

DNA CONSTRUCT AND NUCLEIC ACID COMPOUND TOOLS

5

FIELD OF THE INVENTION

[0001] The present invention relates generally to the field of nucleic acid therapeutics. In particular, the invention relates to a system for characterizing the functional cellular uptake of a nucleic acid.

10

BACKGROUND OF THE INVENTION

[0002] The delivery of nucleic acid therapeutics (NATs) can be quantified *in vitro* and *in vivo* using labelled or unlabelled methods.

[0003] A typical labelled method for quantifying NATs is by conjugating NATs to fluorophores. The fluorescence is then measured in the cells exposed to the NAT. As the fluorophore may interfere with the on-target activity of the NAT, a labelled method cannot discern whether the uptake of the fluorophore-tagged NAT is functional or productive in the cell. Thus, such a method leads to a huge false-positive rate. In addition, the efficiency of cellular uptake cannot be quantitatively compared between experiments and among cell types/tissues, given that there is no baseline fluorescence level to normalize the measured fluorescence.

20

[0004] A conventional unlabelled method is to quantify the uptake of an NAT (where the NAT is targeting the mRNA of a known gene) by measuring the mRNA abundance. However, this method is low-throughput as it requires tedious RNA extraction and amplification steps and does not provide single-cell resolution.

25

[0005] There is thus a need for a system for characterizing functional cellular uptake of nucleic acids that overcome the drawbacks of the prior art. Furthermore, other desirable features and characteristics will become apparent from the subsequent detailed description and the appended claims, taken in conjunction with the accompanying drawings and this background of the disclosure.

30

SUMMARY OF THE INVENTION

[0006] In one aspect of the present invention, there is provided a DNA construct comprising: (a) a starter exon encoding a start codon; (b) a first frame-switch exon located downstream of the starter exon, the first frame-switch exon comprising a first acceptor splice site, wherein inclusion or exclusion of the first frame-switch exon during alternative splicing

35

determines the downstream reading frame of the mature mRNA transcript; and (c) a first specific exon located downstream of the first frame-switch exon, the first specific exon comprising a second acceptor splice site, a first reporter gene and a second reporter gene, wherein the first reporter gene is upstream or downstream of the second reporter gene and
5 wherein the first reporter gene is in a different reading frame from the second reporter gene, wherein the exclusion of the first frame-switch exon causes the first reporter gene to be in-frame with the start codon and wherein the inclusion of the first frame-switch exon causes the second reporter gene to be in-frame with the start codon.

[0007] In one embodiment, the DNA construct as described herein further comprises a
10 second specific exon located downstream of the first specific exon.

[0008] In one embodiment, the DNA construct as described herein further comprises a second frame-switch exon located downstream of the first frame-switch exon and upstream of the first specific exon, the second frame-switch exon comprising a third acceptor splice site, wherein the second frame-switch exon is in a different reading frame from the first frame-switch exon and wherein the combination of inclusion or exclusion of the first frame-switch
15 exon and the second frame-switch exon during alternative splicing determines the downstream reading frame of the mature mRNA transcript.

[0009] In one embodiment, the DNA construct as described herein further comprises a third reporter gene.

20 [0010] In one embodiment, the third reporter gene is in the same reading frame as the first or second reporter genes.

[0011] In one embodiment, the third reporter gene is in a different reading frame from the first and second reporter genes.

[0012] In one embodiment, the DNA construct as described herein further comprises a
25 fourth reporter gene.

[0013] In one embodiment, the DNA construct as described herein comprises the sequence of SEQ ID NO:67.

[0014] In another aspect, there is provided a method of measuring the efficacy of a nucleic acid delivery into a cell by a delivery agent, the method comprising: (a) using the delivery
30 agent to introduce the nucleic acid into the cell, wherein the cell expresses the DNA construct as described herein and wherein the nucleic acid is configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct; and (b) quantifying the functional uptake of the nucleic acid into the cell by measuring the relative expression of two of the reporter genes, wherein the functional uptake

of the nucleic acid is indicative of the efficacy of delivery of the nucleic acid by the delivery agent.

[0015] In one embodiment, step (b) of the method as described herein comprises quantifying the functional uptake of the nucleic acid into the cell by measuring the relative expression of the first reporter gene and the second reporter gene.

[0016] In one embodiment, the nucleic acid is selected from the group consisting of a steric blocker antisense oligonucleotide, a RNaseH-activating GAPmer, an siRNA, a CRISPR/Cas-sgRNA and a snRNA.

[0017] In one embodiment, the steric blocker antisense oligonucleotide is a splice-switching oligonucleotide (SSO).

[0018] In one embodiment, the delivery agent is selected from the group consisting of an intra-molecular chemical modification, an inter-molecular delivery agent, an encapsulation delivery agent, a complexation delivery agent, a viral-based delivery agent, a cellular delivery agent and a sub-cellular delivery agent.

[0019] In one embodiment, the expression of the reporter genes is measured by fluorescence, luminescence, immunohistochemistry or quantitative PCR.

[0020] In another aspect, there is provided a method for measuring the effect of chemical modification on the potency or activity of a nucleic acid, the method comprising: (a) introducing a chemically-modified nucleic acid into a cell, wherein the cell expresses the DNA construct as described herein and wherein the chemically-modified nucleic acid is configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct; and (b) quantifying the potency or activity of the chemically-modified nucleic acid by measuring the relative expression of two of the reporter genes.

