTG
155	pMM458 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAAACAA TGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTAGCAGACTTTTTTTTATTGTGAT ATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGTCTCTAGTCAAGT TTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAAT TTCTCATTTGCAGTG
156	pMM459 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGAGTAAGT

		CTGCCAGCATTATGAAAGTGAATCCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTGAAAGAGACCAGCAGACTTTTT TTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGC TCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCT CTGGTTCTAATTTCTCATTTGCAGTG
157	pMM470 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTAATAGGCAGACCAG CAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACAT ATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTC GTCAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
158	pMM471 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTACTAGTGGTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTACTAGGCAGACGAGCA CTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCA AGCCTCTGGTTCTAATTTCTCATTTGCAGTG
159	pMM472 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTACTAGGCAGACGAGCA CTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCA AGCCTCTGGTTCTAATTTCTCATTTGCAGTG
160	pMM473 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTAGACCAGCAGACTTTTTTTTATT GTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGT CAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTT CTAATTTCTCATTTGCAGTG
161	pMM474 ON-switch	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGAGTAAGT CTGCCAGCATTATGAAAGTGAATCCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAACAAA TGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAAGTTAGGCAGACCAGCAGACTTTTTT TTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGC TCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCT CTGGTTCTAATTTCTCATTTGCAGTG
162	pMM567 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTAAAAATAAAAAATAAACTTTTAAAGTACATTAAG ACTATCAACTTAATTTCTGATCATATTTTGTGAAATAAAATAAGTAAAATGCTTGTGAAAC AAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTA TTTTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTGTAGACAAAATCAAAAAGAAGGAAGT GCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCCTTACTTTTGT AAAACCTTTATGGTTTGTGGAAACAAAATGTTTTTGAACATTTAAAAAGTTTCAGATGTTAGA AAGTTGAAAGGTTAATAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCT

		AGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTGTCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCA GTG
163	pMM568 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTGAGATGTTAGAAAAGTTGAAAGTTAATAGGCAGACCAGCA GACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATAT GAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGT CAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG
164	pMM569 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTGAGATGTTAGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTTGCAGTG
165	pMM570 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTGAGATGTTAGAAAAGTTAGGCAGACCAGCAGACTTTTTTTT ATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
166	pMM591 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG GCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTGAGATGTTAGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTTGCAGTG
167	pMM592 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTGAGATGTTAGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
168	pMM593 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAAAATATATAGGAGTAA GTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACA AATGTTTTTGAACATTTAAAAAGTTGAGATGTTAGAAAAGTTACTAGTAGGCAGACCAGCA GACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATAT GAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGT CAAGCCTCTGGTTCTAATTTCTCATTTGCAGTG

169	pMM594 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCCTTATAAGGAGTAAGTCTGCCAGCATT ATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGAAC ATTTAAAAAGTTTCAAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTATT GTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGT CAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTT CTAATTTCTCATTGTCAGTG
170	pMM595 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG AATAATTAAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTTCAAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTGTCAGTG
171	pMM596 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG TAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAG CATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTTG AACATTTAAAAAGTTTCAAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTT ATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTGTCAGTG
172	pMM597 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCC TGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGAACATTTAAAA AGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATG GGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA ACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT TCATTTGTCAGTG
173	pMM598 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCATT TTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCC AGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTT TGAACATTTAAAAAGTTTCAAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTT TTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCT CTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTC TGGTTCTAATTTCTCATTGTCAGTG
174	pMM599 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTG CAAGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTTCAAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTGTCAGTG
175	pMM600 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT

		TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAACAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTGTCAGTG
176	pMM601 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAGCCAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTGTCAGTG
177	pMM642 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCGCATC TTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAAATACTACTTG CTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATA AAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATTTTATAT ATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTA GACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGCA TTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTTGAA CATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTAT TGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGTCTAG TCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT TCTAATTTCTCATTGTCAGTG
178	pMM643 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAAATACTACTT GCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAAT AAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATA TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTAT TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGTCTAG GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTGTCAGTG
179	pMM709 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAAATACTACTT GCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATA AAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATAT ATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTA GACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGCA TTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTTGA CATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTAT TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGTCTAG TCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT TCTAATTTCTCATTGTCAGTG
180	pMM644 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAAATACTACTT GCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAAT AAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATA TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAAACCTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTAT TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGTCTAG GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTGTCAGTG
181	pMM645 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAAATACTACTT GCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAAT AAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATA TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC

		ATTATGAAAGTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTA TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
182	pMM646 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATTTTGTGAAT AAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATA TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTA TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
183	pMM647 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT AAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATA TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTA TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
184	pMM648 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT AAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATA TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTA TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
185	pMM649 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT TATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTA TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
186	pMM650 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTAAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTT AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGC ATTATGAAAGTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGA ACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTA TTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTGCAGTG
187	pMM651 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATAAAATAAGT AAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTAT CTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAAT CAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGAGTAAGTCTGCCAGCTATTATGAAA GTGAATCTTACTTTTGTAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGAACATTTAAA AAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATAT GGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTT

		AACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCGAGTG
188	pMM634 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAATGTTTTG AACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTT ATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTCGAGTG
189	pMM711 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT AACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTT ATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTCGAGTG
190	pMM635 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTGTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTT ATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTCGAGTG
191	pMM636 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTCAGATGTTAGTACTAGTAGGCAGACCAGCAGACTTTTTTTT ATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTCGAGTG
192	pMM637 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTTT TTTTATTGTGATATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTCGAGTG
193	pMM638 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACCTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTA GTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTTCGAGTG
194	pMM639 ON-switch	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA

		GCATCTTTTCCTGACAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAGTTTCAAGTGTGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCGCACTGTACACTCTGACATATGAAGTGCTCT AGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTGTCAGTG
195	pMM640 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAGTTTCAAGTGTGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCT GTTCTAATTTCTCATTGTCAGTG
196	pMM641 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAGTTTCAAGTGTGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAAATTCGTCAAGCCTCTG GTTCTAATTTCTCATTGTCAGTG
197	pMM712 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGT TTTTGAACATTTAAAAGTTTCAAGTGTGAAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCT GTTCTAATTTCTCATTGTCAGTG
198	pMM579 ON-switch	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCCTGACAATTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATAC TACTTGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGT TGAATAAAATAAGTAAAATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGG ATTTTAGACAAAATCGGTGCTCACATTCTTAAATAGGAGTAAGTCTGCCAGCATTATGA AAGTGAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGAACATTTA AAAAGTTTCAAGTGTGAAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTATTGTGAT ATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGT TAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGTTCTAAT TTCTCATTGTCAGTG
199	pMM618 ON-switch	GGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCAGCATCTTTTCCTGAC AATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAATACTACTTGCTTCTCTCTT TATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAA AATGCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCT ATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCCTCCAGGATTTTAGACAAAATCA AAAAGAAGGAAGGTGCTCACATTCTTAAATAGGAGTAAGTCTGCCAGCATTATGAAAGT GAATCTTACTTTTGTAAAACTTTATGGTTTGTGGAAAACAAATGTTTTTGAACATTTAAAAA GTTTCAAGTGTGAAAAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTATTGTGATG GGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTA ACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGTTCTAATTTCT CATTTGTCAGTG
200	pMM619 ON-switch	ACTGGCTATTATCAGGTAAGTAATCACTCAGCATCTTTTCCTGACAATTTTTGTAGTTA TGTGACTTTGTTTTGTAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTTTAAAAG ACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAAATGCTTGTGAAAC AAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTA

		TTTTTTTAACTTCCTTTATTTTCTTCCAGGATTTTAGACAAAATCAAAAAGAAGGAAGGT GCTCACATTCTTAAATAGGAGTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGT AAAACCTTATGGTTTGTGGAACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGA AAGTTACTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCA TACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGCCACAG AGGACATGGTTTAACTGGAAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCAGAGT
201	CMV promoter	CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCAT TGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTC AATGGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGC CAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAG TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA CCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGACTCACGG GGATTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA CGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCATTTGACGCAAATGGGCGGTAGGCG TGTACGGTGGGAGGTCTATATAAGCAGAGCT
202	CAG promoter	ATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTTCATAGCCCAT TATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACG ACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT TCCATTGACGTCAATGGGTGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAA GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTG GCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATT AGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCT CCCCCCCCCTCCCAACCCCAATTTTGTATTTATTTATTTTATTTTGTGACGCGAT GGGGGCGGGGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGG GCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAATCAGAGCGCGCGCTCC GAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGC GCGGCGGGGCGGAGTCTGCTGCGCGCTGCTTCCGCCCTGCGCCCGCTGCGCGCGGCG CCTCGCGCGCGCCCGCCCGGCTCTGACTGACCGGCTTACTCCACAGGTGAGCGGGC GGGACGGCCCTTCTCCTCGGGGCTGTAAATAGCGCTTGGTTTAAATGACGGCTTGTTC TTTCTGTGGCTGCGTGAAGCCTTGAGGGGCTCCGGAGGGCCCTTTGTGCGGGGGGA GCGGCTCGGGGGGTGCGTGCGTGTGTGTGCGTGCGGAGCGCGCGCTGCGGCTCC GCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCG CAGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTTGCGGGGGGGG TGCGAGCGGAACAAAGCCTGCGTGCGGGGTGTGCGTGCGGGGTGAGCAGGGGG TGTGGGCGCGTGGTGGGCTGCAACCCCGCTGCAACCCCGCTCCCGAGTTGTGA GCACGGCCCGGCTTGGGTGCGGGGCTCCGTACGGGGCGTGGCGCGGGGCTCGCCG TGCCGGGCGGGGGGTGCGGCGAGGTGGGGGTGCCGGGCGGGGCGGGGCGCGCTGT GGCGGGGAGGGCTCGGGGAGGGGCGCGGCGGCCCGCGCGCCCGCGCGCTCG CGAGGCGCGGCGAGCGCGAGCCATTGCTTTTATGGTAATCGTGCGAGAGGGCGCAG GGACTTCTTTGTCCCAATCTGTGCGGAGCGGAAATCTGGGAGGCGCGCGCGCACCC CCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCGGCGAGGAAGGAAATGGGCGGG GAGGGCCTTCTGCGCTGCGCGCGCGCGCTCCCTTCTCCAGCTCGGGGCT GTCCGCGGGGGGACGGCTGCTTCCGGGGGGGACGGGGCAGGGCGGGGTTGCGCTTC TGCGGTGTGACCGGCGCG
203	hEF1a promoter	GCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGG GGGAGGGGTGCGCAATTGAACCGGTGCCAGAGGTGGCGCGGGGTAACTGGG AAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGGGAGAACCCTATATA AGTGCAGTAGTCGCCGTGAACGTTCTTTTTGCAACGGGTTTGCGCCAGAACACAGGT AAGTGCCGTGTGTGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCCTG CCTTGAATTACTTCCACGCCCTTGGCTGCACTACGTGATTCTTGATCCCGAGCTTCGG TTGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCCTTAAGGAGCCCTTCGCTCGTGC TTGAGTTGAGGCCTGGCTTGGGCGCTGGGGCGCGCGCTGCGAATCTGGTGGCACCTT CGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCT GCGACGCTTTTTTCTGGCAAGATAGTCTTGAATGCGGGCCAAGATCTGCACACTGG TATTTCCGTTTTTGGGGCCGCGGGCGGCGACGGGGCCGTGCGTCCAGCGCACATGT TCGGCGAGGCGGGGCTGCGAGCGCGGCCACCGAGAATCGGACGGGGTAGTCTCAA GCTGGCCGGCCTGCTCTGGTGCCTGGCCTGCGCGCGCGCTGTATCGCCCCGCCCTGG GCGGCAAGGCTGGCCCGGTGCGCACCAAGTTGCGTGAGCGGAAAGATGGCCGCTTCC GGCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGT GAGTCACCCACACAAAGGAAAGGGCCTTCCGTCTCAGCGCTCGCTCATGTGACTC CACGAGTACCGGGCGCCGTCCAGGCACCTGATTAGTTCTCAAGCTTTTGAGTACGT CGTCTTTAGGTTGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGT GGAGACTGAAGTTAGGCCAGCTTGCACATTGATGTAATTCCTTGGAAATTTGCCCTTT TGAGTTTGGATCTTGGTTCACTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTCTTCCA TTTCAGGTGTCGTGA
204	hUbiC promoter	GGCCTCCGCGCCGGTTTTTGGCGCCTCCCGCGGCGCGCCCCCTCCTCACGGCGAGCG CTGCCACGTACAGCAAGGGCGCAGCGAGCGTCTGATCCTTCCGCCCGGACGCTCAG GACAGCGGCCCGCTGCTCATAAGACTCGGCCCTTAGAACCCAGTATCAGCAGAAGGAC ATTTTAGGACGGGACTTGGGTGACTCTAGGGCACTGTTTTCTTTCCAGAGAGCGGAAC AGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCTGCGGAGGGATCTCCGTGGGGCGGT

		GAACGCCGATGATTATATAAGGACGCGCCGGGTGTGGCACAGCTAGTTCCGTGCGAGC CGGGATTTGGGTCGCGGTTCTTGTGTTGTGGATCGCTGTGATCGTCACTTGGTGAGTAGC GGGCTGCTGGGCTGGCCGGGGCTTTCGTGGCCGCCGGGCCGCTCGGTGGGACGGAA GCGTGTGGAGAGACCGCCAAGGGCTGTAGTCTGGGTCCGCGAGCAAGGTTGCCCTGAA CTGGGGGTTGGGGGGAGCGCAGCAAAATGGCGGCTGTTCCCGAGTCTTGAATGGAAGA CGCTTGTGAGGCGGGCTGTGAGGTCGTTGAAACAAGGTGGGGGGCATGGTGGGCGGC AAGAACCAGGTCTTGAGGCTTCGCTAATGCGGGAAGCTTATTGGGTGAGATG GGCTGGGGCACCATCTGGGGACCTGACGTGAAGTTTGTCACTGACTGGAGAACTCGG TTTGTCGTCTGTTGCGGGGGCGGCAGTTATGGCGGTGCCGTTGGGCAGTGACCCGTA CCTTTGGGAGCGCGCGCCCTCGTCGTGTCGTGACGTACCCGTTCTGTTGGCTTATAAT GCAGGGTGGGGCCACCTGCCGTAGGTGTGCGGTAGGCTTTTCCCGTCGACGACG CAGGGTTCGGGCCCTAGGGTAGGCTCTCCTGAATCGACAGGCGCCGGACCTCTGGTGAG GGGAGGGATAAGTGAGCGCTCAGTTTCTTTGGTCTGTTTATGTACCTATCTTCTTAAGT AGCTGAAGCTCCGGTTTGAACATGCGCTCGGGGTTGGCGAGTGTGTTTTGTGAAGTT TTTTAGGCACCTTTTGAATGTAAATCATTGGGTCAATATGTAATTTTCAGTGTTAGACTA GTAAATGTCCGCTAAATCTGGCCGTTTTTGGCTTTTTTGTAGAC
205	RSV promoter	AATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAGCAACAT GCCTTACAAGGAGAGAAAAAGCACCGTGATGCCGATTGGTGGAAGTAAGGTGGTACG ATCGTGCCCTTATTAGGAAGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAA TTCCGCTTGCAGAGATATTGATTTAAGTGCCTAGCTCGATACAATAAACGCCATTTGA CCATTCACCACATTGGTGTGCAC
206	TK promoter	AAATGAGTCTTCGGACCTCGCGGGGGCCGCTTAAGCGGTGGTTAGGGTTTGTCTGACG CGGGGGGAGGGGAAGGAACGAAACACTCTCAATTGAGGGCGGGCTCGGGTTTGGTC TTGGTGGCCACGGGCACGCAGAGAGCGCCGCGATCCTCTTAAGCACCCCCCGCCCT CCGTGGAGGCGGGGTTTGGTCGCGCGGGTGGTAACGTGGCGGGCCGCTGACTCGGGC GGGTGCGCGCGCCAGAGTGTGACCTTTTCGGTCTGCTCGCAGACCCCCGGGCGGGC CCGCCGCGGGCGGCGACGGGCTCGCTGGGTCTAGGCTCCATGGGACCGTATACGTG GACAGGCTCTGGAGCATCCGCACGACTGCGGTGATATTACCGGAGACCTTCTGCGGGA CGAGCCGGGTACGCGGGCTGACGCGGAGCGTCCGTTGGGCGACAAACACCAGGACGG GGCACAGGTACACTATCTTGTACCCGGAGGCGCGAGGACTGCAGGAGCTTCAGGGA GTGGCGCAGCTGCTTCATCCCCGTGGCCCGTTGCTCGCGTTTGTGGCGGTGTCCCCG GAAGAAATATATTTGCATGCTTTAGTTCTATGATGACACAAACCCCGCCAGCGTCTTG TCATTGGCGAATTCGAACACGCAGATGCAGTGGGGCGGGCGCGGTCCCAGGTCCACTT CGCATATTAAGGTGACGCGTGTGGCCTCGAACCCGAGCGACCCCTGCAGCGACCCGCT TAA
207	PGK promoter	AAGCTTGGGAGTTCGCTTGGCCCTTTTCCAAGGCAGCCCTGGGTTTGGCAGGGGACGC GGCTGCTCTGGGCGTGGTTCCGGGAAACGCAGCGGCGCCGACCCCTGGGTCTCGCACA TTCTTCACGTCCGTTCCGACGCTCACCCGGATCTTCGCCGCTACCCCTTGGGGCCCCC GGCAGCGCTTCTGCTCCGCCCTTAAGTCGGGAAGGTTCTTGGGTTCCGCGGCTGC CGGACGTGACAAACGGAAGCCGCACGTCTCACTAGTACCCTCGCAGACGACAGCGCC AGGGAGCAATGGCAGCGCGCCGACCGCGATGGGCTGTGGCCAATAGCGCTGCTCAG CAGGGCGCGCCGAGAGCAGCGGCCGGGAAGGGCGGTGCGGAGGCGGGGTGTGCG GGCGGTAGTGTGGGCCCTGTTCTGCGCCGCGCGGTGTTCCGCATTCTGCAAGCCTCCG GAGCGCACGTGCGCAGTGGCTCCCTCGTTGACCGAATCACCGACCTCTCTCCCCAG
208	CAG minimal promoter	GCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCC ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGACTTTCCATTGACG TCAATGGGTGGAGTATTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGATCATAT GCCAAGTACGCCCCCTATTGACGTCAATGACGTAAATGGCCCGCCTGGCATTATGCC AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTA TTACCATGGTCGAGGTGAGCCCAAGTTCTGCTTCACTCTCCCATCTCCCCCCTCC CCACCCCCAATTTTGTATTTATTTATTTTAAATATTTTGCAGCGATGGGGCGGGGG GGGGGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGGA GGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCTTTTAT GGCAGGCGGCGGCGGCGGCGGCTATAAAAAGCGAAGCGCGCGGCGGGCG
209	Nluc coding sequence	GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGA CCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGGGGTGTCCGTAA CTCCGATCCAAAGGATTGTCTGAGCGGTGAAATGGGCTGAAGATCGACATCCATGTC ATCATCCCGTATGAAGGTCTGAGCGGCGACCAATGGCCAGATCGAAAAATTTTAA GGTGGTGTACCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGG TAATCGACGGGGTTACGCCGAACATGATCGACTATTTCCGACGGCCGATGAAGGCATC GCCGTGTTGACGGCAAAAAGATCACTGTAAACAGGGACCTGTGGAACGGCAACAAAT TATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACG GAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCG
210	PEST coding sequence	AATTCTCAGCGCTTTCCGCTGAGGTTGAAGAGCAAGCCGCCGGTACATTGCCTATGTC CTGCGCACAAAGAGCGGTATGGACCGGCACCCAGCCGCTTGTGCTTCAGCTCGCATC AACGTC
211	Nluc-PEST coding sequence	GTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCCGGCTACAACCTGGA CCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTCAGAATCTCGGGGTGTCCGTAA CTCCGATCCAAAGGATTGTCTGAGCGGTGAAATGGGCTGAAGATCGACATCCATGTC ATCATCCCGTATGAAGGTCTGAGCGGCGACCAATGGCCAGATCGAAAAATTTTAA GGTGGTGTACCTGTGGATGATCATCACTTTAAGGTGATCCTGCACTATGGCACACTGG TAATCGACGGGGTTACGCCGAACATGATCGACTATTTCCGACGGCCGATGAAGGCATC GCCGTGTTGACGGCAAAAAGATCACTGTAAACAGGGACCTGTGGAACGGCAACAAAT TATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACG GAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCG

		ATCATCCCGTATGAAGGTCTGAGCGGGCGACCAATGGGCCAGATCGAAAAATTTTTAA GGTGGTGTAACCGTGTGGATGATCATCACTTTAAGGTGATCCTGCACACTGGGCACACTGG TAATCGACGGGGTTACGCCGAACATGATCGACTATTTCCGACGGCCGTATGAAGGCATC GCCGTGTTTCGACGGCAAAAAGATCACTGTAACAGGGACCTGTGGAACGGCAACAAAAT TATCGACGAGCGCTGATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACG GAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGAATTCTCACGGCTTTCCGCC TGAGGTTGAAGAGCAAGCCCGCGGTACATTGCCTATGTCCTGCGCAAGAAAGCGGT ATGGACCGGCACCCAGCCGCTTGTGCTTCAGCTCGCATCAACGTC
212	Bipartite NLS	AAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAGAAGAAGCGGAAAGTC TGATAAAAAGTACTCAATAGGTCTGGACATAGGGACAAATTCAGTTGGTTGGGCAGTGAT CACAGATGAATACAAGGTTCTTCTAAGAAGTTTAAGGTGCTGGGTAATACTGATCGACA CTCTATCAAGAAAACTTGATTGGGGCACTTCTCTTCGACTCCGGGGAGACTGCAGAAG CTACACGGCTGAAGAGAACAGCTCGGCGACGATACACCAGACGAAAAACCGCATTGT TATCTTCAAGAAATCTTTTCAATGAAATGGCAAAAGTGGAATGATAGTTTCTTCCATCGGC TTGAGGAGAGTTTCTTGGTGGAGGAAGATAAGAAGCAGCAAAAGGCCCAATATTGGGA AATATAGTTGATGAGGTAGCCTACCATGAGAAATACCCAATCTACCATCTTCGGAAA AACTTGTGGACTCTACAGATAAGGCTGACTTGAGGCTCATTTATCTGGCACTGGCTCAC ATGATAAAATTCAGAGGTCATTTTTTGATCGAGGGTGACCTGAATCCAGATAACTCCGAT GTCGATAAACTTTTATCCAACCTCGTGACAGACATAATCAAGCTTTTGAGGAAAATCCCA TAAACGCTTCTGGGGTTGACGCCAAAGCCATCCTCTCAGCTCGCCTTTCCAAGTCTCGA AGACTGGAAAATCTCATAGCTCAACTCCAGGGGAAAAGAAAAACGGACTGTTCCGGTAA TCTCATAGCTCTTTCAGTCTGACTCCTAATTTTAAGAGCAATTTTGACCTTGCCGAA GACGCTAAACTCTCAGCTCTCAAAGGACACTTATGACGATGATTTGACAAATTTGCTTGC CAAATTTGGTGACCAAGTACGCTGACCTGTTTCTTCCGCAAAAAATCTTTCAGATGCTATT CTCTTGAGTGACATCCTGAGAGTAAATACCGAAATACCAAGGCCCCACTCAGTGCTTC CATGATTAAGCGCTATGACGAGCATCATCAAGACCTTACTCTGTTGAAGGCATTGGTAAG GCAGCAATTGCCAGAAAAATACAAAGAAATATTTTTCGACCAAGTCAAAAAATGGTTATGC AGGTTACATCGACGGTGGCGCCTCACAGGAGGAGTTTTATAAGTTCATCAAACTATCTT GGAGAAAATGGATGGCACCGAGGAACCTCTTGTCAAATTGAACCGGGAAGATTTGCTTA GGAAACAGCGGACATTTGATAACGGCAGCATACCACACCAATCCATCTGGTGAGCTG CATGCAATTTCTCCGGCGACAGGAGGATTTTTACCCTTTCTCAAAGATAATAGAGAAAAAG ATCGAAAAGATTCTCACTTTCCGGATTCCCTACTATGTTGGGCCCTCGCACGAGGCAAT TCTCGGTTTGCTTGATGACAAGAAAGTCTGAAGAGACCATCACTCCCTGGAATTTTGAA GAAGTCGTAGACAAAGGGGCATCAGCACAGTCTTCATAGAGCGGATGACAAATTTTGA TAAGAACCCTTCAAACGAGAAAAGTCTTCAAACAACTCACTCTTGACGAATACTTTACT GTATATAACGAACTGACAAAAGGTTAAATATGTGACCGAGGGCATGCGAAAACCGCCTT CCTTTCCGGTGAGCAAAAAAGGCTATTGTGATCTGTTGTTTAAGACAAATAGAAAAAGT TACTGTGAACAATGAAGGAGGATTACTTTAAGAAAATCGAATGCTTTGATAGCGTGGGA AATCTCTGGCGTGGAGGACCGCTTCAATGCATCATTGGGAACCTTACCATGACCTGCTCA AGATTATCAAAGATAAAGATTTTCTGGATAATGAAGAAAATGAGGACATTCTTGAAGATAT TGTACTCACTCTGACCCTTTTGAGGACAGGGAGATGATTGAAGAAAGACTGAAGACCTA CGCCCATCTGTTGACGACAAAGTAATGAAGCAATTGAAGCGAAGGATACACAGGT GGGGACGGCTTTCTAGGAAGCTGATAAATGGCATCCGGGATAAACAGTCCGGCAAGAC TATACTGGATTTTCTGAAGTCTGACGGTTTTGCCAATCGCAATTTTCATGCAATTGATACAC GACGATTCAGTACCTTTAAGGAGGATATTCAGAAGGCACAGGTTTCAGGGCAAGGTGA CAGTTTGCACGAGCATATTGCAAAATCTCGCAGGGTCTCCCGCATAAAGAAAGGGGATAC TGCAGACAGTTAAGGTCGTAGATGAGCTGGTGAAGGTGATGGGGCGGCACAAGCCCGA GAATATCGTTATTGAAATGGCACGAGAAAAATCAAACAACCTCAAAGGTCAGAAGAACTC ACGAGAACGGATGAAGCGCATTTGAAGAAGGTATCAAGGAACCTCGGTAGCCAAATATTGA AGGAACACCCGTGTTGAAAACACCCAGCTGCAAAACGAGAAGTTGTACCTGTATTATCTGC AGAATGGTAGAGATATGTATGTAGATCAAGAAGTACGACATTAACCGACTTTCTGATTACG ATGTTGACCATATCGTTCCACAATCCTTTCTTAAGGATGATTCCATTGACAATAAGGTCTT GACACGCTCCGATAAGAATCGAGGTAAGTGAATGTACCTTCAGAGGAGGTGGTGA AGAAGATGAAAACTACTGGCGCCAGCTCCTTAATGCCAAGCTCATTACTCAACGAAAAT TTGACAATTTGACCAAGGCTGAACGGGGAGGACTTTTCAGAGCTGGATAAAGCCGGATTT ATTAAGCGCCAACCTGTAGAACTAGGCAATCACAAAGCACGTTGCCAGATCCTCGAT AGTCGAATGAATACCAAGTATGATGAGAATGACAACTCATTAGAGAGGTTAAGGTTATT ACTCTTAAGAGTAAGCTCGTTTCTGATTTTCAGGAAAGATTTTCAATTTTATAAAGTCCGAG AAATCAACAATTACCACCACGCACACGACGCTTACCTCAATGCAGTAGTCGGACAGCT CTCATCAAGAAATATCCCAATTTGAATCCGAATTTGTTTATGGGATTACAAAGTGTAC GATGTAAGGAAAATGATAGCTAAATCAGAGCAGGAGATTGGCAAGGCTACTGCAAAAGTA TTTCTTCTACTCAAACATAATGAATTTCTTTAAACAGAGATTACTCTCGCCAACGGAGAA ATCCGAAAACGGCTCTCATCGAAACAAACGCGAGACTGGAGAAATCGTTTGGGACAA AGGCCGAGATTTTGCTACTGTTAGAAAGGTACTGTCAATGCCTCAAGTTAATATAGTAAA GAAAACCGAGGTACAAACCGGGGGCTTTTCTAAAGAAATCAATACTCCCTAAGCGAAATTC TGATAAGCTCATTGCCAGAAAGAAAGATTGGGACCCAAGAAAGTATGGCGGGTTCTGACT CTCCTACAGTCGCATACAGCGTGTGGTAGTCGCTAAAGTAGAGAAAGGGAAATCAAAA AACTTAAGTCTGTAAAAGAATTGTTGGGTATCACTATAATGGAACGGTCAAGTTTCGAG AAGAACCCCATAGACTTCTGGAGGCCAAAGGATACAAAGAAAGTCAAGAAAGACCTCAT CATCAAGCTCCCAAAATATTCTCTTTTCAGATTGGAACCGGCCGGAAGAGGATGCTCG
213	Cas9 coding sequence	

		CTAGTGCTGGGGAGCTCCAAAAAGGAAACGAGTTGGCACTCCCCAGTAAGTATGTAAAC TTTCTTTACCTGGCAAGCCACTACGAGAAGTTGAAGGGAAGTCTTGAGGATAACGAGCA GAAACAACCTGTTCTGGAGCAACATAAACATTACCTTGACGAGATAATTGAGCAAATTAG CGAGTTTTCAAAGCGCGTAATCTGGCCGACGCAAACTCCGACAAAGTACTTAGCGCAT ACAACAAACACAGGGACAAACCAATCCGAGAACAAAGCTGAAAAATCATACATTTGTTTA CCCTCACAAACTTGGGTGCACCCGACGATTAAAGTACTTTGACACAAACATTGACAGG AAGAGATATACCAGTACCAAGAAGTACTGGATGCCACCTTGATTACCAATCTATTACC GGTCTGTATGAGACCAGAATAGATCTGAGTCAGCTGGGAGGAGAT
214	Linker coding sequence	GGGAGTGG
215	3xSV40_NLS	GCCCCAAGAAAAAAGGAAGGTTGGTAGCGCAGGCAGTGCCGCGAGGGTCTGGTCCAA AAAAAAGCGGAAGGTCGAGGACCCCTAAGAAGAAACGAAAGTGGAC
216	Bipartite NLS-Cas9- Linker-3xSV40_NLS- STOP	AAACGGACAGCCGACGGAAGCGAGTTCCGAGTCACCAAAGAAGAACGGAAAGTCGGGA GTGGTGATAAAAAAGTACTCAATAGGTCTGGACATAGGGACAAATTCAGTTGGTTGGGCA GTGATCACAGATGAATACAAGGTTCTTCTAAGAAGTTTAAGGTGCTGGGTAACTACTGAT CGACACTCTATCAAGAAAACTTGATTGGGGCACTTCTCTCGACTCCGGGGAGACTGC AGAAGCTACACGGCTGAAGAGAACAGCTCGGCGACGATACACAGACGAAAAAACCGC ATTTGTTATCTTCAAGAAATCTTTTCCAATGAAATGCCAAAAGTGGATGATAGTTTCTTCC ATCGGCTTGAGGAGAGTTTCTTGGTGGAGGAAGATAAGAAGCAGGAAAGGCACCCCAATA TTTGGAAATATAGTTGATGAGGTAGCCTACCATGAGAAATACCACTCTACCATCTT CGGAAAAAATCTGTGGACTCTACAGATAAGGCTGACTTGAGGCTCATTTATCTGGCACTG GCTCACATGATAAAATTCAGAGGTCATTTTTGATCGAGGGTGACCTGAATCCAGATAAC TCCGATGTCGATAAACTCTTTATCCAACCTCGTGCAGACATATAATCAGCTCTTTTGAGGAAA ATCCCATAAACGCTTCTGGGTTGACGCCAAAGCCATCCTCTCAGCTCGCCTTTCCAAG TCTCGAAGACTGGAAAATCTCATAGCTCAACTCCAGGGGAAAAAGAAAAACGGACTGTT CGGTAATCTCATAGCTCTTCACTCGGTCTGACTCCTAATTTAAGAGCAATTTTGACCTT GCCGAAGACGCTAAACTTCAGCTCTCAAAGGACACTTATGACGATGATTTGGACAATTTG CTTGCCCAATTTGGTGACCACTACGCTGACCTGTTTTCTTGGCAGAAAAATCTTTCAGAT GCTATTCTCTTGAGTGACATCCTGAGAGTAAATACCGAAATAACCAAGGCCCACTCAGT GCTTCCATGATTAAGCGCTATGACGAGCATCATCAAGACCTTACTCTGTTGAAGGCATTG GTAAGGCAGCAATTGCCAGAAAAATACAAAGAAATATTTTTCGACCAGTCAAAAAATGGT TATGCAGGTTACATCGACGGTGGCGCCTCACAGGAGGAGTTTTATAAGTTTCATCAAACC TATCTTGAGAAAAATGGATGGCACCCGAGGAACCTCTTGCAAAATTGAACCGGGAAGATT TGCTTAGGAAACAGCGGACATTTGATAACGGCAGCATACCAACACCAATCCATCTGGGT GAGCTGCATGCAATTCCTCGCGACAGGAGGATTTTACCCTTTCTCAAAGATAATAGA GAAAAGATCGAAAAAGATTCTCACTTTCCGGATTCCCTACTATTCTTGGGCCCTCGCACA GGCAATTCTCGGTTTGCTTGATGACAAGAAAGTCTGAAGAGACCATCACTCCCTGGAA TTTTGAAGAAGTCGTAGACAAAGGGGCATCAGCACAGTCTTCATAGAGCGGATGACAA ATTTTGATAAGAACCCTTCAAACGAGAAAGTCTTCAAACACTCACTCTTGACGAATA CTTTACTGTATATAACGAACGTACAAAGGTTAAATATGTGACCGAGGGCATCGCAAAACC AGCCTTCTTTCGGTGAGCAAAAAAGGCTATTGTGATCTGTTGTTAAGACAAATAG AAAAGTTACTGTGAACAATTGAAGGAGGATTACTTTAAGAAAAATCGAATGCTTTGATAG CGTGGAATCTCTGGCGTGGAGGACCGTTCAATCGATCATTTGGGAATTAACCATGACC TGCTCAAGATTATCAAAGATAAAGATTTTCTGGATAATGAAGAAAAATGAGGACATTCTTGA AGATATTGTACTCACTCTGACCTTTTGGAGGACAGGAGATGATTGAAGAAAGACTGAA GACCTACGCCCATCTGTTTCGACGACAAAGTAAATGAAGCAATTGAAGCGAAGGAGGTACA CAGGGTGGGACGGCTTCTAGGAAGCTGATAAATGGCATCCGGATAAACAGTCCGG CAAGACTATACTGGATTTTCTGAAGTCTGACGGTTTTGCCAATCGCAATTTTCATGCAATT GATACACGACGATTCACTGACCTTTAAGGAGGATATTCAGAAGGCACAGGTTTCAGGGC AAGGTGACAGTTTGACGAGCATATTGCAAATCTCGCAGGGTCTCCCGCCATAAAGAAG GGGATACTGCAGACAGTTAAGGTCGTAGATGAGCTGGTGAAGGTGATGGGCGGGCACA AGCCCGAGAATATCGTTATTGAAATGGCACGAGAAAAATCAAACAACCTAAAAAGGTCAGA AGAACTCACGAGAACGGATGAAGCGCATTGAAGAAGGTATCAAGGAACCTCGGTAGCCAA ATATTGAAGGAACACCTGTTGAAACACCCAGCTGCAAAACGAGAAGTTGTACCTGTAT TATCTGCAGAATGGTAGAGATATGTATGATGATCAAGAACCTCGACATTAACCGACTTTCT GATTACGATGTTGACCATATCGTTCCACAATCCTTTCTTAAGGATGATTCCATTGACAATA AGGTCTTGACACGCTCCGATAAGAATCGAGGTAAGGATGATAATGTACCTTCAGAGGAG GTGGTGAAGAAGATGAAAACTACTGGCGCCAGCTCCTTAATGCCAAGCTCATTACTCA ACGAAAAATTTGACAATTTGACCAAGGCTGAACGGGGAGGACTTTTCAGAGCTGGATAAAG CCGGATTTATTAAGCGCCAACCTGTAGAAACTAGGCAAAATCACAAAGCACGTTGCCAG ATCCTCGATAGTCGAATGAATACCAAGTATGATGAGAATGACAAACTCATTAGAGAGGTT AAGGTTATTACTCTTAAGAGTAAGCTCGTTTTCTGATTTTCAGGAAAGATTTCAATTTTATAA AGTCCGAGAAAAATCAACAATTACCACACGACACGACGCTTACCTCAATGCAGTAGTCG GGACAGCTCTCATCAAGAAATATCCCAATTTGAATCCGAATTTGTTTATGGGGATTACA AAGTGATCGATGAAGGAAAAATGATAGCTAAATCAGAGCAGGAGATTGGCAAGGCTACT GCAAAGTATTTCTTCTACTCAACATAATGAATTTCTTTAAAAACAGAGATTACTCTGCCA ACGGAGAAATCCGAAAACGGCCTCTCATCGAAACAAACGGCGAGACTGGAGAAATCGTT TGGGACAAAGGCCGAGATTTTGTACTGTTAGAAAGGTAATGTCAATGCCCTCAAGTTAAT ATAGTAAAGAAAACCGAGGTACAAACCGGGGGCTTTTCTAAAGAAATCAATACTCCCTAAG CGAAATTCGTAGTAAGCTCATTGCCAGAAAGAAAGATTGGGACCAAGAAAGTATGGCGG GTTGACTCTCCTACAGTCGCATACAGCGTGTGGTAGTCGCTAAAGTAGAGAAAGGGA

		AATCAAAAAAAGTCTGTAAGGAATTGTTGGGTATCACTATAATGGAACGGTCAA GTTTCGAGAGAAGCCCATAGACTTCCTGGAGGCCAAAGGATACAAAGAAGTCAAGAAA GACCTCATCATCAAGCTCCCAAAATATTCTCTTTTCGAGTTGGAACGGCCGGAAGAG GATGCTCGCTAGTGTGGGGAGCTCCAAAAAGGAACGAGTTGGCACTCCCCAGTAAG TATGTAAGCTTTCTTACCTGGCAAGCCACTACGAGAAGTTGAAGGGAAGTCTGAGGAT AACGAGCAGAAAACACTGTTGCTGGAGCAACATAACATTACCTTGACGAGATAATTGAG CAAATTAGCGAGTTTTCAAAGCGCGTAATCTGGCCGACGCAAACTCGACAAGTACT TAGCGCATACAACAAACAGGGACAAACCAATCCGAGAACAAGCTGAAAACATCATAC ATTTGTTTACCCTCACAACTTGGGTGCACCCGACGATTTAAGTACTTTGACACAACCA TTGACAGGAAGAGATATACCAGTACCAAGAAGTACTGGATGCCACCTTGATTACCAAT CTATTACCGGTCTGTATGAGACCAGAATAGATCTGAGTCAGCTGGGAGGAGATGCCCC AAGAAAAAAGGAAGGTTGGTAGCGCAGGCAGTCCCGCAGGGTCTGGTCCAAAAAAA AGCGGAAGGTCGAGGACCCTAAGAAGAAACGGAAGTGGACTGAA
217	MeCP2 coding sequence	GCCGCCGCCGCCGCCGCCGCCGCCGAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGA GACTGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAA GTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCGTGC AGCCATCAGCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAGCAGAGACATCAGA AGGGTCAGGCTCCGCCCGGCTGTGCCGGAAGCTTCTGCCCTCCCCAAACAGCGGCG CTCCATCATCGTGACCGGGGACCCATGTATGATGACCCCACTCCCTGAAGGCTGG ACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGATTT GATCAATCCCAAGGGAAGCCCTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAA GGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAACCTGGGAGAGGGA GCCCTCCCGGCGAGAGCAGAAACCACTAAGAAGCCCAATCTCCCAAGCTCCAGG AACTGGCAGAGGCCGGGACGCCCAAGGGAGCGGCACCACGAGACCAAGGCGGC CACGTGAGAGGGTGTGACAGGTGAAAGGGTCTCTGGAGAAAAGTCTGGGAAGCTCCTT GTCAAGATGCCTTTTCAAACCTTCGCCAGGGGGCAAGGCTGAGGGGGTGGGGCCCA CATCCACCAAGTCATGGTGATCAAACGCCCGCGCAGGAAGGCTGAGGCCGA CCCTCAGGCCATTCCCAAGAAACGGGGCGGAAAGCCGGGGAGTGTGGTGCCAGCCGC TGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAG ACCGTACTCCCATCAAGAAGCGCAAGACCCTGGGAGACGGTCAGCATCGAGGTCAAGG AAGTGGTGAAGCCCTGCTGGTGTCACCCCTCGGTGAGAAGAGCGGGAAAGGACTGAA GACCTGTAAGAGCCCTGGGCGGAAAGCAAGGAGAGCAGCCCAAGGGCGCAGCAG CAGCGCTCCTCACCCTCCAAAGAGGAGCACCACCACCTCACCACCTCAGAGTCC CCAAAGGCCCGTGCCACTGCTCCACCCCTGCCCCACCTCCACTGAGCCGAGA GCTCCGAGGACCCACAGCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAA AGAGGAGAAGATGCCAGAGGAGGCTCACTGGAGAGCGACGGCTGCCCAAGGAGCC AGCTAAGACTCAGCCCGCGGTTGCCACCGCGCCACGGCCGAGAAAAGTACAAACAC CGAGGGGAGGAGAGCGCAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAG AGGAGCCTGTGGACAGCCGACGCCCGTGACCGAGAGAGTTAGCTGA
218	pMM599 ON-switch with MeCP2 coding sequence	GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAGGTAAGTAATCACTCA GCATCTTTTCTGACAAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAATTTATAAAAATAC TACTTGCTTCTCTTTTATATTACTTTAAAGACTATCAACTAATTTGATCATATTTTGT TGAATAAAATAAGTAAATGTCTTGTGAAACAAATGCTTTTTAACATCCATATAAAGCTAT CTATATATAGCTATCTATATCTATATAGCTATTTTTTTTAACTTCCTTTATTTCTTCCAGG ATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGAGTAAGTCTG CCAGCATTATGAAAGTGAATCTTACTTTTTGTAACCTTTATGTTTTGTGAAAACAAATGT TTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAGACTTT TTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGT GCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGC CTCTGGTTCTAATTTCTCATTTGCACTGGCCGCCGCCGCCGCCGCCGCGCCGAGCGGA GGAGGAGGAGGAGGCGAGGAGGAGAGACTGGAAGAAAAGTCAGAAGACCAGGACCTC CAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGA GAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCACCACTCTGCTGAGCCCGCA GAGGCAGGCAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCGGCTGTGCCGGAA GCTTCTGCCCTCCCCAAACAGCGCGCTCCATCATCCGTGACCGGGGACCCATGTATG ATGACCCACCCCTGCCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCG CTCTGCTGGGAAGTATGATGTGATTTTATCAATCCCCAGGGAAAAGCCTTCGCTCTAA AGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCCTAATGATT TTGACTTCACGGTAACCTGGAGAGGGAGCCCTCCCGCGAGAGCAGAAACCACTAA GAAGCCCAATCTCCAAAGCTCCAGGAACCTGCAGAGGGCCGGGACGCCCAAGG GAGCGGCACACAGAGACCAAGGCGGCCACGTGAGAGGTGTGCAAGTGAAGAAAGGT CCTGGAGAAAAGTCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACCTTCGCCAGGGG GCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTGATGGTGATCAAACGCC CCGGCAGGAAGCGAAAAGCTGAGGCCGACCTCAGGCCATTTCCCAAGAAACGGGGCC GAAAGCCGGGAGTGTGTGGCAGCCGTGCCGCCGAGGCCCAAAAAAGAAAGCGTGA AGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACC CGGGAGACGGTCAGCATCAGAGTCAAGGAAGTGGTGAAGCCCTGCTGGTGTCCACCC TCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAA GGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTACCCCCCAAGGAGCA CCACCACCATCACCACCACTCAGAGTCCCCAAAGGCCCGCTGCCACTGCTCCACCC

		CTGCCCCACCTCCACCTGAGCCCCGAGAGCTCCGAGGACCCACCAGCCCCCTGAGC CCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCAGAGGAGGCTCACT GGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGTTGCCACCGC CGCCACGGCCGAGAAAAGTACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTT TCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGCCGTGA CCGAGAGAGTTAGCTGA
219	45 nucleotide <i>SMN2</i> exon 6 comprising ATG>GGC at positions 82 to 84, and AT>CA 109 to 110 (<i>SMN2</i> exon 6 variant 4)	GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAG
220	45 nucleotide <i>SMN2</i> exon 6 comprising A>C at position 75, ATG>GGC at positions 82 to 84, and AT>CA 109 to 110 (<i>SMN2</i> exon 6 variant 5)	GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAG
221	<i>SMN2</i> exon 6 variant consensus 3	GCAATTTTCX ₁ TGGTACGGCAGTGGCTATCATACTGGCTATTATCAG wherein X ₁ = A or C
222	<i>SMN2</i> exon 6 variant consensus 4	GCX _{1,2} AATTTTCX ₃ TGGTACGGCAGTGGCTATCATACTGGCTATTX ₄₋₅ X ₆₋₇ G wherein X _{1,2} = absent or TT; X ₃ = A or C; X ₄₋₅ = AT or CA; X ₆₋₇ = AT, CT or CA
223	<i>SMN2</i> exon 7 comprising G>A at position 2, deletion of AAAAAGAAGGAA at positions 17-28, and insertion of A after position 48 (exon 7 variant 4)	GATTTTAGACAAAATCGGTGCTCACATTCCTTAAATATAAGGA
224	<i>SMN2</i> exon 7 comprising G>A at position 2, and insertion of TA after position 48 (exon 7 variant 5)	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATAGGA
225	<i>SMN2</i> exon 7 variant consensus 3	GATTTTAGACAAAATCX ₁₋₁₂ GGTGCTCACATTCCTTAAATAX ₁₃₋₁₅ GGA wherein X ₁₋₁₂ = absent or AAAAAGAAGGAA; X ₁₃₋₁₅ = TAA or absent
226	<i>SMN2</i> exon 7 variant consensus 4	GATTTTXX ₁₋₆ AX ₇₋₈ CAAAATCX ₉₋₂₀ GGTGCTCACATTCCTTAAATAX ₂₁₋₂₃ GGA wherein X ₁₋₆ = absent or GCCACC; X ₇₋₈ = GA or TG; X ₉₋₂₀ = absent, AAAAAGAAGGAA or AAAAGAAGGAA; X ₂₁₋₂₃ = TAA or absent
227	277 nucleotide <i>SMN2</i> intron 7 (intron 7 variant 2)	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTTGGAA AACAAATGTTTTGAACATTTAAAAAGTTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAAT TCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAG
228	260 nucleotide <i>SMN2</i> intron 6, comprising deletion of CCTT at positions 5762-5765 of intron 6, and A>C 4 positions from the 3' end (position 5766 of intron 6) (intron 6 variant 3)	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTT CTGATCATATTTTGTGTAATAAAATAAGTAAATGTCTTGTAACAAAATGCTTTTTTAACA TCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTT TATTTTCCAG
229	<i>SMN2</i> intron 6 variant consensus	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTT CTGATCATATTTTGTGTAATAAAATAAGTAAATGTCTTGTAACAAAATGCTTTTTTAACA TCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTT TATTTTCCX ₁₋₄ AG wherein X ₁₋₆ = absent or TTCC
230	<i>SMN2</i> exon 6 variant 6	GATTCTCTTGATGATGCTGATGCTTTGGGAAGTATGTTAATTTTCATGGTACATGAGTGGC

		TATCATACTGGCTATTATATG
231	SMN2 exon 6 variant 7	AGTATGTTAATTTTCATGGTACATGAGTGGCTATCATACTGGCTATTATATG
232	SMN2 exon 6 variant 8	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGAIGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATG
233	SMN2 exon 6 variant 9	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATTTG
234	SMN2 exon 6 variant 10	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGAIGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATATG
235	SMN2 exon 6 variant 11	ATAATTCCCCCACCACCTCCCATATGTCCAGATTCTCTTGATGATGCTGATGCTTTGGGA AGTGGCTTAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATTAG
236	SMN2 exon 6 variant 12	ACTGGCTATTATCAG
237	SMN2 exon 6 variant 13	GGCAGTGGCTATCATACTGGCTATTATCAG
238	SMN2 intron 6 variant 4	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATAC AACTGTCTGAGGCTTAATTTACTCTTGCATTGTCCCTAAGTATAATTTTAGTTAATTTTAAA AAGCTTTTCATGCTATTGTTAGATTATTTTGAATTGATTGTTCTTTGGGATAACCTTTAAAGT ACATTAAGAACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGT CTTGTAACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATC TATATAGCTATTTTTTAACTTCTTTATTTTCTTACAG
239	SMN2 intron 6 variant 5	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACGTAATGTCTTGTAACAAATGCTTTTTAACATCCATATAAAGC TATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCTTTATTTTCTTCC AG
240	SMN2 intron 6 variant 6	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAATAC AACTGTCTGAGGCTTAATTTACTCTTGCATTGTCCCTAAGTATAATTTTAGTTAATTTTAAA AAGCTTTTCATGCTATTGTTAGATTATTTTGAATTGAAATTTACTTTTT CTAAATAATGTTTTAATCTCTGATTTGAAATTGATTGTAGGGAATGGAAAGATGGGATAA TTTTTCATAAATGAAAAATGAAATCTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTTGC TCTGTTGCCAGGCTGGAGTGCAATGGCGTGATCTTGGCTCACAGCAAGCTCTGCCTCC TGGATTACGCCATTCTCTGCCTCAGCCTCAGAGGTAGCTGGGACTACAGGTGCCTGC CACCACGCCTGGCTAGCTGGGATTAGAGGTCCCAACCATGCCTGGCTAATTTTTTTG TACTTTTCAGTAGAAACGGGGTTTTGCCATGTTGGCCAGGCTGTTCTCGAACTCCTGAGC TCAGGTGATCCAACCTGTCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTG TGCTAGCCTGAGCCACCACGCCGGCCTAATTTTTAAATTTTTTTGTAGAGACAGGGTCTC ATTATGTTGCCAGGGTGGTGTCAAGCTCCAGGTCTCAAGTGATCCCCCTACCTCCGCC TCCCAAGTTGTGGGATTGTAGGCATGAGCCACTGCAAGAAAACCTTAACTGCAGCCTA ATAATTTGTTTCTTTGGGATAACTTTTAAAGTACATTAAAGACTATCAACTTAATTTCTGA TCATATTTTGTGAATAAAATAAGTAAATGTCTTGTAACAAATGCTTTTTAACATCCA TATAAGCTATCTATATAGCTATCTATATCTATATAGCTATTTTTTTTTAACTTCTTATT TTCTTCCAG
241	SMN2 intron 6 variant 7	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTAAAAATGACATTAAGAACTAT CAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTAACAAAT GCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTT TTAACTTCTTTATTTTCTTCCAG
242	SMN2 intron 6 variant 8	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAATTAAGTACATTA AAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTA AAACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATA GCTATTTTTTTAACTTCTTTATTTTCTTCCAG
243	SMN2 intron 6 variant 9	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAACTTTTA AAGTACATTAAGAACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAA ATGCTTTGTGAACAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTA TATCTATATAGCTATTTTTTTAACTTCTTTATTTTCTTCCAG
244	SMN2 intron 6 variant 10	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTAAAAATAAAAAATAAAAAAGG GATAACTTTTAAAGTACATTAAGAACTATCAACTTAATTTCTGATCATATTTTGTGAATA AAATAAGTAAATGTCTTGTAACAAATGCTTTTTAACATCCATATAAAGCTATCTATAT ATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCTTTATTTTCTTCCAG
245	SMN2 intron 6 variant 11	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTT CTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTAACAAATGCTTTTTAAACA TCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTTTAACTTCTT TATTTTCTTCCA
246	SMN2 intron 6 variant 12	GTAAGTAATCGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAAT TATAAAATACTACTTGTCTCTCTTTATATTACTTTAAAGACTATCAACTTAATTTCTGAT CATATTTTGTGAATAAAATAAGTAAATGTCTTGTAACAAATGCTTTTTAACATCCAT

		ATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
247	SMN2 intron 6 variant 13	GTAAGTAATCACTCAGCATCTTTTCAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
248	SMN2 intron 6 variant 14	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGATGTGACTTTGTTTTGTAAATTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
249	SMN2 intron 6 variant 15	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTGTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
250	SMN2 intron 6 variant 16	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
251	SMN2 intron 6 variant 17	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
252	SMN2 intron 6 variant 18	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
253	SMN2 intron 6 variant 19	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
254	SMN2 intron 6 variant 20	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTAAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
255	SMN2 intron 6 variant 21	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTTGTAGTTATGTGACTTTGTTTTGTAAATTTATAAAATACTACTTGGCTTCTCTCTTTATATTACTTTAAAAGACTATCAACTTAATTTCTGATAAAATAAGTAAATGTCTTGTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTCCAG
256	SMN2 exon 7 variant 6	GGTTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
257	SMN2 exon 7 variant 7	GATTTTAGACAAAATCAAAAAGAAGGTAGGTGCTCACATTCCCTTAAATATAAGGA
258	SMN2 exon 7 variant 8	GATTTTAGACAAAATCAAAAAGAAGGATGGTGCTCACATTCCCTTAAATATAAGGA
259	SMN2 exon 7 variant 9	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
260	SMN2 exon 7 variant 10	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
261	SMN2 exon 7 variant 11	GATTTTAGACAAAATCAAAATGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
262	SMN2 exon 7 variant 12	GATTTTAGACAAAATCAAAAAGAGGGAAGGTGCTCACATTCCCTTAAATATAAGGA
263	SMN2 exon 7 variant 13	GATTTTAGACAAAATCAAAAAGAAGGAAGGAGCACACATTCCCTTAAATATAAGGA
264	SMN2 exon 7 variant 14	GATTTTAGAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
265	SMN2 exon 7 variant 15	GATTTTAGAGAAGGAAGGTGCTTAAATATAAGGA
266	SMN2 exon 7 variant 16	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
267	SMN2 exon 7 variant 17	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATGATGA
268	SMN2 exon 7 variant 18	GATTTTAGACAAAATCAAAAAGGTGCTCACATTCCCTTAAATATAAGGA
269	SMN2 exon 7 variant 19	GATTTTAGACAAAATCCTCACATTCCCTTAAATATAAGGA
270	SMN2 exon 7 variant 20	GATTTTAGACTCACATTCCCTTAAATATAAGGA
271	SMN2 exon 7 variant 21	GATTTTAGACAAAATCGTGCTCACATTCCCTTAAATATAAGGA
272	SMN2 exon 7 variant 22	GATTTTAGACAAAATCTGCTCACATTCCCTTAAATATAAGGA
273	SMN2 exon 7 variant 23	GATTTTAGACAAAATCGCTCACATTCCCTTAAATATAAGGA
274	SMN2 exon 7 variant 24	GATTTTAGACAAAATCAAGGTGCTCACATTCCCTTAAATATAAGGA
275	SMN2 exon 7 variant 25	GATTTTAGACAAAATCGAAGGTGCTCACATTCCCTTAAATATAAGGA
276	SMN2 exon 7 variant 26	GATTTTAGACAAAATCGGAAGGTGCTCACATTCCCTTAAATATAAGGA
277	SMN2 exon 7 variant 27	GATTTTAGACAAAATCGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA

278	SMN2 exon 7 variant 28	GATTTTAGACAAAATCTCACATTCCCTTAAATATAAGGA
279	SMN2 exon 7 variant 29	GATTTTAGACAAAATCACATTCCCTTAAATATAAGGA
280	SMN2 exon 7 variant 30	GATTTTAGACAAAATCATTCCCTTAAATATAAGGA
281	SMN2 exon 7 variant 31	GATTTTAGACAAAATCGGTGCTTAAATATAAGGA
282	SMN2 exon 7 variant 32	GATTTTAGACAAAAGGTGCTCACATTCCCTTAAATATAAGGA
283	SMN2 exon 7 variant 33	GATTTTAGACGGTGCTCACATTCCCTTAAATATAAGGA
284	SMN2 exon 7 variant 34	GATTTTAGAGGTGCTCACATTCCCTTAAATATAAGGA
285	SMN2 exon 7 variant 35	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTTAAATATAAGGA
286	SMN2 exon 7 variant 36	GATTTTAGCTTAAATATAAGGA
287	SMN2 exon 7 variant 37	GATTTTAGACAAAATCAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
289	SMN2 exon 7 variant 38	GATTTTAGACAAAATCAAAAAGGCACATTCCCTTAAATATAAGGA
290	SMN2 exon 7 variant 39	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
291	SMN2 exon 7 variant 40	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTTAAATATAAGGA
292	SMN2 exon 7 variant 41	GATTTTAGACAAAATCGGTGCTCACATTCCCTTAAATAGGA
293	SMN2 exon 7 variant 42	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATATAGGA
294	SMN2 exon 7 variant 43	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTTAAATAGGA
295	SMN2 exon 7 variant 44	GAATAATTAAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGA
296	SMN2 exon 7 variant 45	TTTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATAGGA
297	SMN2 exon 7 variant 46	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAACAGGA
298	SMN2 exon 7 variant 47	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAGCCAGGA
299	SMN2 exon 7 variant 48	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCCACCATGCTTAAATATAAGGA
301	SMN2 exon 7 variant 49	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAGCCACCATGATAAGGA
302	SMN2 exon 7 variant 50	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAACAACATGGA
303	SMN2 exon 7 variant 51	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACAGCCACCATGAATATAAGGA
304	SMN2 exon 7 variant 52	GATTTTAGAGCCACCATGCAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
305	SMN2 exon 7 variant 53	GATTTTAGACAAAATAGCACCATGCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
306	SMN2 exon 7 variant 54	GATTTTAGACAAAATAGCACCATGGAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
307	SMN2 exon 7 variant 55	GATTTTAGACAAAATCAAAAATGGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
308	SMN2 exon 7 variant 56	GATTTTAGACAAAATCACCATGGAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
309	SMN2 exon 7 variant 57	GATTTTAGACAAAATATGGAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
310	SMN2 exon 7 variant 58	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
311	SMN2 exon 7 variant 59	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATGGTAAGGA
312	SMN2 exon 7 variant 60	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGGA
313	SMN2 exon 7 variant 61	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTATGGATATAAGGA
314	SMN2 exon 7 variant 62	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATATAAGATGA
315	SMN2 exon 7 variant 63	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATGCTTAAATATAAGGA
316	SMN2 exon 7 variant 64	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATGAGGA
317	SMN2 exon 7 variant 65	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATGATAAGGA
318	SMN2 exon 7 variant 66	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATGGCCTTAAATATAAGGA
319	SMN2 exon 7 variant 67	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTAACATGGCCTTAAATATAAGGA
320	SMN2 exon 7 variant 68	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATATGCCTTAAATATAAGGA
321	SMN2 exon 7 variant 69	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCATGTTAAATATAAGGA
322	SMN2 exon 7 variant 70	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTATGAATATAAGGA
323	SMN2 exon 7 variant 71	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTATGGAATATAAGGA
324	SMN2 exon 7 variant 72	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATTAAGATGA
325	SMN2 exon 7 variant 73	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAAATGAGGA
326	SMN2 exon 7 variant 74	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTGAAATGAGGA
327	SMN2 exon 7 variant 75	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCAAATGAGGA
328	SMN2 exon 7 variant 76	GATTTTAGACAAAATCAAAAAGAAGGAAGGAATAGGA
329	SMN2 exon 7 variant 77	GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCCTTAGCCACCATGGA
330	SMN2 intron 7 variant 3	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAA AACAAATGTTTTGAACGTACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGCCACAG AGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTCAG
331	SMN2 intron 7 variant 4	GTAAGTCTGCTCCTCTTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTAAAAAGTTCAGATGTTAGAAGTTTATTTATTTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTGCAG
332	SMN2 intron 7 variant 5	GTAAGTCTGCAAGGCTTATGAAAGTGAATCTTACTTTTGTAAACTTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTAAAAAGTTCAGATGTTAGAAGTTTATTTATTTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT

		CATTTGCAG
333	SMN2 intron 7 variant 6	GTAAGTCTGCTAGTCTTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTGCAG
334	SMN2 intron 7 variant 7	GTAAGTCTGCTTTTTCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTG GAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATA TGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTT TAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATT TCTCATTTGCAG
335	SMN2 intron 7 variant 8	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTG GAAAACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATA TGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTT TAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATT TCTCATTTGCAG
336	SMN2 intron 7 variant 9	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGGTTAAAGGTTAATGTAA ACAATCAATATTAAGAATTTTGTATGCCAAAACATTAGATAAAAGGTTAATCTACATCCC TACTAGAATTCTCATACTTAACTGGTTGGTTGTGTGGAAGAAACATACTTTTCAATAAAG AGCTTTAGGATATGATGCCATTTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTAT TGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAG TCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT TCTAATTTCTCATTTCCAG
337	SMN2 intron 7 variant 10	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGGTTAAAGGTTAATGTAA ACAATCAATATTAAGAATTTTGTATGCCAAAACATTAGATAAAAGGTTAATCTACATCCC TACTAGAATTCTCATACTTAACTGGTTGGTTGTGTGGAAGAAACATACTTTTCAATAAAG AGCTTTAGGATATGATGCCATTTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTAT TGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAG TCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT CTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTC TGGTTCTAATTTCTCATTTGCAG
338	SMN2 intron 7 variant 11	GTAAGTGGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAAA ACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGGG ATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACT TGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTC ATTTGCAG
339	SMN2 intron 7 variant 12	GTAAGATTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTGCAG
340	SMN2 intron 7 variant 13	GTAAGTGTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTGCAG
341	SMN2 intron 7 variant 14	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTT AACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAG
342	SMN2 intron 7 variant 15	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAG
343	SMN2 intron 7 variant 16	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAG
344	SMN2 intron 7 variant 17	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACAACATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAG
345	SMN2 intron 7 variant 18	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCTTTTTTATTGTGATATGG

		GATAACCTAGGCATACTGCACTGTACACTCTGTGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAG
346	SMN2 intron 7 variant 19	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGACATATGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAG
347	SMN2 intron 7 variant 20	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTATATGAAGTGCTCTAGTCAAGTTTAACTGGT GTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTT GCAG
348	SMN2 intron 7 variant 21	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTATTGTGATATGG GATAACCTAGGCATACTGCACTGTACACTCTGGCATATGAAGTGCTCTAGTCAAGTTTAA CTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTGCAG
349	SMN2 intron 7 variant 22	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAACCTTTTTTATTGTGATATGG GATAACCTAGGCATACTTGAATTAGGAGGGAGGATTTCATTGTGCTCTAGTCAAGTTTAA CTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCT CATTTGCAG
350	SMN2 intron 7 variant 23	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTAGCAGACTTTTTTTTAT TGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAG TCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT TCTAATTTCTCATTTGCAG
351	SMN2 intron 7 variant 24	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTGAAAGAGACCAGCAGA CTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCA AGCCTCTGGTTCTAATTTCTCATTTGCAG
352	SMN2 intron 7 variant 25	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTGAAAGGTTAATAGGCA GACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACT CTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACT GGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAG
353	SMN2 intron 7 variant 26	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTGAAAGGTTAATGTAA CTAGTAGGCAGACCAGCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGC ACTGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACA TGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAG
354	SMN2 intron 7 variant 27	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGCAGACTTTTTTTTATTGTGA TATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAGTCAAG TTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGTTCTAA TTTCTCATTTGCAG
356	SMN2 intron 7 variant 29	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTAGACCAGCAGACTTTT TTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTG CTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCC TCTGGTTCTAATTTCTCATTTGCAG
357	SMN2 intron 7 variant 30	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTTAGGCAGACCAGCAGA CTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCA AGCCTCTGGTTCTAATTTCTCATTTGCAG
358	SMN2 intron 7 variant 31	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAAT TCGTCAAGCCTCTGGTTCTAATTTCTCATTTGCAG
359	SMN2 intron 7 variant 32	GTAAGTCTGCCAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGT AAGCCTCTGGTTCTAATTTCTCATTTGCAG
360	SMN2 intron 7 variant 33	GTAAGTCTGCAAGCATTATGAAAAGTGAATCTTACTTTTGAAAACCTTTATGGTTTGTGGAA

		AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAAT TCGTCAAGCCTCTGGTCTAATTTCTCATTTCAG
361	SMN2 intron 7 variant 34	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA ATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAG ACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATG AAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
362	SMN2 intron 7 variant 35	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAG ACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATG AAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
363	SMN2 intron 7 variant 36	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCAGCAG CTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
364	SMN2 intron 7 variant 37	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGTACTAGTAGGCAGACCAGCAG ACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATG AAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
365	SMN2 intron 7 variant 38	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
366	SMN2 intron 7 variant 39	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAGGCATACTGCACTGTACACTCTGACATATGA AGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
367	SMN2 intron 7 variant 40	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATG AAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
368	SMN2 intron 7 variant 41	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
369	SMN2 intron 7 variant 42	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
370	SMN2 intron 7 variant 43	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGAAAACTTTATGGTTTGTGGAA AACAAATGTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAGGCAGACCA GCAGACTTTTTTTTATTGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGAC ATATGAAGTGCTCTAGTCAAGTTTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTC AAGCCTCTGGTCTAATTTCTCATTTCAG
371	SMN2 exon 8 variant 3	GAAATGCTGGCATAGAG
372	SMN2 exon 8 variant 4	GAAATGCTGGC
373	SMN2 exon 8 variant 5	GAAATGCT
374	SMN2 exon 8 variant 6	GAAAT
375	SMN2 intron 6 variant 22	GTAAGTAATCACTCAGCATCTTTTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGT AAATTTATAAAATACTACTTGCTTCTCTTTTATATTACTTTAAAGACTATCAACTTAATTT CTGATCATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAACAAAATGCTTTTTAACA TCCATATAAAGCTATCTATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTT TATTTTCCCTCCAT
376	SMN2 intron 6 variant 23	GTAAGTAATCACTCAGCATCTCTGACAATTTTTTGTAGTTATGTGACTTTGTTTTGTAAATT TATAAAATACTACTTGCTTCTCTTTTATATTACTTTAAAGACTATCAACTTAATTTCTGAT CATATTTTGTGAATAAAATAAGTAAATGTCTTGTGAACAAAATGCTTTTTAACATCCAT ATAAAGCTATCTATATAGCTATCTATATCTATATAGCTATTTTTTTAACTTCCTTTATTT TCCTTCAG
377	3' 162 positions of SMN2	TTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTT

	intron 6, comprising deletion of TTCC from 6 positions upstream of 3' end	GTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTAT ATAGCTATTTTTTTAACTTCCTTTATTTTCCAG
378	3' 162 positions of <i>SMN2</i> intron 6, consensus	TTAAAAGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCTT GTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCTATATCTAT ATAGCTATTTTTTTAACTTCCTTTATTTTCCX ₁₋₄ AG wherein X ₁₋₄ = absent or TTCC
379	Positions 1 to 102 of <i>SMN2</i> intron 7, comprising C>G at position 7	GTAAGTGTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAA
380	Positions 1 to 102 of <i>SMN2</i> intron 7, consensus	GTAAGTX ₁ TGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAA wherein X ₁ = C or G
381	<i>SMN2</i> intron 7 variant 44	GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGTAACCTTTATGGTTTGTGGAA AACAAATGTTTTTGAACATTTAAAAAGTTCAGATGTTAGAAAGTTGAAAGGTTAATGTAAA ACAATCAATATTAAAGAATTTTGATGCCAAAACATTAGATAAAAGGTTAATCTACATCCC TACTAGAATTCTCATACTTAACTGGTTGGTTGTGTGGAAGAAACATACTTTTCAACAATAAG AGCTTTAGGATATGATGCCATTTATATCACTAGTAGGCAGACCAGCAGACTTTTTTTTAT TGTGATATGGGATAACCTAGGCATACTGCACTGTACACTCTGACATATGAAGTGCTCTAG TCAAGTTAACTGGTGTCCACAGAGGACATGGTTTAACTGGAATTCGTCAAGCCTCTGGT TCTAATTTCTTTTTTTCAG

Numbered statements

The following numbered paragraphs (paras) describe particular aspects and embodiments of the present disclosure:

5

1. A polynucleotide comprising in 5' to 3' order:

(i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:222, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;

10

(ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:378 at its 3' end, and consisting of fewer than 1044 nucleotides;

15

(iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:226;

(iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:380 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;

20

(v) a fifth nucleotide sequence: (a) consisting of the dinucleotide 'GA', 'TG' or 'TT', or (b) comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:19, wherein the nucleotide sequence comprises 'GA', 'TG' or 'TT' at the positions corresponding to positions 1 and 2 of SEQ ID NO:19, or (c) encoding a polypeptide of interest, and comprising 'GA', 'TG' or 'TT' at positions 1 and 2; and

25

(vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a) or (v)(b), a sixth nucleotide sequence encoding a polypeptide of interest; wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

2. The polynucleotide according to para 1, wherein when the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.
- 5 3. The polynucleotide according to para 1 or para 2, wherein the first nucleotide sequence comprises a nucleotide sequence according to SEQ ID NO:222.
- 10 4. The polynucleotide according to any one of paras 1 to 3, wherein the first nucleotide sequence comprises, or consists of, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:106, SEQ ID NO:219 or SEQ ID NO:220.
- 15 5. The polynucleotide according to any one of paras 1 to 4, wherein the second nucleotide sequence comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 or SEQ ID NO:377 at its 3' end.
6. The polynucleotide according to any one of paras 1 to 5, wherein the second nucleotide sequence consists of fewer than 500 nucleotides.
- 20 7. The polynucleotide according to any one of paras 1 to 6, wherein the second nucleotide sequence comprises, or consists of, SEQ ID NO:9, SEQ ID NO:10 or SEQ ID NO:228.
8. The polynucleotide according to any one of paras 1 to 7, wherein the third nucleotide sequence consists of a nucleotide sequence according to SEQ ID NO:226.
- 25 9. The polynucleotide according to any one of paras 1 to 8, wherein the third nucleotide sequence consists of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:223 or SEQ ID NO:224.
10. The polynucleotide according to any one of paras 1 to 9, wherein the fourth nucleotide sequence comprises SEQ ID NO:15 or SEQ ID NO:379 at its 5' end, and comprises SEQ ID NO:16 at its 3' end.
- 30 11. The polynucleotide according to any one of paras 1 to 10, wherein the fourth nucleotide sequence consists of fewer than 500 nucleotides.
12. The polynucleotide according to any one of paras 1 to 11, wherein the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:227 or SEQ ID NO:340.
- 35 13. The polynucleotide according to any one of paras 1 to 12, wherein the fifth nucleotide sequence: (a) consists of the dinucleotide 'GA' or 'TG', or (b) comprises, or consists of, a nucleotide sequence according to SEQ ID NO:19.
- 40

14. The polynucleotide according to any one of paras 1 to 13, wherein the fifth nucleotide sequence: (a) consists of the dinucleotide 'GA' or 'TG', or (b) comprises, or consists of, SEQ ID NO:20 or SEQ ID NO:21.
- 5 15. The polynucleotide according to any one of paras 1 to 14, wherein the polynucleotide comprises a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence selected from the group consisting of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:111, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:164 and SEQ ID NO:171.
- 10 16. The polynucleotide according to any one of paras 1 to 15, wherein the polynucleotide further comprises a promoter sequence 5' to the start codon.
17. The polynucleotide according to any one of paras 1 to 16, wherein the polynucleotide further comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.
- 15 18. The polynucleotide according to any one of paras 1 to 17, wherein the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.
19. A vector comprising a polynucleotide according to any one of paras 1 to 18.
- 20 20. The vector according to para 19, wherein the vector is an adeno-associated virus (AAV) vector.
21. A pharmaceutical composition comprising a polynucleotide according to any one of paras 1 to 18, or a vector according to para 19 or para 20, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.
- 25 22. A cell comprising a polynucleotide according to any one of paras 1 to 18, or a vector according to para 19 or para 20.
- 30 23. The cell according to para 22, wherein the cell further comprises a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
24. A method for modifying a cell to express a polypeptide of interest, comprising:
- 35 (i) introducing into a cell a polynucleotide according to any one of paras 1 to 18, or a vector according to para 19 or para 20; and
- (ii) subsequently contacting the cell with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
- 40 25. A method for expressing a polypeptide of interest in a cell, comprising contacting a cell according to para 22 with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

26. A polynucleotide according to any one of paras 1 to 18, a vector according to para 19 or para 20, or a pharmaceutical composition according to para 21, for use in a method of medical treatment or prophylaxis.

5

27. A polynucleotide according to any one of paras 1 to 18, a vector according to para 19 or para 20, or a pharmaceutical composition according to para 21, for use in treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.

10

28. Use of a polynucleotide according to any one of paras 1 to 18, a vector according to para 19 or para 20, or a pharmaceutical composition according to para 21, in the manufacture of a medicament for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.

15

29. A method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising administering to a subject a polynucleotide according to any one of paras 1 to 18, a vector according to para 19 or para 20, or a pharmaceutical composition according to para 21.

20

30. The polynucleotide, vector, or pharmaceutical composition for use according to para 27, the use according to para 28, or the method according to para 29, wherein treating or preventing the disease or condition further comprises administering to the subject a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

25

31. The polynucleotide, vector, or pharmaceutical composition for use according to para 27 or para 30, the use according to para 28 or para 30, or the method according to para 29 or para 30, wherein the disease or condition is a disease or condition characterised by deficiency of the polypeptide of interest.

30

32. A kit, comprising:

- (i) a polynucleotide according to any one of paras 1 to 18, a vector according to para 19 or para 20, or a pharmaceutical composition according to para 21; and
- (ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

35

33. A polynucleotide comprising in 5' to 3' order:

- (i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:1, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;

40

- (ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 at its 3' end, and consisting of fewer than 1044 nucleotides;

(iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:26, wherein the third nucleotide sequence comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12;

(iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;

(v) a fifth nucleotide sequence: (a) consisting of the trinucleotide 'GAG' or (b) encoding a polypeptide of interest, and comprising 'GAG' at positions 1 to 3; and

(vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a), a sixth nucleotide sequence encoding a polypeptide of interest; wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

34. The polynucleotide according to para 33, wherein when the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.

35. The polynucleotide according to para 33 or para 34, wherein the first nucleotide sequence comprises a nucleotide sequence according to SEQ ID NO:1.

36. The polynucleotide according to any one of paras 33 to 35, wherein the first nucleotide sequence comprises, or consists of, SEQ ID NO:3.

37. The polynucleotide according to any one of paras 33 to 36, wherein the second nucleotide sequence comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 at its 3' end.

38. The polynucleotide according to any one of paras 33 to 37, wherein the second nucleotide sequence consists of fewer than 500 nucleotides.

39. The polynucleotide according to any one of paras 33 to 38, wherein the second nucleotide sequence comprises, or consists of, SEQ ID NO:10.

40. The polynucleotide according to any one of paras 33 to 39, wherein the third nucleotide sequence consists of a nucleotide sequence according to SEQ ID NO:26.

41. The polynucleotide according to any one of paras 33 to 40, wherein the third nucleotide sequence consists of SEQ ID NO:13 or SEQ ID NO:27.

42. The polynucleotide according to any one of paras 33 to 41, wherein the fourth nucleotide sequence comprises SEQ ID NO:15 at its 5' end, and comprises SEQ ID NO:16 at its 3' end.

43. The polynucleotide according to any one of paras 33 to 42, wherein the fourth nucleotide sequence consists of fewer than 500 nucleotides.
- 5 44. The polynucleotide according to any one of paras 33 to 43, wherein the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17.
45. The polynucleotide according to any one of paras 33 to 44, wherein the polynucleotide comprises at least 80% sequence identity to SEQ ID NO:28 or SEQ ID NO:29.
- 10 46. The polynucleotide according to any one of paras 33 to 45, wherein the polynucleotide further comprises a promoter sequence 5' to the start codon.
47. The polynucleotide according to any one of paras 33 to 46, wherein the polynucleotide further comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.
- 15 48. The polynucleotide according to any one of paras 33 to 47, wherein the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.
49. A vector comprising a polynucleotide according to any one of paras 33 to 48.
- 20 50. The vector according to para 49, wherein the vector is an adeno-associated virus (AAV) vector.
51. A pharmaceutical composition comprising a polynucleotide according to any one of paras 33 to 48, or a vector according to para 49 or para 50, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.
- 25 52. A cell comprising a polynucleotide according to any one of paras 33 to 48, or a vector according to para 49 or para 50.
- 30 53. The cell according to para 52, wherein the cell further comprises a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
54. A method for modifying a cell to express a polypeptide of interest, comprising introducing into a cell a polynucleotide according to any one of paras 33 to 48, or a vector according to para 49 or para 50.
- 35 55. A method for inhibiting expression of a polypeptide of interest in a cell, comprising contacting a cell according to para 52 with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
- 40

56. A polynucleotide according to any one of paras 33 to 48, a vector according to para 49 or para 50, or a pharmaceutical composition according to para 51, for use in a method of medical treatment or prophylaxis.
- 5 57. A polynucleotide according to any one of paras 33 to 48, a vector according to para 49 or para 50, or a pharmaceutical composition according to para 51, for use in treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
- 10 58. Use of a polynucleotide according to any one of paras 33 to 48, a vector according to para 49 or para 50, or a pharmaceutical composition according to para 51, in the manufacture of a medicament for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
- 15 59. A method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising administering to a subject a polynucleotide according to any one of paras 33 to 48, a vector according to para 49 or para 50, or a pharmaceutical composition according to para 51.
- 20 60. The polynucleotide, vector, or pharmaceutical composition for use according to para 57, the use according to para 58, or the method according to para 59, wherein the disease or condition is a disease or condition characterised by deficiency of the polypeptide of interest.
61. A kit, comprising:
- 25 (i) a polynucleotide according to any one of paras 33 to 48, a vector according to para 49 or para 50, or a pharmaceutical composition according to para 51; and
- (ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
- 30 1A. A polynucleotide comprising in 5' to 3' order:
- (i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:1, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;
- (ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80%
- 35 sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 at its 3' end, and consisting of fewer than 1044 nucleotides;
- (iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:11, wherein the third nucleotide sequence comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12, and wherein the
- 40 third nucleotide sequence comprises insertion of 'A' immediately after the position corresponding to position 48 of SEQ ID NO:12;

(iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;

(v) a fifth nucleotide sequence: (a) consisting of the dinucleotide 'GA', 'TG' or 'TT', or (b) comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:19, wherein the nucleotide sequence comprises 'GA', 'TG' or 'TT' at the positions corresponding to positions 1 and 2 of SEQ ID NO:19, or (c) encoding a polypeptide of interest, and comprising 'GA', 'TG' or 'TT' at positions 1 and 2; and

(vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a) or (v)(b), a sixth nucleotide sequence encoding a polypeptide of interest; wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

2A. The polynucleotide according to para 1A, wherein when the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.

3A. The polynucleotide according to para 1A or para 2A, wherein the first nucleotide sequence comprises a nucleotide sequence according to SEQ ID NO:1.

4A. The polynucleotide according to any one of paras 1A to 3A, wherein the first nucleotide sequence comprises, or consists of, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:6.

5A. The polynucleotide according to any one of paras 1A to 4A, wherein the second nucleotide sequence comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 at its 3' end.

6A. The polynucleotide according to any one of paras 1A to 5A, wherein the second nucleotide sequence consists of fewer than 500 nucleotides.

7A. The polynucleotide according to any one of paras 1A to 6A, wherein the second nucleotide sequence comprises, or consists of, SEQ ID NO:9 or SEQ ID NO:10.

8A. The polynucleotide according to any one of paras 1A to 7A, wherein the third nucleotide sequence consists of a nucleotide sequence according to SEQ ID NO:11.

9A. The polynucleotide according to any one of paras 1A to 8A, wherein the third nucleotide sequence consists of SEQ ID NO:13 or SEQ ID NO:14.

10A. The polynucleotide according to any one of paras 1A to 9A, wherein the fourth nucleotide sequence comprises SEQ ID NO:15 at its 5' end, and comprises SEQ ID NO:16 at its 3' end.

- 11A. The polynucleotide according to any one of paras 1A to 10A, wherein the fourth nucleotide sequence consists of fewer than 500 nucleotides.
- 12A. The polynucleotide according to any one of paras 1A to 11A, wherein the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17 or SEQ ID NO:18.
- 13A. The polynucleotide according to any one of paras 1A to 12A, wherein the fifth nucleotide sequence: (a) consists of the dinucleotide 'GA', or (b) comprises, or consists of, a nucleotide sequence according to SEQ ID NO:19.
- 14A. The polynucleotide according to any one of paras 1A to 13A, wherein the fifth nucleotide sequence: (a) consists of the dinucleotide 'GA', or (b) comprises, or consists of, SEQ ID NO:20 or SEQ ID NO:21.
- 15A. The polynucleotide according to any one of paras 1A to 14A, wherein the polynucleotide comprises a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence selected from the group consisting of SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24 and SEQ ID NO:25.
- 16A. The polynucleotide according to any one of paras 1A to 15A, wherein the polynucleotide further comprises a promoter sequence 5' to the start codon.
- 17A. The polynucleotide according to any one of paras 1A to 16A, wherein the polynucleotide further comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.
- 18A. The polynucleotide according to any one of paras 1A to 17A, wherein the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.
- 19A. A vector comprising a polynucleotide according to any one of paras 1A to 18A.
- 20A. The vector according to para 19A, wherein the vector is an adeno-associated virus (AAV) vector.
- 21A. A pharmaceutical composition comprising a polynucleotide according to any one of paras 1A to 18A, or a vector according to para 19A or para 20A, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.
- 22A. A cell comprising a polynucleotide according to any one of paras 1A to 18A, or a vector according to para 19A or para 20A.
- 23A. The cell according to para 22A, wherein the cell further comprises a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
- 24A. A method for modifying a cell to express a polypeptide of interest, comprising:

(i) introducing into a cell a polynucleotide according to any one of paras 1A to 18A, or a vector according to para 19A or para 20A; and

(ii) subsequently contacting the cell with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

5

25A. A method for expressing a polypeptide of interest in a cell, comprising contacting a cell according to para 22A with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

10 26A. A polynucleotide according to any one of paras 1A to 18A, a vector according to para 19A or para 20A, or a pharmaceutical composition according to para 21A, for use in a method of medical treatment or prophylaxis.

15 27A. A polynucleotide according to any one of paras 1A to 18A, a vector according to para 19A or para 20A, or a pharmaceutical composition according to para 21A, for use in treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.

20 28A. Use of a polynucleotide according to any one of paras 1A to 18A, a vector according to para 19A or para 20A, or a pharmaceutical composition according to para 21A, in the manufacture of a medicament for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.

25 29A. A method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising administering to a subject a polynucleotide according to any one of paras 1A to 18A, a vector according to para 19A or para 20A, or a pharmaceutical composition according to para 21A.

30 30A. The polynucleotide, vector, or pharmaceutical composition for use according to para 27A, the use according to para 28A, or the method according to para 29A, wherein treating or preventing the disease or condition further comprises administering to the subject a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

35 31A. The polynucleotide, vector, or pharmaceutical composition for use according to para 27A or para 30A, the use according to para 28A or para 30A, or the method according to para 29A or para 30A, wherein the disease or condition is a disease or condition characterised by deficiency of the polypeptide of interest.

32A. A kit, comprising:

40 (i) a polynucleotide according to any one of paras 1A to 18A, a vector according to para 19A or para 20A, or a pharmaceutical composition according to para 21A; and

(ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

33A. A polynucleotide comprising in 5' to 3' order:

- 5 (i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:1, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;
 - (ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 10 80% sequence identity to SEQ ID NO:8 at its 3' end, and consisting of fewer than 1044 nucleotides;
 - (iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:26, wherein the third nucleotide sequence comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12;
 - (iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% 15 sequence identity to SEQ ID NO:15 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;
 - (v) a fifth nucleotide sequence: (a) consisting of the trinucleotide 'GAG' or (b) encoding a polypeptide of interest, and comprising 'GAG' at positions 1 to 3; and
 - (vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a), a sixth 20 nucleotide sequence encoding a polypeptide of interest;
- wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

34A. The polynucleotide according to para 33A, wherein when the polynucleotide is a polyribonucleotide, 25 splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.

35A. The polynucleotide according to para 33A or para 34A, wherein the first nucleotide sequence 30 comprises a nucleotide sequence according to SEQ ID NO:1.

36A. The polynucleotide according to any one of paras 33A to 35A, wherein the first nucleotide sequence comprises, or consists of, SEQ ID NO:3.

37A. The polynucleotide according to any one of paras 33A to 36A, wherein the second nucleotide 35 sequence comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 at its 3' end.

38A. The polynucleotide according to any one of paras 33A to 37A, wherein the second nucleotide sequence consists of fewer than 500 nucleotides.

40 39A. The polynucleotide according to any one of paras 33A to 38A, wherein the second nucleotide sequence comprises, or consists of, SEQ ID NO:10.

- 40A. The polynucleotide according to any one of paras 33A to 39A, wherein the third nucleotide sequence consists of a nucleotide sequence according to SEQ ID NO:26.
- 5 41A. The polynucleotide according to any one of paras 33A to 40A, wherein the third nucleotide sequence consists of SEQ ID NO:13 or SEQ ID NO:27.
- 42A. The polynucleotide according to any one of paras 33A to 41A, wherein the fourth nucleotide sequence comprises SEQ ID NO:15 at its 5' end, and comprises SEQ ID NO:16 at its 3' end.
- 10 43A. The polynucleotide according to any one of paras 33A to 42A, wherein the fourth nucleotide sequence consists of fewer than 500 nucleotides.
- 44A. The polynucleotide according to any one of paras 33A to 43A, wherein the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17.
- 15 45A. The polynucleotide according to any one of paras 33A to 44A, wherein the polynucleotide comprises at least 80% sequence identity to SEQ ID NO:28 or SEQ ID NO:29.
- 46A. The polynucleotide according to any one of paras 33A to 45A, wherein the polynucleotide further
- 20 comprises a promoter sequence 5' to the start codon.
- 47A. The polynucleotide according to any one of paras 33A to 46A, wherein the polynucleotide further comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.
- 25 48A. The polynucleotide according to any one of paras 33A to 47A, wherein the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.
- 49A. A vector comprising a polynucleotide according to any one of paras 33A to 48A.
- 30 50A. The vector according to para 49A, wherein the vector is an adeno-associated virus (AAV) vector.
- 51A. A pharmaceutical composition comprising a polynucleotide according to any one of paras 33A to 48A, or a vector according to para 49A or para 50A, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.
- 35 52A. A cell comprising a polynucleotide according to any one of paras 33A to 48A, or a vector according to para 49A or para 50A.
- 53A. The cell according to para 52A, wherein the cell further comprises a splicing modifier that promotes
- 40 *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

- 54A. A method for modifying a cell to express a polypeptide of interest, comprising introducing into a cell a polynucleotide according to any one of paras 33A to 48A, or a vector according to para 49A or para 50A.
- 5 55A. A method for inhibiting expression of a polypeptide of interest in a cell, comprising contacting a cell according to para 52A with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
- 56A. A polynucleotide according to any one of paras 33A to 48A, a vector according to para 49A or para 10 50A, or a pharmaceutical composition according to para 51A, for use in a method of medical treatment or prophylaxis.
- 57A. A polynucleotide according to any one of paras 33A to 48A, a vector according to para 49A or para 50A, or a pharmaceutical composition according to para 51A, for use in treating or preventing a disease 15 or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
- 58A. Use of a polynucleotide according to any one of paras 33A to 48A, a vector according to para 49A or para 50A, or a pharmaceutical composition according to para 51A, in the manufacture of a medicament 20 for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
- 59A. A method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising 25 administering to a subject a polynucleotide according to any one of paras 33A to 48A, a vector according to para 49A or para 50A, or a pharmaceutical composition according to para 51A.
- 60A. The polynucleotide, vector, or pharmaceutical composition for use according to para 57A, the use according to para 58A, or the method according to para 59A, wherein the disease or condition is a 30 disease or condition characterised by deficiency of the polypeptide of interest.
- 61A. A kit, comprising:
- (i) a polynucleotide according to any one of paras 33A to 48A, a vector according to para 49A or para 50A, or a pharmaceutical composition according to para 51A; and
- 35 (ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

Table A

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
1	pMM112 ON-switch	SEQ ID NO:106	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	GA

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:22)					
2	pMM130 ON-switch (SEQ ID NO:23)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	GA
3	pMM59 ON-switch (SEQ ID NO:24)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:20
4	pMM198 ON-switch (SEQ ID NO:25)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:21
5	pMM193 OFF-switch (SEQ ID NO:28)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	GAG
6	pMM194 OFF-switch (SEQ ID NO:29)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:27	SEQ ID NO:17	GAG
7	pLS41 ON-switch (SEQ ID NO:37)	SEQ ID NO:2	SEQ ID NO:238	SEQ ID NO:256	SEQ ID NO:18	SEQ ID NO:20
8	pLS76 ON-switch (SEQ ID NO:38)	SEQ ID NO:2	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:20
9	pLS159 ON-switch (SEQ ID NO:39)	SEQ ID NO:2	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	SEQ ID NO:20
10	pLS160 ON-switch (SEQ ID NO:40)	SEQ ID NO:2	SEQ ID NO:239	SEQ ID NO:13	SEQ ID NO:330	SEQ ID NO:20
11	pLS167 ON-switch (SEQ ID NO:41)	SEQ ID NO:230	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:20
12	pLS168 ON-switch (SEQ ID NO:42)	SEQ ID NO:231	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:20
13	pLS179 ON-switch (SEQ ID NO:43)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	SEQ ID NO:20
14	pMM70 ON-switch (SEQ ID NO:44)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	SEQ ID NO:20
15	pMM71 ON-switch (SEQ ID NO:45)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	SEQ ID NO:20
16	pMM72 ON-switch (SEQ ID NO:46)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	SEQ ID NO:20
17	pMM73 ON-switch (SEQ ID NO:47)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	SEQ ID NO:20
18	pMM143 ON-switch (SEQ ID NO:48)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:331	SEQ ID NO:20
19	pMM144 ON-switch (SEQ ID NO:49)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:332	SEQ ID NO:20
20	pMM145 ON-switch (SEQ ID NO:50)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:333	SEQ ID NO:20
21	pMM146 ON-switch (SEQ ID NO:51)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:334	SEQ ID NO:20
22	pMM147 ON-switch (SEQ ID NO:52)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:335	SEQ ID NO:20
23	pMM136 ON-switch	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:257	SEQ ID NO:17	SEQ ID NO:20

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:53)					
24	pMM137 ON-switch (SEQ ID NO:54)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:258	SEQ ID NO:17	SEQ ID NO:20
25	pMM138 ON-switch (SEQ ID NO:55)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:259	SEQ ID NO:17	SEQ ID NO:20
26	pMM139 ON-switch (SEQ ID NO:56)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:260	SEQ ID NO:17	SEQ ID NO:20
27	pMM140 ON-switch (SEQ ID NO:57)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:261	SEQ ID NO:17	SEQ ID NO:20
28	pMM141 ON-switch (SEQ ID NO:58)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:262	SEQ ID NO:17	SEQ ID NO:20
29	pMM142 ON-switch (SEQ ID NO:59)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:263	SEQ ID NO:17	SEQ ID NO:20
30	pMM151 ON-switch (SEQ ID NO:60)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:264	SEQ ID NO:17	SEQ ID NO:20
31	pMM152 ON-switch (SEQ ID NO:61)	SEQ ID NO:3	SEQ ID NO:10	SEQ ID NO:265	SEQ ID NO:17	SEQ ID NO:20
32	pLS174 ON-switch (SEQ ID NO:62)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:20
33	pMM56 ON-switch (SEQ ID NO:63)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:299	SEQ ID NO:18	SEQ ID NO:20
34	pMM60 ON-switch (SEQ ID NO:64)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:301	SEQ ID NO:18	SEQ ID NO:20
35	pMM61 ON-switch (SEQ ID NO:65)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:302	SEQ ID NO:18	SEQ ID NO:20
36	pMM62 ON-switch (SEQ ID NO:66)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:329	SEQ ID NO:18	SEQ ID NO:20
37	pMM63 ON-switch (SEQ ID NO:67)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:303	SEQ ID NO:18	SEQ ID NO:20
38	pLS176 ON-switch (SEQ ID NO:68)	SEQ ID NO:234	SEQ ID NO:9	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:20
39	pMM108 ON-switch (SEQ ID NO:69)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:20
40	pMM110 ON-switch (SEQ ID NO:70)	SEQ ID NO:232	SEQ ID NO:9	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:20
41	pMM111 ON-switch (SEQ ID NO:71)	SEQ ID NO:233	SEQ ID NO:9	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:20
42	pMM242 ON-switch (SEQ ID NO:72)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:304	SEQ ID NO:18	SEQ ID NO:20
43	pMM243 ON-switch (SEQ ID NO:73)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:305	SEQ ID NO:18	SEQ ID NO:20
44	pMM244 ON-switch (SEQ ID NO:74)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:306	SEQ ID NO:18	SEQ ID NO:20
45	pMM245 ON-switch	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:308	SEQ ID NO:18	SEQ ID NO:20

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:75)					
46	pMM246 ON-switch (SEQ ID NO:76)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:309	SEQ ID NO:18	SEQ ID NO:20
47	pMM247 ON-switch (SEQ ID NO:77)	SEQ ID NO:5	SEQ ID NO:9	SEQ ID NO:307	SEQ ID NO:18	SEQ ID NO:20
48	pLS175 ON-switch (SEQ ID NO:78)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:20
49	pMM123 ON-switch (SEQ ID NO:79)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:310	SEQ ID NO:18	SEQ ID NO:20
50	pMM124 ON-switch (SEQ ID NO:80)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:311	SEQ ID NO:18	SEQ ID NO:20
51	pMM125 ON-switch (SEQ ID NO:81)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:312	SEQ ID NO:18	SEQ ID NO:20
52	pMM126 ON-switch (SEQ ID NO:82)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:315	SEQ ID NO:18	SEQ ID NO:20
53	pMM127 ON-switch (SEQ ID NO:83)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:313	SEQ ID NO:18	SEQ ID NO:20
54	pMM128 ON-switch (SEQ ID NO:84)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:314	SEQ ID NO:18	SEQ ID NO:20
55	pMM129 ON-switch (SEQ ID NO:85)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:316	SEQ ID NO:18	SEQ ID NO:20
56	pMM201 ON-switch (SEQ ID NO:86)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:318	SEQ ID NO:18	SEQ ID NO:20
57	pMM202 ON-switch (SEQ ID NO:87)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:319	SEQ ID NO:18	SEQ ID NO:20
58	pMM203 ON-switch (SEQ ID NO:88)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:320	SEQ ID NO:18	SEQ ID NO:20
59	pMM204 ON-switch (SEQ ID NO:89)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:321	SEQ ID NO:18	SEQ ID NO:20
60	pMM205 ON-switch (SEQ ID NO:90)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:322	SEQ ID NO:18	SEQ ID NO:20
61	pMM206 ON-switch (SEQ ID NO:91)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:323	SEQ ID NO:18	SEQ ID NO:20
62	pMM248 ON-switch (SEQ ID NO:92)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:324	SEQ ID NO:18	SEQ ID NO:20
63	pMM249 ON-switch (SEQ ID NO:93)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:325	SEQ ID NO:18	SEQ ID NO:20
64	pMM250 ON-switch (SEQ ID NO:94)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:326	SEQ ID NO:18	SEQ ID NO:20
65	pMM251 ON-switch (SEQ ID NO:95)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:327	SEQ ID NO:18	SEQ ID NO:20
66	pMM252 ON-switch (SEQ ID NO:96)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	GA
67	pMM254 ON-switch	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:315	SEQ ID NO:17	GA

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:97)					
68	pMM255 ON-switch (SEQ ID NO:98)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:314	SEQ ID NO:17	GA
69	pMM256 ON-switch (SEQ ID NO:99)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:316	SEQ ID NO:17	GA
70	pMM257 ON-switch (SEQ ID NO:100)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:317	SEQ ID NO:17	GA
71	pMM199 ON-switch (SEQ ID NO:101)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:336	SEQ ID NO:21
72	pMM200 ON-switch (SEQ ID NO:102)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:381	SEQ ID NO:21
73	pMM236 ON-switch (SEQ ID NO:103)	SEQ ID NO:6	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:21
74	pMM237 ON-switch (SEQ ID NO:104)	SEQ ID NO:235	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:21
75	pMM238 ON-switch (SEQ ID NO:105)	SEQ ID NO:233	SEQ ID NO:9	SEQ ID NO:13	SEQ ID NO:18	SEQ ID NO:21
76	pMM273 ON-switch (SEQ ID NO:111)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:17	TG
77	pMM362 ON-switch (SEQ ID NO:112)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:266	SEQ ID NO:338	TG
78	pMM363 ON-switch (SEQ ID NO:113)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:267	SEQ ID NO:339	TG
79	pMM364 ON-switch (SEQ ID NO:114)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:340	TG
80	pMM358 ON-switch (SEQ ID NO:115)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:268	SEQ ID NO:17	TG
81	pMM359 ON-switch (SEQ ID NO:116)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:223	SEQ ID NO:17	TG
82	pMM360 ON-switch (SEQ ID NO:117)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:269	SEQ ID NO:17	TG
83	pMM361 ON-switch (SEQ ID NO:118)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:270	SEQ ID NO:17	TG
84	pMM365 ON-switch (SEQ ID NO:119)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:271	SEQ ID NO:17	TG
85	pMM366 ON-switch (SEQ ID NO:120)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:272	SEQ ID NO:17	TG
86	pMM367 ON-switch (SEQ ID NO:121)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:273	SEQ ID NO:17	TG
87	pMM368 ON-switch (SEQ ID NO:122)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:274	SEQ ID NO:17	TG
88	pMM369 ON-switch (SEQ ID NO:123)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:275	SEQ ID NO:17	TG
89	pMM370 ON-switch	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:276	SEQ ID NO:17	TG

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:124)					
90	pMM371 ON-switch (SEQ ID NO:125)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:277	SEQ ID NO:17	TG
91	pMM372 ON-switch (SEQ ID NO:126)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:278	SEQ ID NO:17	TG
92	pMM373 ON-switch (SEQ ID NO:127)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:279	SEQ ID NO:17	TG
93	pMM374 ON-switch (SEQ ID NO:128)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:280	SEQ ID NO:17	TG
94	pMM375 ON-switch (SEQ ID NO:129)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:281	SEQ ID NO:17	TG
95	pMM376 ON-switch (SEQ ID NO:130)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:282	SEQ ID NO:17	TG
96	pMM377 ON-switch (SEQ ID NO:131)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:283	SEQ ID NO:17	TG
97	pMM378 ON-switch (SEQ ID NO:132)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:284	SEQ ID NO:17	TG
98	pMM379 ON-switch (SEQ ID NO:133)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:264	SEQ ID NO:17	TG
99	pMM380 ON-switch (SEQ ID NO:134)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:285	SEQ ID NO:17	TG
100	pMM381 ON-switch (SEQ ID NO:135)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:265	SEQ ID NO:17	TG
101	pMM382 ON-switch (SEQ ID NO:136)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:286	SEQ ID NO:17	TG
102	pMM386 ON-switch (SEQ ID NO:137)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:287	SEQ ID NO:17	TG
103	pMM387 ON-switch (SEQ ID NO:138)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:27	SEQ ID NO:17	TG
104	pMM436 ON-switch (SEQ ID NO:139)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:289	SEQ ID NO:17	TG
105	pMM437 ON-switch (SEQ ID NO:140)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:290	SEQ ID NO:17	TG
106	pMM438 ON-switch (SEQ ID NO:141)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:291	SEQ ID NO:17	TG
107	pMM477 ON-switch (SEQ ID NO:142)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:341	TG
108	pMM478 ON-switch (SEQ ID NO:143)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:342	TG
109	pMM479 ON-switch (SEQ ID NO:144)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:343	TG
110	pMM480 ON-switch (SEQ ID NO:145)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:344	TG
111	pMM481 ON-switch	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:345	TG

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:146)					
112	pMM482 ON-switch (SEQ ID NO:147)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:346	TG
113	pMM483 ON-switch (SEQ ID NO:148)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:347	TG
114	pMM484 ON-switch (SEQ ID NO:149)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:348	TG
115	pMM485 ON-switch (SEQ ID NO:150)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:349	TG
116	pMM464 ON-switch (SEQ ID NO:151)	SEQ ID NO:219	SEQ ID NO:241	SEQ ID NO:13	SEQ ID NO:350	TG
117	pMM465 ON-switch (SEQ ID NO:152)	SEQ ID NO:219	SEQ ID NO:242	SEQ ID NO:13	SEQ ID NO:351	TG
118	pMM466 ON-switch (SEQ ID NO:153)	SEQ ID NO:219	SEQ ID NO:243	SEQ ID NO:13	SEQ ID NO:352	TG
119	pMM467 ON-switch (SEQ ID NO:154)	SEQ ID NO:219	SEQ ID NO:244	SEQ ID NO:13	SEQ ID NO:353	TG
120	pMM468 ON-switch (SEQ ID NO:155)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:350	TG
121	pMM469 ON-switch (SEQ ID NO:156)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:351	TG
122	pMM470 ON-switch (SEQ ID NO:157)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:352	TG
123	pMM471 ON-switch (SEQ ID NO:158)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:354	TG
124	pMM472 ON-switch (SEQ ID NO:159)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:227	TG
125	pMM473 ON-switch (SEQ ID NO:160)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:356	TG
126	pMM474 ON-switch (SEQ ID NO:161)	SEQ ID NO:219	SEQ ID NO:10	SEQ ID NO:13	SEQ ID NO:357	TG
127	pMM567 ON-switch (SEQ ID NO:162)	SEQ ID NO:220	SEQ ID NO:242	SEQ ID NO:224	SEQ ID NO:352	TG
128	pMM568 ON-switch (SEQ ID NO:163)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:352	TG
129	pMM569 ON-switch (SEQ ID NO:164)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:227	TG
130	pMM570 ON-switch (SEQ ID NO:165)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:357	TG
131	pMM591 ON-switch (SEQ ID NO:166)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:358	TG
132	pMM592 ON-switch (SEQ ID NO:167)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:359	TG
133	pMM593 ON-switch	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:293	SEQ ID NO:227	TG

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:168)					
134	pMM594 ON-switch (SEQ ID NO:169)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:294	SEQ ID NO:227	TG
135	pMM595 ON-switch (SEQ ID NO:170)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:295	SEQ ID NO:227	TG
136	pMM596 ON-switch (SEQ ID NO:171)	SEQ ID NO:220	SEQ ID NO:228	SEQ ID NO:224	SEQ ID NO:227	TG
137	pMM597 ON-switch (SEQ ID NO:172)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:328	SEQ ID NO:227	TG
138	pMM598 ON-switch (SEQ ID NO:173)	SEQ ID NO:220	SEQ ID NO:375	SEQ ID NO:296	SEQ ID NO:227	TG
139	pMM599 ON-switch (SEQ ID NO:174)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:360	TG
140	pMM600 ON-switch (SEQ ID NO:175)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:297	SEQ ID NO:227	TG
141	pMM601 ON-switch (SEQ ID NO:176)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:298	SEQ ID NO:227	TG
142	pMM642 ON-switch (SEQ ID NO:177)	SEQ ID NO:220	SEQ ID NO:246	SEQ ID NO:224	SEQ ID NO:227	TG
143	pMM643 ON-switch (SEQ ID NO:178)	SEQ ID NO:220	SEQ ID NO:376	SEQ ID NO:224	SEQ ID NO:227	TG
144	pMM709 ON-switch (SEQ ID NO:179)	SEQ ID NO:220	SEQ ID NO:247	SEQ ID NO:224	SEQ ID NO:227	TG
145	pMM644 ON-switch (SEQ ID NO:180)	SEQ ID NO:220	SEQ ID NO:248	SEQ ID NO:224	SEQ ID NO:227	TG
146	pMM645 ON-switch (SEQ ID NO:181)	SEQ ID NO:220	SEQ ID NO:249	SEQ ID NO:224	SEQ ID NO:227	TG
147	pMM646 ON-switch (SEQ ID NO:182)	SEQ ID NO:220	SEQ ID NO:251	SEQ ID NO:224	SEQ ID NO:227	TG
148	pMM647 ON-switch (SEQ ID NO:183)	SEQ ID NO:220	SEQ ID NO:250	SEQ ID NO:224	SEQ ID NO:227	TG
149	pMM648 ON-switch (SEQ ID NO:184)	SEQ ID NO:220	SEQ ID NO:252	SEQ ID NO:224	SEQ ID NO:227	TG
150	pMM649 ON-switch (SEQ ID NO:185)	SEQ ID NO:220	SEQ ID NO:253	SEQ ID NO:224	SEQ ID NO:227	TG
151	pMM650 ON-switch (SEQ ID NO:186)	SEQ ID NO:220	SEQ ID NO:254	SEQ ID NO:224	SEQ ID NO:227	TG
152	pMM651 ON-switch (SEQ ID NO:187)	SEQ ID NO:220	SEQ ID NO:255	SEQ ID NO:224	SEQ ID NO:227	TG
153	pMM634 ON-switch (SEQ ID NO:188)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:361	TG
154	pMM711 ON-switch (SEQ ID NO:189)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:362	TG
155	pMM635 ON-switch	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:363	TG

Row	Switch Name	Col. A 1 st Nt Sequence	Col. B 2 nd Nt Sequence	Col. C 3 rd Nt Sequence	Col. D 4 th Nt Sequence	Col. E 5 th Nt Sequence
	(SEQ ID NO:190)					
156	pMM636 ON-switch (SEQ ID NO:191)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:364	TG
157	pMM637 ON-switch (SEQ ID NO:192)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:365	TG
158	pMM638 ON-switch (SEQ ID NO:193)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:366	TG
159	pMM639 ON-switch (SEQ ID NO:194)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:367	TG
160	pMM640 ON-switch (SEQ ID NO:195)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:368	TG
161	pMM641 ON-switch (SEQ ID NO:196)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:369	TG
162	pMM712 ON-switch (SEQ ID NO:197)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:370	TG
163	pMM579 ON-switch (SEQ ID NO:198)	SEQ ID NO:220	SEQ ID NO:10	SEQ ID NO:292	SEQ ID NO:227	TG
164	pMM618 ON-switch (SEQ ID NO:199)	SEQ ID NO:237	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:227	TG
165	pMM619 ON-switch (SEQ ID NO:200)	SEQ ID NO:236	SEQ ID NO:10	SEQ ID NO:224	SEQ ID NO:227	TG

The present disclosure includes the combination of the aspects and preferred features described except where such a combination is clearly impermissible or expressly avoided.

5

The section headings used herein are for organisational purposes only and are not to be construed as limiting the subject matter described.

Aspects and embodiments of the present disclosure will now be illustrated, by way of example, with reference to the accompanying figures. Further aspects and embodiments will be apparent to those skilled in the art. All documents mentioned in this text are incorporated herein by reference.

10

Throughout this specification, including the claims which follow, unless the context requires otherwise, the word 'comprise,' and variations such as 'comprises' and 'comprising,' will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

15

As used herein, a 'peptide' refers to a chain of two or more amino acid monomers linked by peptide bonds. A peptide typically has a length in the region of about 2 to 50 amino acids. A 'polypeptide' is a polymer chain of two or more peptides. Polypeptides typically have a length greater than about 50 amino acids. Reference herein to peptides, polypeptides and proteins also includes

20

glycopeptides/glycopolypeptides/glycoproteins, lipopeptides/lipopolypeptides/lipoproteins, nucleopeptides/nucleopolypeptides/nucleoproteins, *etc.*

As used herein, an amino acid sequence, or a region of a polypeptide, which 'corresponds' to a specified reference amino acid sequence or region of a polypeptide has at least 60%, *e.g.* one of at least $\geq 65\%$, $\geq 70\%$, $\geq 75\%$, $\geq 80\%$, $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100% sequence identity to the amino acid sequence of the amino acid sequence/polypeptide/region. An amino acid sequence/region/position of a polypeptide/amino acid sequence which 'corresponds' to a specified reference amino acid sequence/region/position of a polypeptide/amino acid sequence can be identified by sequence alignment of the subject sequence to the reference sequence, *e.g.* using sequence alignment software such as ClustalOmega (Söding, J. 2005, Bioinformatics 21, 951-960).

Similarly, a nucleotide sequence, or a region of a polynucleotide, which 'corresponds' to a specified reference nucleotide sequence or region of a polynucleotide has at least 60%, *e.g.* one of at least $\geq 65\%$, $\geq 70\%$, $\geq 75\%$, $\geq 80\%$, $\geq 85\%$, $\geq 90\%$, $\geq 91\%$, $\geq 92\%$, $\geq 93\%$, $\geq 94\%$, $\geq 95\%$, $\geq 96\%$, $\geq 97\%$, $\geq 98\%$, $\geq 99\%$ or 100% sequence identity to the amino acid sequence of the nucleotide sequence/polynucleotide/region. A polynucleotide/region/position of a polynucleotide/nucleotide sequence which 'corresponds' to a specified reference nucleotide sequence/region/position of a polynucleotide/nucleotide sequence can be identified by sequence alignment of the subject sequence to the reference sequence, *e.g.* using sequence alignment software such as ClustalOmega (Söding, J. 2005, Bioinformatics 21, 951-960).

As used herein, an amino acid sequence (*e.g.* the amino acid sequence of a peptide/polypeptide/domain/region) which is 'derived from' a reference amino acid sequence (*e.g.* the amino acid sequence of a reference peptide/polypeptide/domain/region) comprises, or consists of, an amino acid sequence having at least 60%, *e.g.* one of at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the reference amino acid sequence. Similarly, a nucleotide sequence (*e.g.* a nucleotide sequence of a polynucleotide) which is 'derived from' a reference nucleotide sequence comprises, or consists of, a nucleotide sequence having at least 60%, *e.g.* one of at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the reference nucleotide sequence.

It must be noted that, as used in the specification and the appended claims, the singular forms 'a,' 'an,' and 'the' include plural referents unless the context clearly dictates otherwise. Ranges may be expressed herein as from 'about' one particular value, and/or to 'about' another particular value. When such a range is expressed, another embodiment includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by the use of the antecedent 'about,' it will be understood that the particular value forms another embodiment.

Where a nucleic acid sequence is disclosed herein, the reverse complement thereof is also expressly contemplated.

Methods described herein may preferably be performed *in vitro*. The term '*in vitro*' is intended to encompass procedures performed with cells in culture whereas the term '*in vivo*' is intended to encompass procedures with/on intact multi-cellular organisms.

5 Brief Description of the Figures

Embodiments and experiments illustrating the principles of the present disclosure will now be discussed with reference to the accompanying figures.

Figure 1. Schematic representation of the design of the In-frame ON-switch. The cassette comprising the in-frame switch is constructed using the exon 6-intron 6-exon 7-intron 7-exon 8 region of human *SMN2* followed by the coding sequence of the gene of interest (GOI) downstream of exon 8. The translation start site is at the 5' end of exon 6. Intron 6 is modified to reduce its length and exon 7 contains an insertion of the nucleotide A after position 48. The sequence of exon 8 is only the initial 23 nucleotides. In the absence of risdiplam splicing modifier, exon 7 is excluded resulting in the generation of a frameshift in the coding sequence of the GOI. Conversely, risdiplam treatment leads to exon 7 inclusion and the correct translation of the GOI open reading frame and its protein production. Star indicates the insertion of nucleotide "A" causing a frame-shift.

Figure 2. Small molecule-mediated expression of GFP from the In-frame ON-switch 48h post-treatment with risdiplam analogue. FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the in-frame ON-switch pLS41. The graph shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

Figure 3. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pLS41 and pLS76 ON-switches. (A) Schematic representation indicating the changes in the modified (Mod) in-frame switch (*i.e.* pLS76) relative to the wildtype (WT) sequence of *SMN2* intron 6 and exon 7. (B-D) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS41 or pLS76 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 3B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows the results of analysis of the products of splicing by RT-PCR performed on HEK293 stably expressing pLS76, and treated with low (100nM) or high (1 μ M) concentrations of risdiplam, or DMSO.

Figure 4. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pLS159 and pLS160 ON-switches, which have reduced intron 6 and intron 7 lengths relative to the pLS76 ON-switch. (A) Schematic representation of the differences between the pLS76, pLS159 and pLS160 ON-switches. (B-D) FACS-based measurements of GFP expression

upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS76, pLS159, or pLS160 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 4B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 5. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pLS167 and pLS168 ON-switches, which have reduced exon 6 lengths relative to the pLS76 ON-switch. (A) Schematic representation of the differences between the pLS76, pLS167 and pLS168 ON-switches. The modified region is marked with a dashed box. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS76, pLS167, or pLS168 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 5B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 6. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pLS179 ON-switch, which combines the intron truncations of pLS159 and the exon 6 truncation of pLS168. (A) Schematic representation indicating the changes in the modified (Mod) in-frame switch (*i.e.* pLS76) relative to the wildtype (WT) sequence of SMN2 intron 6 and exon 7. (B-D) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS76, pLS159, pLS168 or pLS179 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 6B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows the results of analysis of the products of splicing by RT-PCR performed on HEK293 stably expressing pLS179, and treated with low (100nM) or high (1 μ M) concentrations of risdiplam, or DMSO.

Figure 7. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM70, pMM71, pMM72, pMM73 and pMM112 ON-switches, which have reduced exon 8 lengths relative to the pLS179 ON-switch. (A) Schematic representation of the differences between the pLS179, pMM70, pMM71, pMM72, pMM73 and pMM112 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of

risdiplam in HEK293 cells stably expressing the pLS179, pMM70, pMM71, pMM72, pMM73 or pMM112 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 7B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 8. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM130 ON-switch, in which the Kozak sequence and start codon present in pMM112 are removed. (A) Schematic representation of the differences between the pMM112 and pMM130 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS179, pMM112 or pMM130 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 8B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 9. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM143, pMM144, pMM145, pMM146 and pMM147 ON-switches, in which the ISS-N1 region of intron 7 is mutated to reduced leakiness. (A) Schematic representation of the differences between the pLS179, pMM143, pMM144, pMM145, pMM146 and pMM147 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS179, pMM143, pMM144, pMM145, pMM146 or pMM147 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 9B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 10. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM136, pMM137, pMM138, pMM139, pMM140, pMM141 and pMM142 ON-switches, in which the ESE2 region (marked in grey based on Sivaramakrishnan *et al.*, Nat Commun 8, 1476 (2017)) of exon 7 is mutated to improve switch performance. (A) Schematic representation of the differences between the pLS179, pMM136, pMM137, pMM138, pMM139, pMM140, pMM141 and pMM142 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS179, pMM136, pMM137, pMM138, pMM139, pMM140, pMM141 or pMM142 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-

treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 10B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1μM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 11. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM151 and pMM152 ON-switches, which comprise deletions in exon 7. (A) Schematic representation of the differences between the pMM112, pMM151 and pMM152 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM112, pMM151 or pMM152 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 11B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1μM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 12. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM56, pMM59, pMM60, pMM61, pMM62 and pMM63 ON-switches, in which ATG codons in exon 6 are removed, and in which comprise modifications in exon 7. (A) Schematic representation of the differences between (Ai) pLS76 and pLS174, and (Aii) pLS174, pMM56, pMM59, pMM60, pMM61, pMM62 and pMM63 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS76, pMM56, pMM59, pMM60, pMM61, pMM62 and pMM63 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 12B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1μM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 13. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM108, pMM110, and pMM111 ON-switches, in which ATG codons in exon 6 are removed. (A) Schematic representation of the differences between the pLS76, pLS176, pMM59, pMM108, pMM110, and pMM111 ON-switches. Insertions relative to the sequence of pLS174 are shown in black and in bold, and mutations relative to the sequence of pLS174 are shown in black, italics and bold. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pLS176, pMM59, pMM108, pMM110, or pMM111 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 13B, as the total median GFP signal without normalization to DMSO control. (D) Shows the

percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

5 **Figure 14.** Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM242, pMM243, pMM244, pMM245, pMM246, and pMM247 ON-switches, in which ATG codons are introduced into exon 7. Insertions relative to the sequence of pLS174 are shown in black and in bold, and mutations relative to the sequence of pLS174 are shown in black, italics and bold. (A) Schematic representation of the differences between the pLS174, pMM59, pMM59,
10 pMM242, pMM243, pMM244, pMM245, pMM246, and pMM247 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM242, pMM243, pMM244, pMM245, pMM246, or pMM247 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure
15 14B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

20 **Figure 15.** Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM123, pMM124, pMM125, pMM126, pMM127, pMM128 and pMM129 ON-switches, in which ATG codons are introduced into exon 7 downstream of ESE2. (A) Schematic representation of the differences between the pLS175, pMM123, pMM124, pMM125, pMM126, pMM127, pMM128 and pMM129 ON-switches. Insertions relative to the sequence of pLS175
25 are shown in black and in bold, and mutations relative to the sequence of pLS175 are shown in black, italics and bold. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM123, pMM124, pMM125, pMM126, pMM127, pMM128 or pMM129 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam
30 treatment. (C) Shows the data of Figure 15B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

35 **Figure 16.** Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM201, pMM202, pMM203, pMM204, pMM205 and pMM206 ON-switches, in which ATG codons are introduced into exon 7 downstream of ESE2. (A) Schematic representation of the differences between the pLS175, pMM126, pMM201, pMM202, pMM203, pMM204,
40 pMM205 and pMM206 ON-switches. Insertions relative to the comparator sequence are shown in black and in bold, and mutations relative to the comparator sequence are shown in black, italics and bold. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of

risdiplam in HEK293 cells stably expressing the pMM126, pMM201, pMM202, pMM203, pMM204, pMM205 or pMM206 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 16B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 17. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM248, pMM249, pMM250 and pMM251 ON-switches, in which ATG codons are introduced into exon 7 downstream of ESE2. (A) Schematic representation of the differences between the pLS175, pMM129, pMM248, pMM249, pMM250 and pMM251 ON-switches. Insertions relative to the comparator sequence are shown in black and in bold, and mutations relative to the comparator sequence are shown in black, italics and bold. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM129, pMM248, pMM249, pMM250 or pMM251 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 17B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 18. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM252, pMM254, pMM255, pMM256 and pMM257 ON-switches, in which ATG codons are introduced into exon 7 downstream of ESE2. (A) Schematic representation of the differences between the pMM130, pMM252, pMM254, pMM255, pMM256 and pMM257 ON-switches. (B-D). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM126, pMM129, pMM130, pMM252, pMM254, pMM255, pMM256 or pMM257 constructs. (B) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the data of Figure 18B, as the total median GFP signal without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 19. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM198, pMM199 and pMM200 ON-switches, in which an ATG codon is introduced, spanning exons 7 and 8. (A) Schematic representation of the differences between the pLS76, pLS175, pMM198, pMM199 and pMM200 ON-switches. The start codon spanning exons 7 and 8

is shown in bold and italics. (B) Schematic representation of the predicted splicing of the split start codon ON-switches in the absence (DMSO control) or presence of risdiplam. (C-E). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM112, pMM198, pMM199 or pMM200 constructs. (C) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (D) Shows the data of Figure 19C, as the total median GFP signal without normalization to DMSO control. (E) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1μM). (F) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 20. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), for the pMM236, pMM237 and pMM238 ON-switches, in which ATG codons in exon 6 are removed, a weak Kozak sequence is introduced in exon 7, and an ATG codon is introduced in exon 8. (A) Schematic representation of the differences between the pLS76, pMM236, pMM237 and pMM238 ON-switches. (B) Schematic representation of the predicted splicing of the pMM236, pMM237 and pMM238 ON-switches in the absence (DMSO control) or presence of risdiplam. (C-E). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM112, pMM236, pMM237 or pMM238 constructs. (C) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (D) Shows the data of Figure 20C, as the total median GFP signal without normalization to DMSO control. (E) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1μM). (F) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 21. Analysis of constitutive reporter expression (*i.e.* in the absence of risdiplam), and risdiplam-induced suppression of reporter expression for the pMM193 and pMM194 OFF-switches. (A) Schematic representation of splicing of *SMN2* exon 6 to exon 8-derived ON and OFF-switches. (B) Schematic representation of the differences between the pMM130 ON-switch and the pMM193 OFF-switch. (C) Schematic representation of the differences between the pMM193 and pMM194 OFF-switches. (D-F). FACS-based measurements of GFP expression upon treatment of HEK293 cells stably expressing the pMM130, pMM193, or pMM194 constructs with increasing concentrations of risdiplam. (D) Shows median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (E) Shows the data of Figure 21D, as the total median GFP signal without normalization to DMSO control. (F) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells, at the highest concentration of risdiplam treatment (1μM). (G) Shows the level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 22. Analysis of risdiplam-induced reporter expression, and background reporter expression (*i.e.* in the absence of risdiplam), after replacing IRES-blasticidin sequence with BGH polyA signal for in-frame

switches pMM259 (pMM112), pMM260 (pMM130), pMM262 (pLS168), pMM263 (pLS76), pMM272 (pMM198) and Monteys *et al.* Nature (2021) 596: 291-295 construct (pMM274). (A) Schematic representation of plasmid backbone modification. IRES-blasticidin-SV40 sequence was replaced with BGH polyA signal. In-frame switch sequences remain unmodified. (B) List of in-frame switches used for the experiment with their original numbers (C-E). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM259, pMM260, pMM262, pMM263, pMM272 and Monteys *et al.* Nature (2021) 596: 291-295 construct (pMM274) (C) Shows induction rate calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment, GFP-positive cells. (D) Shows the data of Figure 19C, expressed as the total median GFP signal without normalization to DMSO control. (E) Shows the percentage of GFP-positive (% GFP+) cells divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1µM). (F) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (G) Shows a fragment of Fig. 22C with induction rate at 125 nM risdiplam concentration calculated as median GFP expression normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment, GFP-positive cells.

Figure 23. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM259, pMM263, pMM273 and pMM274. (A) Schematic representation of the differences between the pMM259 and pMM273 ON-switches. (B) Schematic representation of the predicted splicing of the split start codon ON-switches in the absence (DMSO control) or presence of risdiplam. (C-E). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM259, pMM263, pMM273 and pMM274 constructs. (C) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (D) Shows the total median GFP signal of GFP positive cells without normalization to DMSO control. (E) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1µM). (F) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam.

Figure 24. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM273, pMM362, pMM363, and pMM364. (A) Schematic representation of the differences between the pMM273, pMM362, pMM363, and pMM364 ON-switches. (B-F). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM273, pMM362, pMM363, and pMM364 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1µM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or

in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

5 **Figure 25.** Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM273, pMM358, pMM359, pMM360, and pMM361. (A) Schematic representation of the differences between the pMM273, pMM358, pMM359, pMM360, and pMM361 ON-switches. (B-F). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM273, pMM358, 10 pMM359, pMM360, and pMM361 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of 15 risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the results of analysis of the products of splicing by PCR performed on HEK293 20 stably expressing pMM273, pMM358, pMM359 and pMM360 and treated with increasing concentrations of risdiplam, or DMSO.

Figure 26. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM273 and pMM365 - pMM375. (A) Schematic 25 representation of the differences between the pMM273 and pMM365 - pMM375 (B-G). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM273 and pMM365 - pMM375 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total 30 median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression 35 at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M).

Figure 27. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter 40 expression (*i.e.* in the absence of risdiplam) for pMM273 and pMM376 - pMM387. (A) Schematic representation of the differences between the pMM273 and pMM376 - pMM387 (B-G). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293

cells stably expressing the pMM273 and pMM376 - pMM387 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M).

Figure 28. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM273, pMM436, pMM437, and pMM438. (A) Schematic representation of the differences between the pMM273, pMM436, pMM437, and pMM438 (B-F). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM273, pMM436, pMM437, and pMM438 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

Figure 29. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM273 and pMM477 - pMM485. (A) Schematic representation of the differences between the pMM273 and pMM477 - pMM485. (B-F). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM273 and pMM477 - pMM485 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

Figure 30. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM273 and pMM464 - pMM474 (A) Schematic representation of the differences between the pMM273 and pMM464 - pMM474 as well as a table describing intron modifications in every construct. Intron sequences added to the constructs were taken from the SMN2 transcript (B-F). FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM273 and pMM464 - pMM474 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

Figure 31. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM466, pMM470, pMM472, pMM474 and pMM567 - pMM570 (A) Schematic representation of the differences between the pMM466, pMM470, pMM472, pMM474 and pMM567 - pMM570. Two modifications were introduced to pMM567 - pMM570, single A>C mutation in exon 6 to remove an ATG start site and TAA deletion in exon 7. (B-F) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM466, pMM470, pMM472, pMM474 and pMM567 - pMM570 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

Figure 32. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM569 and pMM591 - pMM601. (A) Schematic representation of the differences between the pMM569 and pMM591 - pMM601. Two regions of pMM596 were modified: exon 7 region and intron 7 region. (B-G) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM569 and pMM591 - pMM601 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows induction of GFP expression calculated as median GFP expression at 250 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment.

GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM).

Figure 33. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM569, pMM642 - pMM651, and pMM709. (A) Schematic representation of the intron 6 deletion screen performed using pMM569 construct. 5 nt step by step deletions spanning the whole intron 6 were introduced and tested. (B) Selected intron 6 mutations tested in constructs pMM642 - pMM651, and pMM709. (C-H) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM569, pMM642 - pMM651, and pMM709. (C) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (D) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (E) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (F) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (G) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (H) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM).

Figure 34. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM569, pMM634 - pMM641, pMM711 and pMM712. (A) Schematic representation of the intron 7 deletion screen performed using pMM569 construct. 5 nt step by step deletions spanning the whole intron 7 were introduced and tested. (B) Selected intron 7 mutations tested in constructs pMM634 - pMM641, pMM711 and pMM712. (C-H) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM569, pMM634 - pMM641, pMM711 and pMM712 constructs. (C) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (D) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (E) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (F) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (G) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control

cells, 48h after risdiplam treatment. (H) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM).

Figure 35. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM569, pMM579, pMM595, and pMM596. (A) Schematic representation of the differences between the pMM569, pMM579, pMM595, and pMM596. Constructs pMM579, pMM595 and pMM596 were selected due to their low GFP expression in DMSO (low leakiness) (B-G) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM569, pMM579, pMM595, and pMM596 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM).

Figure 36. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM569, pMM618, and pMM619. (A) Schematic representation of the differences between the pMM569, pMM618, and pMM619. pMM618 has 13 nt deletion at the 5' side of exon 6, while pMM619 has 30 nt deletion of exon 6 5' side. (B-G) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM569, pMM618, and pMM619 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM).

Figure 37. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM569 switch with different promoters (pMM569, pMM581 - pMM584, pMM589, pMM590, and pMM562, and pMM619). (A) Schematic representation of the pMM569 switch and the list of all tested promoters. (B-F) FACS-based measurements of GFP

expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM569, pMM581 - pMM584, pMM589, pMM590, and pMM562, and pMM619 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (500 nM).

Figure 38. Analysis of risdiplam-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdiplam) for pMM263, pMM274, pMM273, pMM359, and pMM569. (A) Schematic representation of the differences between the pMM263, pMM274, pMM273, pMM359, and pMM569. These constructs represent various iterations of the switch design. (B-G) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM263, pMM274, pMM273, pMM359, and pMM569 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total median GFP signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdiplam divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype HEK293 cells (WT), or in cells expressing the indicated constructs, in the absence of risdiplam. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM concentration of risdiplam, normalized to median GFP expression in DMSO-treated control cells, 48h after risdiplam treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdiplam treatment (1 μ M).

Figure 39. Analysis of risdiplam-induced NanoLuc (NLuc) luciferase reporter expression, and background NLuc reporter expression (*i.e.* in the absence of risdiplam) for pMM568, pMM569, pMM570 and pMM579. (A) Schematic representation of the differences between the pMM568, pMM569, pMM570 and pMM579. (B-G) Luminescence-based measurements of NLuc luciferase expression upon induction with increasing concentrations of risdiplam in HEK293 cells stably expressing the pMM568, pMM569, pMM570 and pMM579 constructs. (B) Shows induction of NLuc expression calculated as NLuc signal expression at each concentration of risdiplam, normalized to NLuc signal in DMSO-treated control cells, 48h after risdiplam treatment. (C) Shows the total NLuc signal without normalization to DMSO control. (D) Shows the total NLuc signal for WT HEK293 cells as well as cells expressing indicated constructs, in the absence of risdiplam (DMSO treated).

Figure 40. Risdipram-mediated expression of Cas9 in HEK293T stable cell lines assessed by CD81 staining. (A) Schematic representation of the differences between the constructs pMM263, pMM274, pMM273, pMM359, and pMM569, representing various iterations of the switch design (B) Experimental setup. HEK293T with stably incorporated pMM263, pMM274, pMM273, pMM359, and pMM569 switches

5 controlling Cas9 expression were nucleoporated with 3 different sgRNAs targeting CD81 locus, each at 150 nM concentration, seeded at 96-well plates and right after treated with various concentrations of risdipram. After 72h cells were stained using APC-conjugated CD81 antibodies and analysed by FACS. (C) Percentage of CD81 positive cells in cells expressing pMM263 switch in either non-treated cells or cells treated with sgRNA and increasing concentrations of risdipram. (D) Percentage of CD81 positive

10 cells in cells expressing pMM274 switch in either non-treated cells or cells treated with sgRNA and increasing concentrations of risdipram. (E) Percentage of CD81 positive cells in cells expressing pMM273 switch in either non-treated cells or cells treated with sgRNA and increasing concentrations of risdipram (F) Percentage of CD81 positive cells in cells expressing pMM359 switch in either non-treated cells or cells treated with sgRNA and increasing concentrations of risdipram. (G) Percentage of CD81 positive

15 cells in cells expressing pMM569 switch in either non-treated cells or cells treated with sgRNA and increasing concentrations of risdipram. (H) Percentage of CD81 positive cells in cells expressing pMM596 switch in either non-treated cells or cells treated with sgRNA and increasing concentrations of risdipram.

Figure 41. Schematic representation of risdipram-inducible switch application for CAR-T therapy. (A)

20 Patient's isolated T cells would be isolated and modified to introduce inducible CAR construct either alone, or combined with a standard CAR expressing construct. Next, cells would be infused into patients, who would activate CAR expression using risdipram. The switch could be used either to express a single CAR receptor, or in combination with another CAR, to induce second CAR expression. Picture created with BioRender.com.

25

Figure 42. Analysis of risdipram-induced GFP reporter expression, and background GFP reporter expression (*i.e.* in the absence of risdipram) for pMM569, pMM579, pMM595, and pMM596 in Jurkat leukemic T-cell immortalized cell line. (A) Schematic representation of the differences between the pMM569, pMM579, pMM595, and pMM596. Constructs pMM579, pMM595 and pMM596 were selected

30 due to their low GFP expression in DMSO (low leakiness) (B-G) FACS-based measurements of GFP expression upon induction with increasing concentrations of risdipram in Jurkat cells stably expressing the pMM569, pMM579, pMM595, and pMM596 constructs. (B) Shows induction of GFP expression calculated as median GFP expression at each concentration of risdipram, normalized to median GFP expression in DMSO-treated control cells, 48h after risdipram treatment. (C) Shows the total median GFP

35 signal of GFP-positive cells without normalization to DMSO control. (D) Shows the percentage of GFP-positive (% GFP+) cells at each concentration of risdipram divided by the %GFP+ cells at the highest concentration of risdipram treatment (1 μ M). (E) Shows the background level of total median GFP expression in wildtype Jurkat cells (WT), or in cells expressing the indicated constructs, in the absence of risdipram. (F) Shows induction of GFP expression calculated as median GFP expression at 62.5 nM

40 concentration of risdipram, normalized to median GFP expression in DMSO-treated control cells, 48h after risdipram treatment. (G) Shows the percentage of GFP-positive (% GFP+) cells at DMSO divided by the %GFP+ cells at the highest concentration of risdipram treatment (1 μ M).

Figure 43. Risdiplam-mediated expression of GFP in AAV-transduced neurons. (A) illustration of the in-frame switch constructs used in this experiment: pLS76 the original switch and pMM359 switch, the more optimized counterpart. (B) hESC-derived neurons were differentiated from the neuronal stem cell (NSC)-state for 28 days, transduced with AAV7.m8 containing the sequence illustrated in A. After three days of AAV-transduction, Risdiplam was added to the cell culture at varying concentrations as indicated. Medium was exchanged every 3 days with fresh addition of risdiplam and presented images were acquired 7 days after risdiplam treatment.

Figure 44. Optimized in-frame switch pMM569 allows for controlled expression of MeCP2. (A) Schematic representation of the in-frame switch based on the sequence from pMM569. The gene of interest sequence of GFP was replaced by the MeCP2 coding sequence and cloned into piggybac plasmid backbone under CMV promoter. (B) H4 cells were used as wild type as a control for native MeCP2 expression in contrast to MeCP2-knockout (KO) H4 cells, which were generated as a single cell clone using CRISPR/Cas9-mediated knockout of MeCP2. H4-KO cells were transfected with pMM569, encoding for MeCP2 as a therapeutic protein. (C) Transfected H4-KO cells were treated with indicated concentrations of risdiplam for 24h and imaging was conducted using anti-MeCP2 antibody (cell signaling technology, cat#3456). Images were acquired and used to quantify the MeCP2 expression level in the respective groups. (D) Representative images are shown that were used to generate the quantification in C.

Examples

Example 1: Design and construction of the in-frame ON-switch

The in-frame ON-switch comprises exon 6, exon 7, and a partial sequence of exon 8 of *SMN2* (survival of motor neuron 2), along with intervening intronic regions (intron 6 and intron 7). GFP is used as a reporter gene. Exon 7 contains a single nucleotide insertion (A at position 48). This allows in-frame translation of GFP upon inclusion of exon 7 (in the presence of risdiplam) by inactivation of the translation termination codon at the 3' end of exon 7¹⁷.

The entirety of exon 6 was used, while the downstream intronic region was reduced to 414 nucleotides (from its original length of 5769 nucleotides) comprising the first 202 and the last 212 nucleotides of intron 6. The entirety of intron 7 was used, followed by the first 23 nucleotides of exon 8 (Figure 1). A Kozak sequence (GCCACC) as well as an ATG codon were placed upstream of exon 6 for the initiation of translation. The gene expression was driven by the human cytomegalovirus (CMV) promoter. In the presence of increasing concentrations of the small molecule splicing modifier (SMSM), risdiplam, splicing occurs, causing inclusion of exon 7 in the final mRNA transcript, and in-frame translation of the GFP reporter gene.

Example 2: Inducible reporter gene expression using the in-frame ON-switch in the presence of small molecule splice modifier (risdiplam analogue)

Mammalian HEK293 cells were cultured and transfected as described in Example 23 with plasmids that encode either the in-frame, or frame-shifted gene cassette (see Figure 1 for schematic). Cells were seeded in 96-well plates at 10,000 cells/well and treated with varying amounts of risdiplam analogue (RO7021707-000-007). After 48h of incubation, cells were collected for flow-cytometry-based analysis using the CytoFLEX Flow Cytometer from Beckman Coulter (Figure 2). The fluorescence intensity of GFP (for in-frame switch) was measured and analysis performed using Flowjo software. Median GFP intensity of each of the treated conditions was normalized to DMSO-treated cells to result in a relative fold induction rate of reporter gene expression, informing on the performance (induction rate) of the switch. We measured increased GFP expression upon induction with the risdiplam analogue at nanomolar concentrations, which resulted in toxicity at high concentrations (>1.25 μ M), leading to reduced GFP expression level (Figure 2).

Example 3: Optimization of the in-frame switch: Reduction in leakiness (background expression)

An improved version of the in-frame switch was generated by introducing two point mutations within the 3' splice site of exon 7: acagGGT (WT) to ccagGAT (mutant), which has been shown in a previous study to promote skipping of exon 7¹⁸ (Figure 3A). This construct was termed pLS76 and was used as a starting construct for further modifications as described below. In order to characterize the construct, HEK293 cells were transfected with WT (pLS41) and mutated (pLS76) constructs and treated with increasing amounts of risdiplam compound. The nucleotide sequences of the pLS41 and pLS76 ON-switches are shown in SEQ ID NOs:37 and 38, respectively. Fluorescence activated cell sorting (FACS)-based readouts revealed reduced background expression in the modified sequence (ccagGAT) compared to WT (acagGGT), with a comparable induction rate (Figure 3B-3D). Figure 3B, shows relative fold change in median GFP expression level normalized to DMSO control 48h after risdiplam treatment at the indicated concentrations. The total GFP signal is shown in Figure 3C, whereas Figure 3D represents the % of GFP-positive cells normalized to the condition with the highest risdiplam treatment. Untransfected, wildtype HEK293 cells serve thereby for gating out GFP-negative cells. This type of data presentation allows the demonstration of the increase in the number of GFP-positive cells rather than the GFP intensity upon induction with risdiplam. Next, we assessed the leakiness of the cells by comparing total GFP expression in cells treated with DMSO only (Figure 3E) and found lower background GFP expression level in the construct that contains the acagGGT> ccagGAT mutation. The splicing efficiency was finally analyzed by RT-PCR experiments showing partial (at 100 nM) or complete (at 1 μ M) inclusion of exon 7 (Figure 3F) in the improved construct (pLS76).

Example 4: Optimization of the in-frame switch: Reduction in intron length

One of the main limitations of AAV-mediated gene therapy is the size constraint of the whole expression cassette of 4.7 kb. Thus, usage of small regulatory elements in AAV vectors is highly desirable. In order to optimize the size of the in-frame switch (pLS76), the length of the introns separating exon 6 and 7, and exon 7 and 8 were gradually reduced.

In a first step, the central region of intron 6 was removed, leaving 102 nucleotides of the 5' end and 162 nucleotides of the 3' end of intron 6. Similarly, the central region of intron 7 was removed, leaving 101 nucleotides at the 5' end and 151 nucleotides at the 3' end. In both cases, only the central intronic regions were removed, in order to minimise the chances of disrupting splicing regulatory sequences, which are typically located in close proximity to 5' and 3' splice sites (*E* adjacent to their upstream and downstream flanking exons). This construct, termed pLS159, reduced the total length of the in-frame switch by 314 nucleotides. The length of the 5' end and 3' end of intron 6 and intron 7 were then further reduced to: intron 6 (72 nucleotides at 5' end and 112 nucleotides at 3' end); intron 7 (77 nucleotides at 5' end and 100 nucleotides at 3' end). This construct, termed pLS160, led to a total reduction of 492 nucleotides of the in-frame switch length. The nucleotide sequences of the pLS159 and pLS160 ON-switches are shown in SEQ ID NOs:39 and 40, respectively.

The effect of the different modifications on the in-frame switch performance as assessed with FACS revealed that a moderate reduction in intron length by 314 nucleotides (pLS159) leads to a smaller construct with overall better performance. It was characterized by a higher fold change in GFP expression level upon induction with risdiplam relative to DMSO-treated cells (Figure 4B), an increase in total level of median GFP intensity (Figure 4C), as well as an increase in the number of GFP-positive cells (Figure 4D), while background expression level remained unchanged (Figure 4E). In contrast, extensive reduction in the length of introns by 492 nucleotides (pLS160) did not lead to increased GFP expression levels (Figure 4B-4D).

Example 5: Optimization of the in-frame switch: Reduction in exon 6 length.

As exonic splicing regulatory sequences are usually localized close to the 5'SS, the inventors investigated whether it was possible to further reduce the size of the in-frame switch, by gradually reducing its length from the 5' end of exon 6. To do so, two variants of the switch were generated that were lacking 30 (pLS167) or 60 (pLS168) nucleotides of the 5' end of exon 6 (illustrated in Figure 5A). Results showed increased overall expression of the reporter gene (GFP) as measured by FACS-based analysis. The nucleotide sequences of the pLS167 and pLS168 ON-switches are shown in SEQ ID NOs:41 and 42, respectively.

In these experiments, HEK293 cells were stably transfected, followed by treatment with risdiplam at the indicated concentrations for 48(H)Original in-frame switch construct pLS76 served as a control. In comparison to pLS76, partial truncation of exon 6 by 30 bp (pLS167) significantly outperformed the original sequence, in terms of fold change increase (Figure 5B) as well as total increase in median GFP signal (Figure 5C). Although less pronounced than in pLS167, the pLS168 construct also showed increased GFP expression in comparison to pLS76 (Figure 5B-5C). Increased GFP expression from pLS167 and pLS168 constructs relative to pLS76 may be a consequence of the polypeptide added to GFP encoded by the pLS167 and pLS168 constructs. Increased GFP expression from pLS167 relative to pLS168 may be a consequence of pLS167 comprising a 'G' immediately after the Kozak and start codon sequences, instead of 'A' (as in the pLS168 construct).

The number of GFP-positive cells in cultures expressing the original construct, pLS76, increased proportionally with increasing concentration of risdiplam. Similar observations were made with cells expressing pLS168 (Figure 5D). However, the number of GFP-positive cells in cultures expressing pLS167 was very high in cells treated with DMSO control, suggesting higher leakiness in these cells.

- 5 Furthermore, absolute median GFP intensity between the various cells was compared, showing highest background expression level in cultures expressing pLS167 (Figure 5E). Therefore, for further optimization, pLS168 was chosen due to lower leakiness and improvement in the induction rate.

Example 6: Combined modifications of reduced intron length and reduced exon length in one construct.

- 10 Next, the modifications of pLS159 and pLS168 were combined into one construct. The minimal required intron length for optimal induction (pLS159) was combined with the minimal required length of exon 6 (pLS168). This combination resulted in construct pLS179 that contained a shorter sequence of exon 6 and a shorter sequence of intron 6 and intron 7. The nucleotide sequence of the pLS179 ON-switch is
15 shown in SEQ ID NO:43.

- In total, the length of the switch was reduced from 1056 to 639 nucleotides (total reduction of 417 nucleotides) (Figure 6A). Furthermore, an ATG start codon present within the exon 6 sequence, downstream of the synthetic Kozak/Start codon was mutated (TAC**ATG**AGT>TAC**GGC**AGT) in order to
20 reduce alternative translation start sites. Similarly, 6 nucleotides at the 5' end of the remaining exon 6 sequence (as modified in pLS168) were deleted: AGTATG, resulting in a shorter final sequence length of exon 6 of 45 bp. Two additional point mutations were introduced, A>G and T>C to improve the KOZAK sequence. To test these constructs, HEK293 were stably transfected. Cells were then treated with increasing amounts of risdiplam and GFP expression level was assessed using FACS-based approaches
25 (Figure 6B-6E) or RT-PCR analysis (Figure 6F). The new construct that combined the exon 6 truncation with the intron 6 and intron 7 truncations resulted in an in-frame switch with significantly higher sensitivity to low nanomolar risdiplam concentrations. Results showed higher levels of GFP expression relative to DMSO control (Figure 6B), higher total GFP expression level (Figure 6C), as well as increased number of GFP-positive cells (Figure 6D). The latter is in line with the overall increase in background expression
30 level of GFP (in DMSO-treated cells, Figure 6E).

- Furthermore, splicing efficiency was assessed by RT-PCR experiments. RNA was extracted from HEK293 cells stably transfected with pLS179 and treated with risdiplam at low (100nM) or high (1μM) concentrations. Results showed partial inclusion of exon 7 at low concentrations of risdiplam and full
35 inclusion at high concentrations of risdiplam, correlating with increase in GFP signal (Figure 6F). In summary, results showed increased protein expression level upon appropriate truncation in the introns and exon 6 region.

Example 7: Reduction in nucleotide sequence of exon 8.

- 40 Based on the design of the in-frame switch and the position of the translation start site at the 5' end of exon 6, the final amino acid sequence expressed using the in-frame switch is composed of the transgene of interest linked to a fragment of SMN2 exon 6, exon 7, and exon 8 due to inclusion of these exons after

RNA splicing. This results in the generation of a tag at the N-terminal part of the expressed transgene, which might destabilize the protein or trigger an immune reaction. In order to reduce the tag length as much as possible, the length of the nucleotide sequence of exon 8 was reduced in a stepwise manner. As depicted in Figure 7, 5 different constructs were generated that were based on the optimized construct pLS179 (Figure 6) with varying lengths of exon 8 (starting construct pLS179: GAAATGCTGGCATAGAGCAGCAC; pMM70: GAAATGCTGGCATAGAG; pMM71: GAAATGCTGGC; pMM72: GAAATGCT; pMM73: GAAAT; and pMM112: GA). The first two nucleotides (GA) of exon 8 were kept unchanged in construct pMM112 in order to retain the 3' splice site consensus motif. The nucleotide sequences of the pLS179, pMM70, pMM71, pMM72, pMM73 and pMM112 ON-switches are shown in SEQ ID NOs:43, 44, 45, 46, 47 and 22, respectively.

HEK293 cells were stably transfected with these constructs, and treated with increasing amounts of risdiplam. FACS-based analysis showed that the reduction in exon 8 nucleotide sequence did not lead to significant changes in the relative (Figure 7B) or in the total (Figure 7C) GFP expression level in comparison to untruncated exon 8 (pLS179). Furthermore, no significant changes in the number of GFP-positive cells were observed (Figure 7D) and only a slight increase in the background expression level was seen (Figure 7E), suggesting that the truncations of exon 8 did not have a strong effect on the splicing-dependent GFP expression, while reducing the length of the tag by 6 amino acids (pMM112).

Example 8: Reduction of the tag sequence by using ATG start site contained at the 3' end of exon 6.

Exon 6 inherently contains an ATG at the 3' end (5'...TATTATATG'3). It was hypothesized that this ATG start site would be sufficient to drive translation and could potentially replace the synthetic Kozak/start site that had been introduced at the 5' end of exon 6. A successful translation driven by an ATG in closer proximity to the transgene sequence will reduce the length of any amino acids that will be attached to the final transgene output due to exon inclusion.

A new construct (pMM130) was generated based on pMM112, with the Kozak sequence/start site removed from the 5' end of exon 6. The design of the constructs and their respective sequences are illustrated in Figure 8A. The nucleotide sequence of the pMM130 ON-switch is shown in SEQ ID NO:23. HEK293 cells were stably transfected with the constructs and analysed by FACS to assess relative (Figure 8B) and total (Figure 8C) GFP expression levels. The induction rate of the in-frame switch in response to varying concentrations of risdiplam could be maintained by the pMM130 construct that lacks a canonical Kozak (*i.e.* conforming to SEQ ID NO:33) and relies on the ATG site at the 3' end of exon 6. In addition, no changes were observed in the number of GFP-positive cells as well as the background expression of GFP in DMSO-treated cells when comparing pMM130 to starting construct pMM112 or pLS179. With the modification in pMM130 that relies on ATG start site at the 3' end of exon 6 the length of the tag added to the final transgene output could be reduced by 15 amino acids. In combination, the various different modifications reduced the length of the N-terminal tag from 64 amino acids in pLS76 to 20 amino acids in pMM130. In addition, the overall size of the switch was reduced from 1056 nucleotides for pLS76, to 618 nucleotides for pMM130.

Example 9: Mutations in the Intronic Splicing Silencer-N1 (ISS-N1) region in order to reduce leakiness.

ISS-N1 is a 15-nucleotide long motif located immediately downstream of the 5' splice site of exon 7 of the SMN2 gene ((5'..GA|GTAAGTCTGCCAGCATTATGAAAGTGA..'3), ISS-N1 motif marked as bold). ISS-N1 was described as a major inhibitory element and a master checkpoint of SMN2 exon 7 splicing because its deletion fully restored exon 7 inclusion (Singh *et al.* RNA (2010) 16(6):1167-81). Modifications within the inhibitory region could help reduce background splicing events (leading to inclusion of exon 7 in the absence of the SMSM) driven by the native SMN2 sequence. Therefore, mutations were introduced within the ISS-N1 region that were found to lead to increased inhibition of exon 7 inclusion¹⁹. The modifications included nucleotide mutations: CAGCA>TCCTC (pMM143), CAGCA>AAGGC (pMM144), CACGA>TAGTC (pMM145); nucleotide insertions: TTT after position 10 of intron 7 (pMM146), and CCC after position 15 of intron 7 (pMM147). The starting construct for these modifications was pLS179. The nucleotide sequences of the pMM143, pMM144, pMM145, pMM146 and pMM147 ON-switches are shown in SEQ ID NOs:48, 49, 50, 51 and 52, respectively.

The modified constructs were stably transfected into HEK293 cells. Risperidone-dependent GFP expression was assessed using FACS-based readouts. In comparison to pLS179, modifications in the ISS-N1 region maintained comparable fold-change increase in relative GFP expression levels (Figure 9B), while total GFP expression levels were reduced (Figure 9C). Furthermore, results showed a reduction in the GFP-positive cells in DMSO-treated cells (Figure 9D) as well as reduced total median GFP intensity (Figure 9E). In conclusion, reduced leakiness results in reduced overall reporter gene expression, while relative reporter gene expression (to DMSO-treated cells) remains unchanged.

Example 10: Mutations in the ESE2 region to improve switch performance

The ESE2 region (AAAAAGAAAGGAA) was previously described as a risperidone binding site¹⁵. To further optimize the in-frame switch performance by means of increasing induction rate, single point mutations were introduced, within the ESE2 region as highlighted in bold in Figure 10A. HEK293 cells were stably transfected with constructs carrying the respective mutations (pMM136-pMM142) and treated with increasing concentrations of risperidone. The nucleotide sequences of the pMM136, pMM137, pMM138, pMM139, pMM140, pMM141 and pMM142 ON-switches are shown in SEQ ID NOs:53, 54, 55, 56, 58 and 59, respectively.

No increase in the relative (Figure 10B) or total (Figure 10C) GFP expression level was observed. For the majority of the modifications (pMM136, pMM137, pMM138, pMM139, pMM142), increased GFP expression in DMSO-treated cells was seen, suggesting that the risperidone-mediated splice regulation was hampered (Figure 10C). Similarly, the number of GFP-positive cells (Figure 10D) as well as the background expression level in DMSO-treated cells (Figure 10E) also indicated that the risperidone-mediated splice regulation was impaired for the majority of modifications. However, two modifications (pMM140 and pMM141) showed comparable (pMM140) or a slightly lower (pMM141) leakiness in terms of both the number of GFP-positive cells (Figure 10D) and median GFP expression in DMSO (Figure 10E) compared to pLS179. In summary, the majority of indicated modifications within the ESE2 region do

not lead to improved performance of the in-frame switch, with only the pMM141 construct causing slightly reduced leakiness.

Example 11: Partial deletions in the exon 7 region do not lead to increased induction rate but cause slight reduction in leakiness.

The risdiplam-mediated effect is dependent on exon 7 and was previously described to involve the ESE2 region as well as the 5' splice site¹⁵. With the aim of reducing the length of the amino acid tag on the gene of interest, partial deletions in the exon 7 sequence were performed as described in Figure 11A (pMM151: deletion of 12bp, pMM152: deletion of 12bp and 9bp). The nucleotide sequences of the pMM151 and pMM152 ON-switches are shown in SEQ ID NOs:60 and 61, respectively. HEK293 cells were stably transfected with the original construct pMM112 and with the modified constructs pMM151 and pMM152 and treated with risdiplam at increasing concentrations. After 48h of treatment, FACS-based analysis was used to assess GFP intensity. The reduction in the exon 7 region resulted in lower induction levels as measured by median GFP signal relative to DMSO (Figure 11B) as well as the total median GFP signal (Figure 11C). In line with previous observations linking lower induction rate to lower background signal, results showed a lower number of cells expressing GFP in the absence of risdiplam (Figure 11D) and lower median GFP signal in DMSO (Figure 11E) in cells transfected with pMM151 and pMM152 in comparison to pMM112 (containing the complete exon 7 sequence). In conclusion, reduction in the exon 7 region reduces the leakiness of the in-frame switch.

Example 12: Introducing translation start site within exon 7

The current construct design leads to the generation of a residual amino acid chain that is fused to the N-terminus of the gene of interest due to the positioning of the translation start site (Kozak and ATG) at the 5' end of exon 6. With the aim of reducing the tag length, the Kozak and ATG start site at exon 6 were removed and all alternative ATGs within the exon 6 sequence were mutated as described in Figure 12A(i) (ATG>GGC). This construct was named pLS174 and served as the starting construct to introduce Kozak and ATG start sites within the exon 7 sequence. The nucleotide sequence of the pLS174 ON-switch is shown in SEQ ID NO:62.

As described in Figure 12A(ii), 6 different constructs were generated, in which a Kozak/ATG site was created by single point mutations or nucleotide insertions. The first construct (pMM56) contained mutations downstream of the ESE2 region to recreate a Kozak and ATG start site (GCTCACAT- ->GCCACCATG), construct pMM59 contained an insertion of Kozak sequence and mutation of the sequence to resemble ATG upstream of the ESE2 region (-----AGA > GCCACCATG), pMM60 was modified by 2 point mutations and an insertion (AT----- > GCCACCATG), pMM61 contained mutations and a single nucleotide insertion to create a Kozak-like sequence that has previously been shown to enhance translation initiation site²⁰, (T-ATAAG > CAACATG), pMM62 contained mutations and a single nucleotide insertion (AATATAA-G > GCCACCATG), pMM63 contained mutations and 2 nucleotide insertions (TT—CCTTA > GCCACCATG). The nucleotide sequences of the pMM56, pMM59, pMM60, pMM61 and pMM62 ON-switches are shown in SEQ ID NOs:63, 64, 65, 66 and 67, respectively.

As previously described, HEK293 cells were stably transfected with the respective constructs and blasticidin-selected cell pool was treated with increasing amounts of risdiplam as indicated in Figure 12B-12D. The original construct pLS76 (Figure 3) was used as a control construct. Using FACS-based measurements, pMM56 and pMM60-63 were found to result in high levels of GFP-positive cells and

5 constitutive GFP expression in the absence of risdiplam (Figure 12C–12E) leading to lack of control of transgene expression with risdiplam (Figure 12B). In contrast, cells expressing pMM59 showed an increase in the GFP induction rate to much higher level than cells expression pLS76 (Figure 12B), as well as an increase in the total median GFP signal with increasing risdiplam concentrations (Figure 12C). The total GFP-positive cells in pMM59 were significantly higher than in pLS76, but to a lower extent than

10 pMM56 and pMM61-63 (Figure 12D). Furthermore, the background expression of GFP in the absence of risdiplam was only slightly higher in pMM59 in comparison to pLS76, however, significantly lower than in pMM56 and pMM61-63. In pMM59, the Kozak/ATG translation start site is provided upstream of ESE2, whereas in pMM56 and pMM60-63 the Kozak/ATG translation start site is provided downstream of ESE2. The presented data suggest that pMM59 with a Kozak sequence/start site introduced upstream of the

15 ESE2 region provides very high induction rate as well as low leakiness and can be used to reduce the size of the tag on the final transgene output. Additionally, pMM59 represents a new in-frame switch design, where the translation start site is inside the inducible exon itself.

Example 13: Modifications in the 5' splice site of exon 6 in order to remove the ATG start site.

20 Alternative ATG translation start sites within exon 6 were mutated in pLS176 as indicated in Figure 13A and highlighted in bold. The previous construct pMM59 carried a modification at the 5' splice site of exon 6 from ATG to CTG. With the aim of testing various modifications and their impact on the 5' splice site usage of exon 6, 3 additional constructs to pMM59 were generated, in which the ATG sequence at the 3' end of exon 6 was mutated to CAG (pMM108), GTG (pMM110), TTG (pMM111) in comparison to CTG

25 (pMM59). The nucleotide sequences of the pMM108, pMM110 and pMM111 ON-switches are shown in SEQ ID NOs:69, 70 and 71, respectively.

HEK293 cells were stably transfected with pMM59, pMM108, pMM110 and pMM111 and treated with different concentrations of risdiplam for 48h. FACS-based analysis on the median GFP expression

30 relative to DMSO (Figure 13B) as well as the total GFP expression (Figure 13C) revealed no major differences between pMM108 and pMM59 in-frame switches, leading to a conclusion that both CTG and CAG can be used in the 5' splice site. However, when mutating the 3' end of exon 6 to TTG (pMM111) or GTG (pMM110), the GFP induction rate is reduced. In line with previous results showing that lower leakiness follows lower induction rate, results showed reduced total number of GFP-positive cells (Figure

35 13D) as well as reduced median GFP expression in DMSO-treated cells stably expressing pMM110 and pMM111 (Figure 13E). For the following modifications, pMM59 was used as a starting construct.

Example 14: Introduction of translation start site upstream of the ESE2 sequence in exon 7.

Based on the previous results using pMM59, a construct in which Kozak and ATG start site were

40 positioned upstream of the ESE2 sequence by point mutations and insertions, various constructs that contain Kozak and ATG at different positions upstream of the ESE2 region were generated, as illustrated in Figure 14A. Constructs were generated containing: Kozak-ATG (GCCACCATG, in pMM59 and

pMM242); Kozak-like/ATG (GCACCATG in pMM243 and pMM244); another Kozak-like/ATG sequence (CACCATG, pMM245); only ATG insertion (pMM246); and ATG insertion within the ESE2 region (pMM247). In addition, pMM244, pMM245, and pMM246 contained a C>G mutation that was placed downstream of the inserted ATG. The nucleotide sequences of the pMM242, pMM243, pMM244, pMM245, pMM246 and pMM247 ON-switches are shown in SEQ ID NOS:72, 73, 74, 75, 76 and 77, respectively.

Results showed risdiplam-regulated induction only in the constructs pMM243 and pMM247, whereas all the other constructs with the described modifications did not show any risdiplam-dependent GFP expression as measured using FACS (Figure 14B-14C). Instead, results showed a high number of GFP positive cells (Figure 14D) as well as high median GFP expression (Figure 14E) in DMSO-treated cells, indicating very high leakiness and lack of regulation through these mutations (pMM242, pMM244, pMM245, pMM246). There are several important conclusions from the experiment. Firstly, a comparison between pMM59 and pMM242 shows that moving the Kozak sequence by merely 3 nucleotides leads to constitutive GFP expression in pMM242 compared to pMM59. Furthermore, comparison between pMM243 and pMM244, which differ only by a single point mutation (C to G), shows that the region around ESE2 is highly sensitive to changes. Thus, it is unexpected that the insertion of ATG in the middle of ESE2 sequence leads to a fully functional switch, with very desirable properties (low leakiness and high induction rate). This observation suggests that the current understanding of the role and sequence requirements of ESE2 might not be complete.

Example 15: Introduction of translation start site within exon 7 and downstream of the ESE2 sequence.

Different outcomes have been seen from different modifications and positioning of the translation start site in relation to the ESE2 region. Therefore, it was decided to further evaluate the effect of positioning the ATG start site downstream of the ESE2 sequence. pLS175 was used as the starting construct that contained mutations of ATG start site (indicated in bold in Figure 15A) as well as ATG>CAG mutation in the 5' end of exon 6. From this starting construct, a wide variety of constructs were generated: pMM123 (insertion of T at position 53 of exon 7), pMM124 (insertion of G at position 49 of exon 7, a mutation A>G at position 50), pMM125 (insertion of AT at position 31 of exon 7, a single point mutation of T>G at position 32 to resemble ATG, and an insertion of G at position 34 of exon 7), pMM126 (insertion of G at position 40 of exon 7), pMM127 (insertion of A at position 43 and two point mutations AA>GG at position 42), pMM128 (insertion of AT at position 54 of exon 7), and pMM129 (deletion of AT at position 47 and mutation of A>G at position 51). The nucleotide sequences of the pLS175, pMM123, pMM124, pMM125, pMM126, pMM127, pMM128 and pMM129 ON-switches are shown in SEQ ID NOS:78, 79, 80, 81, 82, 83, 84 and 85, respectively.

Results showed risdiplam-dependent increase in relative GFP induction rate (Figure 15B) as well as increase in total level of median GFP (Figure 15C) in the construct pMM126. All the other constructs were less sensitive in the induction rate, possibly due to high leakiness as shown by measuring the total GFP-positive cells as well as total median GFP intensity in DMSO-treated cells (Figure 15D-15E) as seen in pMM123, pMM125, pMM127, and pMM128. The two constructs pMM124 and pMM129 also showed low

levels of leakiness (Figure 15E), however, no induction was seen in pMM124 and relatively low induction in pMM129. In the case of pMM124, the lack of induction was due to the formation a cryptic splice site (GT) splicing to exon 8, as described previously (WO 2009/151546 A2).

5 **Example 16: Additional modifications to pMM126 in order to improve performance.**

Following up on the successful insertion of an ATG site downstream of the ESE2 sequence in pMM126, pMM126 was further modified in order to increase the induction rate by generating a stronger translation start site. Six new constructs (illustrated in Figure 16A) were designed as follows: pMM201 (mutation T>G), pMM202 (C>A, and T>G), pMM203 (generation of a novel ATG start site: G>A and C>G, T>C),
 10 pMM204 (GT>TC, C>A, and insertion of TG to reconstitute ATG), pMM205 (deletion of G at position 40, insertion of TG at position 46), pMM206 (deletion of G at position 40, insertion of ATGGA at position 45). The nucleotide sequences of the pMM201, pMM202, pMM203, pMM204, pMM205 and pMM206 ON-switches are shown in SEQ ID NOs:86, 87, 88, 89, 90 and 91, respectively.

15 The described constructs were transfected into HEK293 cells for stable expression and treated at increasing risdiplam concentrations for 48(H)Although FACS-based measurements showed that introducing the modifications described in Figure 16A did not result in improved GFP induction rate (Figure 16B) or total GFP median expression level (Figure 16C), constructs pMM204 and pMM205
 20 showed reduced leakiness compared to pMM126 (lower median GFP signal in DMSO) with a comparable induction rate. Moreover, results showed an increased number of GFP-positive cells in the DMSO-treated cells, stably expressing pMM201, pMM202 and pMM206 (Figure 16C-16E). Interestingly, single or double nucleotide mutations in constructs pMM201 and pMM202 lead to constitutive GFP expression, while bigger changes as in pMM203 had much smaller impact, suggesting again that the region downstream of ESE2 is very sensitive to modifications.

25

Example 17: Engineering of a stronger translation start site in pMM129

As described in Figure 15, pMM129 did show improved performance in terms of low leakiness and still contained acceptable induction rate, although to a significantly lower extent than pMM126. Nevertheless, as pMM129 has significantly reduced tag size compared to pMM126, the sequence context of the ATG
 30 start site was modified in order to create a strong translation start site, with the aim of increasing the GFP expression rate in the presence of risdiplam. Based on the work by Noderer *et al.*²⁰, the following mutations and/or insertions were introduced: pMM248 (insertion of TAA at position 49, G>T mutation at position 57), pMM249 (T>A at position 44), pMM250 (T>G at position 44), pMM251 (T>C mutation at position 42, deletion of TT at position 43). The nucleotide sequences of the pMM248, pMM249, pMM250,
 35 and pMM251 ON-switches are shown in SEQ ID NOs:92, 93, 94 and 95, respectively.

The above-described constructs (illustrated in Figure 17A) were transfected into HEK293 cells for stable expression and treated at increasing risdiplam concentrations for 48h. No improved performance was observed through the introduction of these modifications, *i.e.* no increase in GFP induction rate (Figure
 40 17B), increase in total level of median GFP including the DMSO-treated control (Figure 17C-17E), suggesting that the introduced modifications led to increased leakiness of the constructs and no improvements in risdiplam-controlled GFP expression. Again, increase in leakiness in constructs pMM249

and pMM250, having just single point mutations compared to pMM129, shows the sensitivity of this region to any changes.

Example 18: Implementing learnings on translation start site (Kozak/ATG) within exon 7 on the short in-frame switch pMM130

Next, it was decided to apply the learnings from moving the translation start site to exon 7 on the shortest version of the in-frame switch, pMM130 (Figure 8). As illustrated in Figure 18A, the ATG site in construct pMM130 was firstly removed to create pMM252, with all ATG sites mutated. The modifications of the previously described pMM126, pMM128, and pMM129 were added to pMM252, resulting in pMM254, pMM255, and pMM256, respectively. One additional modification was added resulting in the construct pMM257, which contained an ATG start site by insertion of a G at position 49. The nucleotide sequences of the pMM252, pMM254, pMM255, pMM256 and pMM257 ON-switches are shown in SEQ ID NOs:96, 97, 98, 99 and 100, respectively.

These modifications result in a significantly reduced length of the amino acid tag to the transgene of interest (see table "Tag length" in Figure 18A). HEK293 cells were transfected with the indicated constructs and selected for stable expression under blasticidin selection marker. Cells were treated with varying concentrations of risdiplam for 48h before GFP expression was measured by FACS. Results showed a general increase in leakiness from each of the modifications that were introduced into the shorter in-frame switch pMM130. pMM254 (short in-frame switch with translation start site in exon 7) showed higher leakiness and hence lower induction rate than the control construct pMM126 (long in-frame switch with translation start site in exon 7). Similar observations were made with pMM255 and its control construct pMM128, as well as pMM256 and its control construct pMM129. These observations were true for GFP induction rate (Figure 18B), total level of median GFP (Figure 18C), as well as the number of GFP-positive cells (Figure 18D). The effect of increased background expression level (GFP expression in DMSO-treated cells) is further shown in Figure 18E. However, important to note is that despite the increased leakiness, the newly modified constructs result in a shorter tag length (see Figure 18A). Finally, pMM257 did not lead to increase in leakiness; however, induction rate was slightly reduced (Figure 18A-18E).

Example 19: Engineering of splice-induced formation of ATG translation start site between exon 7 and exon 8

All the previously described constructs had a short peptide added to the expressed transgene. In order to remove the tag entirely, we generated a split-ATG approach that was modified to contain an A at the 3' end of exon 7 and a TG at the 5' end of exon 8. With no risdiplam added, splicing occurs between exon 6 and exon 8, and no translation start site is formed. Upon treatment with risdiplam, exon 7 is spliced in, forming the ATG site together with exon 8 and driving translation of the gene of interest. No additional tag is added in this case (Figure 19B). In order to achieve this, constructs were designed based on pLS175, a construct in which alternative ATG start sites in exon 6 were mutated (Figure 19A). As the construct pLS175 contains 23 nucleotide long exon 8 sequence, an 8 amino acid tag still remains. However, as demonstrated in the Figure 7, exon 8 can be truncated to just 2 nucleotides, thus completely removing the tag. As described in Figure 19A, 3 constructs were generated with different mutations: pMM198 contained

a mutation GA>TG in the 5' end of exon 8, pMM199 and pMM200 contained the same mutation as pMM198 but were additionally modified in the intron 7 region in order to increase the strength of polypyrimidine tract with the aim of increasing splicing efficiency between exon 7 and exon 8 and consequently compensating for the generation of a weak splice site by mutating the 5' end of exon 8 from GA (consensus) to TG (non-optimal splice site). The described constructs were transfected into HEK293 cells for stable expression and treated with risdiplam at increasing concentrations for 48h before they were assessed for GFP expression by FACS-based measurements. pMM112, the short improved in-frame switch was used as a control construct to compare the performance of the split-ATG constructs. The nucleotide sequences of the pMM198, pMM199 and pMM200 ON-switches are shown in SEQ ID NOs:25, 101 and 102, respectively.

Results showed that all the split-ATG-based constructs (pMM198-pMM200) showed sensitivity to risdiplam treatment as seen by induced GFP expression relative to DMSO control (Figure 19C), as well as the total level of median GFP expression (Figure 19D), although to a lower extent than pMM112. The increase in the number of GFP-positive cells upon increase in the concentration of risdiplam was also measured (Figure 19E). The leakiness of the split-ATG switch was slightly higher than in pMM112 (Figure 19F), however, the modifications that were introduced in intron 7 did not lead to significant changes in the GFP expression level (Figure 19C-19E). In conclusion, the in-frame switch can be successfully engineered to form an ATG translation start site dependent on exon 7 inclusion without compromising the risdiplam-mediated regulation of GFP expression.

Example 20: Introducing ATG start site in exon 8

In another attempt to reduce the length of the amino acid tag fused the gene of interest, the placement of the ATG start site at the 5' end of exon 8 sequence was investigated. To test this approach, pLS76 was used as a starting construct. The Kozak and ATG sites were removed from the 5' end of exon 6 and alternative ATG start sites within exon 6 were mutated (Figure 20A). Additionally, the 3' end of exon 6 was mutated in each of the modified constructs as follows: pMM236 (ATG>CAG), pMM237 (ATG>TAG), pMM238 (ATG>TTG). The mutations in combination with the upstream sequence result in a suboptimal Kozak-like sequence. The exon 7 sequence remained as described for pLS76. In order to introduce a translation start site within exon 8, a mutation was introduced at the 5' end of exon 8 in pMM236-pMM238 (G-A>ATG) as illustrated in Figure 20A. The risdiplam-mediated gene expression control is retained by the presence of a Kozak-like sequence within the native sequence of exon 7 (highlighted in bold in Figure 20B). The nucleotide sequences of the pMM236, pMM237 and pMM238 ON switches are shown in SEQ ID NOs:103, 104 and 105, respectively.

The presented concept relies on the presence of risdiplam to allow inclusion of exon 7 and hence the introduction of an efficient Kozak-like sequence contained within exon 7 (TAAGGA) to enhance translation driven by the downstream ATG start codon present in exon 8 (at the 5' end of pMM236-pMM238). In the absence of risdiplam, only a weak Kozak-like sequence (provided by the mutations at the 3' end of exon 6, see Figure 20A) will be present and the translation start will not be enhanced. The presented design was tested in HEK293 cells that were stably expressing pMM236-pMM238 and compared to the improved in-frame switch pMM112. Cells were treated with increasing amounts of

risdiplam for 48 h, followed by FACS-based analysis. Results showed that GFP expression by the present constructs can be induced with increasing amounts of risdiplam as shown by assessment of the induction rate (Figure 20C) as well as the total median expression level of GFP (Figure 20D), suggesting that the weak/strong Kozak-like concept in combination with exon 7 inclusion allows a certain control over GFP expression level. However, high GFP expression levels were also measured in the uninduced, DMSO-treated cells as shown by the total cell number of GFP-positive cells (Figure 20E) as well as the median GFP signal (Figure 20F). As hypothesized, the weaker the Kozak-like sequence in exon 8, the lower the GFP expression in DMSO and higher the induction rate. In summary, the Kozak sequence-based switch does allow control over gene expression, however, the background expression is elevated.

Example 21: Engineering a SMN2-based OFF-switch

Inducible gene expression systems could not only be developed as an ON-switch to turn on gene expression but also as an OFF-switch to inhibit gene expression. In this context, a risdiplam-inducible OFF-switch was developed that would respond to various concentrations of risdiplam by suppressing transgene translation. The design strategy of the presented ON-switch was used and a construct was engineered to allow inhibition of translation by introducing a single nucleotide change (Figure 21A). To create an OFF-switch, the design was based on the pMM130 construct having reduced intron length (intron 6 and 7), reduced exon length (exon 6 and 8), as well as reduced transgene tag size (by using the ATG codon positioned at the 3' end of exon 6 as the translation start site) (Figure 21B). Presence of ATG at the 3' end of exon 6 allows for complete tag removal in the OFF-switch. pMM130 contains only 2 nucleotides of exon 8 (GA); by adding an additional nucleotide (G) downstream of the sequence (Figure 21B), a frame-shift in the transgene sequence (positioned downstream of exon 8) can be introduced upon inclusion of exon 7 by risdiplam, giving rise to an OFF-switch. In contrast, in the absence of risdiplam (exclusion of exon 7), the transgene can be correctly translated.

In order to reduce the length of a peptide produced in the OFF state and drive transcript degradation by nonsense mediated decay mechanism, the presented OFF-switch (pMM193, Figure 21B) was further modified by introducing premature termination codon in exon 7. To do that, a nucleotide deletion at position 19 of exon 7 was created (TCAAAAA>TCAAAA-), (Figure 21C), which will lead to the generation of a premature STOP codon downstream of the mutation within exon 7. In the presence of risdiplam, exon 7 will be included within the mature RNA transcript and with the presence of the premature STOP codon, the translation of the downstream-encoded transgene will be interrupted and transcript degraded (Figure 21C, pMM194). The nucleotide sequences of the pMM193 and pMM194 OFF-switches are shown in SEQ ID NOs:28 and 29, respectively.

As for previous designs, HEK293 cells were used, stably expressing the constructs pMM193, pMM194 and pMM130 as a control, cells were treated for 48h with increasing amounts of risdiplam and FACS used to measure changes in GFP expression as a response to risdiplam treatment. As shown in Figure 21D-21E, GFP expression was significantly reduced in pMM193-, and pMM194-transfected cells with increasing concentrations of risdiplam, while cells transduced with pMM130 (the ON-switch) showed increased GFP expression level. The number of GFP-positive cells in the cells expressing the OFF-

(pMM193, pMM194) or the ON-switch (pMM130) was also measured. A complete reduction in GFP-positive cells in the population that carries the OFF-switch was not observed (Figure 21F). It is possible that incubation of risdiplam for 48h is not sufficient to measure complete reduction of the GFP due to the enhanced GFP version used in these assays. Total GFP expression level in untreated cells was also measured. As expected, GFP expression is very high in pMM193 and pMM194-transfected cells in the absence of risdiplam (Figure 21G).

Example 22: Constructs lacking selection antibiotic resistance gene

All previously tested constructs had IRES-blasticidin-SV40 sequence added after the GFP encoding sequence leading to expression of antibiotic resistance gene (blasticidin), which was used for stable cell line selection. However, as the in-frame switch is intended to be used without antibiotic selection in AAV viral vectors, the inventors decided to replace the whole IRES-blasticidin-SV40 fragment with BGH polyA (Figure 22A). Five different constructs representing various modifications of the in-frame switch were selected: pMM263 (pLS76) SMN2 in-frame switch, pMM262 (pLS168) in-frame switch with 60 nucleotide reduction of exon 6, pMM259 (NB pMM112 is the same construct as pMM259) improved in-frame switch combining truncated introns, exon 6 and exon 8, pMM260 (pMM13) improved in-frame switch with reduced tag sequence and pMM272 (pMM198) split ATG in-frame switch. In addition, a previously published ON-switch was added for comparison pMM274 (Monteys *et al.* Nature (2021) 596: 291-295) (Figure 22B). The described constructs were transfected into HEK293 cells for stable expression and treated with risdiplam at increasing concentrations for 48 h before they were assessed for GFP expression by FACS-based measurements. The nucleotide sequences of the pMM263, pMM262, pMM259, pMM260, pMM272, pMM274 switches are shown in SEQ ID NOs: 38, 42, 22, 23, 25 and 107, respectively.

Results showed that all the BGH polyA constructs exhibited comparable induction rates to their IRES-blasticidin-SV40 counterparts (Figure 22C), however with increased median GFP signal and GFP leakiness in DMSO, which is most likely caused by increased transcript stability driven by removal of the long 3'UTR comprised of IRES-blasticidin and usage of a stronger polyA signal sequence (BGH vs SV40) (Figure 22D to 22F). Importantly, a comparison of the in-frame switches of the present disclosure with those of the previously published switch (Monteys *et al.* Nature (2021) 596: 291-295) shows that the present in-frame switches have significantly higher (up to 13x higher, when compared pMM259 vs pMM274) induction rates at all low risdiplam concentrations (Figure 22G).

Example 23: Generation of short split ATG switch [pMM273]

Despite significantly smaller size and better induction rate compared to pMM263 and pMM274 switches, pMM259 still has a short peptide 'tag' added to the expressed transgene. In order to remove the tag entirely, we employed a split-ATG approach, as in Example 19, wherein the 5' end of exon 8 was modified to contain 'TG' instead of 'GA', to create pMM273 (Figure 23A). With no risdiplam added, splicing occurs between exon 6 and exon 8, and no translation start site is formed. Upon treatment with risdiplam, exon 7 is spliced in, forming the ATG site together with exon 8 and driving translation of the gene of interest. No additional tag is added in this case (Figure 23B). In order to achieve this, all remaining ATG start sites were removed in construct pMM263 and exon 8 was mutated to TG resulting in

the construct pMM273 (Figure 23A). As the construct pMM273 contains only a 2 nucleotide-long exon 8, which forms an active ATG start site together with the exon 7, there is no amino acid tag added. The described constructs were transfected into HEK293 cells for stable expression and treated with risdiplam at increasing concentrations for 48h before they were assessed for GFP expression by FACS-based measurements. pMM274, pMM263 and pMM259 were used as control constructs to compare the performance of the split-ATG constructs. The nucleotide sequence of the pMM273 ON-switch is shown in SEQ ID NO:111.

The results show that split-ATG-based construct pMM273 showed sensitivity to risdiplam treatment as seen by induced GFP expression relative to DMSO control (Figure 23C), as well as the total level of median GFP expression (Figure 23D), although to a lesser extent than pMM259. The increase in the number of GFP-positive cells upon increase in the concentration of risdiplam was also measured (Figure 23E), however both pMM259 and pMM273 have almost 100% GFP positive cells in DMSO. Although almost all the cells express GFP without risdiplam treatment, the median GFP expression in DMSO of the split-ATG switch pMM273 was slightly lower compared to pMM259 (Figure 23F). In conclusion, the in-frame switch can be successfully engineered to form an ATG translation start site dependent on exon 7 inclusion, without compromising the risdiplam-mediated regulation of GFP expression.

Example 24: Modifications of exon 7 5'splice site sequences in pMM273

Based on the results obtained with pMM273, we investigated the effect of modifications within the 5' splice site (5'ss) within the exon 7 sequence. Three new constructs were generated with the modifications indicated in Figure 24: pMM362 (G>A at position -3 in exon 7, a deletion of nucleotide C at position +7 of intron 7; and a mutation T>G at position +8 of intron 7), pMM363 (A>G and G>T at positions -5 and -3, respectively, and TC>AT mutation within intron 7), and pMM364 (C>G mutation at position +7 within intron 7). HEK293 cells were stably transfected with each one the constructs or pMM273 (as a control), and treated with the indicated risdiplam concentrations for 48h. The modification of pMM362 disrupted splicing regulation and rendered the switch constitutively active without dependency on the presence or absence of risdiplam (Figure 24 A-C). By contrast, pMM363 and pMM364 clearly showed reduced leakiness in comparison to pMM273 (Figure 24E). However, the level of expression of the transgene was reduced from these constructs compared to pMM273 (Figure 24C). Furthermore, the total expression of the transgene at low risdiplam concentrations was significantly higher for pMM364 than it was for pMM363 (Figure 24C), indicating that the modifications in pMM364 outperformed pMM273 and could be used for further optimization.

Example 25: ESE2 region deletions in pMM273 to reduce leakiness in DMSO

In order to further reduce the background expression in the absence of risdiplam, we hypothesized that deletion of the ESE2 region and adjacent sequence within the exon 7 region would lead to reduced risdiplam-independent inclusion of exon 7. In addition, two of the mutations evaluated (in constructs pMM358 and pMM359) have previously been shown to reduce background *SMN2* splicing levels (Ishigami *et al.*²²). We designed four different constructs (Figure 25A) based on the original construct pMM273 having the original ESE2 sequence (AAAAAGAGGAA): pMM358 (AAAAAGAGGAA>GAAGGA), pMM359 (complete deletion of ESE2 sequence), pMM360 (complete deletion of ESE2 sequence, and

deletion of the downstream nucleotides GGTG), and pMM361 (complete deletion of the ESE2 sequence, deletion of the upstream nucleotides AAAATC, and deletion of the downstream nucleotides GGTGC). Our results revealed a dramatic reduction in median GFP leakiness for all the constructs harboring ESE2 deletions (pMM358-pMM361), as well as reduced number of GFP-positive cells in untreated cells compared to pMM273 (Figure 25D-E). However, only constructs pMM358 and pMM359 were able to respond to risdiplam at concentrations below 500 nM (Figure 25A), suggesting that the strong truncation of exon 7 as designed in pMM360 and pMM361 disrupts splicing. This was further confirmed by PCR amplification of the constructs, which revealed reduced splicing efficiency upon risdiplam treatment for constructs pMM358 and pMM359, but no exon 7 inclusion for construct pMM360 (Figure 25G). Lower basal splicing of constructs pMM358 and pMM359 resulted in lower leakiness, but also lower median GFP expression levels upon risdiplam induction (Figure 25B). Despite lower median GFP expression (Figure 25C), construct pMM359 shows a significantly better induction rate (Figure 25A, 25F) and lower leakiness (Figure 25E) compared to pMM273, making it a better-suited ON-switch for applications requiring low leakiness.

Example 26: Testing ESE2 region deletions in pMM273

Based on the results obtained by removing the ESE2 sequence (Example 25), we hypothesized that a gradual reduction of the ESE2 region and adjacent sequences could result in a construct with reduced leakiness levels that maintains high expression and induction rate. We investigated the effect of deleting the ESE2 region and the adjacent sequences, in a step-by-step manner. First, we designed 11 constructs where the ESE2 region (AAAAAGAAGGAAGG, underlined hereafter) region was either partially or completely removed (Figure 26A). pMM365 has almost complete ESE2 region deletion (-AAAAAGAAGGAAG) and pMM366 has complete ESE2 region deletion (-AAAAAGAAGGAAGG). For pMM367, pMM372, pMM373, pMM374, the ESE2 region and downstream nucleotides were also deleted: pMM367 (-AAAAAGAAGGAAGGT), pMM372 (-AAAAAGAAGGAAGGTGC), pMM373 (-AAAAAGAAGGAAGGTGCTC), pMM374 (-AAAAAGAAGGAAGGTGCTCAC). In Example 25, removal of GGTG sequence in pMM360 resulted in a complete loss of splicing as compared to pMM359. Thus, in pMM375, we introduced the same deletion as in pMM374 (-AAAAAGAAGGAAGGTGCTCAC), but additionally mutated 4 downstream nucleotides from ATTC to GGTG. Finally, we also tested only partial ESE2 region deletions in pMM368 (-AAAAAGAAGG), pMM369 (-AAAAAGAAG), pMM370 (-AAAAAGAA) and pMM371 (-AAAAA).

HEK293 cells, stably transfected with one of the various different constructs or pMM273 (as a control), and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for most of the constructs (Figure 26B-C, 26F), with the exception of the constructs having large deletions (pMM372, pMM373, pMM374), where induction was low and further decreased with increasing size of the deletion. Interestingly, pMM375 – which has the same deletion as inactive pMM374, but with an additional ATTC>GGTG mutation downstream of the ESE2 region – had completely rescued risdiplam induction of transgene expression, which was higher than pMM273 (Figure 26F). These data suggest that the GGTG sequence has an important regulatory role in the construct's splicing and response to risdiplam. Median GFP expression in DMSO (leakiness) was substantially decreased in all of the constructs tested, except

for pMM369, pMM370 and pMM371 (Figure 26E). However, all constructs resulted in more than 95% GFP+ positive cells in DMSO (Figure 26G). Step by step deletions of the ESE2 sequence allowed us to see a striking difference between constructs pMM369 (-AAAAAGAAG, high leakiness, approx. 3x reduced compared to pMM273) and pMM368 (-AAAAAGAAGG, very low leakiness, approx. 10x reduced compared to pMM273), where deletion of one additional nucleotide (G) dramatically reduced leakiness (Figure 26E).

Introducing gradual deletions to the ESE2 region, we identified several mutations that led to lower median GFP leakiness, but maintained high induction rate, such as pMM368, pMM369, pMM375.

Example 27: Testing ESE2 region deletions in pMM273

Based on the results obtained in the gradual ESE2 region deletions (Example 26), we tested several other deletions, where the ESE2 (AAAAAGAAGGAAGG, underlined hereafter) region was either partially or completely removed, while keeping the GGTG sequence intact and instead focusing on deletions of the region upstream of the ESE2 region (Figure 27A). In pMM376, pMM377 and pMM378 we tested complete ESE2 region deletion combined with additional upstream deletions: pMM376 (-TCAAAAAGAAGGAA), pMM377 (-AAAATCAAAAAGAAGGAA) and pMM378 (-CAAAATCAAAAAGAAGGAA). For pMM379, the ESE2 region was only partially deleted (-ACAAAATCAAAA). pMM380 tested deletion in a downstream region (-CTCACATTC), while pMM381 combined pMM379 (-ACAAAATCAAAA) and pMM380 deletions (-CTCACATTC). Finally, pMM382 deleted the majority of SMN2 exon 7 (-ACAAAATCAAAAAGAAGGAGGTGCTCACATTC), while pMM386 (-AAA) and pMM387 (-A) tested small ESE2 region deletions not tested in Example 26.

HEK293 cells, stably transfected with one of the various different constructs or pMM273 (as a control), and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP expression in response to risdiplam for most of the constructs (Figure 27B-C, 27F), except pMM382 (which has a large deletion, leading to no splicing) and pMM386 (where (AAA) deletion led to constitutive splicing of the construct). Constructs pMM376, pMM377, pMM379, pMM380 and pMM381 showed increased induction rate compared to pMM273 at 250 nM concentration of risdiplam (Figure 27F). However these constructs displayed reduced median GFP signal compared to pMM273 (Figure 27C). Relative to pMM273, median GFP expression in DMSO (leakiness) was decreased the most for constructs pMM376, pMM378, pMM381 and pMM382, with a modest decrease in pMM377 and pMM379, and no change or upregulation for pMM380, pMM386 and pMM387 (Figure 27E). However, all the constructs resulted in more than 95% GFP+ positive cells in DMSO (Figure 27G). Compared to pMM273, despite lower median GFP expression, construct pMM381 shows significantly lower leakiness and higher induction rate, making it a better-suited ON-switch for applications requiring low leakiness.

Example 28: Testing exon 7 deletions in pMM273

Based on the results obtained after GGTG deletion combined with the ESE2 region deletion (Examples 25 and 26), the GGTG sequence appears to have a regulatory role in exon 7 splicing. We designed three mutations to test its role in isolation from the ESE2 region sequence: pMM436 harboring partial ESE2

(AGAAGG) deletion and partial GGTG (T~~G~~CT) deletion, pMM437 where GGTG was deleted and pMM438 where GGTGCT sequence was deleted (Figure 28A).

HEK293 cells, stably transfected with one of the various different constructs or pMM273 (as a control), and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for all the constructs, except pMM436 which was inducible only at higher risdiplam concentrations and showed very low median GFP expression (Figure 28B-C, 28F). Deletion of GGTG (pMM437) or GGTGCT (pMM438) resulted in significantly lower leakiness compared to pMM273 (Figure 28E), as well as higher induction rate at risdiplam concentrations above 125 nM (Figure 28A). The percentage of GFP+ cells relative to pMM273 was lower for pMM436, but comparable for pMM437 and pMM438 (Figure 28D). Lower median GFP leakiness in DMSO, translated to lower median expression upon induction for pMM437 and pMM438 making them better-suited for applications requiring low leakiness and lower total transgene expression.

15 **Example 29: Intron 7 deletions in pMM273**

In a recently-published study, Recinos *et al.*²³ identified a region of *SMN2* intron 7 [334, 358] as being important for *SMN2* exon 7 splicing enhancement. In order to test whether deletion of this region can decrease exon 7 incorporation in DMSO (leakiness) we designed multiple deletions of the region (pMM477-pMM483) using the pMM273 ON-switch as a starting point (Figure 29A). In addition, we introduced a single point A>G mutation in construct pMM484 and randomized mutation GCACTGTACACTCTGACATATGAA>TGAATTAGGAGGGGAGGATTCATT in pMM485 (Figure 29A).

HEK293 cells, stably transfected with one of the various different constructs or pMM273 (as a control), and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP expression in response to risdiplam, although to various extents, for all the constructs (Figure 29B-C, 29F). pMM485 yielded very high GFP expression levels in DMSO (Figure 29E) and displayed a low induction rate (Figure 29B). Compared to pMM273 – except pMM482 – all the remaining constructs had comparable leakiness or higher leakiness (Figure 29E), comparable induction at 250 nM risdiplam (Figure 29F) and greater percentage of GFP+ cells in DMSO (Figure 29D). Only construct pMM482 showed moderately (approx. 50%) decrease leakiness and comparable induction rate (Figure 29A, 29F) relative to pMM273. In conclusion, we could not confirm any significant role of the previously described intron 7 region on exon 7 basal splicing in our splice switch.

35 **Example 30: Intron length testing in pMM273**

Intron length may influence regulation and efficiency of splicing. In order to test whether *SMN2* exon 7 inclusion without risdiplam treatment could be further reduced, we designed 11 different constructs (pMM464-pMM474) based on pMM273, where we introduced back previously deleted (Example 4) intronic sequences. Fragments were introduced in the middle of each intron, the length of each fragment and the total intron length are summarised in table shown in Figure 30A.

40

HEK293 cells, stably transfected with one of the various different constructs or pMM273 (as a control), and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP

expression levels. Results showed increased GFP signal in response to risdiplam for all the constructs (Figure 30B-C, 30F). Although the percentage of GFP+ cells was comparable for all the constructs (Figure 30D), the median GFP signal was significantly reduced for all the tested constructs compared to pMM273, with the highest GFP signal reduction in pMM466, pMM470, pMM472 and pMM474 (Figure 30E). Similarly, the induction rate was increased relative to pMM273 for all the tested constructs at all risdiplam concentrations (Figure 30B). At 250 nM concentration of risdiplam, constructs pMM466, pMM470, pMM472 and pMM474 showed the highest induction (Figure 30F) and were selected for further evaluation.

10 **Example 31: Exon 6 and 7 mutations in the improved constructs**

We further modified constructs selected in the Example 30 (pMM466, pMM470, pMM472, pMM474) by mutating a remaining ATG site in exon 6 (ATG>CTG) and deleting a TAA sequence in exon 7 (Figure 31A). The TAA sequence was identified to increase induction rate in a screen where every single, two, three, four or five nucleotides of exon 7 were deleted in construct pMM273 and analyzed by FACS (results not shown).

HEK293 cells, stably transfected with one of the various different constructs and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP expression in response to risdiplam for all the constructs (Figure 30B-C, 30F). While the percentage of GFP+ positive cells in DMSO was slightly decreased in all modified constructs (pMM567-pMM570) (Figure 31D), the median GFP expression was comparable (Figure 31E). However, deletion of TAA sequence significantly increased induction rate at low risdiplam concentrations (Figure 31B), with approx. 8-fold induction for pMM567-570 compared to 6-fold for the constructs from Example 30 (Figure 31F). pMM569 was selected as a particularly promising construct for further evaluation, as it combined low leakiness with high induction rate at low risdiplam concentrations.

25 **Example 32: Exon 6 and 7 mutations in pMM569**

We next ran a randomized mutation screen to identify novel mutations decreasing pMM569 leakiness and potentially improving the induction rate. 10 constructs, pMM591 and pMM593 to pMM601 were selected for further evaluation (Figure 32A). In addition, we tested one of the previous intron 7 mutation constructs (pMM482, Example 29).

HEK293 cells, stably transfected with one of the various different constructs, and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for the majority of constructs, except pMM593 and pMM597, which exhibited only a minimal GFP expression induction (Figure 32B-C, 32F). Constructs pMM594 and pMM599 had significantly improved GFP expression induction at increasing risdiplam concentrations relative to pMM569 (Figure 32B), with an approximate 50% increase at 62.5 nM risdiplam (Figure 32F). However, like pMM600, pMM594 was found to be significantly more leaky in DMSO, both in terms of the percentage of GFP-expressing cells (up to 3x more GFP+ cells) (Figure 32G) and median GFP expression (Figure 32E). Constructs pMM593, pMM595, pMM596, pMM597 and pMM601 had

decreased percentage of GFP+ cells in DMSO compared to pMM569, but only pMM595, pMM596 and pMM601 had an acceptable induction rate at 62.5 nM of risdiplam (Figure 32F).

To summarize, we identified pMM599 (C>A mutation in intron 7) as promising construct for increasing induction rate, without increasing leakiness. We also identified the pMM595 (TTTTAGAC>ATAATTAA mutation at position 3-10 of exon 7), pMM596 (CCTT deletion at position 257-260 of intron 6) and pMM601 (AAT>GCC at position 46 of exon 7) mutations as being variations leading to reduced leakiness levels compared to pMM569.

Example 33: Intron 6 deletion screen in pMM569

While Example 30 suggests that increasing intron length significantly reduces leakiness, we wanted to identify any remaining intronic sequences regulating *SMN2* exon 7 splicing and to this end we performed a whole intron 6 deletion screen, with 5-nt long non-overlapping deletions (Figure 33A). We next selected 11 intron 6 deletion constructs (pMM642-651, pMM709) for further evaluation (Figure 33B).

HEK293 cells, stably transfected with one of the various different constructs, and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for all the constructs (Figure 33C-D, 33G).

Constructs pMM642, pMM645, pMM649 and pMM651 showed increased induction rate at all risdiplam concentrations (Figure 33C), with approximately 25-30% higher induction for pMM642, pMM649 and pMM651 and 2-fold higher for pMM645 at 62.5 nM risdiplam concentration. Leakiness measured as the percentage of GFP+ positive cells in DMSO (Figure 33H) and median GFP expression in DMSO (Figure 33F) was slightly increased or comparable for all the tested constructs, with the exception of increased median GFP expression for pMM650 (Figure 33F). Overall, mutations pMM642, pMM645, pMM649 and pMM651 showed slightly better induction rate with a comparable leakiness levels to pMM569.

Example 34: Intron 7 deletion screen in pMM569

We next performed a whole intron 7 deletion screen, with 5-nt long non-overlapping deletions (Figure 34A). We next selected 10 intron 7 deletion constructs (pMM634-641, pMM711, pMM712) for further evaluation (Figure 34B).

HEK293 cells, stably transfected with one of the various different constructs, and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for all the constructs (Figure 34C-D, 34G).

However, no construct showed increased induction rate compared to pMM569 (Figure 34C). The only exception was pMM635 at 62.5 nM risdiplam (Figure 34G), however that was an outlier, as it was not consistent across concentrations (Figure 34C). Leakiness measured as the percentage of GFP+ positive cells in DMSO (Figure 33H) and median GFP expression in DMSO (Figure 33F) was slightly decreased for deletions pMM636, pMM637 and pMM638, while increased for pMM639 and pMM711, with pMM711 being the most leaky construct.

Example 35: Comparison of leakiness-reducing mutations in pMM569

In order to design a construct tailored for application where low leakiness levels are essential, we combined pMM569 switch with selected mutations shown to reduce leakiness. We selected three mutations: pMM579 having ESE2 region deletion (as in pMM359), pMM595 with mutated exon 7 and pMM596 with intron 6 deletion (Figure 35A).

HEK293 cells, stably transfected with one of the various different constructs, and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for all the constructs (Figure 35B-C, 35F).

Although none of the constructs displayed reduced median GFP expression compared to pMM569 (Figure 35E), pMM579 and pMM596 significantly reduced the percentage of GFP+ cells compared to pMM569 (Figure 35G). While the leakiness of pMM579 and pMM596 was similar, the induction rate and median GFP expression was significantly better for pMM596 (Figure 35B-C), with approximately 5-fold induction at 62.5 nM of risdiplam for pMM596, compared to 2-fold for pMM579 (Figure 35F). Design pMM596 was selected as a promising low leakiness variant of pMM569.

Example 36: Exon 6 deletions in pMM569

To further reduce the size of the pMM569 ON-switch, we tested two exon 6 deletions: from 45 nt in pMM569 to 30 nt in pMM618 and 15 nt in pMM619 (Figure 36A).

HEK293 cells, stably transfected with one of the various different constructs, and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Both induction rate and median GFP signal were comparable for all three constructs at all risdiplam concentrations (Figure 36B-C), with 8-12-fold induction at 62.5 nM risdiplam for all constructs. A slight increase in the percentage of GFP+ cells (Figure 36D, G) as well as median GFP signal in DMSO (Figure 36E) for pMM618 and pMM619 revealed a modest increase in leakiness with a shorter exon 6. To summarize, both exon 6 deletions combine reduced size with similar performance to pMM569.

Example 37: Testing mammalian promoters with pMM569

All previous experiments were performed using a construct employing a CMV promoter. In order to test performance of the risdiplam-inducible switch with other commonly-used promoters we combined pMM569 design with 7 other promoters: pMM581 (full length (FL) CAG promoter), pMM582 (hEF1a promoter), pMM583 (hUbiC promoter), pMM584 (RSV promoter), pMM589 (TK promoter), pMM590 (PGK promoter) and pMM562 (minimal CAG promoter) (Figure 37A).

HEK293 cells, stably transfected with one of the various different constructs, and treated with indicated risdiplam concentrations for 48h, were analyzed by FACS to measure GFP expression levels. Results showed increased GFP signal in response to risdiplam for all the constructs, except pMM589 (TK promoter) and pMM590 (PGK promoter), which showed a very modest increase only at the highest risdiplam concentrations investigated (Figure 37B-C). At 62.5 nM risdiplam concentration, constructs pMM582 (hEF1a) and pMM652 (minimal CAG) had comparable induction rate to pMM569 (CMV), while pMM581 (FL-CAG), pMM583 (hUbiC), pMM584 (RSV), although still inducible, were 2-3-fold lower

(Figure 37F). Construct pMM581 (FL-CAG promoter) showed very high leakiness measured both by median GFP expression and % of GFP+ cells (Figure 39E, G), while minimal CAG promoter (pMM652) and hEF1a (pMM582) leakiness was comparable to CMV. Constructs pMM583 (hUbiC) and pMM584 (RSV) had reduced % of GFP+ cells, while constructs pMM589 (TK) and pMM590 (PGK) had almost no GFP+ cells in DMSO (Figure 31G). It is important to note that although constructs pMM589 (TK) and pMM590 (PGK) had very low induction, they both reached 100% GFP+ cells at 500 nM of risdiplam meaning that the splice switch was still active, only it was masked by low GFP expression levels.

Example 38: Comparison of various iterations of risdiplam-regulated splice switch

Based on previous data generated with various iterations of in-frame-based, risdiplam-regulated splice switches, we performed an experiment comparing the original in-frame splice switch variant based on sequence described in pLS76 (pMM263), a construct based on a splice switch published by Monteys *et al.* Nature (2021) 596: 291-295 (pMM274), the split ATG short switch (pMM273), a pMM273-based switch with a truncation within the ESE2 region, as well as the split ATG switch (pMM569) (Figure 38A). We used HEK293 cells to transfect the constructs described above in order to compare their performance side-by-side. pMM569 showed superior performance in terms of induction rate compared to published constructs as well as various versions of the improved in-frame switch (Figure 38B). At low risdiplam concentration (62.5 nM), pMM569 showed approx. 17-fold induction rate compared to 7-fold for pMM273 and 4-fold for pMM359. While the total level of GFP expression remained unchanged between pMM273, pMM359 and pMM569 at the highest risdiplam concentrations, pMM569 had significantly higher median GFP expression at lower risdiplam concentrations compared to pMM359 (Figure 38C). Background expression, *i.e.* in the absence of risdiplam, exon 7 inclusion and hence median GFP expression was highest in construct pMM273, whereas all other constructs were comparable (Figure 38E). At the same time, the percentage of GFP positive cells was lowest for the least inducible constructs (pMM263 and pMM274) and highest for pMM273, while pMM359 and pMM569 had approximately 50% GFP-positive cells in DMSO (Figure 38D, 38G). While the pMM569 switch achieved almost 100% GFP-positive cells at 15 nM risdiplam, the starting construct (pMM263) as well as the published (pMM274) required much higher risdiplam concentrations. In conclusion, we identified pMM569 as an ON-switch construct that outperformed other variations in terms of induction rate, and having background GFP expression levels comparable to the least leaky constructs tested.

Example 39: Risdiplam-regulated NanoLuciferase (NLuc) expression

In order to test the applicability of the optimized in-frame switch, we assessed whether a different reporter gene can be controlled similarly to GFP by risdiplam. Therefore, we used NanoLuc (NLuc) as a reporter and cloned it into several variations of the switch selected as well performing in Example 31: pMM568, pMM569 (front-runner, reduced length of intron 7 in comparison to pMM568), pMM570 (reduced length of intron 7 in comparison to pMM568), and pMM579 (low leaky version of pMM569, additional deletion of the ESE2 region sequence), (Figure 39A). HEK293 cells stably transfected with one of the various different constructs were treated with indicated risdiplam concentrations for 48h and cells were lysed and analyzed for NLuc expression (Figure 39B-D). We observed increased induction rate in NLuc expression in all of the constructs, however, the deletion of the ESE2 region sequence in pMM612 (switch pMM579 with NLuc as a reporter) resulted in reduced induction rates at low nM concentrations of risdiplam in

comparison to the other switch variants (Figure 39B), in line with reduced total median GFP signal (Figure 39C). However, we did observe slight reduction in the background expression level in construct pMM611 (switch pMM570 with NLuc as a reporter).

5 **Example 40: Risdipam-regulated Cas9 expression for gene editing applications**

As a potential application for regulated gene expression systems, we used Cas9 as a protein of interest under the control of risdipam. We used the same switch designs as in Example 38, and replaced the GFP reporter sequence with a sequence that encodes for Cas9 protein (Figure 40A). We introduced one additional construct (pMM596), which has a deletion within the intron 6 sequence of TCCT as illustrated in Figure 40A. This construct has been shown previously (Example 32) to have lower background expression level, a criteria required for optimal regulation of Cas9 protein expression. In order to test the performance of the various versions of the in-frame switch, we stably transfected HEK293 cells with the above-mentioned constructs using piggybac method. Having established stable HEK293T cell lines, we have used nucleofection-mediated delivery of 3 different sgRNA (UUGGCUUCCUGGGCUGCUA, GCAGCCCUCCACUCCCAUGG, GGCGCUGUCAUGAUUCGU) that target the CD81 gene, encoding for a cell-surface protein. Risdipam was added at the indicated concentrations and cells were incubated for 72h before they were stained with anti-CD81 antibody to assess reduction of expression of the protein upon Cas9 induction with risdipam using FACS-based analysis (Figure 40B). Cells that were not transfected with sgRNA mix served as a negative control. In all of the constructs described in Figure 40A, we observed a risdipam dose-dependent reduction in CD81 expression (Figure 40B-H), with a maximum decrease of 50% for the previously published construct (pMM274) (Figure 40D) and up to 12% for pMM596 (Figure 40H). However, we also observed risdipam-independent reduction of CD81 expression, driven by leaky Cas9 expression, in particular in constructs based on pMM263 (Figure 40C), pMM273 (Figure 40E), pMM359 (Figure 40F), and pMM569 (Figure 40G). However, the risdipam-independent reduction of CD81 expression was significantly lower in the construct having the pMM596 ON-switch (Figure 40H), reaching levels observed for the previously published sequence pMM274 (Figure 40D). In conclusion, the improved ON-switch sequence pMM596, with the significantly reduced leakiness profile, is a promising candidate to be used as a regulation system for inducible Cas9 expression.

30 **Example 41: Risdipam-regulated CAR expression in CAR-T therapy**

A potential therapeutic application for small molecule-regulated gene expression system is illustrated in Figure 41, which is based on the hypothesis that risdipam could be used to induce the expression of chimeric antigen receptors (CARs) in T cells. An option for regulated expression of CARs in T cells could translate into huge benefits for patients, as regulated CAR expression could increase safety of the therapy, allowing some of the side effects observed in the CAR-T therapy space to be overcome, e.g. strong immune reaction and the well-established cytokine storm effect. Furthermore, it could allow for using more aggressive constructs, or using second inducible CAR-expressing constructs to prevent development of cancer resistance during CAR-T therapy.

40 In order to assess the suitability of the risdipam-inducible ON-switch for such applications, we investigated the switch in a more relevant cellular model, namely Jurkat cells, an immortalized cell line

derived from leukemic T-cells. In order to test the in-frame switch in these cells, we used pMM569 and introduced some mutations in order to test the effect on the leakiness in these cells.

We introduced a deletion in the ESE2 region (pMM579), a modification that was previously shown to reduce background expression, pMM595 with the indicated mutations highlighted with bold letters within the exon 7 sequence in Figure 42A, and pMM596 (deletion of TCCT within the intron 6 sequence) – see Figure 42A). Jurkat cells were stably transfected with these constructs and GFP was measured using FACS-based readouts 48h after addition of risdiplam at the indicated concentrations (Figure 42B-G). Similar to our observations in HEK293 cells, we observed high induction rates and low leakiness for the pMM596 switch compared to the other constructs (Figure 42B, 42E-G). In comparison, pMM595 showed higher leakiness (Figure 42E, 42G), although the induction rate was not reduced (Figure 42B, 42E-G), as we have observed that the total level of median GFP signal (Figure 42C) as well as the number of GFP expressing cells (Figure 42D) was high accordingly. In conclusion, we have established that the in-frame, risdiplam-inducible ON-switch performs to a similar standard in Jurkat cells, a T-cell model with relevance for testing CAR-T cells and we propose the use of this switch regulated by risdiplam as a therapeutic approach to regulate CAR expression in patients.

Example 42: Risdiplam-regulated expression of GFP in AAV-transduced neurons.

The in-frame switch constructs pLS76 and its optimized counterpart pMM359 (Figure 43A) were incorporated into an adeno-associated virus (AAV) vector genome, subsequently packaged into AAV vectors with the AAV2.7m8 capsid, and employed to transduce human embryonic stem cell (hESC)-derived neurons at a multiplicity of infection (MOI) of 20,000 genome-containing vector particles per cell. Prior to AAV infection, neurons underwent a differentiation period of 21 days. At three days following AAV transduction, varying concentrations of risdiplam were added to the cultures, with DMSO serving as the negative control. The Incucyte live cell imaging analysis system was utilized to assess the expression of the reporter gene GFP in neurons transduced with AAV vectors containing the different in-frame switches, providing insights into both transduction efficiency and the functionality of the in-frame switches within the AAV context. Consistent with our findings with cell lines transfected with the ON-switch constructs, induction of GFP expression was observed with the AAV genome containing the pLS76 switch as depicted in Figure 43B. Notably, the GFP expression mediated by the optimized pMM359 sequence was significantly enhanced at lower concentrations of risdiplam. These results indicate that the performance of the in-frame switch is directly transferable to an AAV vectored context, confirming the compatibility of the in-frame switch with AAV-based gene delivery applications.

Example 43: Risdiplam-regulated expression of MeCP2 protein

We next sought to confirm that the in-frame switches characterised herein are relevant for the risdiplam-inducible expression of other proteins of interest, e.g. therapeutically-relevant proteins. The GFP reporter in the pMM569 switch was substituted with the MeCP2 coding sequence (as depicted in Figure 44A). We utilized MeCP2-knockout H4 cells as a model system, which were transfected with the piggyBac system to introduce the MeCP2-expressing plasmid controlled by the in-frame switch. Wildtype H4 cells, expressing endogenous levels of MeCP2 and non-transfected MeCP2-knockout cells served as controls (shown in Figure 44B). The Operetta imaging system was employed to detect MeCP2 expression, with

cells stained using an anti-MeCP2 antibody and DAPI to visualize nuclei. Following transfection, cells were selected with puromycin to isolate cells stably-expressing the constructs, which were then exposed to various concentrations of risdiplam for 24 hours (illustrated in Figures 44C-D). In the MeCP2-knockout cells, no protein expression was observed in the absence of risdiplam; however, MeCP2 expression was restored and enhanced in a risdiplam concentration-dependent manner when controlled by the in-frame switch (Figures 44C-D). These findings demonstrate that the pMM569 in-frame switch is not only effective for regulating reporter gene (GFP) expression, but is also capable of modulating the expression of therapeutically-relevant proteins of interest, such as MeCP2. The switch could potentially be used to regulate MeCP2 expression for the treatment of Rett Syndrome.

Example 44: Materials and Methods

Cell culture:

HEK293 [American Type Culture Collection (ATCC): CRL-1573] and H4 [ATCC: HTB-148] cells were cultivated in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (FBS; cat. no. #97068-085, lot no. #065K19, VWR) and 1% (v/v) penicillin/streptomycin solution (Sigma-Aldrich) later referred to as DMEM Complete Medium (DMEM). Jurkat cells (#BA131204, Hoffmann-La Roche, Nutley, NJ) were cultured in RPMI 1640 Medium with GlutaMAX™ Supplement (#61870-036, Gibco) supplemented with 10% FBS (#A3160402, Gibco) and 1% Penicillin-Streptomycin (10,000 U/mL) (#15140122, Gibco). Cells were maintained at $0.5-1.5 \times 10^6$ cells/ml. In order to select the antibiotic resistant cells, 0.8 - 1 µg/ml of puromycin (#A1113803, Gibco) or 20 µg/mL blasticidin (Invitrogen, cat. no. #A1113902) was added to the media. Cells were cultivated in standard cell culture conditions, at 37°C in a humidified atmosphere containing 5% CO₂.

Neuronal cell culture and differentiation:

Human embryonic stem cell (hESC)-derived neural progenitor cells were cultured in laminin (Cat#MSPPLN521-05, VWR)-coated plates in Maintenance medium: DMEM/F12 medium (Cat#31331028, ThermoFischer), mixed at a 1:1 ratio with Neurobasalmedium (Cat# 321103049, ThermoFischer), supplemented with B27 (Cat#12587010), M2 (Cat# 17502048, ThermoFischer), b-Mercaptoethanol (Cat#31350010, ThermoFischer), and Penicillin/Streptomycin (Cat#15140122, ThermoFischer). Cells were cultured and expanded in this Medium and then transferred to predifferentiation medium for 1 week: Above-described maintenance medium, supplemented with sonic hedgehog (Cat#100-45, Peprotech), FGF8b (Cat# 100-25, Peprotech), and Aa2-P (Cat#A8960, SigmaAldrich). After seven days of culture in pre-differentiation medium, cells were then transferred to differentiation medium (BGAA), which is composed of maintenance medium as described above, supplemented with BDNF (Cat# AF-450-02, Peprotech), GDNF (Cat# AF-450-10, Peprotech), Aa2-P (Cat#A8960, SigmaAldrich), cAMP (Cat# D009, BioLog), and Laminin (Cat#11243217001, SigmaAldrich). Cells were cultured in this medium for a total of 21 days before transduction with AAVs, medium was thereafter exchanged every 3 days.

Transfection:

HEK293 and H4 cells were transfected using the piggyBac transposase system and Lipofectamine 2000 (Invitrogen, cat. no. #11668019) used according to manufacturer's instructions. 0.35×10^6 cells were seeded per well in a 6-well plate 24 hours (h) before transfection. Cells were then co-transfected with a

PiggyBac helper plasmid encoding for the transposase enzyme and the donor plasmid that contains the gene of interest under the control of the various splice switch versions at a 3:1 ratio (donor:helper plasmid). 24h post-transfection, cells underwent blasticidin (Invitrogen, cat. no. #A1113902) selection for 10 days at the concentration of 20 µg/mL or puromycin (Gibco, #A1113803) selection for 5 days at the concentration of 0.8 µg/mL. After the selection, cells were washed with DPBS, trypsinized with Trypsin-EDTA (0.05%), phenol red (#25300054, Gibco), resuspended in culture media and seeded on 96-well plates (50 000 cells/well). After 24h cells were treated with varying concentrations of risdiplam or DMSO as a negative control.

10 Nucleofection:

Nucleofection of HEK293 and Jurkat cells was performed using 4D-Nucleofector® X Unit (#AAF-1003X, Lonza) according to the manufacturer's protocol. In order to nucleofect Jurkat cells, 1 million cells per reaction was collected, washed once with DPBS, no calcium, no magnesium (#14190144, Gibco), spun down and resuspended in 100 µL of SE Cell Line Nucleofector™ Solution (#V4XC-1024, Lonza).

15 Resuspended cells were mixed with 2 µg of the donor plasmid that contains the gene of interest under the control of the various splice switch versions and 0.65 µg of PiggyBac encoding plasmid (3:1 ratio of donor:helper plasmids) and transferred to Nucleocuvette™ Vessel. Vessels were transferred to 4D-Nucleofector® X Unit and nucleoporated using the CL-120 program. Subsequently, 400 µL of culture media (RPMI 1640) was added to the vessels and cells were incubated for 5 min. at RT. Next, the whole volume was transferred to 12-well plates and an additional 0.5 mL of fresh media was added. Cells were incubated for 48h after which they were selected with puromycin (0.8 µg/ml) for 5 days. After the selection, cells were washed with DPBS, resuspended in culture media, seeded on 96-well plates (60 000 cells/well), and treated with varying concentrations of risdiplam or DMSO as a negative control.

For HEK293 nucleoporation with sgRNAs, 1 million HEK293 cells stably expressing splice-switch inducible Cas9 enzyme was collected, washed with DPBS, spun down and resuspended in 100 µL of SF Cell Line Nucleofector™ Solution (#V4XC-2032, Lonza) containing three sgRNAs at 150 nM concentration (sgRNA1: UUGGCUUCCUGGGCUGCUA, sgRNA2: GCAGCCCUCCACUCCCAUGG, sgRNA3: GGCGCUGUCAUGAUGUUCGU). Next, cells were transferred to Nucleocuvette™ Vessels and nucleoporated in 4D-Nucleofector® X Unit using CM-130 program. Subsequently, 400 µL of culture media (DMEM High-glucose (#11965092, Gibco) was added to the vessels and cells were incubated for 5 min. at RT. Finally, cells were resuspended in culture media, seeded on 96-well plates and treated with varying concentrations of risdiplam or DMSO as a negative control.

AAV transduction:

35 For transducing neuronal cells with AAVs vectoring GFP transgene under the control of a risdiplam-inducible ON-switch (Example 42), AAVs were produced using standard production protocols. HEK293T cells were cultured in DMEM supplemented with 10% FBS until they reached 50 to 70% confluency. The cells were then triple transfected with pHelper, pAAV ITR-expression, and pAAV Rep-Cap plasmids using PEI reagent in 1x cellFactory5. After 72 hours post-transfection, AAV particles were harvested from the cells. The crude lysate was prepared by lysing the cell pellet with 0.1% Triton and digesting with Benzonase. Primary capture of AAV particles was performed using Poros CaptureSelect AAV-X

(ThermoFisher). A secondary purification step was conducted using iodixanol gradient centrifugation to enrich for full capsids.

- For quality control, vector genome titer was determined by performing qPCR with ITR-specific primers (ITR_qPCR_F: CGGCCTCAGTGAGCGA, qPCR_R: ggaaccctagtgatggagt). The purity of the final product was assessed via 10% SDS-PAGE using Stain technology (Bio-Rad). A purity level of more than 95% was confirmed if only VP1, VP2, and VP3 proteins were detectable. Endotoxin levels were measured using the Endosafe-nexgen-PTS Spectrometer from Charles River Laboratories, with all constructs showing endotoxin levels below the detection limit.
- We selected the capsid AAV2.7m8 as these capsids have previously shown to transduce neurons (data not shown). The multiplicity of infection was titrated and the viral genome of 20,000 genomic copies per cell showed balanced transduction efficiency versus cell survival. Six hours after AAV transduction, medium was removed and fresh BGAA medium was added to the cells. Cells were allowed to rest for 3 days before adding risdiplam at varying concentrations. GFP expression was monitored using Incucyte Live Imaging device and images were acquired for 7 days. Every 3 days, fresh BGAA medium was exchanged alongside risdiplam.

Cloning of the constructs

All the tested constructs were ordered and cloned using Gibson cloning at GenScript.

Assessment of splicing efficiency

20 *Flow Cytometry*

- Reporter gene expression of either EGFP or mCHERRY in the presence of varying amounts of inducer compound (risdiplam (2-(2,8-dimethylimidazo[1,2-b]pyridazin-6-yl)-7-(4,7-diazaspiro[2.5]octan-7-yl)-4H-pyrido[1,2-a]pyrimidin-4-one) or risdiplam analogue ((R)-7-(hexahydropyrrolo[1,2-a]pyrazin-2(1H)-yl)-2-(2-methylimidazo[1,2-b]pyridazin-6-yl)-4H-pyrido[1,2-a]pyrimidin-4-one)) was monitored by measuring the intensity of the fluorescent proteins using flow cytometry. Therefore, cells were washed with PBS, trypsinized and resuspended in autoMACS Running Buffer (Miltenyi Biotec, cat. No. #130-091-221). The analysis was done using CytoFLEX LX flow cytometer (Beckman Coulter). The measurement was analysed using FlowJo software (version 10.8.1) and exported to CSV file and later analysed in Excel (Microsoft). Fold changes in reporter gene expression were calculated based on the median fluorescence intensity measured in the compound-treated cells relative to DMSO-treated cells.

RNA isolation

- HEK293 cells with stably incorporated splice constructs were seeded on 6-well plates (0.5×10^6 cells/well) and treated with DMSO or risdiplam for 24h. After 24h cells were washed with PBS (-/-), trypsinized, resuspended in DMEM CM and transferred to 1.5 mL Eppendorf tubes. Next, cells were centrifuged at $300 \times g$ for 5 min at room temperature, washed once with PBS (-/-) and cell pellets were frozen at -80°C . RNA was isolated using RNeasy Mini Kit (Qiagen, cat. no. #74104) according to manufacturer's protocol. After the isolation RNA concentration was measured and all samples were diluted with UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, cat. no. #10977015) to the same

concentration. The isolated RNA was stored at -80°C and subsequently used for cDNA synthesis by reverse transcriptase (RT)-PCR.

cDNA synthesis

2 µg of the isolated total RNA were used for the cDNA synthesis using SuperScript IV VIL0 (Invitrogen, cat. no. #18090050) according to manufacturer's protocol. In short, RNA was either mixed with SuperScript VIL0 reverse transcriptase (RT) (+RT) or without RT (-RT), 4x master mix and water in 0.2 mL 8-Strip Non-Flex PCR Tubes (Starlab, cat. no. #I1402-3700) to the final volume of 20 µL. Tubes were incubated at 25°C for 10 min, followed by 50°C for 30 min and 80°C for 5 min. Samples were diluted to 100 µL with UltraPure™ DNase/RNase-Free Distilled Water (Invitrogen, cat. no. #10977015) to the final cDNA concentration of 20 ng/µL and stored at -20°C.

PCR

Synthesized cDNA was used for PCRs to analyze splicing efficiency of the tested constructs. PCRs were set up using 10 µL of Q5 Hot Start HF 2x Master Mix (NEB, cat. No. #M0494L), 1 µL forward (MM16, TTGACCTCCATAGAAGACACC) and 1 µL reverse primer (MM40, GCAGATGAACTTCAGGGTCAG) to the final concentration of 0.5 µM each, 6 µL H₂O and 2 µL cDNA to the total volume of 20 µL in 0.2 mL 8-Strip Non-Flex PCR Tubes. PCR thermocycling conditions were as follows: 98°C for 1 min; 15 cycles of (98°C for 15 seconds, 64°C for 15 seconds, 72°C for 30 seconds); 72°C for 2 min; 4°C hold. PCR products were resolved for 20 min on 2% E-Gel™ 48-well Agarose Gels (Invitrogen, cat. no. #G820802) with E-Gel 50 bp DNA Ladder (Invitrogen, cat. no. #10488099) and imaged using E-Gel Power Snap Electrophoresis System (G8300).

Immunocytochemistry staining

In order to stain for cells expression MeCP2 protein (Example 43), H4 cells were transfected with a knockout in MeCP2 with plasmid containing the in-frame switch according to sequence pMM569 and the MeCP2 sequence (plasmid pMM586) using piggybac system according to described protocol. Cells were selected for stable expressers using puromycin. The stable cell pool was cultured in DMEM, supplemented with FBS and P/S and induced with varying concentrations of risdiplam as indicated or DMSO as a negative control. After 24h, cells were used for staining by a washing step with PBS, followed by fixation in paraformaldehyde (4%) for 15min at room temperature (RT) and an additional washing step with PBS. Fixed cells were then treated with PBS, supplemented with 0.3% Triton X-100 for 15 min at RT, washed with PBS and then incubated for 1 hour at RT with a blocking solution, composed of 10% goat serum (G9023) and 0.3% Triton X-100 in PBS. In a next step, cells were incubated with anti-MeCP2 antibody (Cat#3456, Cell Signaling Technologies) at a dilution rate of 1:200 in blocking solution overnight at 4°C before 3 washing steps followed with PBS. A secondary antibody (Goat anti-Rabbit, Cat# A32733, ThermoFischer) was then used at a 1:1000 dilution and incubated for 1 hour at RT, followed by 3 washing steps with PBS. To visualize the nucleus, cells were incubated with DAPI at a 1:1000 for 5min, washed with PBS and images were acquired using the Operetta CLS High-Content Analysis System.

NanoLuc (NLuc) measurement

NLuc measurements (Example 39) were performed using Nano-Glo® Luciferase Assay System (Promega, #N1120) according to the manufacturer protocol. In short, Nano-Glo reagent was prepared at

room temperature by combining one volume of Nano-Glo® Luciferase Assay Substrate with 50 volumes of Nano-Glo® Luciferase Assay Buffer. Next, the reagent was added to each measured well using a volume of reagent equal to that of the cell culture media in each well (100 µL per one well of 96-well plate). Plates were incubated for 3 minutes and measured using PHERAstar FSX Microplate Reader (BMG Labtech).

Cas9 experiments

CD81 staining:

HEK293 cells were nucleoporated with CD81-targeting sgRNAs as described in the nucleofection section hereinabove, seeded on 96-well plates and treated with varying concentrations of risdiplam or DMSO as a negative control for 3 days. To analyse the amount of CD81+ on cell surface, cells were first detached with 40 µL of Trypsin-EDTA (0.05%), phenol red (#25300054, Gibco) and next resuspended in the additional volume of 100 µL DBPS to a total volume of 140 µL. Cells were transferred to fresh 96-well Microplates, PP, U-bottom, 0,3 ml (#701330, Brand), spun down and washed once with DPBS. Next, cells were resuspended in 50 µL of 1:10 mix of APC-conjugated CD81 (#561958, BD) and DBPS and incubated for 15 min. at 4°C. Cells were spun down (180xg, 3 min.), the staining solution was removed, and cells were washed three times with DPBS. Finally, cells were resuspended in DBPS and analyzed by flow cytometry using CytoFLEX LX Flow Cytometer (BeckmanCoulter). The populations of CD81+ cells were defined using CD81-stained HEK293 WT cells (no sgRNA). The measurement was analysed using FlowJo software (version 10.8.1), exported to CSV file and later analysed in Excel (Microsoft).

References

1. Lombardi, L. M., Baker, S.(A)& Zoghbi,(H)Y. MECP2 disorders: from the clinic to mice and back. J Clin Invest 125, 2914-2923, doi:10.1172/JCI78167 (2015).
- 25 2. Sceniak, M. P. *et al.* Mechanisms of Functional Hypoconnectivity in the Medial Prefrontal Cortex of Mecp2 Null Mice. Cereb Cortex 26, 1938-1956, doi:10.1093/cercor/bhv002 (2016).
3. Tillotson, R. & Bird,(A)The Molecular Basis of MeCP2 Function in the Brain. J Mol Biol, doi:10.1016/j.jmb.2019.10.004 (2019).
4. Tillotson, R. *et al.* Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. 30 Nature 550, 398-401, doi:10.1038/nature24058 (2017).
5. Gadalla, K. K. *et al.* Improved survival and reduced phenotypic severity following AAV9/MECP2 gene transfer to neonatal and juvenile male Mecp2 knockout mice. Mol Ther 21, 18-30, doi:10.1038/mt.2012.200 (2013).
6. Sinnett, S.(E)*et al.* Improved MECP2 Gene Therapy Extends the Survival of MeCP2-Null Mice without Apparent Toxicity after Intracisternal Delivery. Mol Ther Methods Clin Dev 5, 106-115, 35 doi:10.1016/j.omtm.2017.04.006 (2017).

7. Garg, S. K. *et al.* Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. *J Neurosci* 33, 13612-13620, doi:10.1523/JNEUROSCI.1854-13.2013 (2013).
8. Matagne, V. *et al.* A codon-optimized Mecp2 transgene corrects breathing deficits and improves survival in a mouse model of Rett syndrome. *Neurobiol Dis* 99, 1-11, doi:10.1016/j.nbd.2016.12.009 (2017).
9. Van Esch, (H) in *GeneReviews*((R)) (eds M. P. Adam *et al.*) (1993).
10. Gadalla, K. K., Bailey, M.(E) & Cobb, S. R. MeCP2 and Rett syndrome: reversibility and potential avenues for therapy. *Biochem J* 439, 1-14, doi:10.1042/BJ20110648 (2011).
- 10 11. Ginhoux, (F) *et al.* HLA-A*0201-restricted cytolytic responses to the rTA transactivator dominant and cryptic epitopes compromise transgene expression induced by the tetracycline on system. *Mol Ther* 10, 279-289, doi:10.1016/j.ymthe.2004.05.012 (2004).
12. Hoyng, S.(A) *et al.* Developing a potentially immunologically inert tetracycline-regulatable viral vector for gene therapy in the peripheral nerve. *Gene Ther* 21, 549-557, doi:10.1038/gt.2014.22 (2014).
- 15 13. Tickner, Z. J. & Farzan, M. Riboswitches for Controlled Expression of Therapeutic Transgenes Delivered by Adeno-Associated Viral Vectors. *Pharmaceuticals (Basel)* 14, doi:10.3390/ph14060554 (2021).
14. Karagiannis, P., Fujita, Y. & Saito, (H) RNA-based gene circuits for cell regulation. *Proc Jpn Acad Ser B Phys Biol Sci* 92, 412-422, doi:10.2183/pjab.92.412 (2016).
- 20 15. Sivaramakrishnan, M. *et al.* Binding to SMN2 pre-mRNA-protein complex elicits specificity for small molecule splicing modifiers. *Nat Commun* 8, 1476, doi:10.1038/s41467-017-01559-4 (2017).
16. Campagne, S. *et al.* Structural basis of a small molecule targeting RNA for a specific splicing correction. *Nat Chem Biol* 15, 1191-1198, doi:10.1038/s41589-019-0384-5 (2019).
- 17 Zhang, M. L., Lorson, (C) L., Androphy, (E) J. & Zhou, J. An *in vivo* reporter system for measuring increased inclusion of exon 7 in SMN2 mRNA: potential therapy of SMA. *Gene Ther* 8, 1532-1538, doi:10.1038/sj.gt.3301550 (2001).
- 25 18. Cho, S. *et al.* 3' Splice site sequences of spinal muscular atrophy related SMN2 pre-mRNA include enhancers for nearby exons. *ScientificWorldJournal* 2014, 617842, doi:10.1155/2014/617842 (2014).
- 30 19. Gao, Y. *et al.* Systematic characterization of short intronic splicing-regulatory elements in SMN2 pre-mRNA. *Nucleic Acids Res* 50, 731-749, doi:10.1093/nar/gkab1280 (2022).
20. Noderer, W. L. *et al.* Quantitative analysis of mammalian translation initiation sites by FACS-seq. *Mol Syst Biol* 10, 748, doi:10.15252/msb.20145136 (2014).

21. Naryshkin, N.(A)*et al.* Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* 345, 688-693, doi:10.1126/science.1250127 (2014).
22. Ishigami, Y. *et al.*, *Nature Communications* 15: 1880, doi: 10.1038/s41467-024-46090-5 (2024).
- 5 23. Recinos, Y. *et al.*, *bioRxiv* [Preprint]. 2023 Aug 21:2023.08.21.554109. doi: 10.1101/2023.08.21.554109.

Claims:

1. A polynucleotide comprising in 5' to 3' order:

(i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:222, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;

(ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:378 at its 3' end, and consisting of fewer than 1044 nucleotides;

(iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:226;

(iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:380 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;

(v) a fifth nucleotide sequence: (a) consisting of the dinucleotide 'GA', 'TG' or 'TT', or (b) comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:19, wherein the nucleotide sequence comprises 'GA', 'TG' or 'TT' at the positions corresponding to positions 1 and 2 of SEQ ID NO:19, or (c) encoding a polypeptide of interest, and comprising 'GA', 'TG' or 'TT' at positions 1 and 2; and

(vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a) or (v)(b), a sixth nucleotide sequence encoding a polypeptide of interest; wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

2. The polynucleotide according to claim 1, wherein when the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.

3. The polynucleotide according to claim 1 or claim 2, wherein the first nucleotide sequence comprises a nucleotide sequence according to SEQ ID NO:222.

4. The polynucleotide according to any one of claims 1 to 3, wherein the first nucleotide sequence comprises, or consists of, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:106, SEQ ID NO:219 or SEQ ID NO:220.

5. The polynucleotide according to any one of claims 1 to 4, wherein the second nucleotide sequence comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 or SEQ ID NO:377 at its 3' end.

6. The polynucleotide according to any one of claims 1 to 5, wherein the second nucleotide sequence consists of fewer than 500 nucleotides.

7. The polynucleotide according to any one of claims 1 to 6, wherein the second nucleotide sequence comprises, or consists of, SEQ ID NO:9, SEQ ID NO:10 or SEQ ID NO:228.
8. The polynucleotide according to any one of claims 1 to 7, wherein the third nucleotide sequence
5 consists of a nucleotide sequence according to SEQ ID NO:226.
9. The polynucleotide according to any one of claims 1 to 8, wherein the third nucleotide sequence consists of SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:223 or SEQ ID NO:224.
10. The polynucleotide according to any one of claims 1 to 9, wherein the fourth nucleotide sequence comprises SEQ ID NO:15 or SEQ ID NO:379 at its 5' end, and comprises SEQ ID NO:16 at its 3' end.
11. The polynucleotide according to any one of claims 1 to 10, wherein the fourth nucleotide sequence consists of fewer than 500 nucleotides.
15
12. The polynucleotide according to any one of claims 1 to 11, wherein the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:227 or SEQ ID NO:340.
13. The polynucleotide according to any one of claims 1 to 12, wherein the fifth nucleotide sequence: (a)
20 consists of the dinucleotide 'GA' or 'TG', or (b) comprises, or consists of, a nucleotide sequence according to SEQ ID NO:19.
14. The polynucleotide according to any one of claims 1 to 13, wherein the fifth nucleotide sequence: (a)
25 consists of the dinucleotide 'GA' or 'TG', or (b) comprises, or consists of, SEQ ID NO:20 or SEQ ID NO:21.
15. The polynucleotide according to any one of claims 1 to 14, wherein the polynucleotide comprises a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence selected from the group consisting of SEQ ID NO:164, SEQ ID NO:171, SEQ ID NO:116, SEQ ID NO:114, SEQ ID NO:111,
30 SEQ ID NO:25, SEQ ID NO:23, SEQ ID NO:24 and SEQ ID NO:22.
16. The polynucleotide according to any one of claims 1 to 15, wherein the polynucleotide further comprises a promoter sequence 5' to the start codon.
17. The polynucleotide according to any one of claims 1 to 16, wherein the polynucleotide further
35 comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.
18. The polynucleotide according to any one of claims 1 to 17, wherein the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.
40
19. A vector comprising a polynucleotide according to any one of claims 1 to 18.

20. The vector according to claim 19, wherein the vector is an adeno-associated virus (AAV) vector.
21. A pharmaceutical composition comprising a polynucleotide according to any one of claims 1 to 18, or a vector according to claim 19 or claim 20, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.
22. A cell comprising a polynucleotide according to any one of claims 1 to 18, or a vector according to claim 19 or claim 20.
23. The cell according to claim 22, wherein the cell further comprises a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
24. A method for modifying a cell to express a polypeptide of interest, comprising:
- (i) introducing into a cell a polynucleotide according to any one of claims 1 to 18, or a vector according to claim 19 or claim 20; and
- (ii) subsequently contacting the cell with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
25. A method for expressing a polypeptide of interest in a cell, comprising contacting a cell according to claim 22 with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
26. A polynucleotide according to any one of claims 1 to 18, a vector according to claim 19 or claim 20, or a pharmaceutical composition according to claim 21, for use in a method of medical treatment or prophylaxis.
27. A polynucleotide according to any one of claims 1 to 18, a vector according to claim 19 or claim 20, or a pharmaceutical composition according to claim 21, for use in treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
28. Use of a polynucleotide according to any one of claims 1 to 18, a vector according to claim 19 or claim 20, or a pharmaceutical composition according to claim 21, in the manufacture of a medicament for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
29. A method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising administering to a subject a polynucleotide according to any one of claims 1 to 18, a vector according to claim 19 or claim 20, or a pharmaceutical composition according to claim 21.

30. The polynucleotide, vector, or pharmaceutical composition for use according to claim 27, the use according to claim 28, or the method according to claim 29, wherein treating or preventing the disease or condition further comprises administering to the subject a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

5

31. The polynucleotide, vector, or pharmaceutical composition for use according to claim 27 or claim 30, the use according to claim 28 or claim 30, or the method according to claim 29 or claim 30, wherein the disease or condition is a disease or condition characterised by deficiency of the polypeptide of interest.

10 32. A kit, comprising:

(i) a polynucleotide according to any one of claims 1 to 18, a vector according to claim 19 or claim 20, or a pharmaceutical composition according to claim 21; and

(ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

15

33. A polynucleotide comprising in 5' to 3' order:

(i) a first nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:1, wherein the first nucleotide sequence does not comprise SEQ ID NO:2;

20

(ii) a second nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:7 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:8 at its 3' end, and consisting of fewer than 1044 nucleotides;

(iii) a third nucleotide sequence consisting of a nucleotide sequence having at least 80% sequence identity to a nucleotide sequence according to SEQ ID NO:26, wherein the third nucleotide sequence comprises 'A' at the position corresponding to position 2 of SEQ ID NO:12;

25

(iv) a fourth nucleotide sequence, comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:15 at its 5' end, and comprising a nucleotide sequence having at least 80% sequence identity to SEQ ID NO:16 at its 3' end;

(v) a fifth nucleotide sequence: (a) consisting of the trinucleotide 'GAG' or (b) encoding a polypeptide of interest, and comprising 'GAG' at positions 1 to 3; and

30

(vi) where the fifth nucleotide sequence is a nucleotide sequence according to (v)(a), a sixth nucleotide sequence encoding a polypeptide of interest; wherein the polynucleotide comprises a start codon 5' to the nucleotide sequence encoding a polypeptide of interest.

35

34. The polynucleotide according to claim 33, wherein when the polynucleotide is a polyribonucleotide, splicing of the polyribonucleotide in the absence of a splicing modifier that promotes *SMN2* exon 7 inclusion yields polyribonucleotides substantially lacking the third nucleotide sequence.

40

35. The polynucleotide according to claim 33 or claim 34, wherein the first nucleotide sequence comprises a nucleotide sequence according to SEQ ID NO:1.

36. The polynucleotide according to any one of claims 33 to 35, wherein the first nucleotide sequence comprises, or consists of, SEQ ID NO:3.
37. The polynucleotide according to any one of claims 33 to 36, wherein the second nucleotide sequence
5 comprises SEQ ID NO:7 at its 5' end, and comprises SEQ ID NO:8 at its 3' end.
38. The polynucleotide according to any one of claims 33 to 37, wherein the second nucleotide sequence consists of fewer than 500 nucleotides.
- 10 39. The polynucleotide according to any one of claims 33 to 38, wherein the second nucleotide sequence comprises, or consists of, SEQ ID NO:10.
40. The polynucleotide according to any one of claims 33 to 39, wherein the third nucleotide sequence consists of a nucleotide sequence according to SEQ ID NO:26.
15
41. The polynucleotide according to any one of claims 33 to 40, wherein the third nucleotide sequence consists of SEQ ID NO:13 or SEQ ID NO:27.
42. The polynucleotide according to any one of claims 33 to 41, wherein the fourth nucleotide sequence
20 comprises SEQ ID NO:15 at its 5' end, and comprises SEQ ID NO:16 at its 3' end.
43. The polynucleotide according to any one of claims 33 to 42, wherein the fourth nucleotide sequence consists of fewer than 500 nucleotides.
- 25 44. The polynucleotide according to any one of claims 33 to 43, wherein the fourth nucleotide sequence comprises, or consists of, SEQ ID NO:17.
45. The polynucleotide according to any one of claims 33 to 44, wherein the polynucleotide comprises at least 80% sequence identity to SEQ ID NO:28 or SEQ ID NO:29.
30
46. The polynucleotide according to any one of claims 33 to 45, wherein the polynucleotide further comprises a promoter sequence 5' to the start codon.
47. The polynucleotide according to any one of claims 33 to 46, wherein the polynucleotide further
35 comprises a polyadenylation sequence 3' to the nucleotide sequence encoding a polypeptide of interest.
48. The polynucleotide according to any one of claims 33 to 47, wherein the polynucleotide comprises an inverted terminal repeat (ITR) sequence at its 5' end, and an ITR sequence at its 3' end.
- 40 49. A vector comprising a polynucleotide according to any one of claims 33 to 48.
50. The vector according to claim 49, wherein the vector is an adeno-associated virus (AAV) vector.

51. A pharmaceutical composition comprising a polynucleotide according to any one of claims 33 to 48, or a vector according to claim 49 or claim 50, and a pharmaceutically-acceptable carrier, diluent, excipient or adjuvant.
52. A cell comprising a polynucleotide according to any one of claims 33 to 48, or a vector according to claim 49 or claim 50.
53. The cell according to claim 52, wherein the cell further comprises a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
54. A method for modifying a cell to express a polypeptide of interest, comprising introducing into a cell a polynucleotide according to any one of claims 33 to 48, or a vector according to claim 49 or claim 50.
55. A method for inhibiting expression of a polypeptide of interest in a cell, comprising contacting a cell according to claim 52 with a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing modifier that promotes *SMN2* exon 7 inclusion is risdiplam.
56. A polynucleotide according to any one of claims 33 to 48, a vector according to claim 49 or claim 50, or a pharmaceutical composition according to claim 51, for use in a method of medical treatment or prophylaxis.
57. A polynucleotide according to any one of claims 33 to 48, a vector according to claim 49 or claim 50, or a pharmaceutical composition according to claim 51, for use in treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
58. Use of a polynucleotide according to any one of claims 33 to 48, a vector according to claim 49 or claim 50, or a pharmaceutical composition according to claim 51, in the manufacture of a medicament for treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest.
59. A method of treating or preventing a disease or condition that would derive therapeutic or prophylactic benefit from an increase in the level of expression of the polypeptide of interest, comprising administering to a subject a polynucleotide according to any one of claims 33 to 48, a vector according to claim 49 or claim 50, or a pharmaceutical composition according to claim 51.
60. The polynucleotide, vector, or pharmaceutical composition for use according to claim 57, the use according to claim 58, or the method according to claim 59, wherein the disease or condition is a disease or condition characterised by deficiency of the polypeptide of interest.

61. A kit, comprising:

(i) a polynucleotide according to any one of claims 33 to 48, a vector according to claim 49 or claim 50, or a pharmaceutical composition according to claim 51; and

(ii) a splicing modifier that promotes *SMN2* exon 7 inclusion, optionally wherein the splicing
5 modifier that promotes *SMN2* exon 7 inclusion is risdiplam.

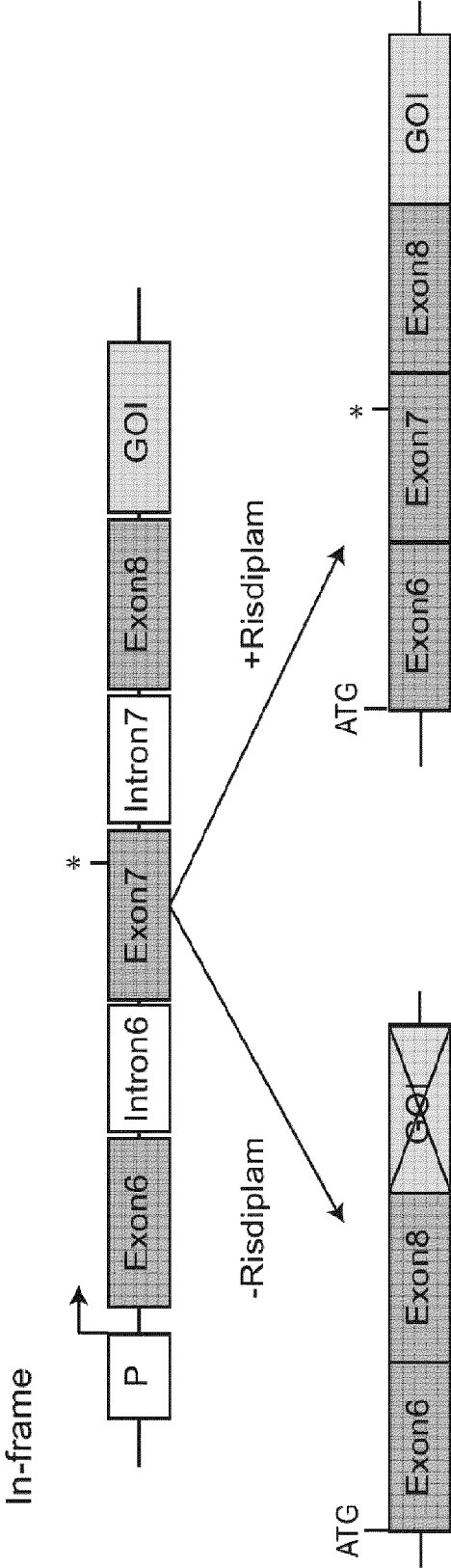


Figure 1

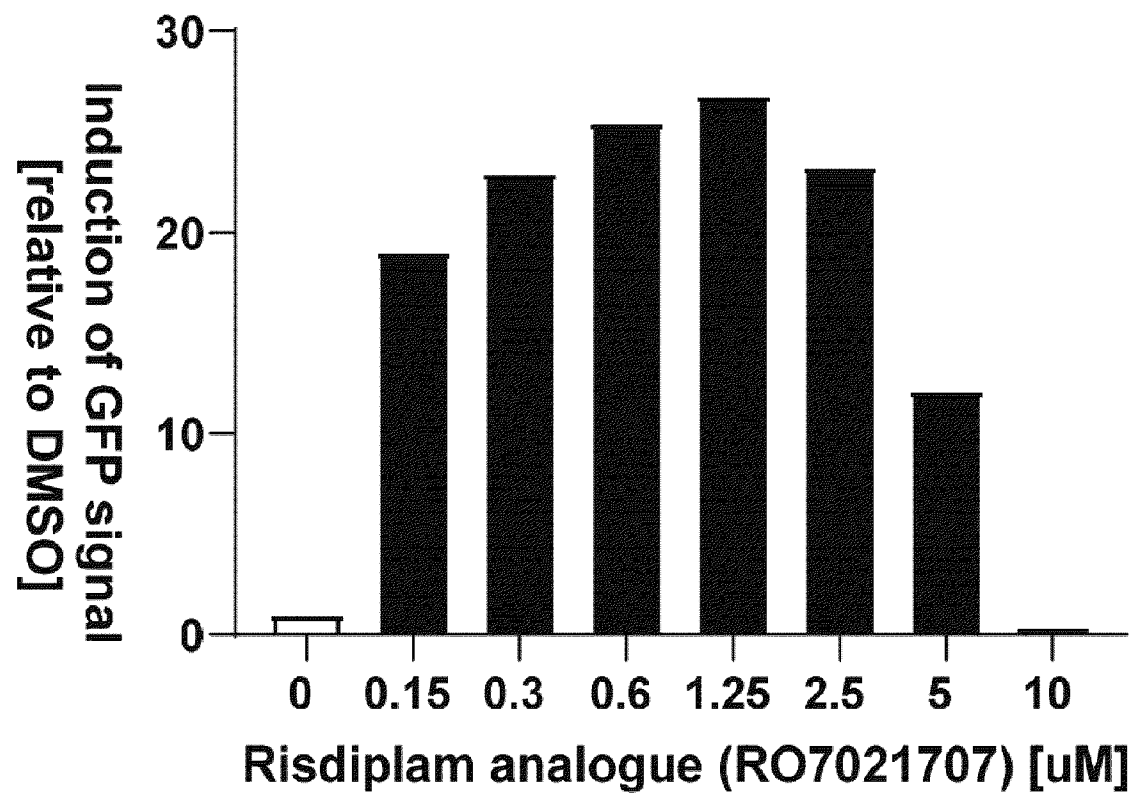
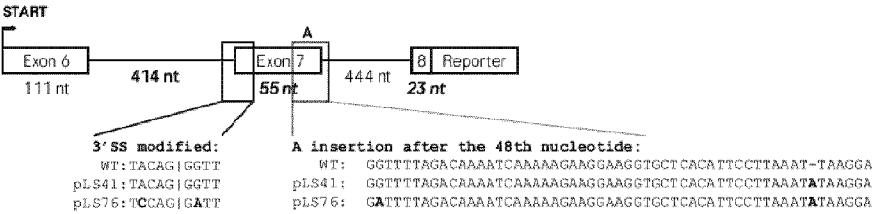


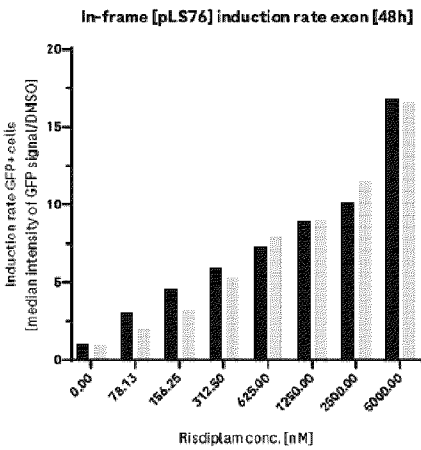
Figure 2

A

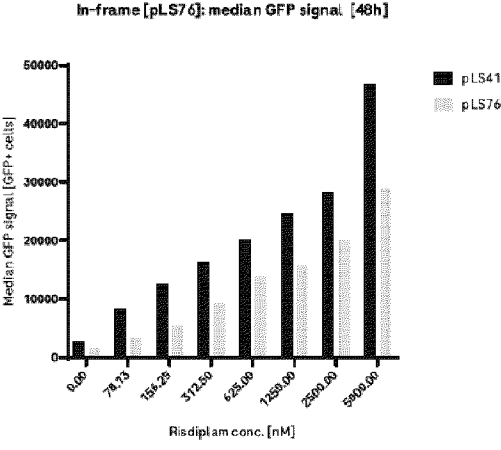
SMN2 in-frame switch [pLS76]



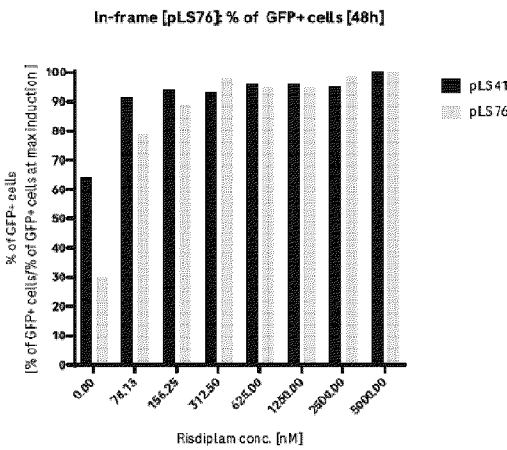
B



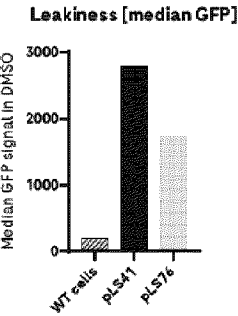
C



D



E



F

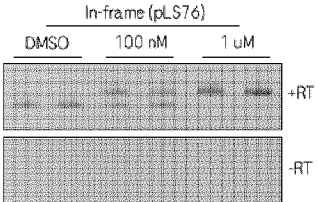
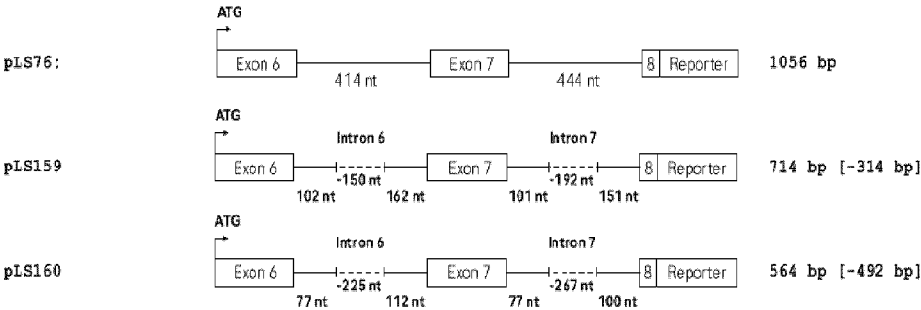


Figure 3

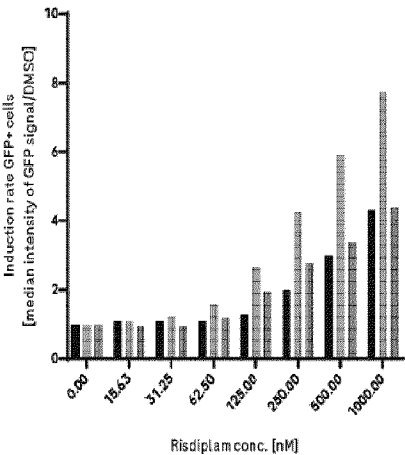
A

Reducing the size of introns (SMN2 in-frame switch):



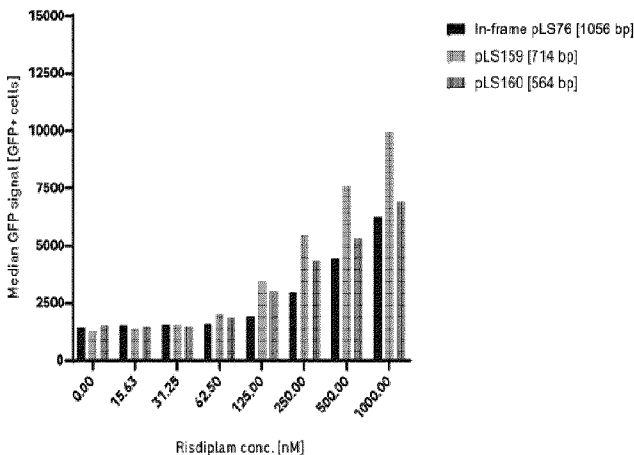
B

Intron reduction: induction rate [48h]



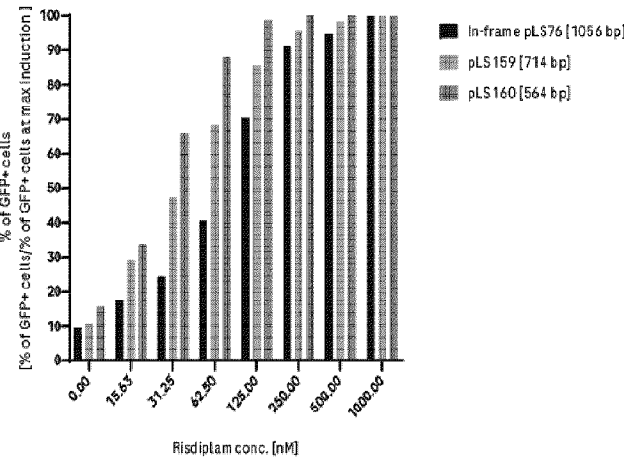
C

Intron reduction: median GFP signal [48h]



D

Introns reduction: % of GFP+ cells [48h]



E

Leakiness [median GFP]

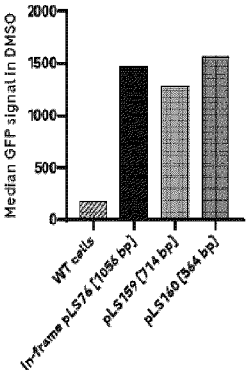
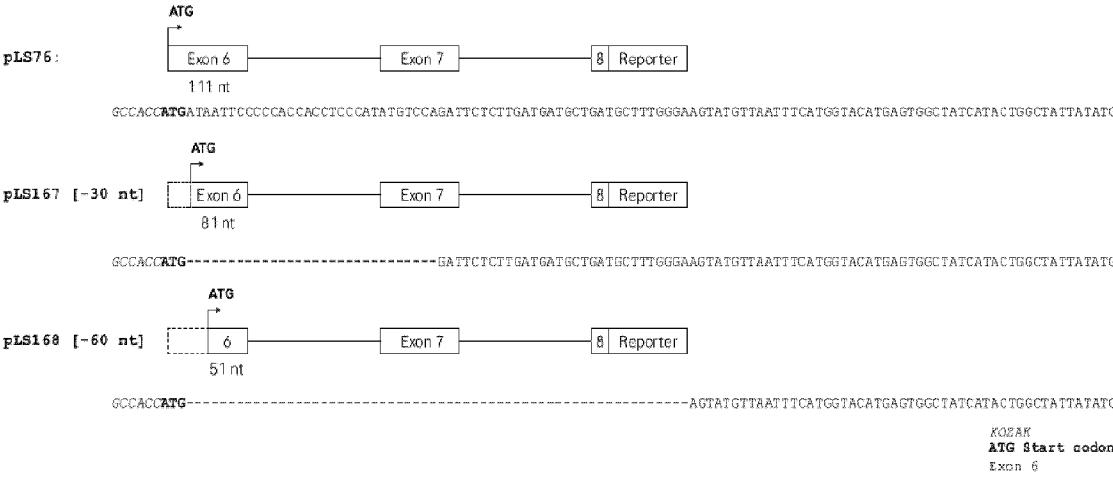


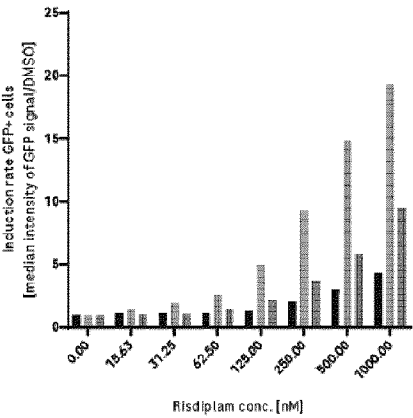
Figure 4

A Reducing the size of exon 6 (SMN2 in-frame switch):



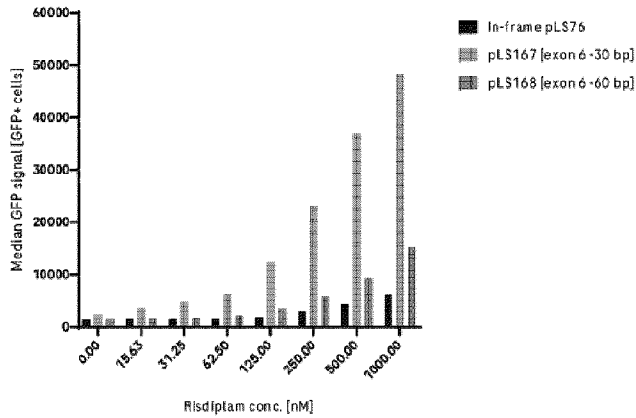
B

Exon 6 reduction: induction rate [48h]



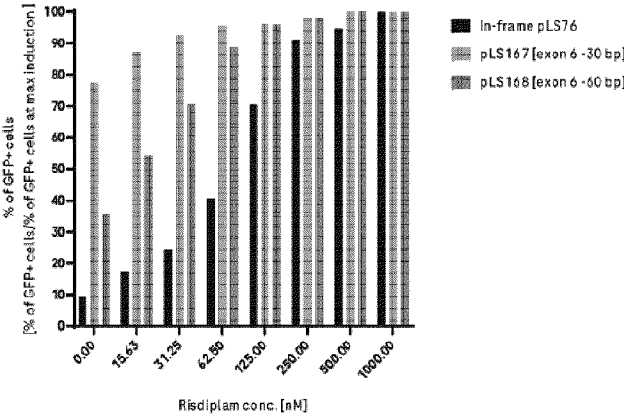
C

Exon 6 reduction: median GFP signal [48h]



D

Exon 6 reduction: % of GFP+ cells [48h]



E

Leakiness [median GFP]

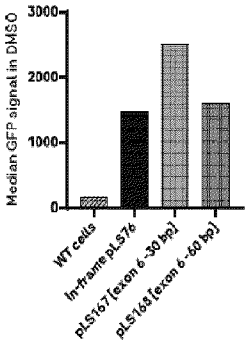
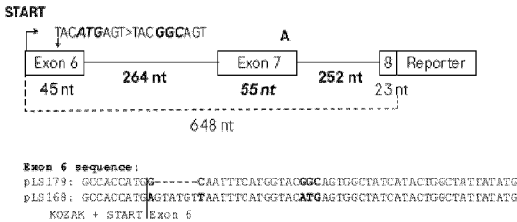


Figure 5

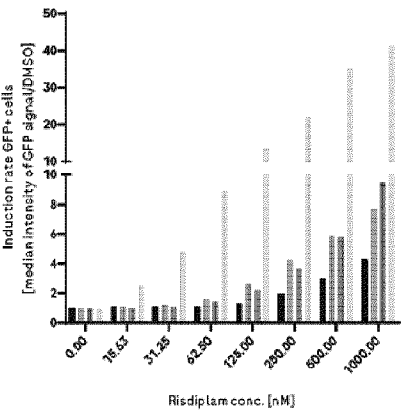
A

pLS179: Improved SMN2 in-frame switch based on pLS76 with pLS159 intron truncation and modified pLS168 exon 6 truncation



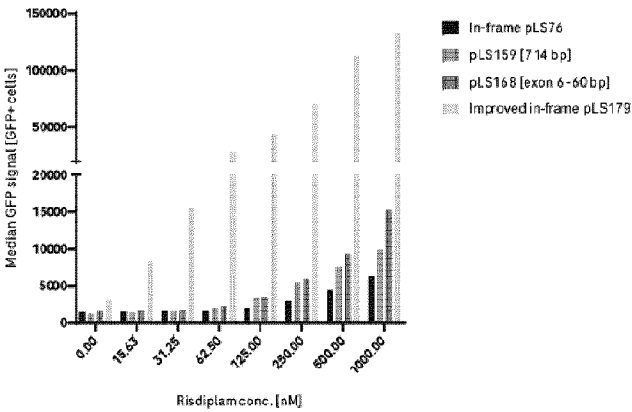
B

Improved in-frame [pLS179]: induction rate [48h]



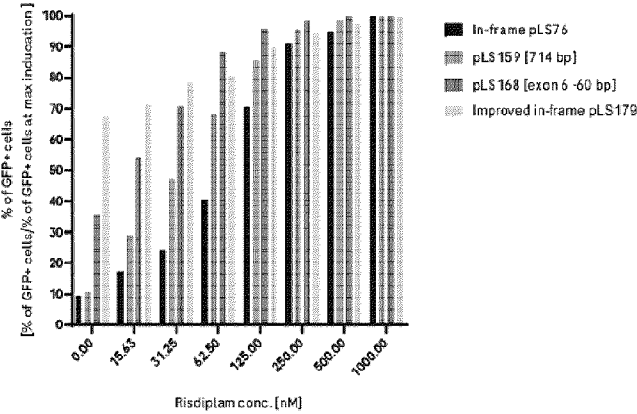
C

Improved in-frame [pLS179]: median GFP signal [48h]



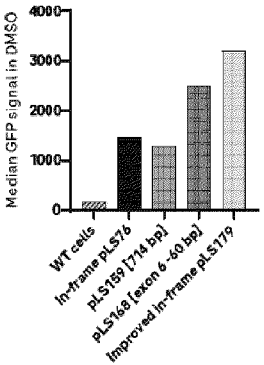
D

Improved in-frame [pLS179]: % of GFP+ cells [48h]



E

Leakiness [median GFP]



F

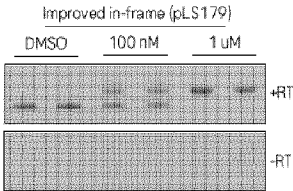
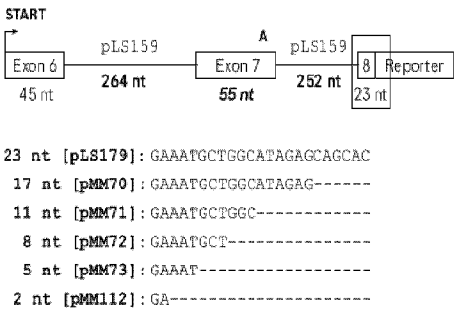


Figure 6

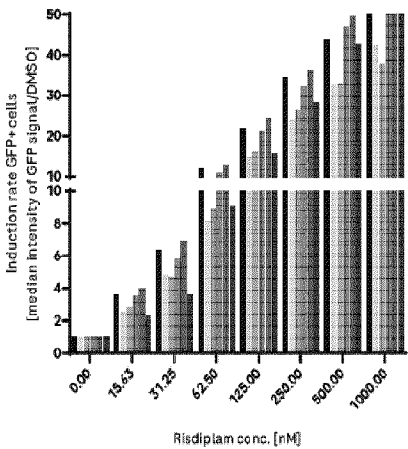
A

pMM112: Shortening exon 8 length to reduce the tag size in pLS179:



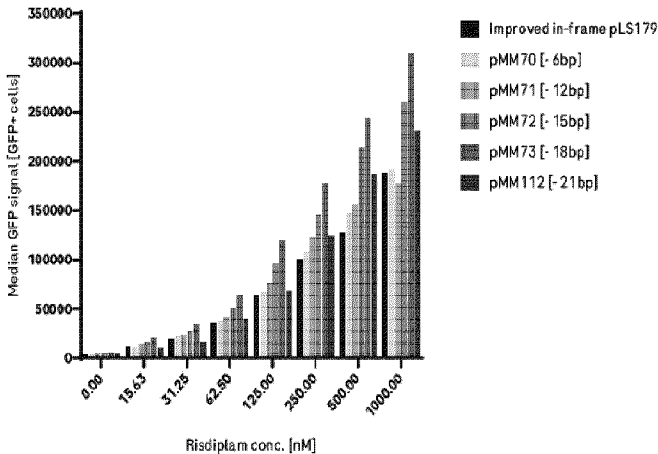
B

Exon 8 truncations: induction rate [48h]



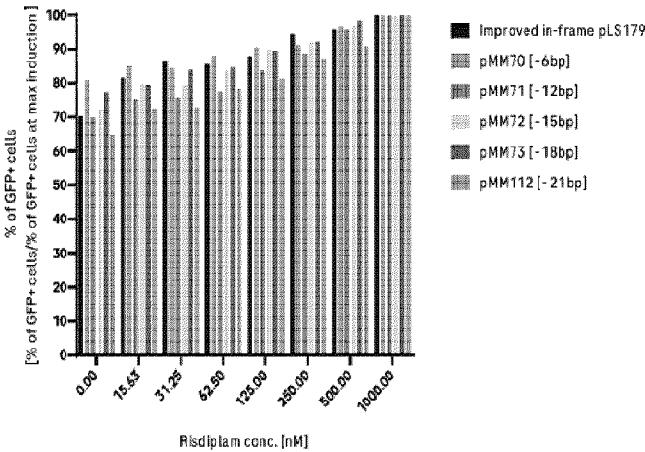
C

Exon 8 truncations: median GFP signal [48h]



D

Exon 8 truncations: % of GFP+ cells [48h]



E

Leakiness [median GFP]

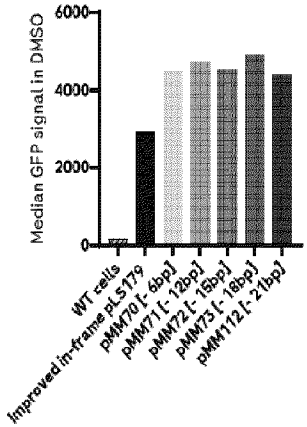


Figure 7

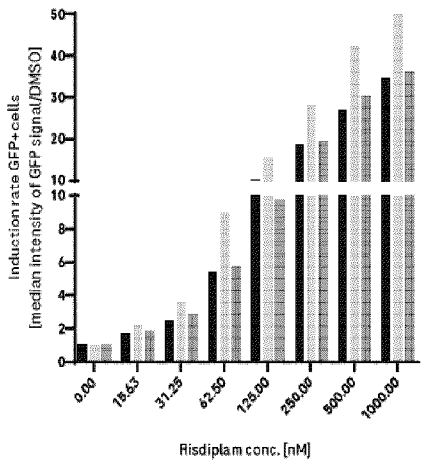
A

pMM130: Using ATG in exon 6 splice site to reduce the tag size [based on pMM112]:



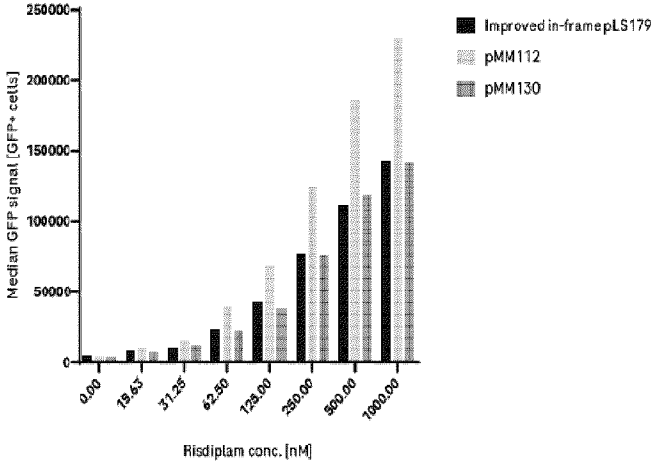
B

Exon 6 splice site ATG: induction rate [48h]



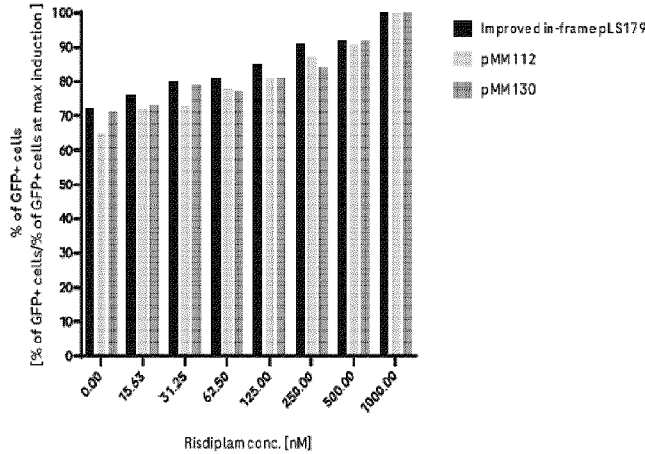
C

Exon 6 splice site ATG: median GFP signal [48h]



D

Exon 6 splice site ATG: % of GFP+ cells [48h]



E

Leakiness [median GFP]

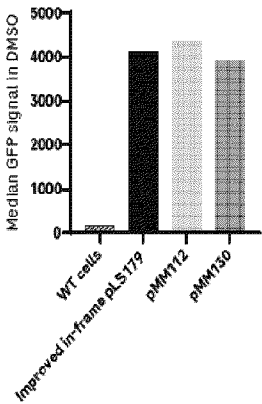
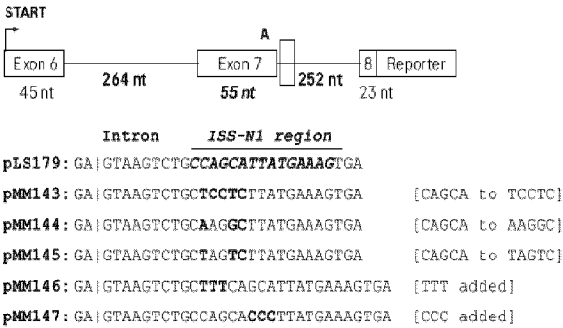


Figure 8

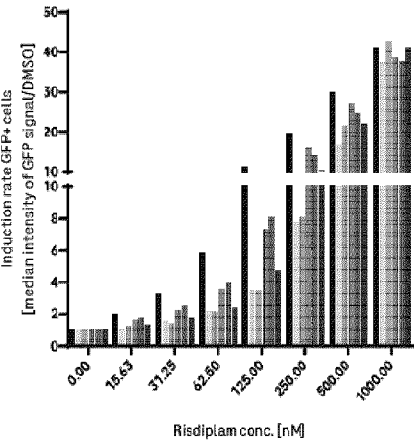
A

pLS179: Mutating ISS-N1 region to reduce leakiness [based on Gao et al, NAR, 2022]



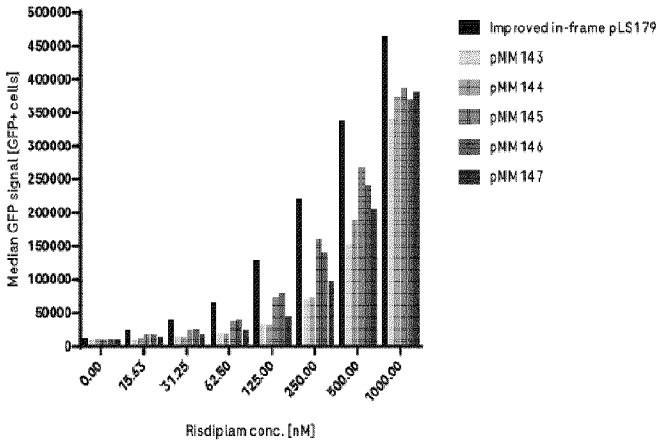
B

Improved in-frame [pLS179] ISS-N1 mutations
induction rate [48h]



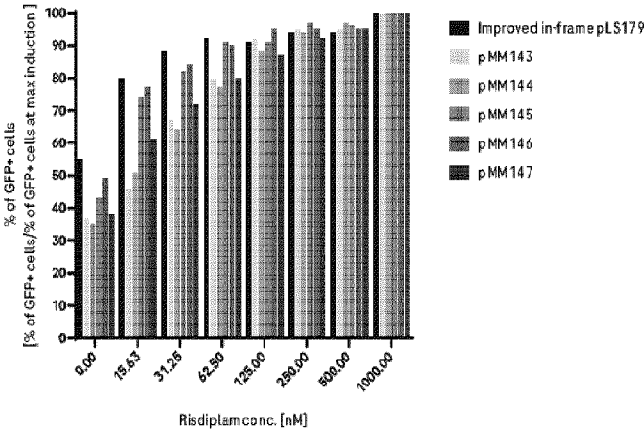
C

Improved in-frame [pLS179] ISS-N1 mutations
median GFP signal [48h]



D

Improved in-frame [pLS179] ISS-N1 mutations
% of GFP+ cells [48h]



E

Leakiness [median GFP]

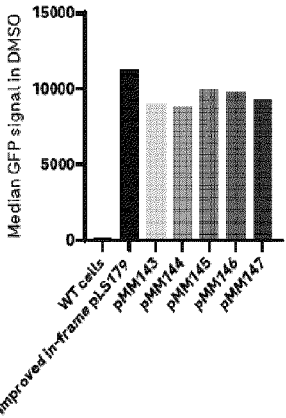
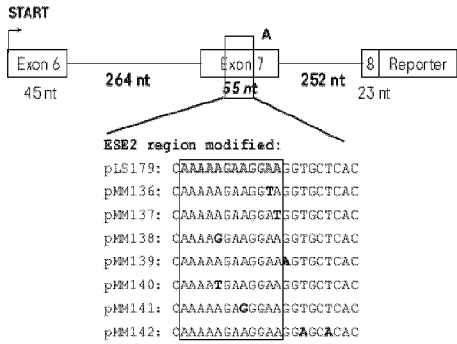


Figure 9

10/44

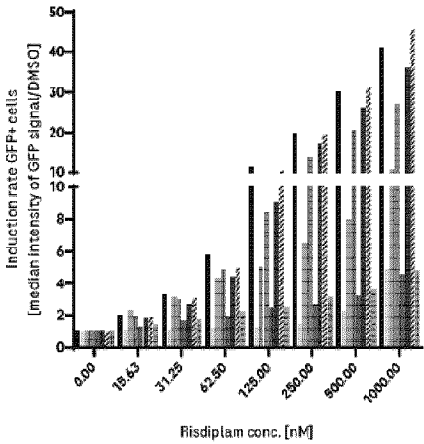
A

pLS179: Mutating ESE2 region to improve switch performance



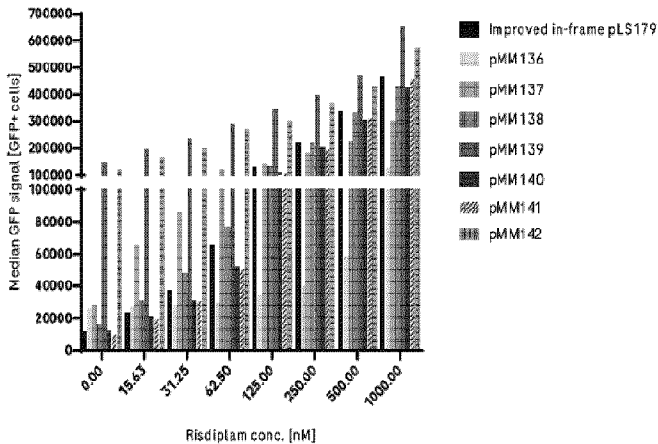
B

Improved in-frame [pLS179] ESE2 mutations
induction rate [48h]



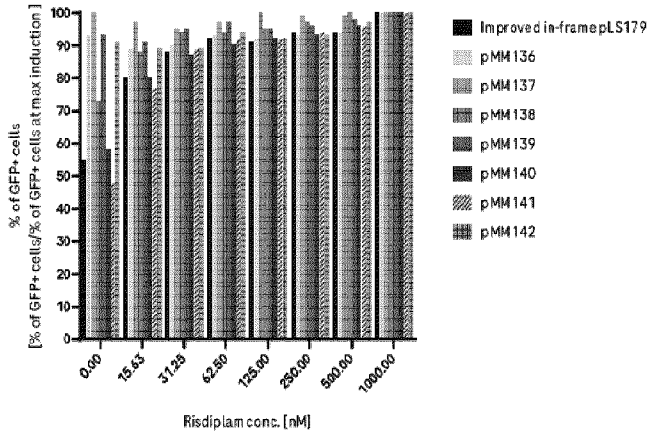
C

Improved in-frame [pLS179] ESE2 mutations
median GFP signal [48h]



D

Improved in-frame [pLS179] ESE2 mutations
% of GFP+ cells [48h]



E

Leakiness [median GFP]

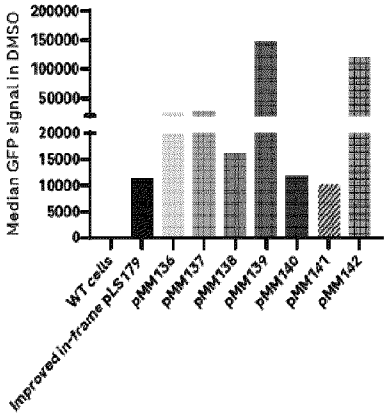
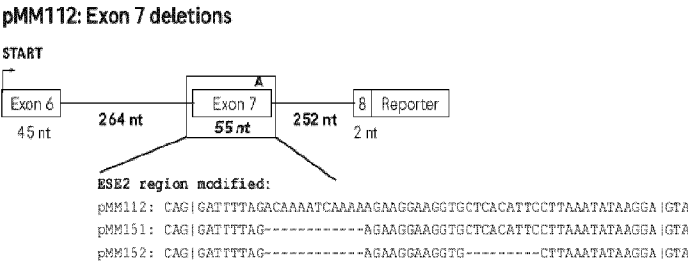


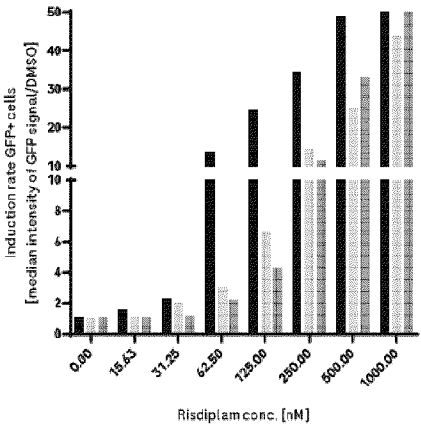
Figure 10

A



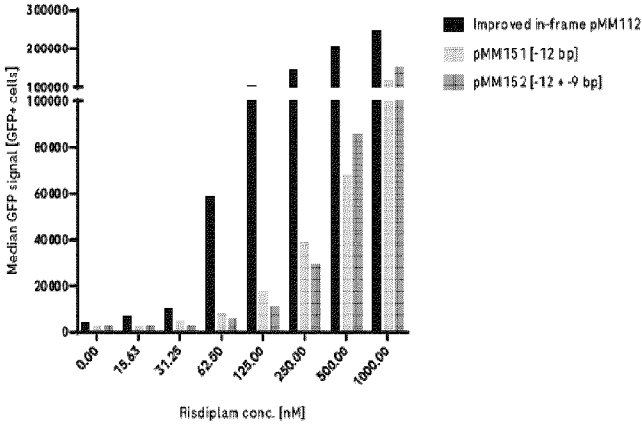
B

**Improved in-frame [pMM112] exon7 deletions
induction rate [48h]**



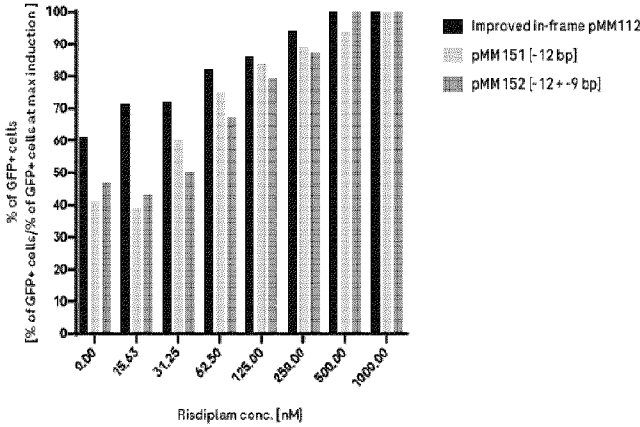
C

**Improved in-frame [pMM112] exon7 deletions
median GFP signal [48h]**



D

**Improved in-frame [pMM112] exon7 deletions
% of GFP+ cells [48h]**



E

Leakiness [median GFP]

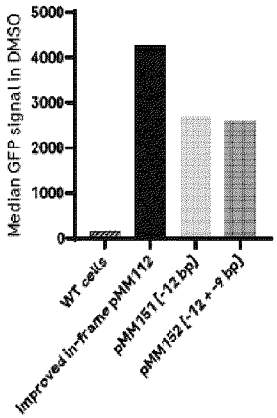
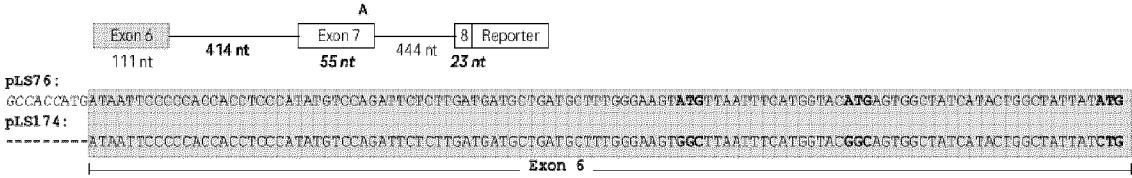


Figure 11

A(i)

Removal of KOZAK-ATG + in-frame ATGs in exon 6



A(ii)

In-frame switch [pLS76]: START in exon 7, based on pLS174

Exon 7 modifications: ESE2 region:

pLS174: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACAT-TCCTTAAATATAAGGA|GTA

pMM56: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATGTCCTTAAATATAAGGA|GTA

pLS174: CAG|GATTTT-----AGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAATATAAGGA|GTA

pMM59: CAG|GATTTTGGCACCATGCAAAAATCAAAAAGGAAGGTGCTCACATTCTTAAATATAAGGA|GTA

pLS174: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAAT-----ATAAGGA|GTA

pMM60: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAAGCACCATGATAAGGA|GTA

pLS174: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAAT-ATAAGGA|GTA

pMM61: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAAGCACCATGGA|GTA

pLS174: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAATATAA-GGA|GTA

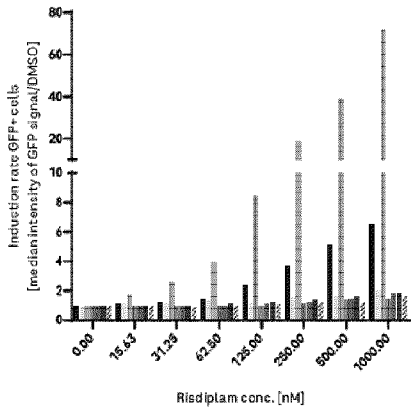
pMM62: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATTCTTAAAGCACCATGGA|GTA

pLS174: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACATT--CCTTAAATATAAGGA|GTA

pMM63: CAG|GATTTTAGACAAAATCAAAAAGGAAGGTGCTCACAGGCCACCATGAATATAAGGA|GTA

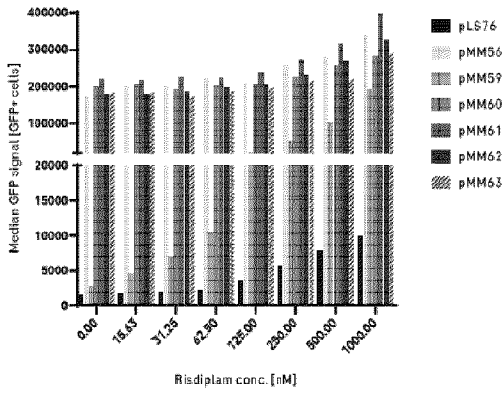
B

START in exon 7 [pLS76]: induction rate [48h]



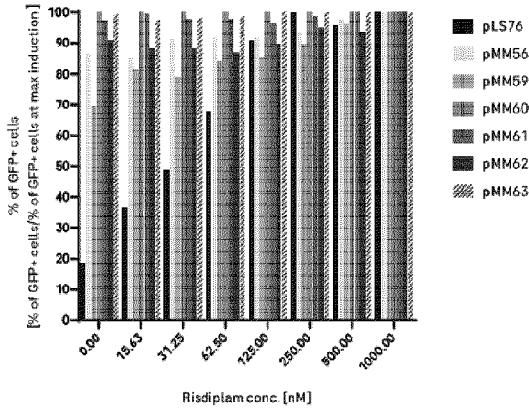
C

START in exon 7 [pLS76]: median GFP signal [48h]



D

START in exon 7 [pLS76]: % of GFP+ cells [48h]



E

Leakiness [median GFP]

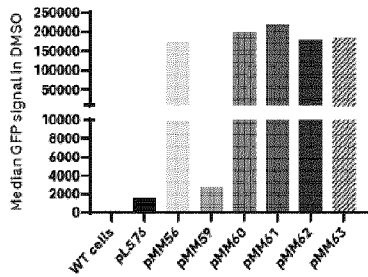
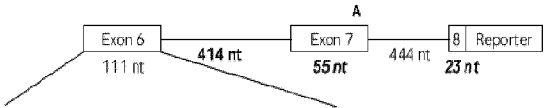


Figure 12

A

In-frame switch [pLS176]: START in exon 7, modifications of exon 6 splice site



Exon 6 - removal of KOZAK-ATG and in-frame ATGs, mutation of the last ATG:

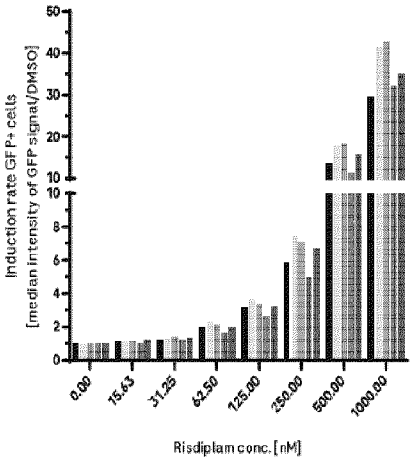
pLS176:
GCCACCATGATAAATCCCCACCCTCCCATATGTCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGT**ATG**TTAATTTTCATGGTAC**ATG**AGTGGCTATCATACTGGCTATTAT**ATG**
pLS176:
-----ATAAATCCCCACCCTCCCATATGTCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGT**GGC**TTAATTTTCATGGTAC**GGC**AGTGGCTATCATACTGGCTATTAT**ATG**
pMM59:
-----ATAAATCCCCACCCTCCCATATGTCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGT**GGC**TTAATTTTCATGGTAC**GGC**AGTGGCTATCATACTGGCTATTAT**CTG**
pMM108:
-----ATAAATCCCCACCCTCCCATATGTCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGT**GGC**TTAATTTTCATGGTAC**GGC**AGTGGCTATCATACTGGCTATTAT**CAG**
pMM110:
-----ATAAATCCCCACCCTCCCATATGTCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGT**GGC**TTAATTTTCATGGTAC**GGC**AGTGGCTATCATACTGGCTATTAT**GTG**
pMM111:
-----ATAAATCCCCACCCTCCCATATGTCAGATTCTCTTGATGATGCTGATGCTTTGGGAAGT**GGC**TTAATTTTCATGGTAC**GGC**AGTGGCTATCATACTGGCTATTAT**TTG**

New translation start site in exon 7:

pLS176: CAG|GATTTT-----AGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA|GTA
pLS176: CAG|GATTT**GCCACCATG**CAAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA|GTA
pMM59: CAG|GATTT**GCCACCATG**CAAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA|GTA

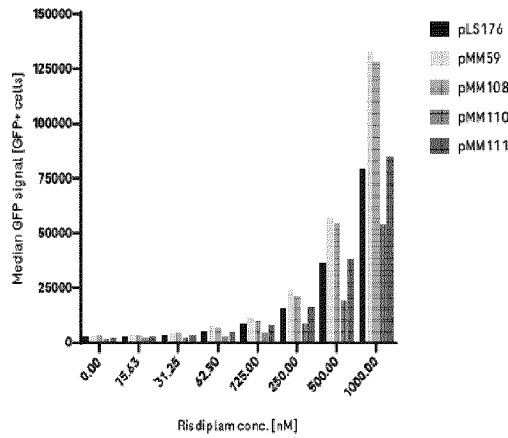
B

START in exon 7 [pLS176]: induction rate [48h]



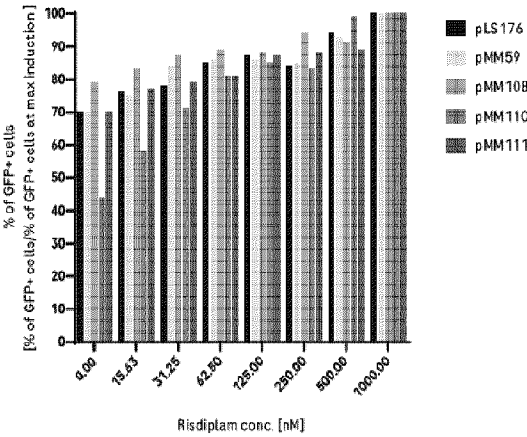
C

START in exon 7 [pLS176]: median GFP signal [48h]



D

START in exon 7 [pLS176]: % of GFP+ cells [48h]



E

Leakiness [median GFP]

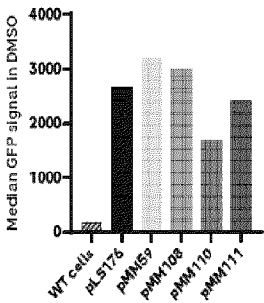


Figure 13

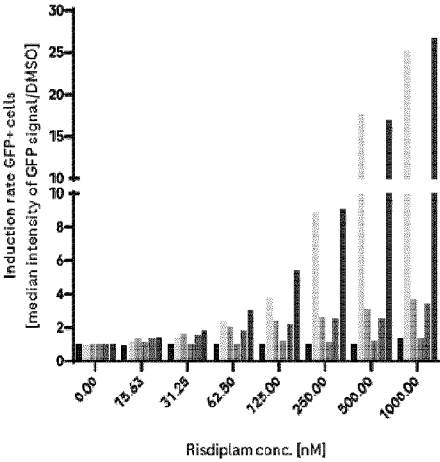
A

In-frame switch [pLS76]: START in exon 7, upstream of ESE2

Exon 7 modifications: ESE2 region:
pLS174: CAG|GATTTT-----AGACAAAATCAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM59: CAG|GATTTT**GCCACCA**TGCAAAAATCAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA Not included
pLS174: CAG|GATTTT**AG**-----CAAAAATCAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM242: CAG|GATTTT**AGGCCACCA**TGCAAAAATCAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pLS174: CAG|GATTTT**AG**CAAAAAT-----CAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM243: CAG|GATTTT**AG**CAAAAATAGCACC**ATG**CAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pLS174: CAG|GATTTT**AG**CAAAAAT-----CAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM244: CAG|GATTTT**AG**CAAAAATAGCACC**ATGG**AAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pLS174: CAG|GATTTT**AG**CAAAAAT-----CAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM245: CAG|GATTTT**AG**CAAAAATAGCACC**ATGG**AAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pLS174: CAG|GATTTT**AG**CAAAAAT---CAAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM246: CAG|GATTTT**AG**CAAAAAT**ATGG**AAAAAGAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pLS174: CAG|GATTTT**AG**CAAAAATCAAAAA---GAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA
pMM247: CAG|GATTTT**AG**CAAAAATCAAAAA**ATG**GAAGGAAGGTGCTCACCATTCTTAAATATAAGGA|GTA

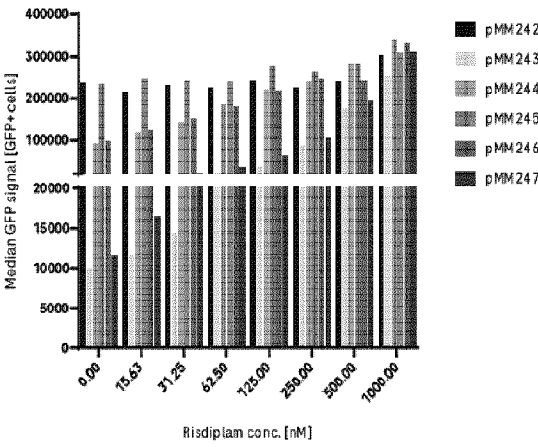
B

START in exon 7 [pLS76]: induction rate [48h]



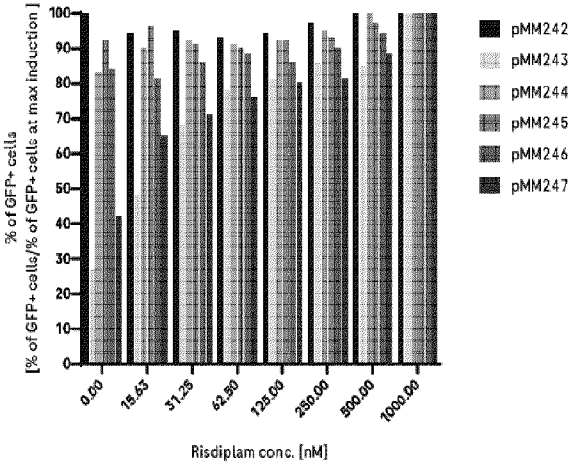
C

START in exon 7 [pLS76]: median GFP signal [48h]



D

START in exon 7 [pLS76]: % of GFP+ cells [48h]



E

Leakiness [median GFP]

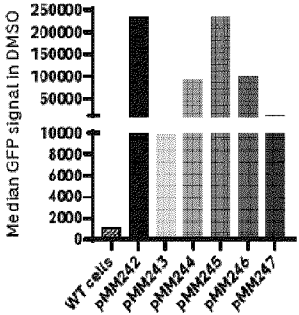


Figure 14

A

In-frame switch [pLS76]: START in exon 7, based on pLS175, downstream of ESE2

Exon 7 modifications:

pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAA-GGA|GTA
pMM123: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAA**AT**GGA|GTA

pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAAT-ATAAGGA|GTA
pMM124: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAA**AT**GTAAGGA|GTA

pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGG--TGC-TCACATTCCTTAAATATAAGGA|GTA
pMM125: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGG**ATGGCGT**CACATTCCTTAAATATAAGGA|GTA

pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACAT-TCCTTAAATATAAGGA|GTA
pMM126: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCAC**ATG**TCCTTAAATATAAGGA|GTA

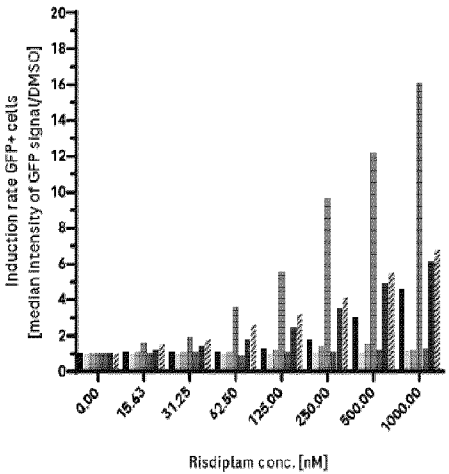
pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCC-TAAATATAAGGA|GTA
pMM127: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCC**TATGG**ATATAAGGA|GTA

pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAG--GA|GTA
pMM128: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAG**ATGA**|GTA

pLS175: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA|GTA
pMM129: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAA--**ATG**AGGA|GTA

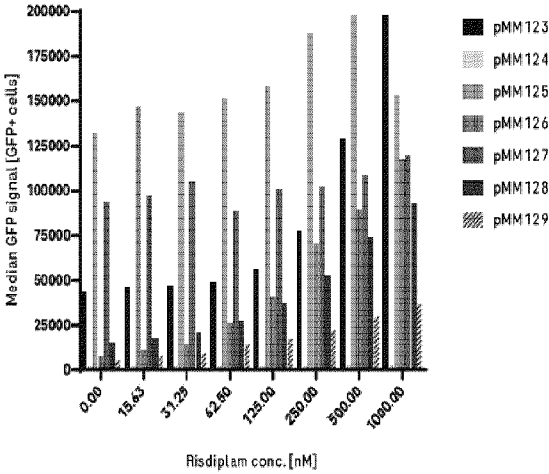
B

START in exon 7 [pLS175]: induction rate [48h]



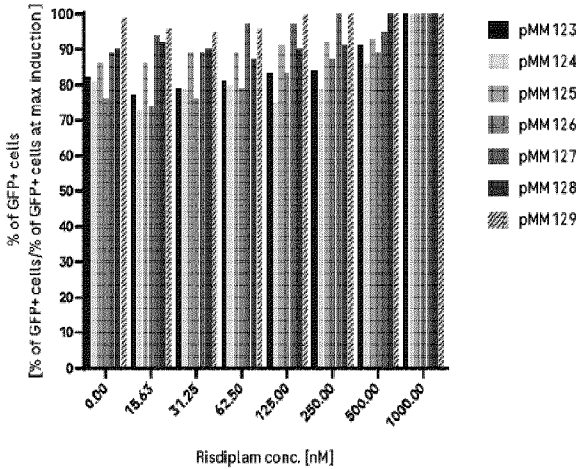
C

START in exon 7 [pLS175]: median GFP signal [48h]



D

START in exon 7 [pLS175]: % of GFP+ cells [48h]



E

Leakiness [median GFP]

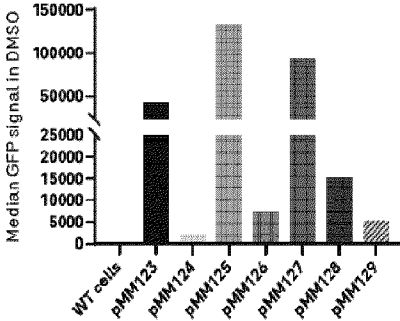
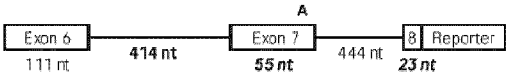


Figure 15

A

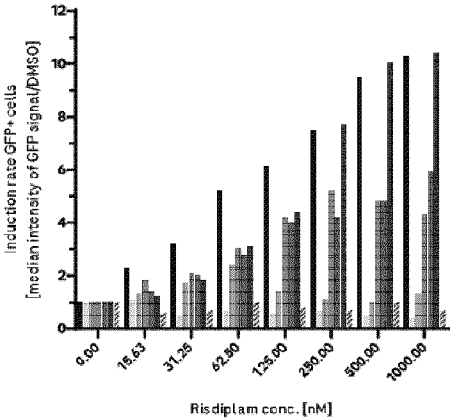
In-frame switch [pLS76]: START in exon 7, based on pMM126, downstream of ESE2



Exon 7 modifications: ESE2 region:
pLS175: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACAT-TCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM201: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM202: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM203: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM204: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM205: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM126: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA
pMM206: CAG | GATTTTAGACAAAATCAAAAAGAGGAAGGTGCTCACATGTCCTTAATATAAGGA | GTA

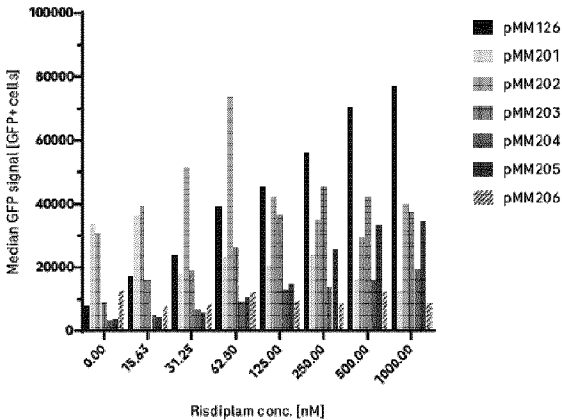
B

START in exon 7 [pMM126]: induction rate [24h]



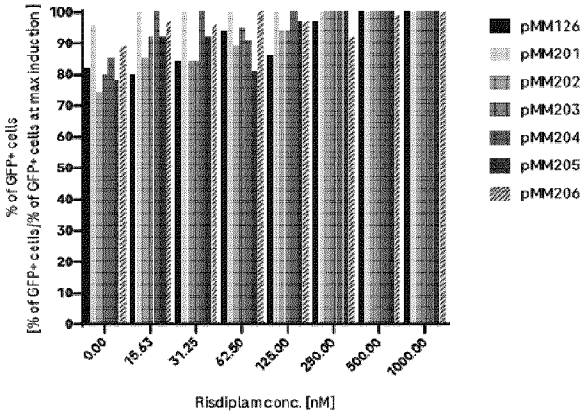
C

START in exon 7 [pMM126]: median GFP signal [24h]



D

START in exon 7 [pMM126]: % of GFP+ cells [24h]



E

Leakiness [median GFP]

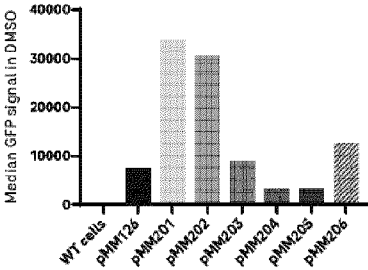


Figure 16

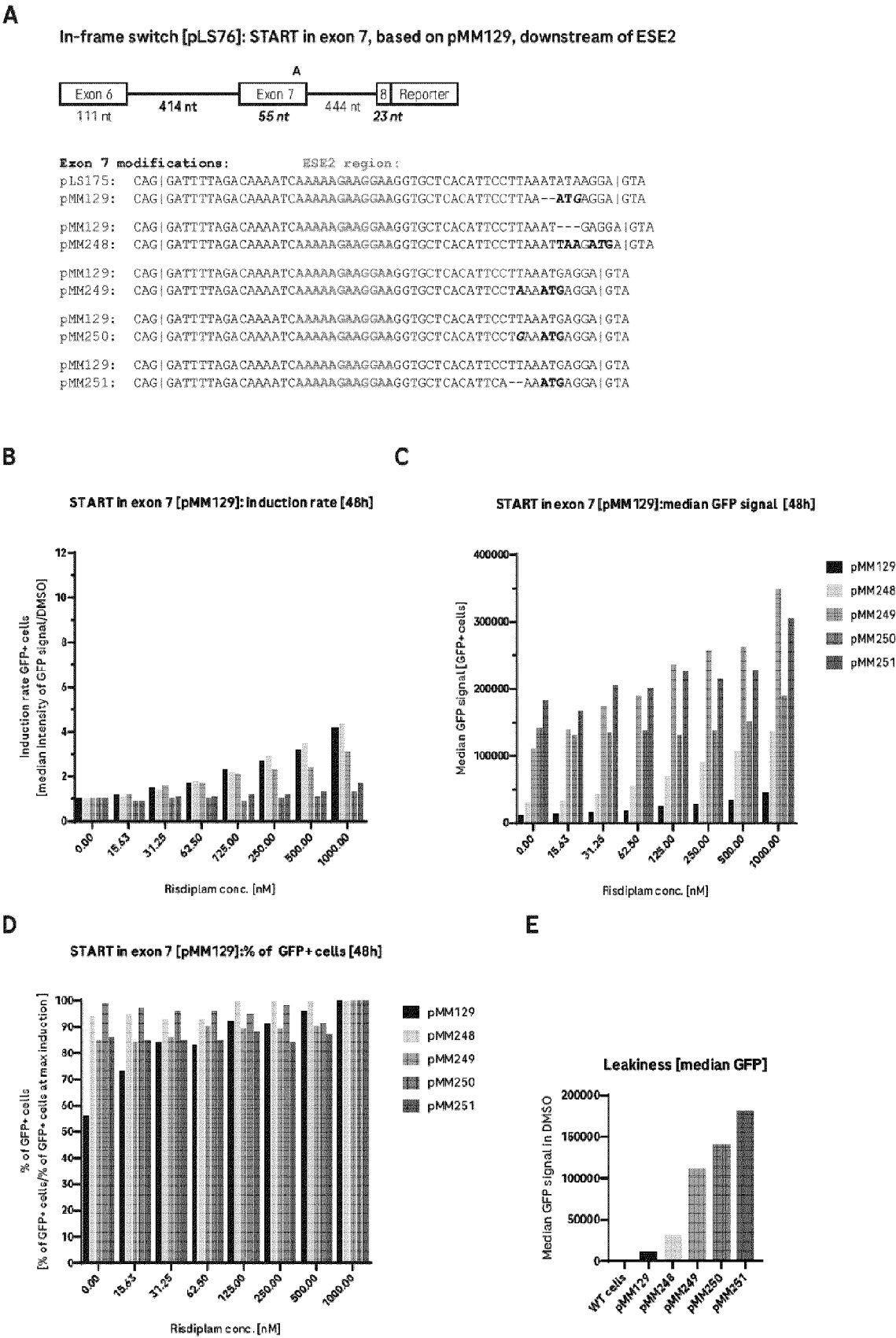
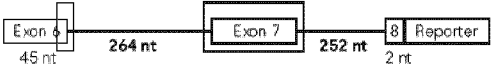


Figure 17

A

In-frame switch: START in exon 7, pMM130 backbone based



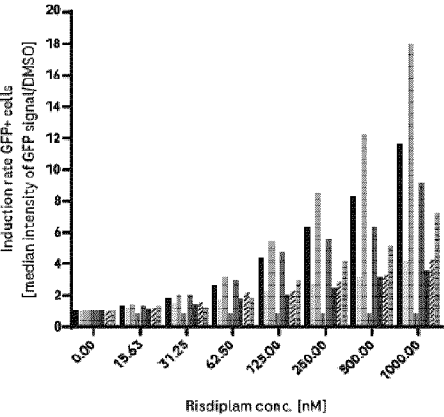
Exon 6 modification:
pMM130: GCAATTTCATGGTACGGCAGTGGCTATCATACTGGCTATTAT**ATG**
pMM252: GCAATTTCATGGTACGGCAGTGGCTATCATACTGGCTATTAT**CAG**

Exon 7 modifications: ESE2 region:
pMM252: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACAT-TCCTTAAATATAAGGA|GTA
pMM254 (as pMM126): CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACAT**GTCC**TTAAATATAAGGA|GTA
pMM252: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAG--GA|GTA
pMM255 (as pMM128): CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAG**ATGA**|GTA
pMM252: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA|GTA
pMM256 (as pMM129): CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAA--AT**GAGGA**|GTA
pMM252: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAAT-ATAAGGA|GTA
pMM257: CAG|GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAAT**GATAAGGA**|GTA

Tag length:
pMM254: 6 aa
pMM255: 1 aa
pMM256: 2 aa
pMM257: 3 aa

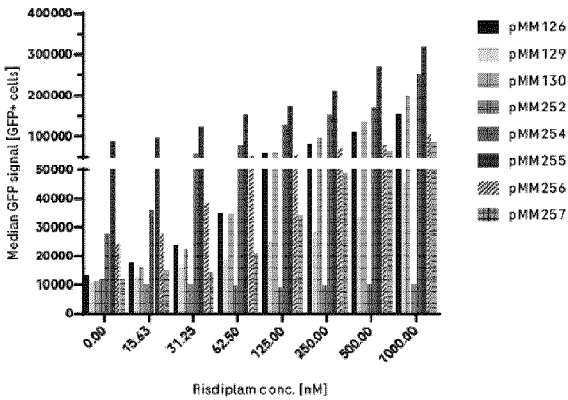
B

START in exon 7 [pMM252]: induction rate [48h]



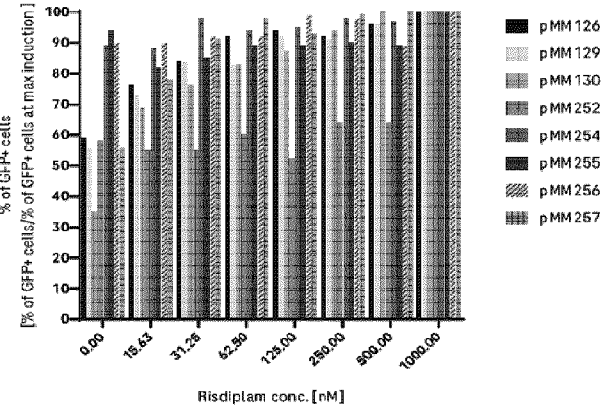
C

START in exon 7 [pMM252]: median GFP signal [48h]



D

START in exon 7 [pMM252]: % of GFP+ cells [48h]



E

Leakiness [median GFP]

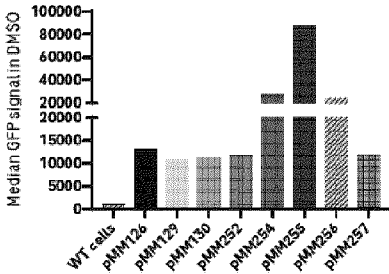
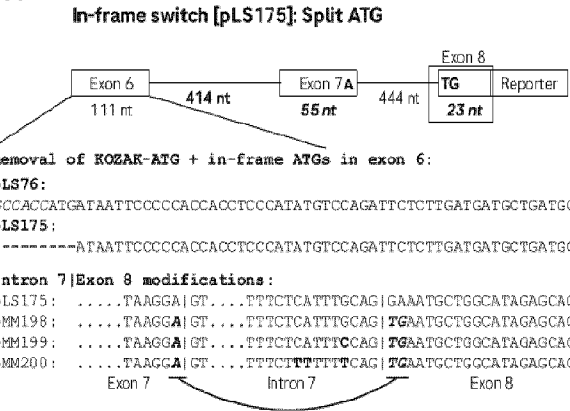
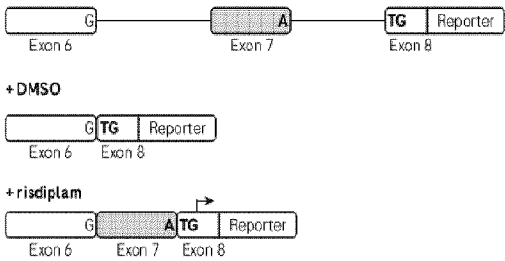


Figure 18

A

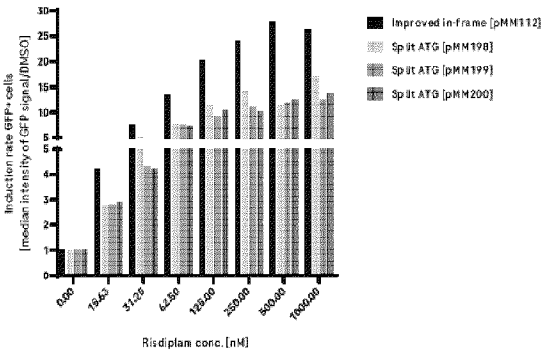


B



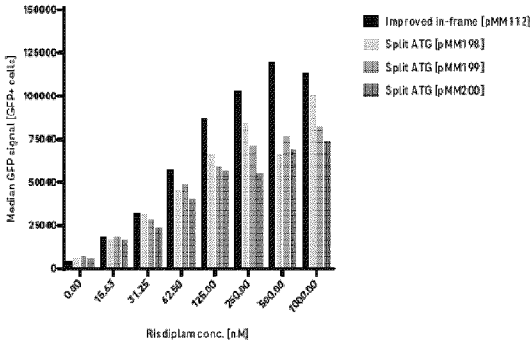
C

In-frame switch split ATG: induction rate [24h]



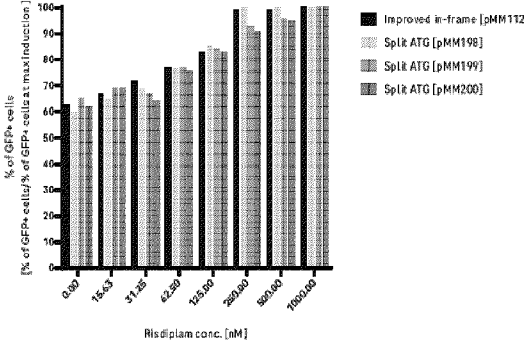
D

In-frame switch split ATG: median GFP signal [24h]



E

In-frame switch split ATG: % of GFP+ cells [24h]



F

Leakiness [median GFP]

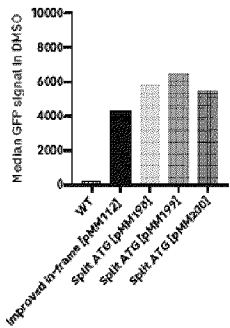
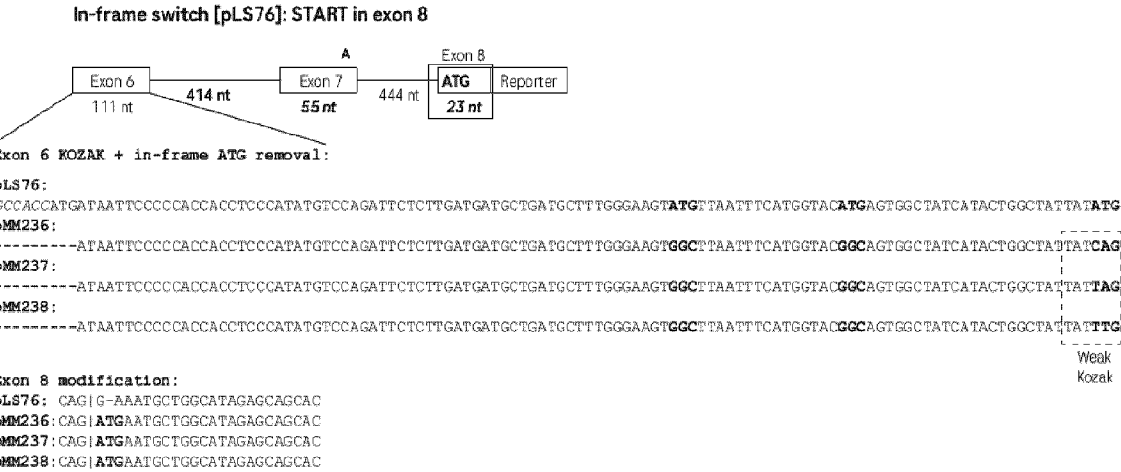
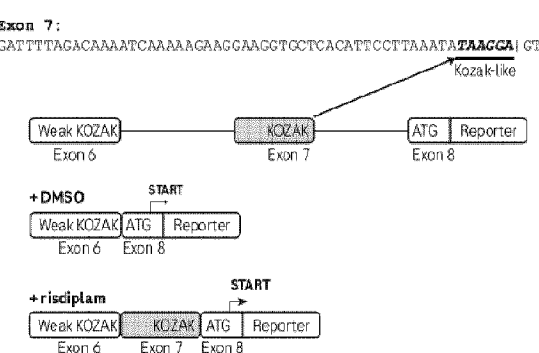


Figure 19

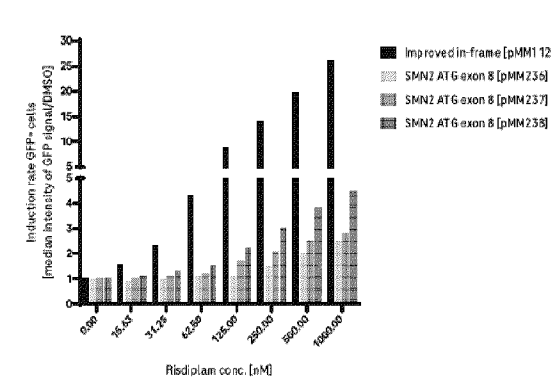
A



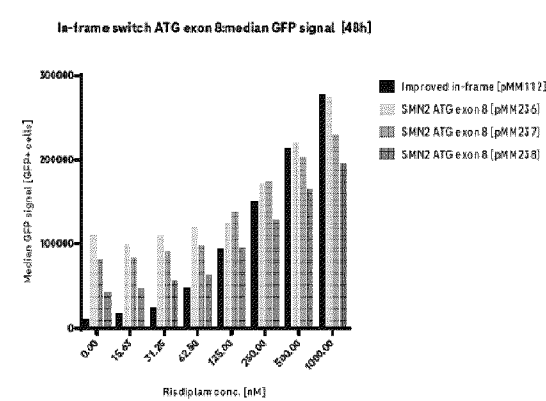
B



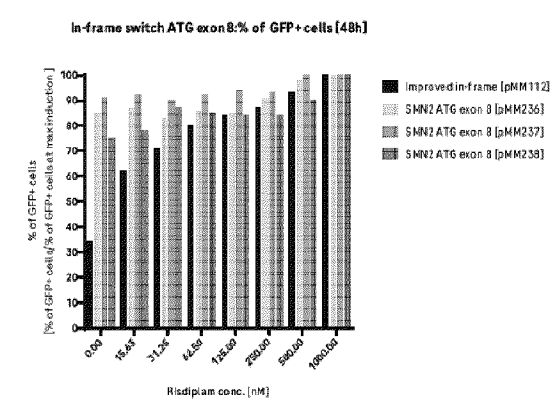
C



D



E



F

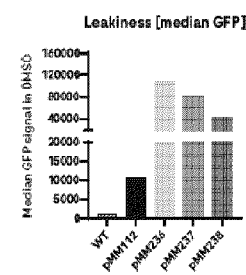


Figure 20

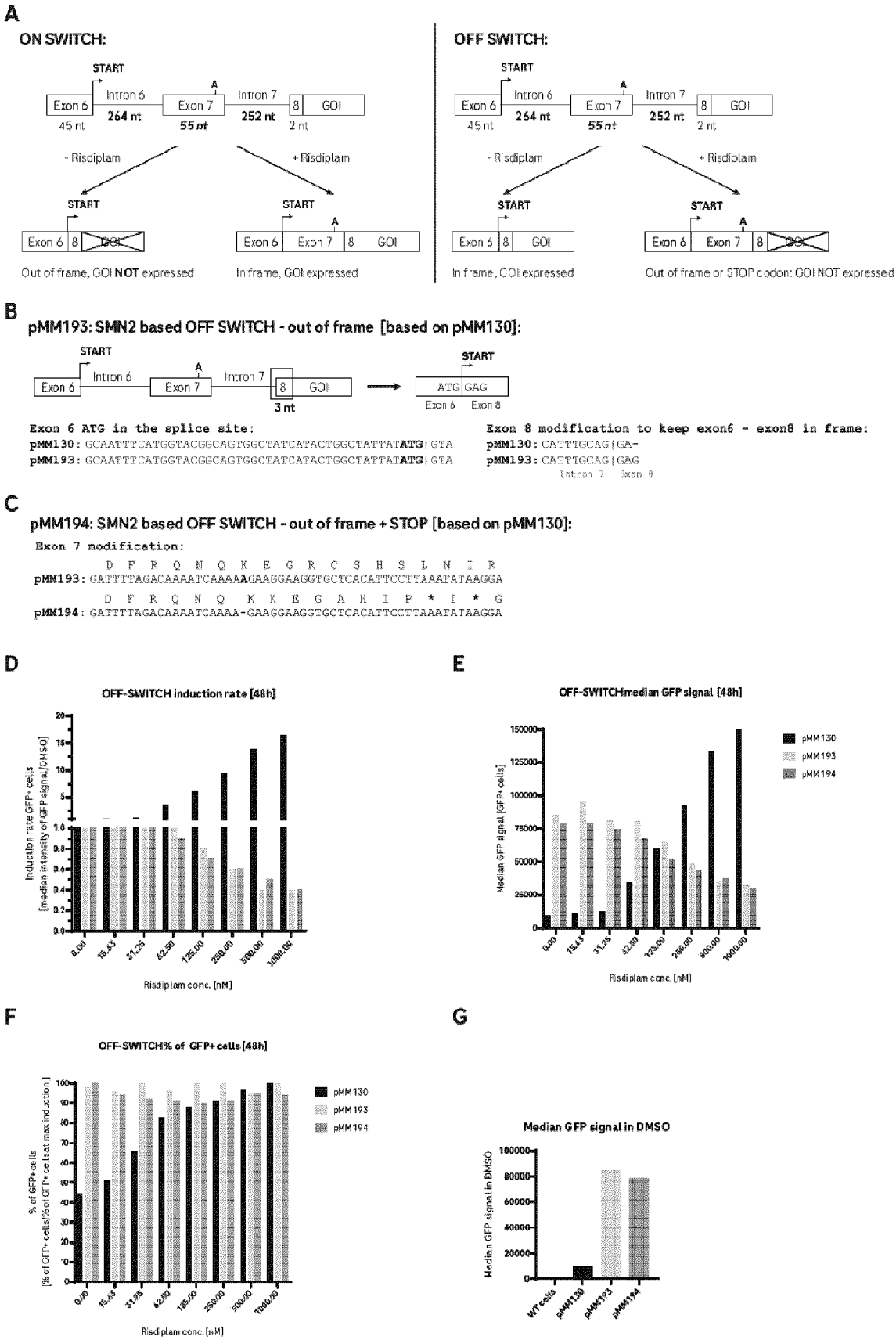
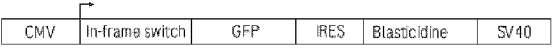


Figure 21

IRES-blasticidine replaced by BGH polyA signal, constructs compared with Monteys et al. construct:

A

Old construct design:



New construct design:



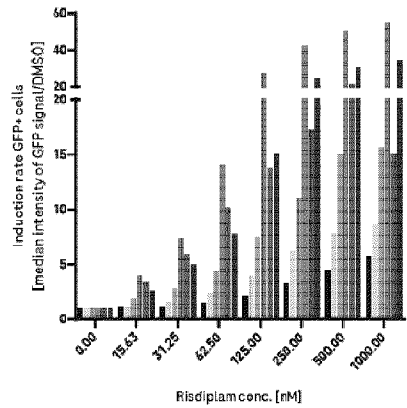
B

Source of switches for the corresponding constructs:

pMM274 = Monteys et al. construct
pMM263 = pLS76
pMM262 = pLS168
pMM259 = pMM112
pMM260 = pMM130
pMM272 = pMM198

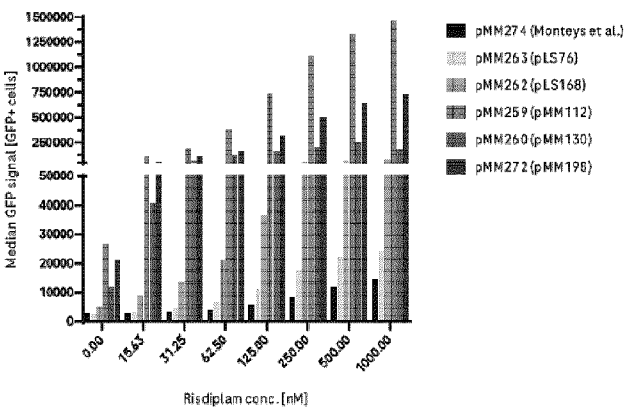
C

IRES replaced by BGH polyA: induction rate [48h]



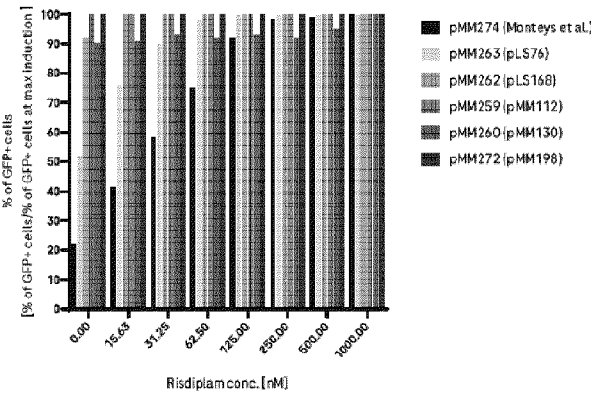
D

IRES replaced by BGH polyA: median GFP signal [48h]



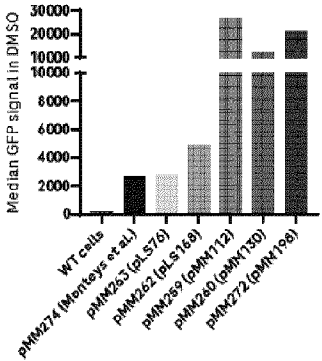
E

IRES replaced by BGH polyA: % of GFP+ cells [48h]



F

Leakiness [median GFP]



G

Induction rate at 125 nM [48h]

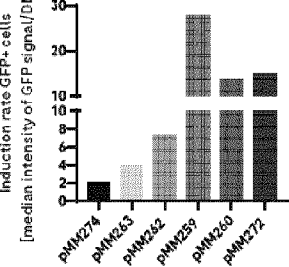


Figure 22

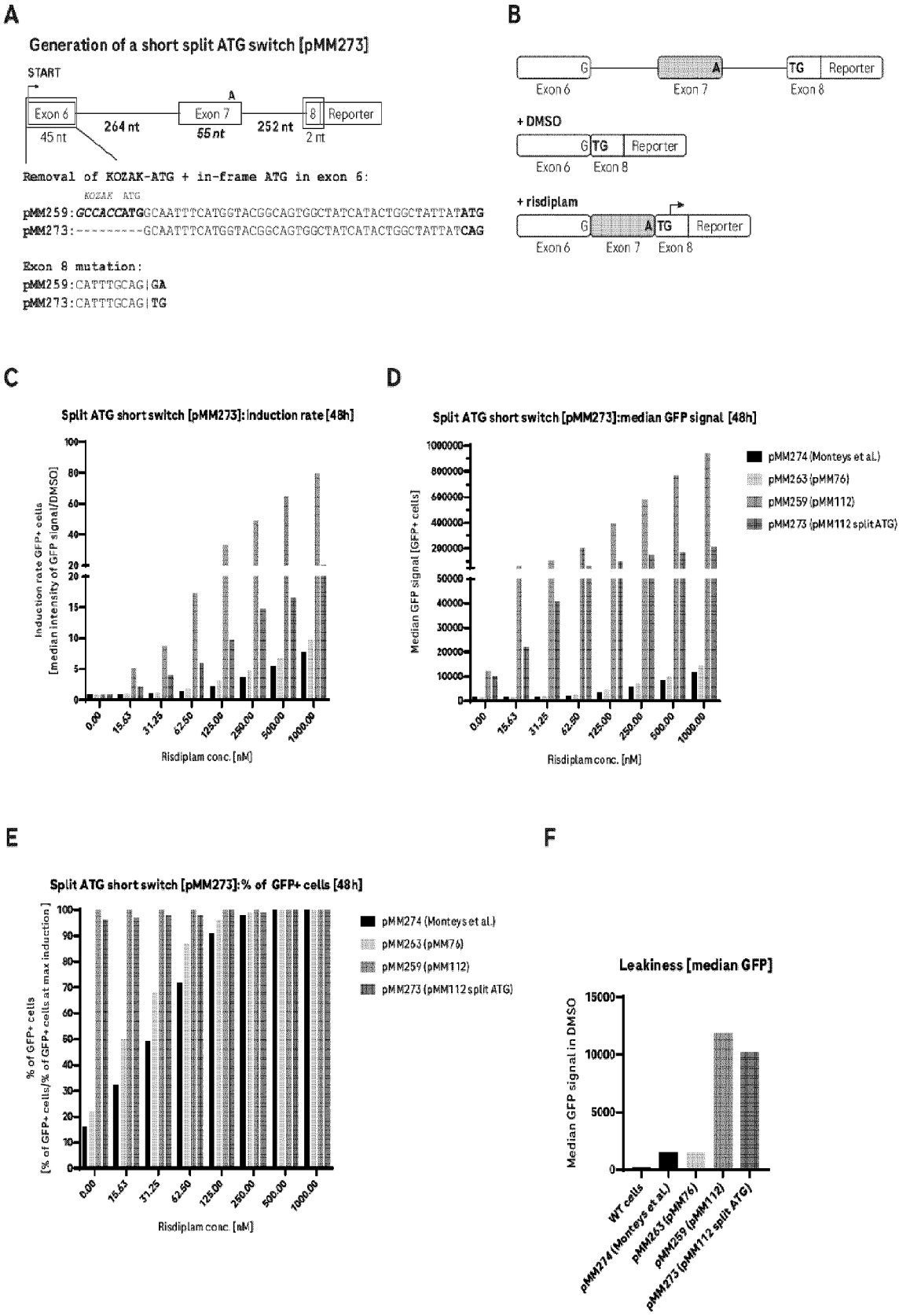


Figure 23

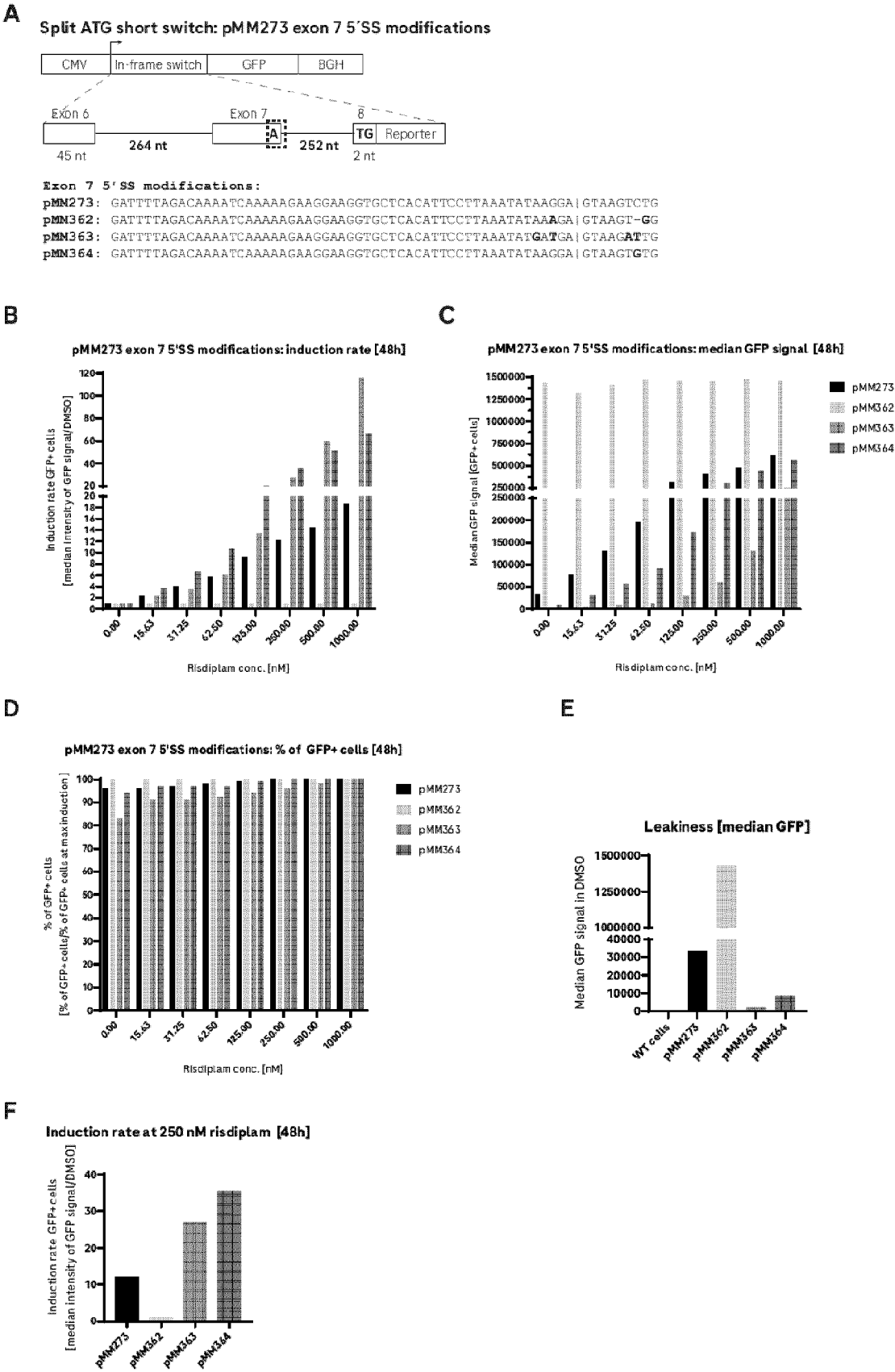
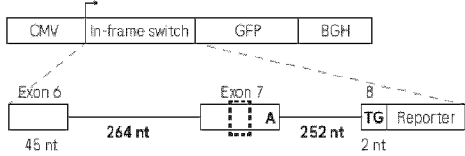


Figure 24

A

Split ATG short switch -ESE2 [pMM359]: ESE2 region deletions in pMM273 to reduce leakiness



Exon 7 ESE2 deletions:

pMM273: GATTTAGACAAAATCAAAAAGAAAGGTGCTCACATTCCTTAAATATAAGGA

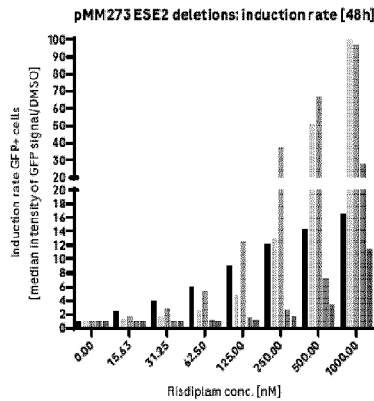
pMM358: GATTTAGACAAAATCAAAA-----AGGTGCTCACATTCCTTAAATATAAGGA

pMM359: GATTTAGACAAAATC-----GGTGCTCACATTCCTTAAATATAAGGA

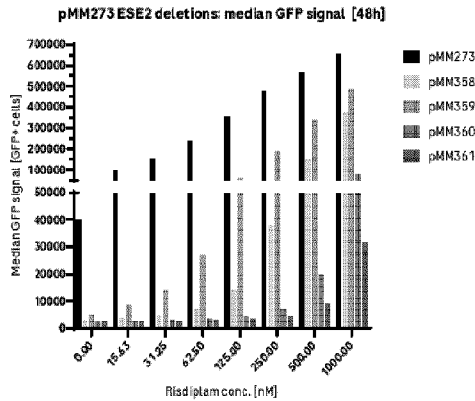
pMM360: GATTTAGACAAAATC-----CTCACATTCCTTAAATATAAGGA

pMM361: GATTTAGAC-----TCACATTCCTTAAATATAAGGA

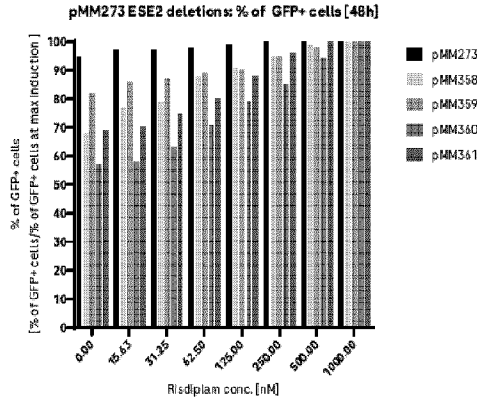
B



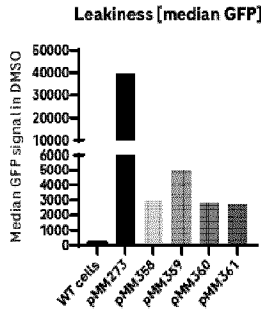
C



D

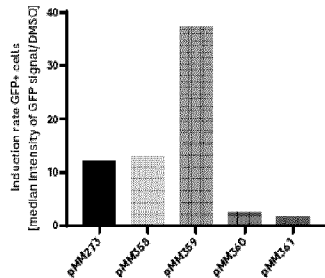


E



F

pMM273 ESE2 deletions: induction rate at 250 nM risdiplam [48h]



G

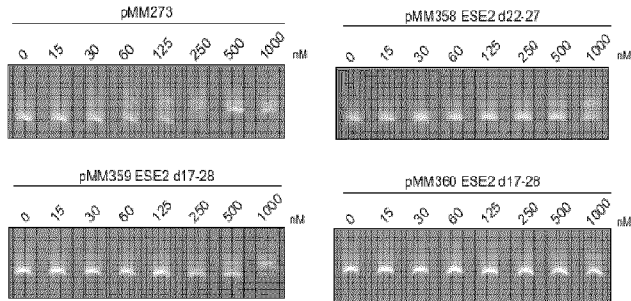


Figure 25

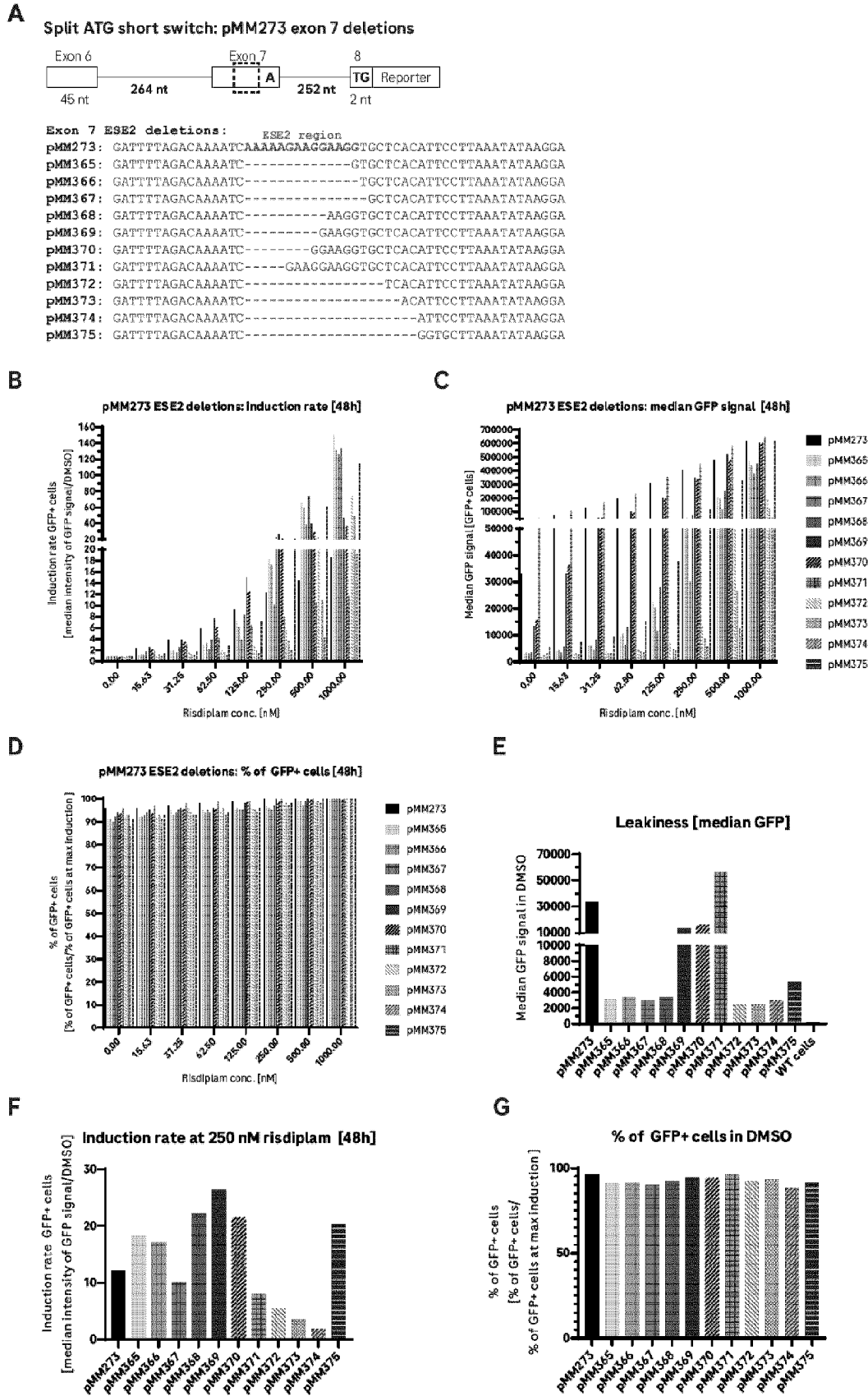
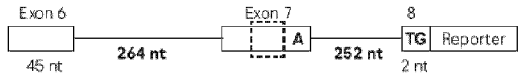


Figure 26

A

Split ATG short switch: pMM273 exon 7 deletions



Exon 7 ESE2 deletions: ESE2 region

pMM273: GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA

pMM376: GATTTTAGACAAAA-----GGTGTCTCATTCCCTTAAATATAAGGA

pMM377: GATTTTAGAC-----GGTGTCTCATTCCCTTAAATATAAGGA

pMM378: GATTTTAGA-----GGTGTCTCATTCCCTTAAATATAAGGA

pMM379: GATTTTAG-----AGAAGGAAGGTGCTCATTCCCTTAAATATAAGGA

pMM380: GATTTTAGACAAAATCAAAAAGAAGGAAGGTG-----CTTAAATATAAGGA

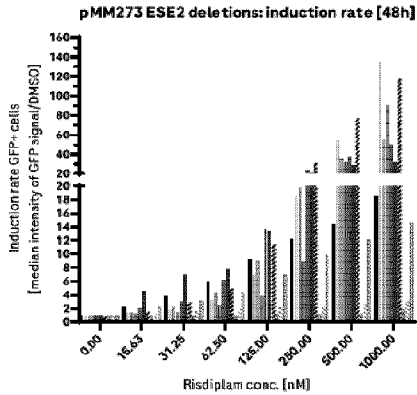
pMM381: GATTTTAG-----AGAAGGAAGGTG-----CTTAAATATAAGGA

pMM382: GATTTTAG-----CTTAAATATAAGGA

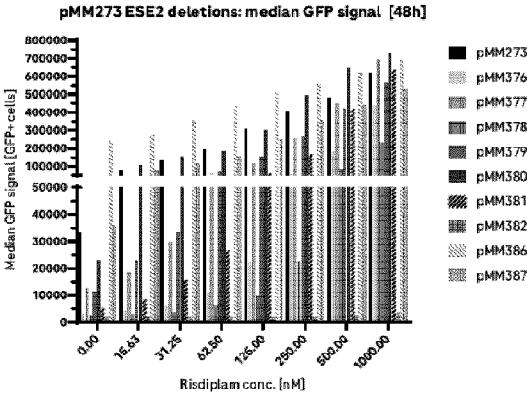
pMM386: GATTTTAGACAAAATC---AAGAAGGAAGGTGCTCATTCCCTTAAATATAAGGA

pMM387: GATTTTAGACAAAATC-AAAAGAAGGAAGGTGCTCATTCCCTTAAATATAAGGA

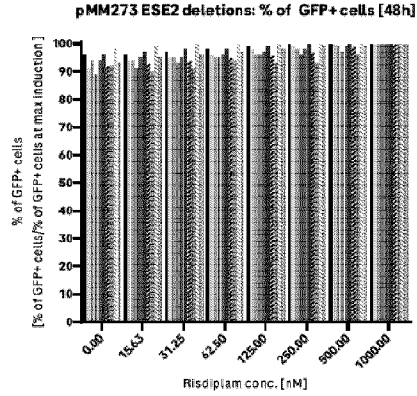
B



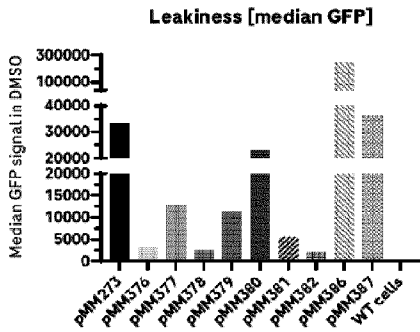
C



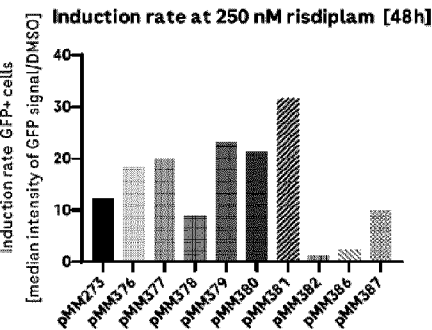
D



E



F



G

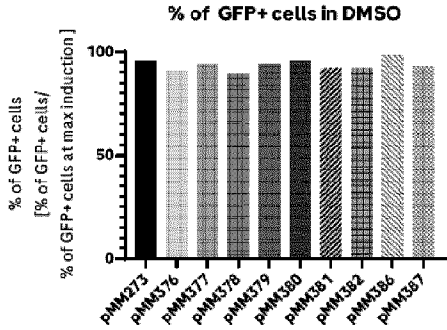
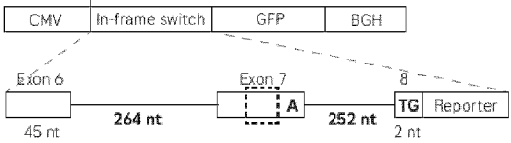


Figure 27

A

Split ATG short switch: pMM273 exon 7 deletions



Exon 7 ESE2 deletions:

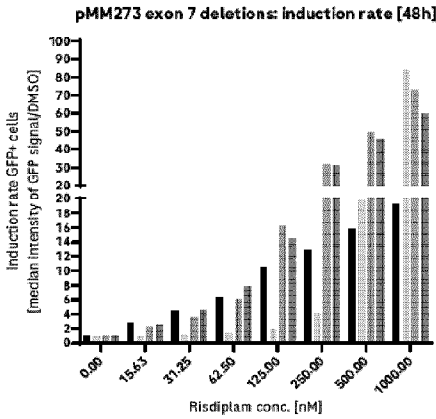
pMM273: GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCTTAAATATAAGGA

pMM436: GATTTTAGACAAAATCAAAA-----AAGG-----CACATTCTTAAATATAAGGA

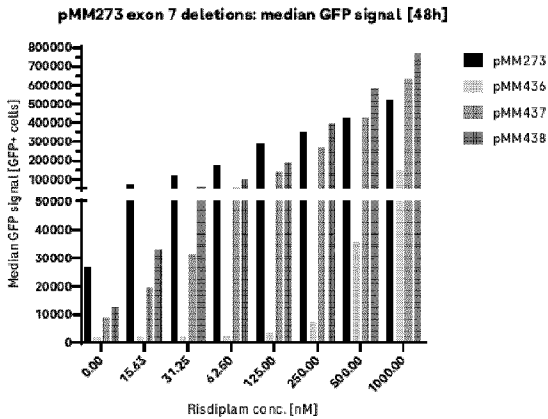
pMM437: GATTTTAGACAAAATCAAAAAGAAGGA-----CTCACATTCTTAAATATAAGGA

pMM438: GATTTTAGACAAAATCAAAAAGAAGGA-----CACATTCTTAAATATAAGGA

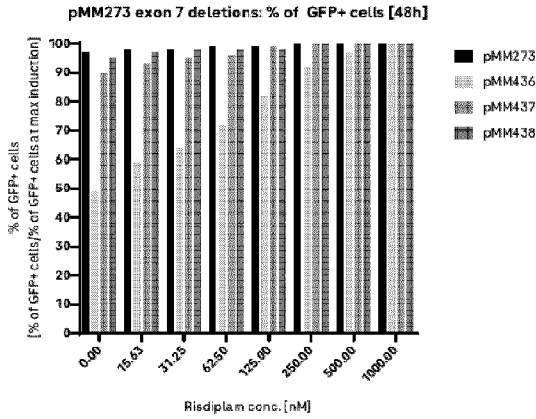
B



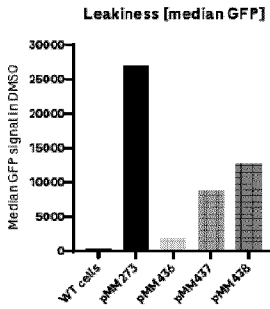
C



D



E



F

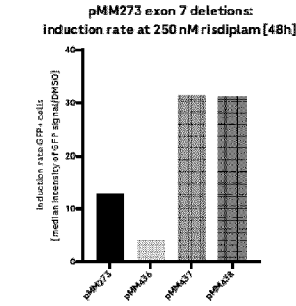


Figure 28

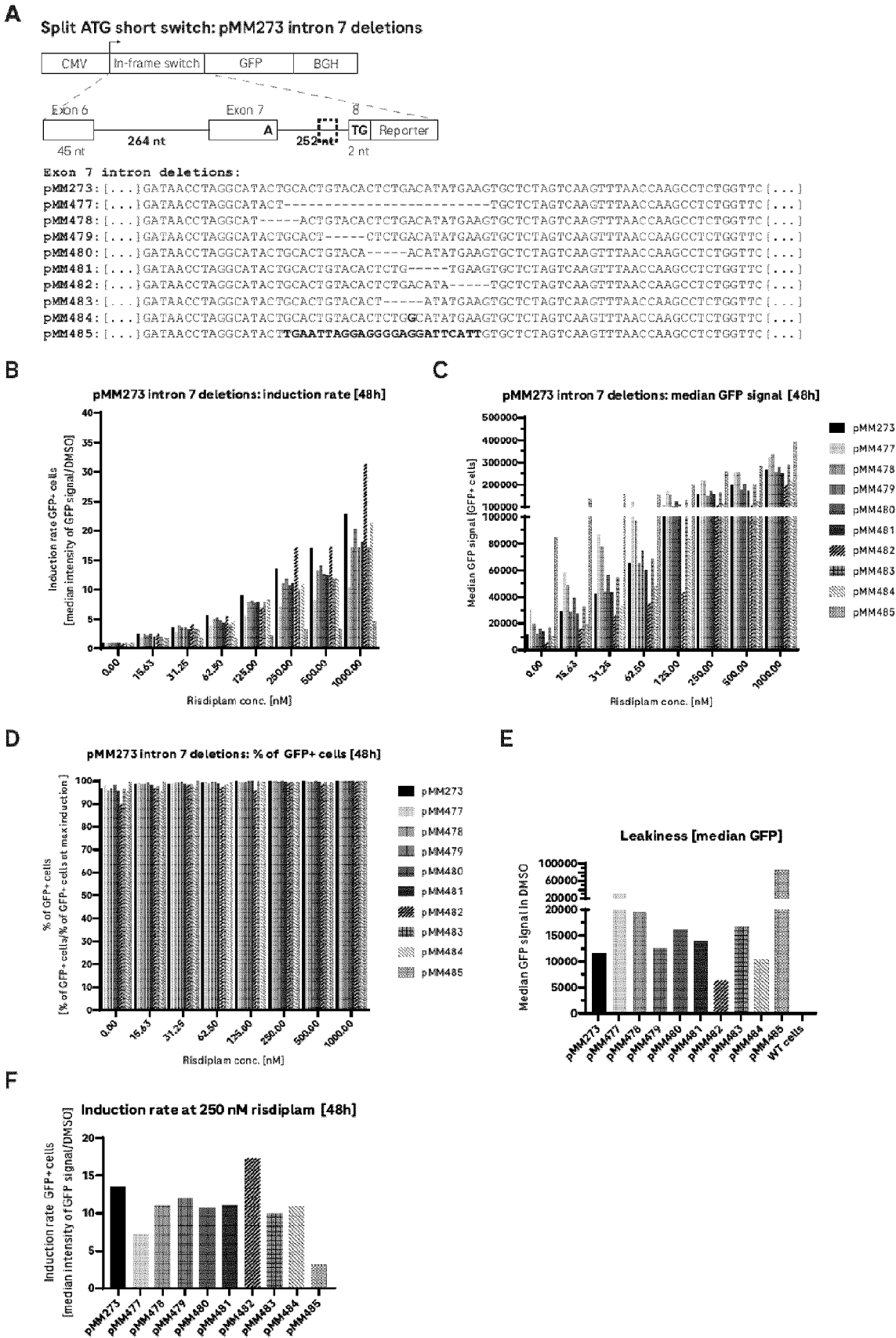
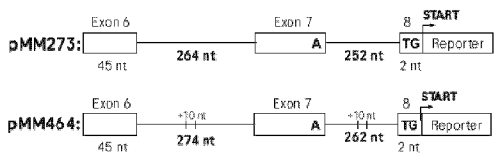


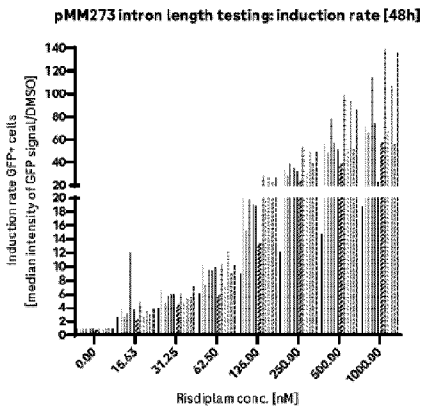
Figure 29

A Split ATG short switch [pMM273]: intron length testing

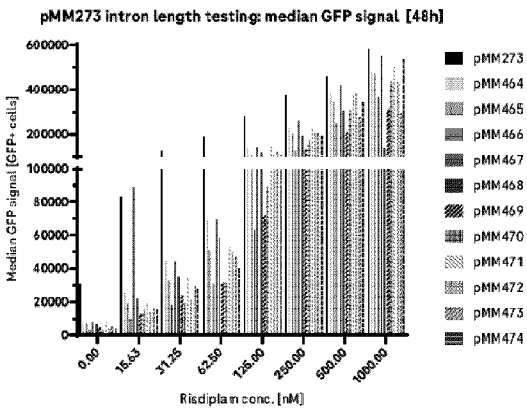


	Intron 7			Intron 8		
	First half [nt]	Second half [nt]	Intron length [nt]	First half [nt]	Second half [nt]	Intron length [nt]
pMM273	102	162	264	102	150	252
pMM464	+5	+5	274	+5	+5	262
pMM465	+10	+10	284	+10	+10	272
pMM466	+15	+15	294	+15	+15	282
pMM467	+20	+20	304	+20	+20	292
pMM468	-	-	264	+5	+5	262
pMM469	-	-	264	+10	+10	272
pMM470	-	-	264	+15	+15	282
pMM471	-	-	264	-	+5	257
pMM472	-	-	264	+5	+20	277
pMM473	-	-	264	+5	+10	267
pMM474	-	-	264	+5	+15	272

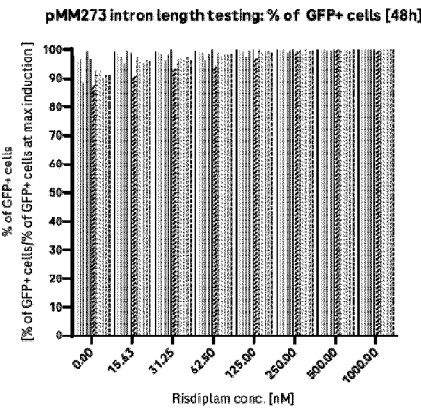
B



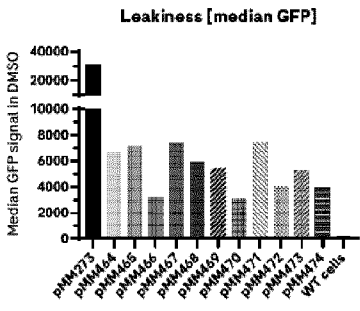
C



D



E



F

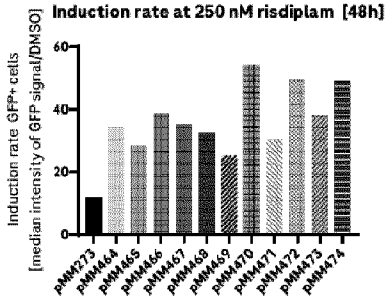
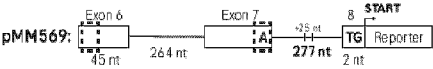


Figure 30

A

Improved split ATG short switch: exon 6 and 7 mutations

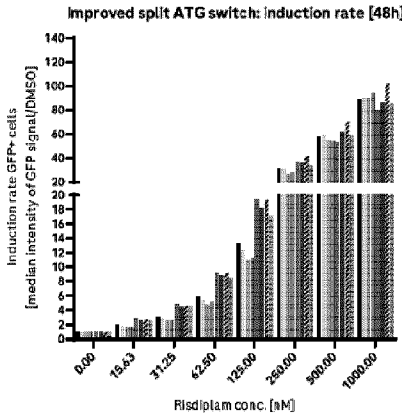


Exon 6 ATG mutation in pMM567-570:
pMM466/470/472/474: GCAATTTTCATGGTACGGCAGTGGCTATCATACTGGCTATTATCAG | GTA
pMM567-570: GCAATTTTCCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAG | GTA

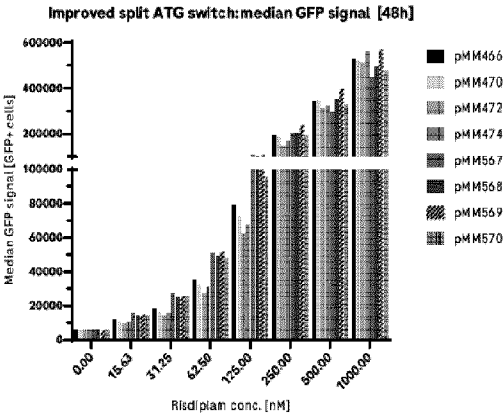
Exon 7 TAA deletion in pMM567-570:
pMM466/470/472/474: GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATATAAGGA
pMM567-570: GATTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATA---GGA

		Intron 7			Intron 8		
		First half [nt]	Second half [nt]	Intron length [nt]	First half [nt]	Second half [nt]	Intron length [nt]
pMM567	pMM466	+15	+15	294	+15	+15	282
pMM568	pMM470	-	-	264	+15	+15	282
pMM569	pMM472	-	-	264	+5	+20	277
pMM570	pMM474	-	-	264	+5	+10	257

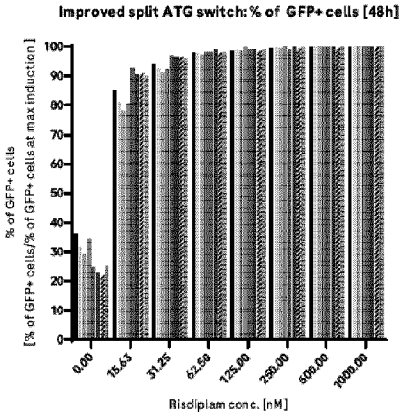
B



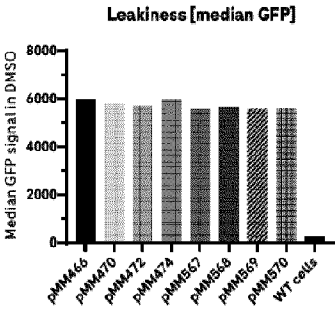
C



D



E



F

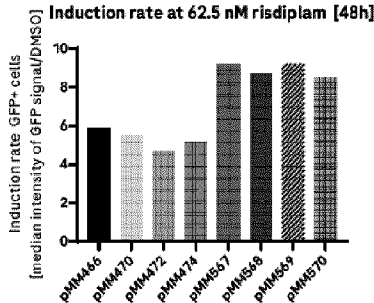
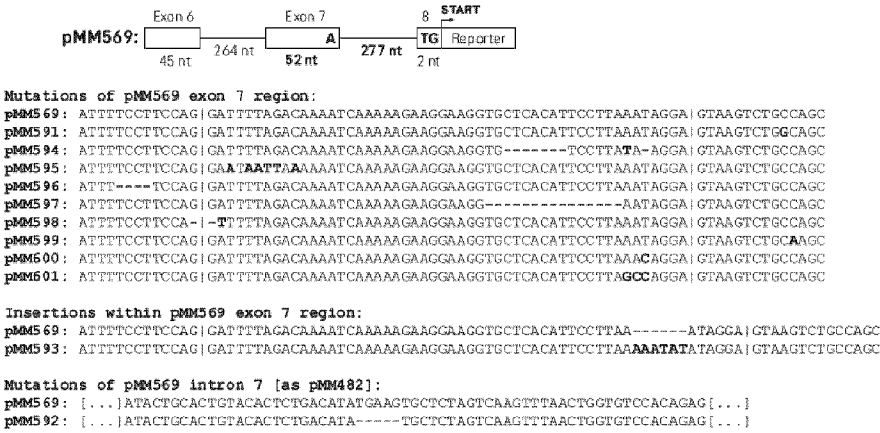


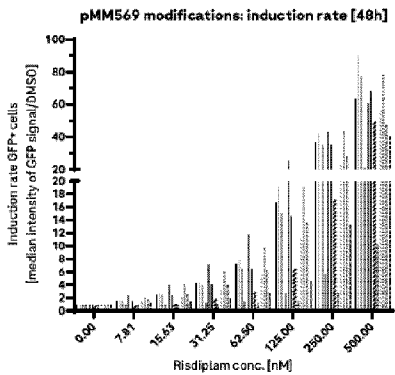
Figure 31

A

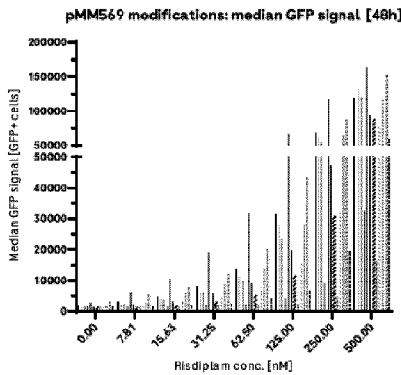
Improved split ATG switch [pMM569]: exon 7 region and intron 7 modifications



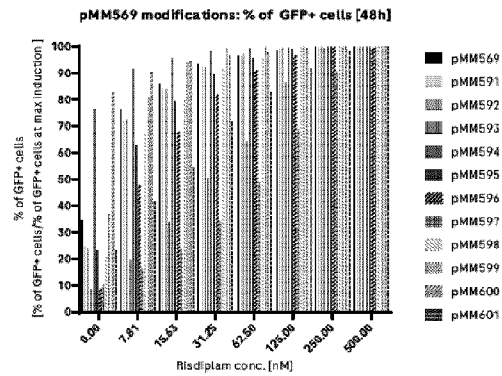
B



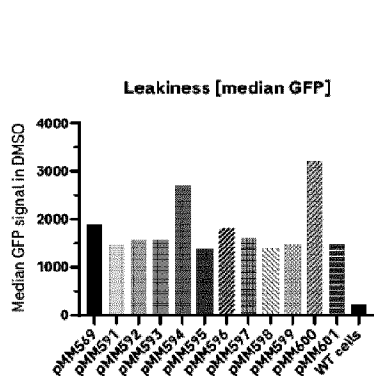
C



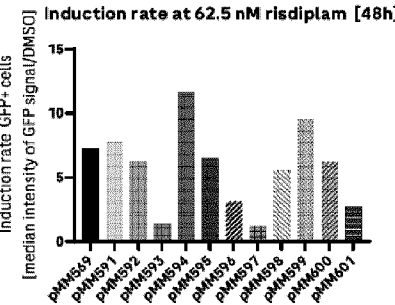
D



E



F



G

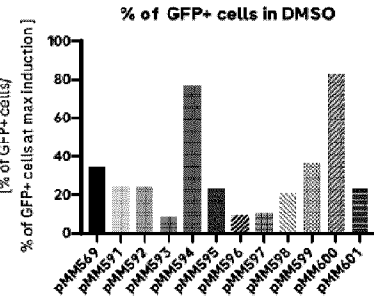
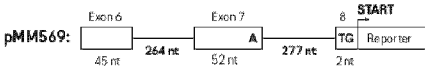


Figure 32

A

Improved split ATG switch [pMM569]: intron 6 deletion screen, 5 nt non-overlapping deletions



Intron 6 [264 nt]:



Example of intron 6 [264 nt] 5 nt deletions:

WT: GTAAGTAATCACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
Del1: ----TATCACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
Del2: GTAAG----ACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
Del3: GTAAGTAATC----GCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]

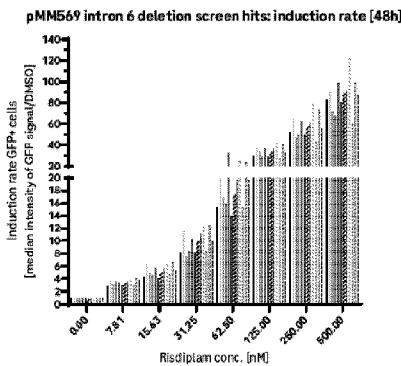
B

Selected intron 6 deletions:

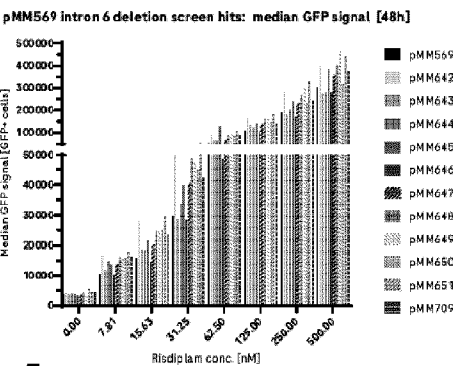
pMM569: GTAAGTAATCACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
pMM642: GTAAGTAATC----GCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
pMM643: GTAAGTAATCACTCAGCATC----CTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
pMM709: GTAAGTAATCACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
pMM644: GTAAGTAATCACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTTGTTTGTAAATTTAT [...]
pMM645: GTAAGTAATCACTCAGCATCTTTCTGACAAATTTTGTAGTTATGTGACTTT----GTAAATTTAT [...]

pMM569: [...] TTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCT [...]
pMM646: [...] TTCTGAT----TTTGTGAATAAAATAAGTAAATGTCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCT [...]
pMM647: [...] TTCTGATCATAT----TGAATAAAATAAGTAAATGTCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCT [...]
pMM648: [...] TTCTGATCATATTTTGT----AAAATAAGTAAATGTCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCT [...]
pMM649: [...] TTCTGATCATATTTTGTGAATAAAATAAGTA----TCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCT [...]
pMM650: [...] TTCTGATCATATTTTGTGAATAAAATAAGTAAATGTCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTA----AGCTATCT [...]
pMM651: [...] TTCTGAT-----AAAATAAGTAAATGTCCTTGAAACAAAATGCTTTTAAACATCCATATAAAGCTATCTATATATAGCTATCT [...]

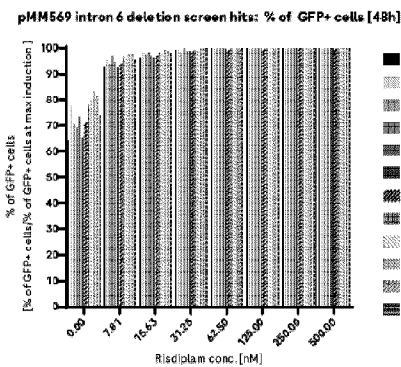
C



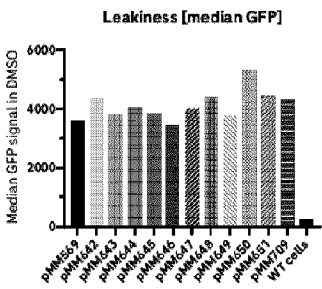
D



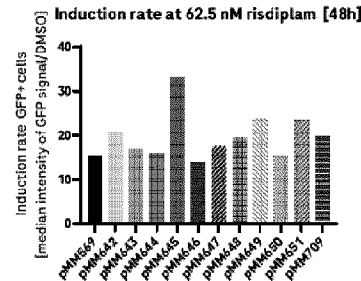
E



F



G



H

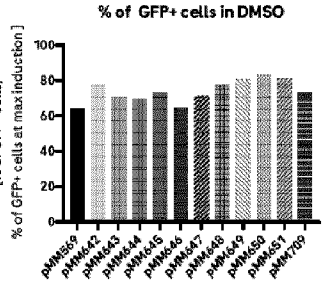


Figure 33

A

Improved split ATG switch [pMM569]: intron 7 deletion screen, 5 nt non-overlapping deletions



Intron 7 [277 nt]:



Example of Intron 7 [264 nt] 5 nt deletions:

WT: GTAAAGTCGCCACATATGAAAGTGAATCTTACTTTGTAAACCTTTATGGTTGTGGG (...)
Del1: -----TCTGCCACATATGAAAGTGAATCTTACTTTGTAAACCTTTATGGTTGTGGG (...)
Del2: GTAAAGTCGCCACATATGAAAGTGAATCTTACTTTGTAAACCTTTATGGTTGTGGG (...)
Del3: GTAAAGTCGCCACATATGAAAGTGAATCTTACTTTGTAAACCTTTATGGTTGTGGG (...)

B

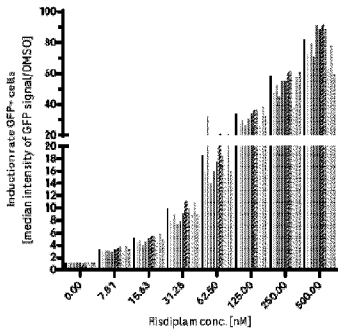
Selected intron 7 deletions:

pMM569: [...] TGTGGAAACAAATGTTTGAACATTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAG [...]
pMM634: [...] TGTGGG-----AATGTTTGAACATTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAG [...]
pMM711: [...] TGTGGAAACAAATGT-----AACATTAAAAAGTTCAGATGTTAGAAAGTTACTAGTAG [...]
pMM635: [...] TGTGGAAACAAATGTTTGAACATTAAAAAGTT-----GTTAGAAAGTTACTAGTAG [...]
pMM636: [...] TGTGGAAACAAATGTTTGAACATTAAAAAGTTCAGATGTTAG-----TACTAGTAG [...]

pMM569: [...] ATATGGGATAACCTAGGCATACTGCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAACTGGAATT [...]
pMM637: [...] ATA-----TAACCTAGGCATACTGCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAACTGGAATT [...]
pMM638: [...] ATATGGG-----TAGGCATACTGCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAACTGGAATT [...]
pMM639: [...] ATATGGGATAACCTAGGC-----GCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAACTGGAATT [...]
pMM640: [...] ATATGGGATAACCTAGGCATACTGCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAACTGGAATT [...]
pMM641: [...] ATATGGGATAACCTAGGCATACTGCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTTTAACTGGTGTCACAGAGGACATGGTTTAA-----ATT [...]
pMM712: [...] ATATGGGATAACCTAGGCATACTGCACGTACACTCTGACATATGAAGTGCTCTAGTCAAGTT-----GGTGTCACAGAGGACATGGTTTAACTGGAATT [...]

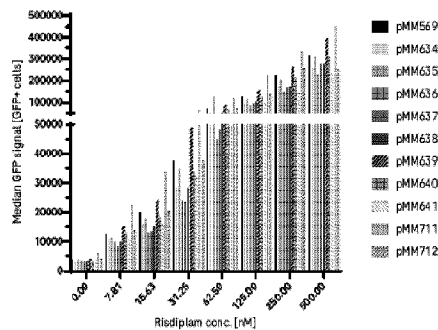
C

pMM569 intron 7 deletion screen hits: induction rate [48h]



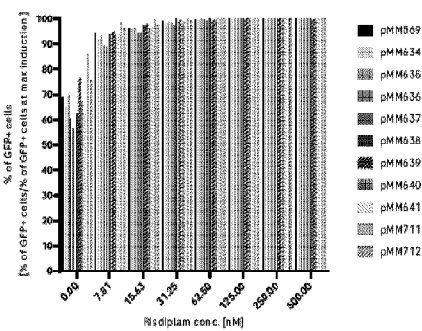
D

pMM569 intron 7 deletion screen hits: median GFP signal [48h]



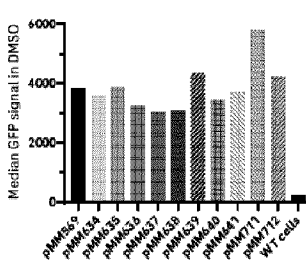
E

pMM569 intron 7 deletion screen hits: % of GFP+ cells [48h]



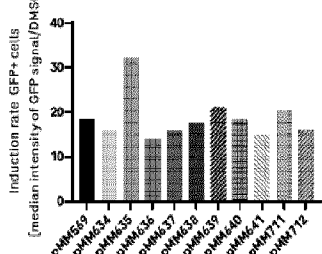
F

Leakiness [median GFP]



G

Induction rate at 62.5 nM risdiplam [48h]



H

% of GFP+ cells in DMSO

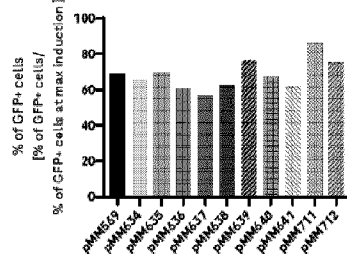
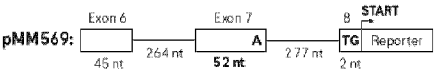


Figure 34

A

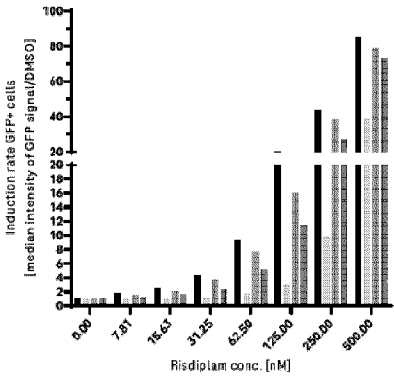
Improved split ATG switch [pMM569]: mutations reducing leakiness in DMSO



pMM569 exon 7 mutations to reduce leakiness:
pMM569: ATTTTCCTTCCAG | GATTTTAGACAAAATCAAAAAGAAGGAGGTGCTCACAATTCCTTAAATAGGA | GTAAGTCTGCCAGC
pMM579: ATTTTCCTTCCAG | GATTTTAGACAAAATC-----GGTGCTCACAATTCCTTAAATAGGA | GTAAGTCTGCCAGC
pMM595: ATTTTCCTTCCAG | GATTAATTAATAAATCAAAAAGAAGGAGGTGCTCACAATTCCTTAAATAGGA | GTAAGTCTGCCAGC
pMM596: ATTT-----TCCAG | GATTTTAGACAAAATCAAAAAGAAGGAGGTGCTCACAATTCCTTAAATAGGA | GTAAGTCTGCCAGC

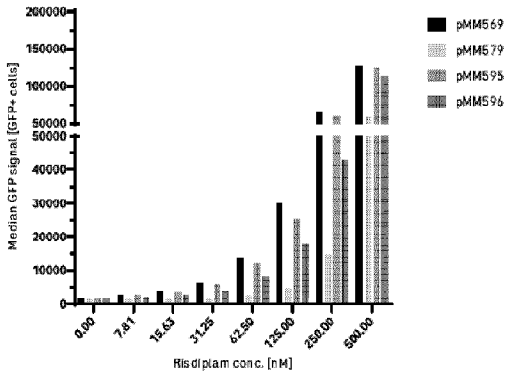
B

Low leakiness pMM569 modifications: induction rate [48h]



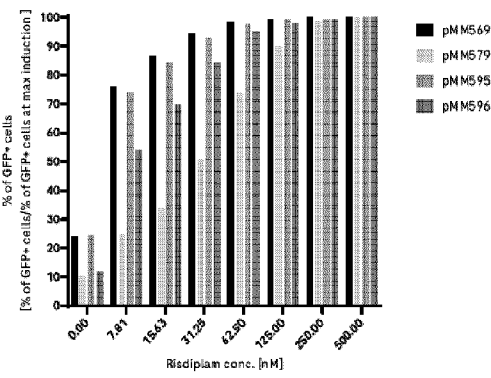
C

Low leakiness pMM569 modifications: median GFP signal [48h]



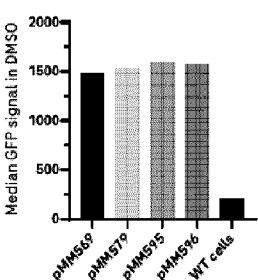
D

Low leakiness pMM569 modifications: % of GFP+ cells [48h]



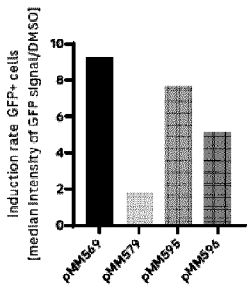
E

Leakiness [median GFP]



F

Induction rate at 62.5 nM risdiplam [48h]



G

% of GFP+ cells in DMSO

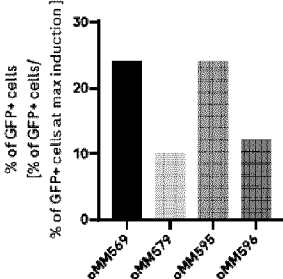
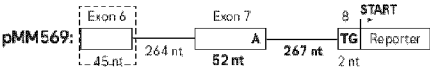


Figure 35

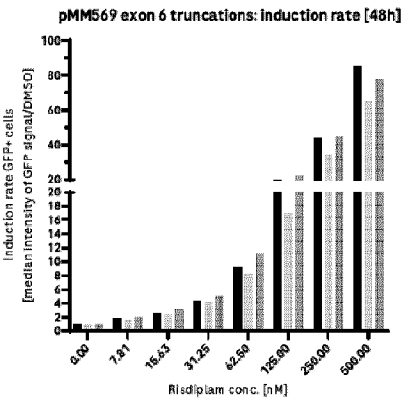
A

Improved split ATG switch [pMM569]: exon 6 truncations

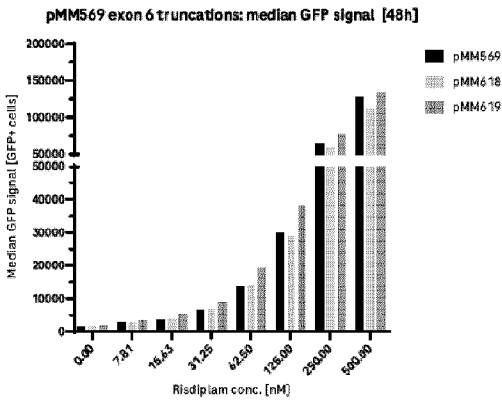


pMM569 exon 6 truncations:
pMM569: GCAATTTCTGGTACGGCAGTGGCTATCATACTGGCTATTATCAG|GT
pMM518: -----GGCAGTGGCTATCATACTGGCTATTATCAG|GT
pMM519: -----ACTGGCTATTATCAG|GT

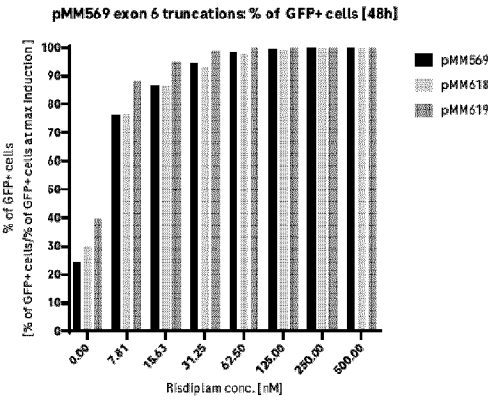
B



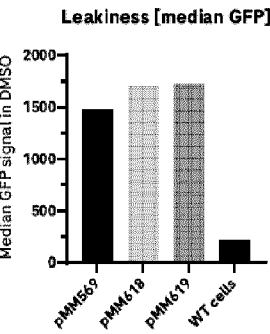
C



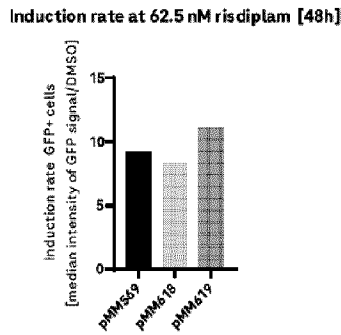
D



E



F



G

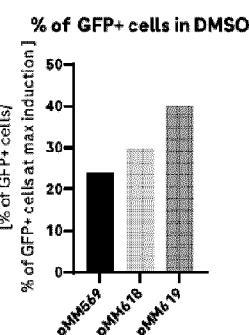
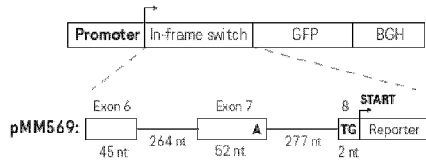


Figure 36

A

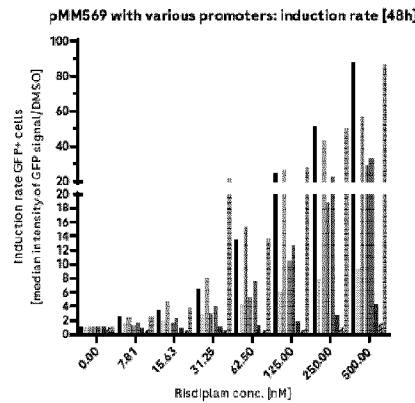
Improved split ATG switch [pMM569]: testing various promoters



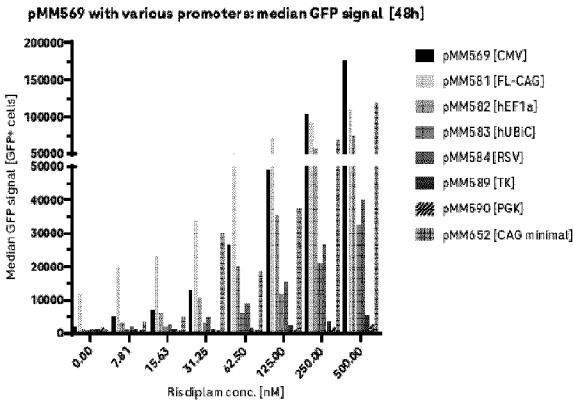
Promoters tested:

pMM569 [CMV]: 508 nt
pMM581 [FL-CAG]: 1624 nt
pMM582 [hEF1a]: 1182 nt
pMM583 [hUBiC]: 1212 nt
pMM584 [RSV]: 261 nt
pMM589 [TK]: 753 nt
pMM590 [PGK]: 517 nt
pMM652 [CAG minimal]: 580 nt

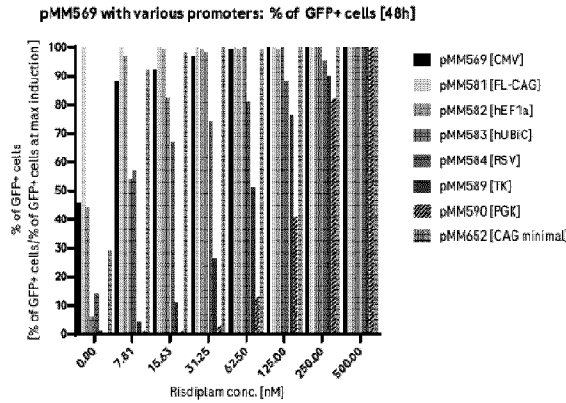
B



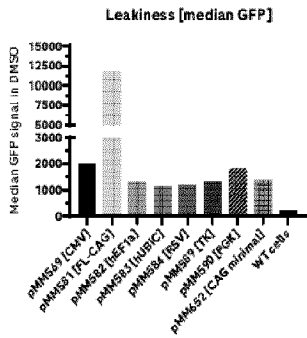
C



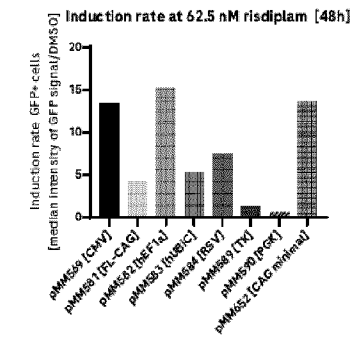
D



E



F



G

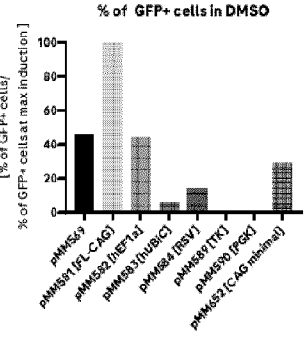


Figure 37

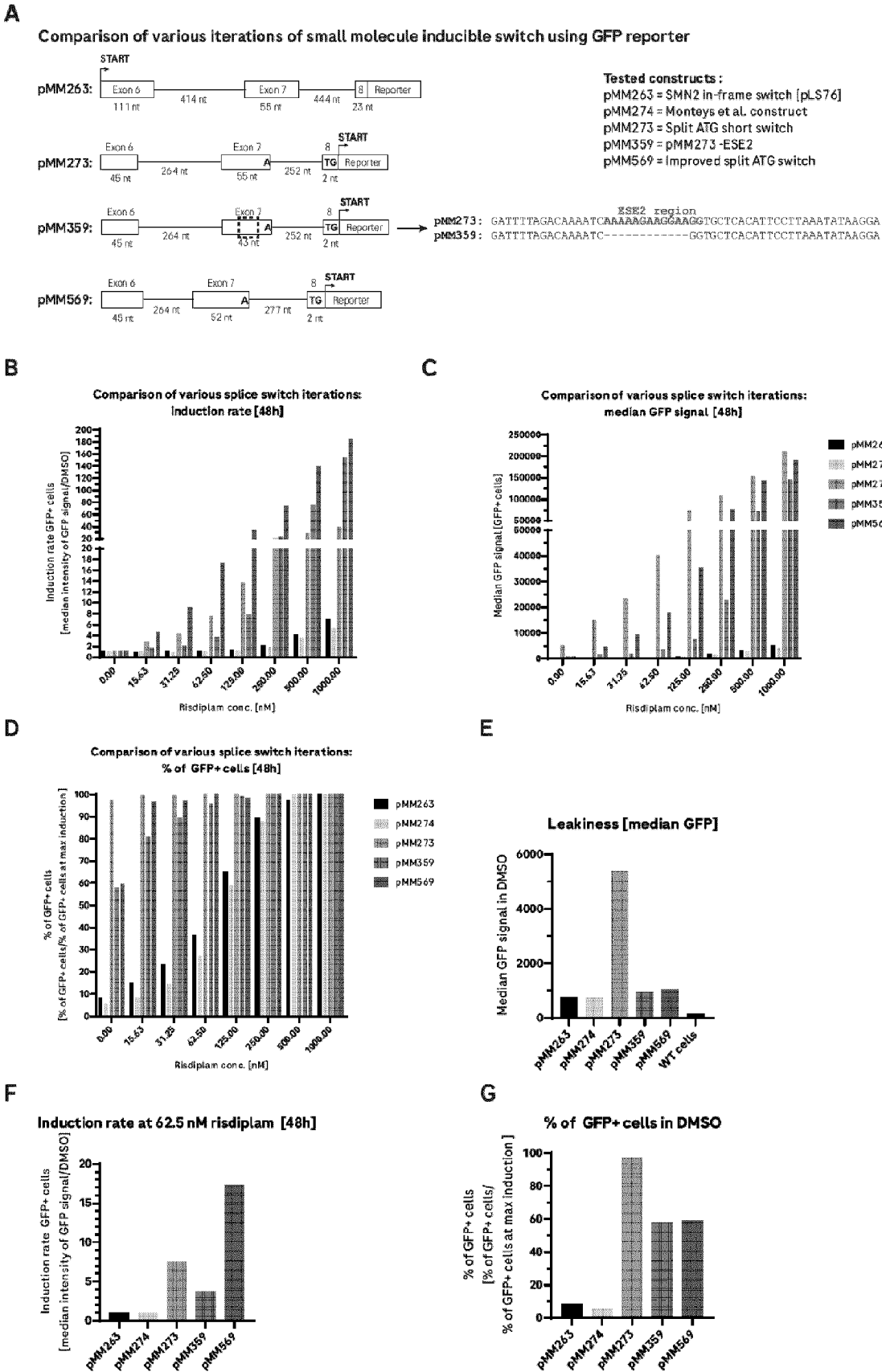
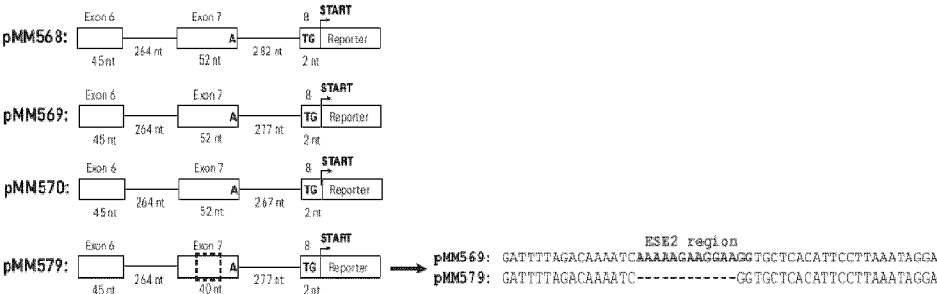


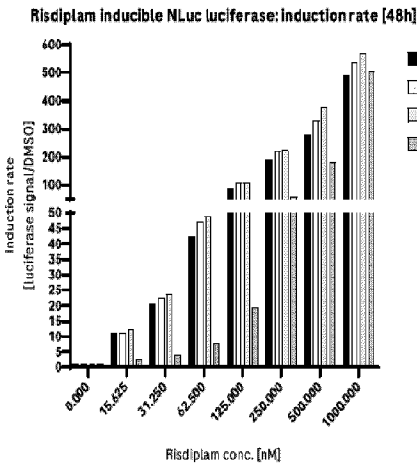
Figure 38

A

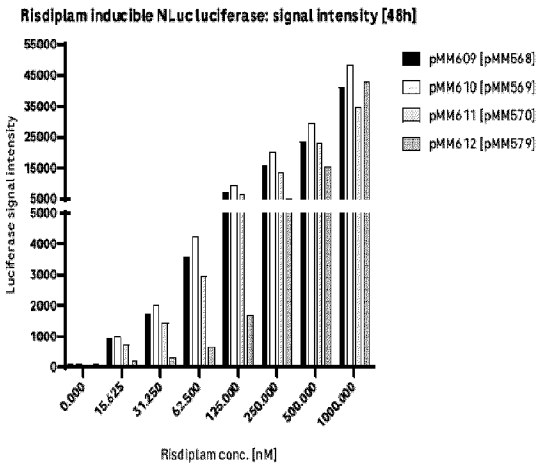
NanoLuc (NLuc) luciferase reporter induction using small molecule splice switch



B



C



D

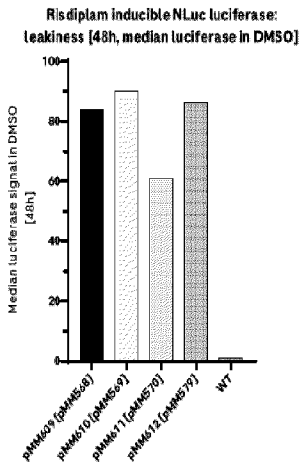
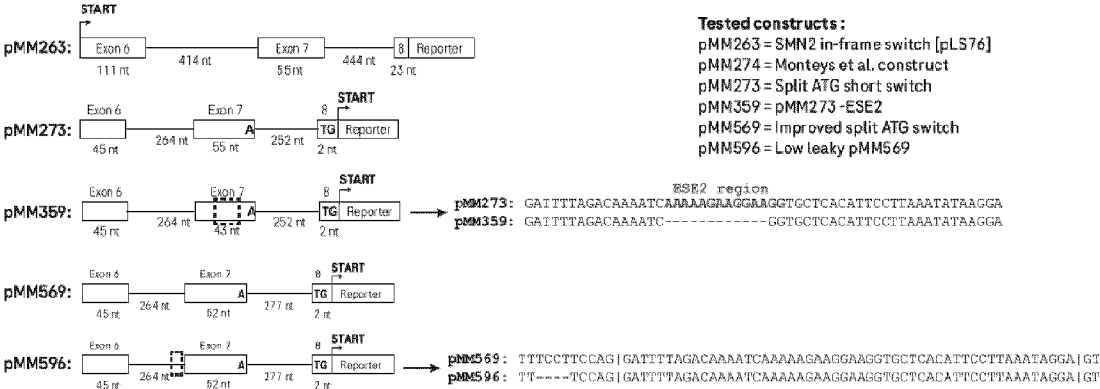


Figure 39

A

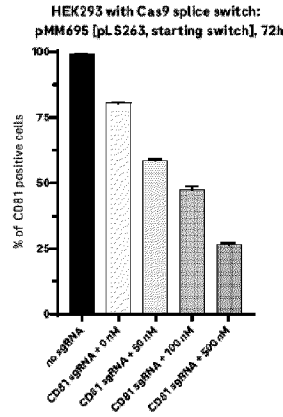
Using small molecule inducible switch to control Cas9 expression



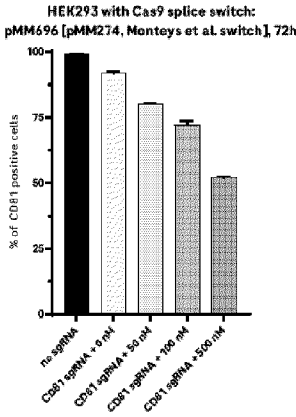
B

HEK293T cells + Inducible Cas9 (PiggyBac) + Nucleofection with 3 x 150 nM CD81 gRNA + Risdipiam $\xrightarrow{72h}$ CD81 FACS

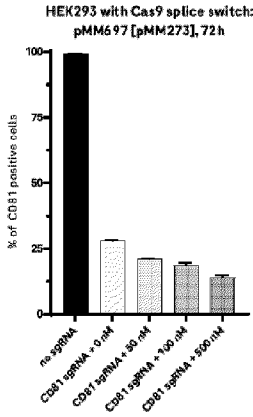
C



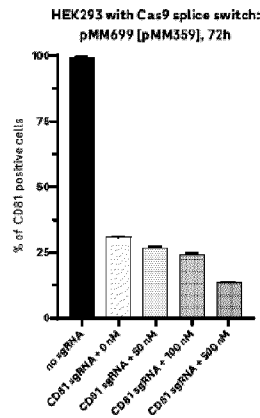
D



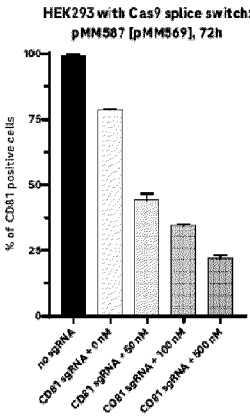
E



F



G



H

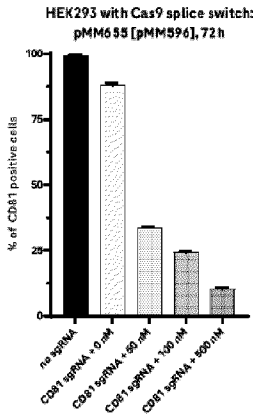
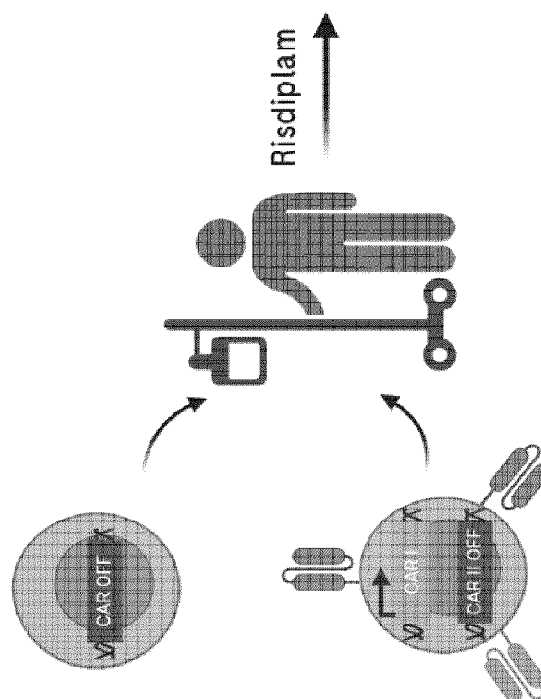


Figure 40

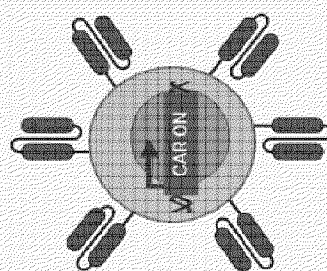
4

Using small molecule inducible switch to control chimeric antigen receptors (CAR) expression in CAR-T therapy

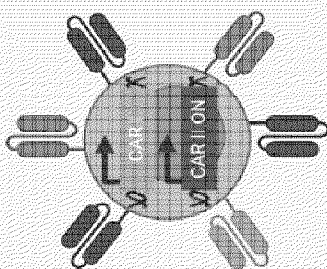
Modified T cells



CAR expressed

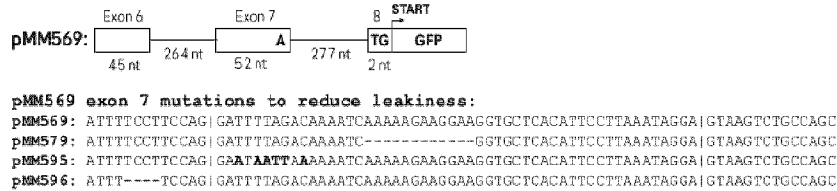


Second CAR expressed



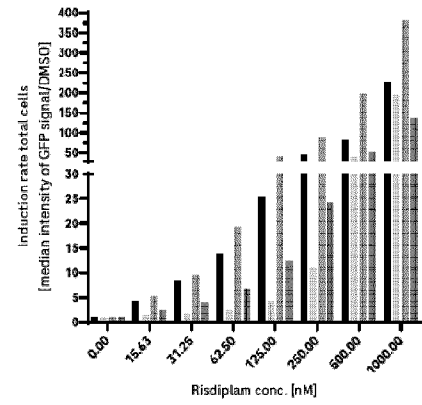
A

Small molecule inducible switch validation in a T cell model (Jurkat leukemic T-cell immortalized cell line)



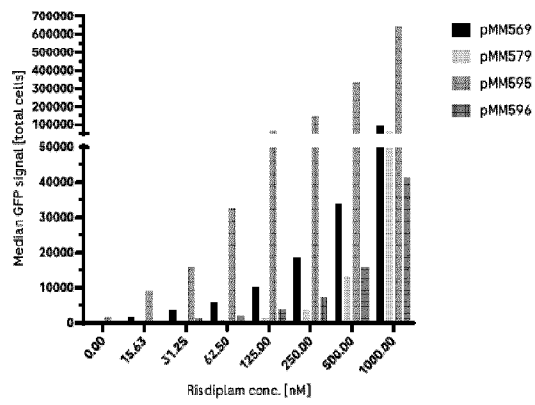
B

Small molecule inducible switch in Jurkat cells: induction rate [48h]



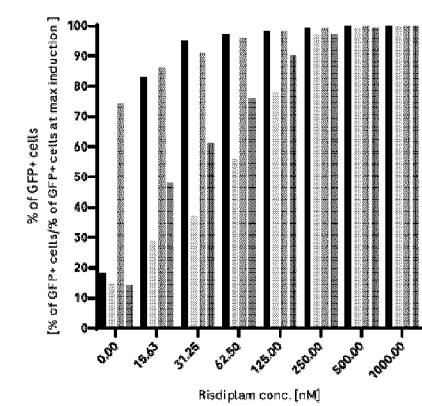
C

Small molecule inducible switch in Jurkat cells: median GFP signal [48h]



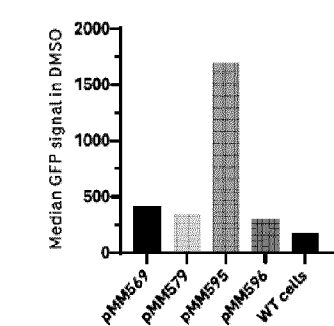
D

Small molecule inducible switch in Jurkat cells: % of GFP+ cells [48h]



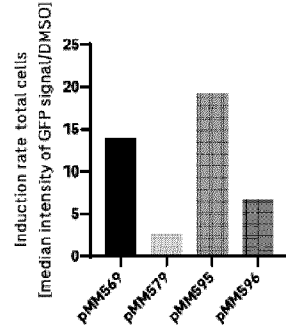
E

Leakiness [median GFP]



F

Induction rate at 62.5 nM risdipiam [48h]



G

% of GFP+ cells in DMSO

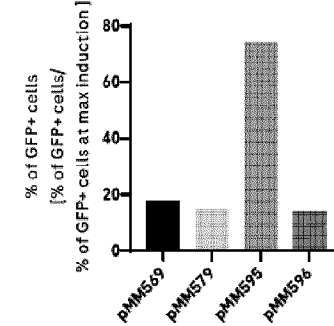


Figure 42

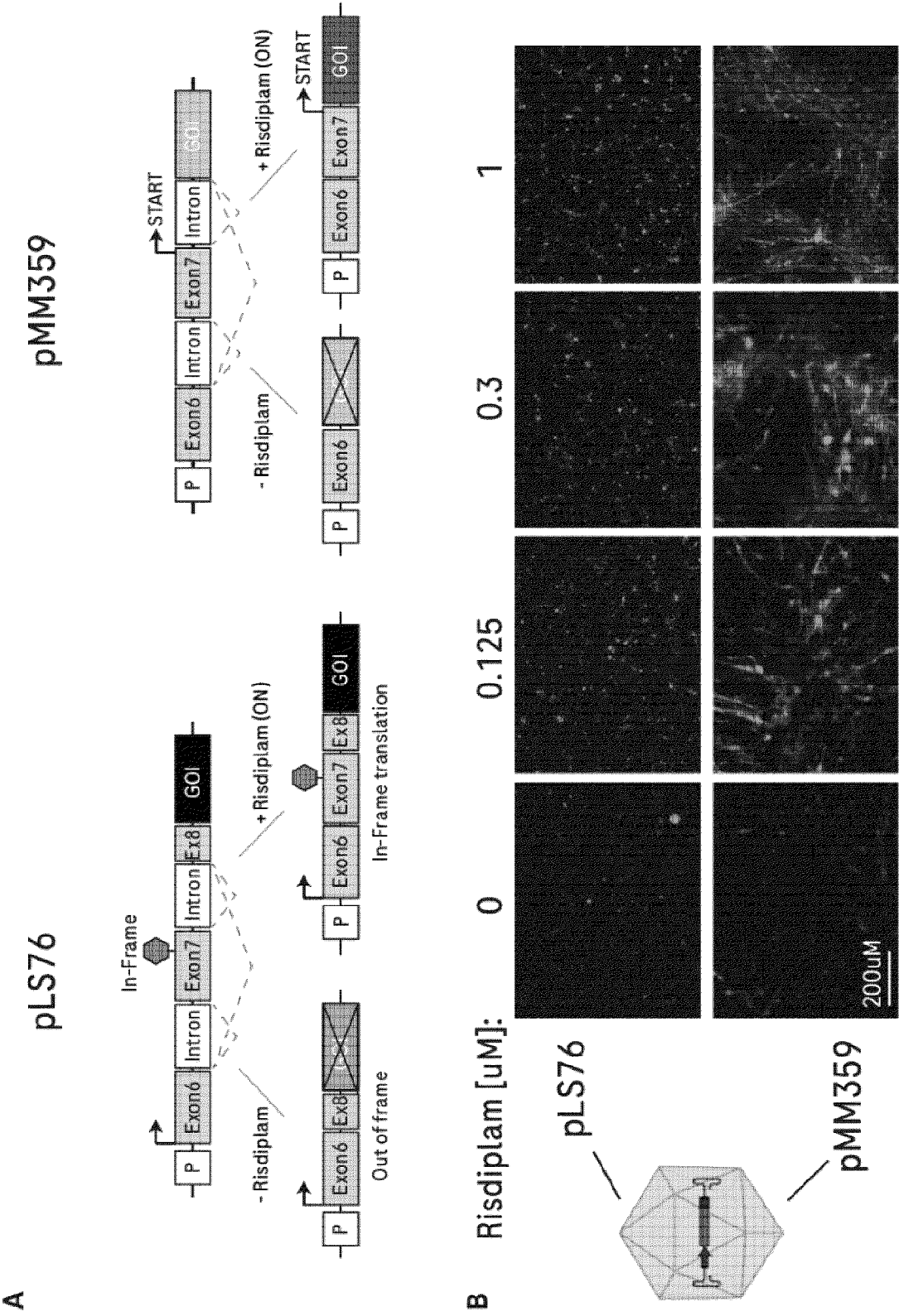


Figure 43

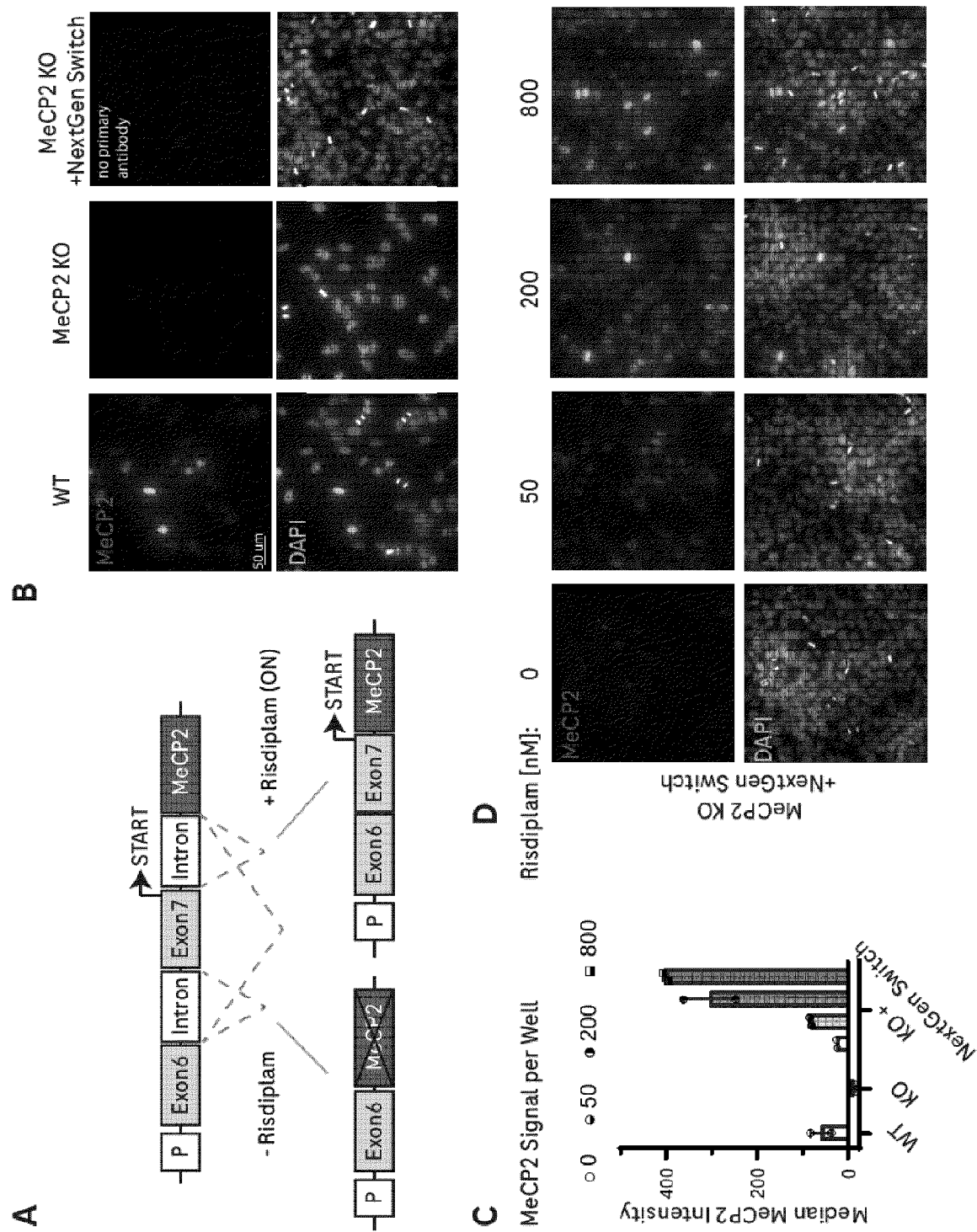


Figure 44

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2024/065689

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N15/63 C12N15/113 A61K48/00 C12N15/11 C12N15/86
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, FSTA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2017 122097 A (PTC THERAPEUTICS INC; HOFFMANN LA ROCHE) 13 July 2017 (2017-07-13) paragraphs [0052], [1798]; figure 2a; sequence 21 -----	1-54, 56-61
X	JP 2015 504057 A (PTC THERAPEUTICS INC - HOFFMANN LA ROCHE) 5 February 2015 (2015-02-05) paragraphs [0050], [1199]; figure 2a; sequence 21 -----	1-54, 56-61
X	JP 2015 509956 A (PTC THERAPEUTICS INC - HOFFMANN LA ROCHE) 2 April 2015 (2015-04-02) paragraphs [0052], [1019]; figure 2a; sequence 21 ----- -/-	1-54, 56-61

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

21 August 2024

Date of mailing of the international search report

23/09/2024

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Rodríguez Estévez, M

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2024/065689

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2013/112788 A1 (PTC THERAPEUTICS INC [US]) 1 August 2013 (2013-08-01) paragraphs [0051], [00685]; figure 2a; sequence 21 -----	1-54, 56-61
X	JP 2015 517988 A (PTC THERAPEUTICS INC) 25 June 2015 (2015-06-25) paragraphs [0052], [1346]; sequence 21 -----	1-54, 56-61
X	US 8 986 935 B2 (PAUSHKIN SERGEY V [US]; NARYSHKIN NIKOLAI A [US] ET AL.) 24 March 2015 (2015-03-24) column 35, lines 39-53 column 96, lines 24-31; figure 1; sequence 6 -----	1-54, 56-61
A	WO 2021/163556 A1 (CHILDRENS HOSPITAL PHILADELPHIA [US]) 19 August 2021 (2021-08-19) the whole document -----	55
A	HASANE RATNI ET AL: "Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA)", JOURNAL OF MEDICINAL CHEMISTRY, [Online] vol. 61, no. 15, 25 July 2018 (2018-07-25) , pages 6501-6517, XP055610353, US ISSN: 0022-2623, DOI: 10.1021/acs.jmedchem.8b00741 Retrieved from the Internet: URL:https://pubs.acs.org/doi/10.1021/acs.jmedchem.8b00741> [retrieved on 2024-08-20] the whole document -----	1-61
A	MONTEYS ALEX MAS ET AL: "Regulated control of gene therapies by drug-induced splicing", NATURE,, [Online] vol. 596, no. 7871, 28 July 2021 (2021-07-28), pages 291-295, XP037535482, DOI: 10.1038/s41586-021-03770-2 Retrieved from the Internet: URL:https://www.nature.com/articles/s41586-021-03770-2> [retrieved on 2024-08-20] the whole document ----- -/-	1-61

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2024/065689

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KLETZL HEIDEMARIE ET AL: "The oral splicing modifier RG7800 increases full length survival of motor neuron 2mRNA and survival of motor neuron protein: Results from trials in healthy adults and patients with spinal muscular atrophy", NEUROMUSCULAR DISORDERS, ELSEVIER LTD, GB, [Online] vol. 29, no. 1, 30 October 2018 (2018-10-30), pages 21-29, XP085595139, ISSN: 0960-8966, DOI: 10.1016/J.NMD.2018.10.001 Retrieved from the Internet: URL:https://www.sciencedirect.com/science/article/pii/S096089661830539X?via%3Dihub> [retrieved on 2024-08-20] the whole document -----	1-61
A	ALICIA J. ANGELBELLO ET AL: "Small molecule targeting of RNA structures in neurological disorders", ANNALS OF THE NEW YORK ACADEMY OF SCIENCES, [Online] vol. 1471, no. 1, 9 April 2019 (2019-04-09), pages 57-71, XP055767614, US ISSN: 0077-8923, DOI: 10.1111/nyas.14051 Retrieved from the Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6785366/> [retrieved on 2024-08-20] the whole document -----	1-61
A	PALACIOS DANIEL S.: "Drug Hunting at the Nexus of Medicinal Chemistry and Chemical Biology and the Discovery of Novel Therapeutic Modalities", JOURNAL OF MEDICINAL CHEMISTRY, [Online] vol. 65, no. 20, 7 October 2022 (2022-10-07), pages 13594-13613, XP093064426, US ISSN: 0022-2623, DOI: 10.1021/acs.jmedchem.2c01491 Retrieved from the Internet: URL:https://pubs.acs.org/doi/pdf/10.1021/acs.jmedchem.2c01491> [retrieved on 2024-08-20] the whole document ----- -/-	1-61

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2024/065689

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2022/204471 A1 (PTC THERAPEUTICS INC [US]) 29 September 2022 (2022-09-29) cited in the application the whole document -----	1-61
A	WO 2015/173181 A1 (HOFFMANN LA ROCHE [CH]; HOFFMANN LA ROCHE [US] ET AL.) 19 November 2015 (2015-11-19) cited in the application the whole document -----	1-61
A	MANASWINI SIVARAMAKRISHNAN ET AL: "Binding to SMN2 pre-mRNA-protein complex elicits specificity for small molecule splicing modifiers", NATURE COMMUNICATIONS, [Online] vol. 8, no. 1, 14 November 2017 (2017-11-14), XP055520527, DOI: 10.1038/s41467-017-01559-4 Retrieved from the Internet: URL:https://www.nature.com/articles/s41467-017-01559-4> [retrieved on 2024-08-20] cited in the application the whole document -----	1-61
A	SCHNEIDER-POETSCH TILMAN ET AL: "Splicing modulators: on the way from nature to clinic", THE JOURNAL OF ANTIBIOTICS, [Online] vol. 74, no. 10, 3 August 2021 (2021-08-03), pages 603-616, XP093196285, UK / JP ISSN: 0021-8820, DOI: 10.1038/s41429-021-00450-1 Retrieved from the Internet: URL:https://www.nature.com/articles/s41429-021-00450-1> [retrieved on 2024-08-20] cited in the application the whole document ----- -/-	1-61

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2024/065689

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	RAMDAS SITHARA ET AL: "New treatments in spinal muscular atrophy: an overview of currently available data", EXPERT OPIN PHARMACOTHER, [Online] vol. 21, no. 3, 24 January 2020 (2020-01-24), pages 307-315, XP055891962, London, UK ISSN: 1465-6566, DOI: 10.1080/14656566.2019.1704732 Retrieved from the Internet: URL:https://www.tandfonline.com/doi/pdf/10.1080/14656566.2019.1704732?needAccess=true [retrieved on 2024-08-20] the whole document -----	1-61

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2024/065689

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed.
 - b. ☒ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*:1(a)).
☒ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2024/065689

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 2017122097 A	13-07-2017	AR 092794 A1	06-05-2015
		AU 2013216870 A1	28-08-2014
		AU 2017204248 A1	13-07-2017
		BR 112014019750 A2	20-06-2017
		CA 2863874 A1	15-08-2013
		CL 2014002100 A1	23-10-2015
		CN 104349777 A	11-02-2015
		CN 108299314 A	20-07-2018
		CO 7061082 A2	19-09-2014
		CR 20140376 A	23-01-2015
		DK 2812004 T3	15-10-2018
		EA 201491505 A1	30-01-2015
		EA 201792465 A1	31-08-2018
		EC SP14017269 A	30-09-2015
		EP 2812004 A2	17-12-2014
		EP 3406252 A1	28-11-2018
		ES 2697174 T3	22-01-2019
		HK 1202077 A1	18-09-2015
		HU E039779 T2	28-02-2019
		IL 233959 A	30-11-2017
		JP 6092897 B2	08-03-2017
		JP 6363744 B2	25-07-2018
		JP 2015508075 A	16-03-2015
		JP 2017122097 A	13-07-2017
		KR 20140121482 A	15-10-2014
		KR 20200093066 A	04-08-2020
		MA 35920 B1	01-12-2014
		MX 354074 B	12-02-2018
		NZ 628186 A	31-03-2016
		PE 20142364 A1	10-01-2015
		PH 12014501786 A1	10-11-2014
		PH 12018501711 A1	23-09-2019
		PL 2812004 T3	31-01-2019
		SG 10201609188W A	29-12-2016
		SG 11201404713P A	26-09-2014
		TR 201813877 T4	21-11-2018
		TW 201336842 A	16-09-2013
		TW 201722942 A	01-07-2017
		UA 116981 C2	11-06-2018
		US 2015005289 A1	01-01-2015
		US 2017129885 A1	11-05-2017
		US 2018105526 A1	19-04-2018
		US 2019375750 A1	12-12-2019
		US 2021276999 A1	09-09-2021
		US 2024067646 A1	29-02-2024
		WO 2013119916 A2	15-08-2013
JP 2015504057 A	05-02-2015	BR 112014016287 A2	04-07-2017
		CA 2861609 A1	04-07-2013
		CN 104244944 A	24-12-2014
		EA 201491296 A1	30-01-2015
		EP 2797592 A1	05-11-2014
		HK 1202059 A1	18-09-2015
		JP 6193881 B2	06-09-2017
		JP 2015504057 A	05-02-2015
		KR 20140128310 A	05-11-2014
		MX 352861 B	13-12-2017
		US RE47689 E	05-11-2019

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2024/065689

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		US 2015119380 A1	30-04-2015
		WO 2013101974 A1	04-07-2013

JP 2015509956 A	02-04-2015	BR 112014021531 A2	20-06-2017
		CA 2865957 A1	06-09-2013
		CN 104302181 A	21-01-2015
		EA 201491617 A1	30-01-2015
		EP 2819519 A1	07-01-2015
		HK 1200056 A1	31-07-2015
		JP 6092264 B2	08-03-2017
		JP 2015509956 A	02-04-2015
		KR 20140131384 A	12-11-2014
		MX 352962 B	15-12-2017
		US 2015166575 A1	18-06-2015
		WO 2013130689 A1	06-09-2013

WO 2013112788 A1	01-08-2013	BR 112014018027 A2	20-06-2017
		CA 2862084 A1	01-08-2013
		CN 104203239 A	10-12-2014
		EA 201491412 A1	30-12-2014
		EP 2809322 A1	10-12-2014
		ES 2733644 T3	02-12-2019
		HK 1202069 A1	18-09-2015
		JP 6193888 B2	06-09-2017
		JP 2015511224 A	16-04-2015
		JP 2017171667 A	28-09-2017
		KR 20140127271 A	03-11-2014
		MX 357834 B	26-07-2018
		PL 2809322 T3	29-11-2019
		US 2015126515 A1	07-05-2015
		WO 2013112788 A1	01-08-2013

JP 2015517988 A	25-06-2015	BR 112014023483 A2	20-06-2017
		CA 2868026 A1	26-09-2013
		CN 104470909 A	25-03-2015
		EA 201491766 A1	31-03-2015
		EP 2828247 A1	28-01-2015
		HK 1202527 A1	02-10-2015
		JP 6219921 B2	25-10-2017
		JP 2015517988 A	25-06-2015
		KR 20140135840 A	26-11-2014
		MX 358514 B	24-08-2018
		US 2015080383 A1	19-03-2015
		WO 2013142236 A1	26-09-2013

US 8986935 B2	24-03-2015	US 2011172284 A1	14-07-2011
		US 2011213005 A1	01-09-2011
		US 2015283115 A1	08-10-2015
		US 2015344877 A1	03-12-2015
		WO 2010019236 A1	18-02-2010
		WO 2010019243 A1	18-02-2010

WO 2021163556 A1	19-08-2021	AU 2021218807 A1	15-09-2022
		CA 3167836 A1	19-08-2021
		CN 115605596 A	13-01-2023
		EP 4103717 A1	21-12-2022
		IL 295542 A	01-10-2022
		JP 2023514242 A	05-04-2023

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2024/065689

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		KR 20220146501 A	01-11-2022
		US 2023340498 A1	26-10-2023
		WO 2021163556 A1	19-08-2021

WO 2022204471	A1	29-09-2022	NONE

WO 2015173181	A1	19-11-2015	AR 100442 A1
		AU 2015261046 A1	05-10-2016
		BR 112016026205 A2	27-10-2016
		CA 2948561 A1	15-08-2017
		CL 2016002836 A1	19-11-2015
		CN 106459092 A	21-04-2017
		CR 20160518 A	22-02-2017
		DK 3143025 T3	21-02-2017
		EA 201692280 A1	09-12-2019
		EA 202090486 A2	31-07-2018
		EP 3143025 A1	30-06-2020
		EP 3663296 A1	22-03-2017
		EP 4241772 A2	10-06-2020
		ES 2761423 T3	13-09-2023
		ES 2949660 T3	19-05-2020
		FR 21C1039 I1	02-10-2023
		HR P20192159 T1	15-10-2021
		HR P20230637 T1	21-02-2020
		HU E046491 T2	29-09-2023
		HU S2100037 I1	30-03-2020
		JP 6236173 B2	28-10-2021
		JP 2017515863 A	22-11-2017
		KR 20170003687 A	15-06-2017
		KR 20210014219 A	09-01-2017
		LT PA2021010 I1	08-02-2021
		MA 39995 B1	11-10-2021
		MA 51988 A	29-11-2019
		MX 371050 B	20-01-2021
		MY 174284 A	14-01-2020
		NL 301128 I1	01-04-2020
		NO 2021035 I1	22-09-2021
		NZ 725008 A	10-09-2021
		PE 20170128 A1	29-11-2019
		PH 12016502081 A1	16-03-2017
		PL 3143025 T3	09-01-2017
		PL 3663296 T3	31-03-2020
		PT 3143025 T	07-08-2023
		SG 11201609497T A	03-12-2019
		SI 3143025 T1	29-12-2016
		SI 3663296 T1	31-01-2020
		TW 201609738 A	31-08-2023
		UA 119670 C2	16-03-2016
		US 2017197990 A1	25-07-2019
		WO 2015173181 A1	13-07-2017
		ZA 201607026 B	19-11-2015
			25-05-2022