[0021] In one embodiment, step (b) of the method as described herein comprises quantifying the potency or activity of the chemically-modified nucleic acid by measuring the relative expression of the first reporter gene and the second reporter gene.

[0022] In another aspect, there is provided a transgenic animal that expresses the DNA construct as described herein.

[0023] In one embodiment, the animal is a rodent.

[0024] In another aspect, there is provided a system for characterizing the functional cellular uptake of a nucleic acid, the system comprising: (a) a DNA construct comprising at least two reporter genes; and (b) a nucleic acid configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct such that one of the at least two reporter genes is preferentially expressed.

[0025] In one embodiment, the DNA construct comprises the DNA construct as described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

5 [0026] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying drawings, in which:

[0027] **Fig. 1** shows an embodiment of the CALIMERO gene. **A:** CALIMERO reporter gene schematics depicted at the top. EGFP isoform (state 0 in Table 1): major transcript expressed
10 in the absence of a modulating NAT. Exons 1, FS and 3A are all spliced in with the transcript in frame for EGFP. tRFP isoform (state 1 in Table 1): exons 1 and 3A are spliced in and the resulting transcript is in frame for tRFP. It is first minor transcript expressed in the absence of a modulating NAT, and can be induced upon splice-modulation of FS by AVA1 SSO or CRISPR guide sg1. This transcript is not degraded by GAPmer1 and siRNA1/2. Luciferase
15 isoform (state 2 in Table 1): exons 1, FS and 3B are spliced in and the resulting transcript is in frame for Luciferase. It is second minor transcript expressed in the absence of a modulating NAT, and can be induced upon skipping Exon 3A with AVA2 SSO or sg2 guide RNA. HisTag isoform (state 3 in Table 1): exons 1 and 3B are spliced in and the resulting transcript is in
20 frame for HisTag, and can be induced upon splice-modulation of FS by AVA1 SSO and skipping of Exon 3A with AVA2 SSO simultaneously. **B:** TaqMan qPCR values from different cell lines transfected with the CALIMERO plasmid.

[0028] **Fig. 2** shows the compound tools for CRISPR/Cas-sgRNA modality. **A:** Top panel: Flow cytometry profile of CALIMERO positive CT26 cells untreated (State 0, left), after transfection with RNAiMAX the Alt+R system containing Cas9, tracrRNA, and AVA1-sgRNA
25 (State 1, middle), and after transduction and selection with lentiCRISPRv2 viral particles as carrier for AVA1-sgRNA (State 1, right). Middle and bottom panels: Quantification of the protein PSI (pPSI; percentage spliced in) for EGFP on single-cells (middle panel) from flow cytometry of CALIMERO-transfected CT26 cells (UTR) transduced with lentiCRISPRv2-sg1 and selected with puromycin (sg1) for 5 days (bottom panel). pPSI (EGFP) measures the
30 fraction of EGFP protein that is obtained from the mean fluorescence intensities (MFIs) of EGFP and tRFP proteins, i.e. $pPSI(EGFP) = \frac{MFI(EGFP)}{[MFI(EGFP) + MFI(tRFP)]}$. **B:** pPSI for EGFP after transfection with Alt+R system using different transfection reagents (columns 2 to 5) or, at 2 and 5 days after transduction with lentiCRISPRv2 viral particles as carrier for AVA1-sgRNA (columns 6 and 7). **C:** Left panel: Correlation between percentage of uncut genomic DNA (gDNA) and average FS pPSI of CALIMERO-positive CT26 cells transfected
35

with Cas9/tracrRNA/sg1 using different commercial transfection reagents as in Figure 2B. Non-linear regression was performed (variable slope, four parameters, least squares regression). Right panel: Correlation between pPSI (at protein level) and the tPSI (at mRNA transcript level). tPSI measures the fraction of FS-containing mature transcripts, which is derived from Taqman qPCR by calculating the abundance of the qPCR probe spanning FS to Exon 3A (PR2-3A, Fig. 3C) over the sum of all the probes (PR2-3A + PR2-3B + PR1-3A, Fig. 3C). D: Luminescence signal of two CALIMERO positive cell lines after incubation of the cells with luciferin; cells were transfected with Alt+R system with AVA2-sgRNA (labelled as sg3). E: Median Fluorescence Intensity (MFI) of EGFP and tRFP signal of B16F10 cells after AAV2-sgRNA Alt+R transfection.

[0029] Fig. 3 shows the quantitation of functional delivery among delivery agents for CRISPR/Cas-sgRNA modality. A: Flow cytometry of Calimero-positive cell line CT26 transfected with IDT Alt+R CRISPR system using multiple transfection reagents or by lentivirus particles. B: Quantification of percentage of cells of the graphs in panel A; C: Taqman PCR scheme with primers and probe locations; D: Taqman PSI (tPSI) for exon 2 (FS exon) calculated on the cDNA from the samples in panel A, refer to Figure 2C legend.

[0030] Fig. 4 shows the compound tools for steric blocker antisense oligonucleotide modality. A: Confocal high-content image of a 96-well plate for CT26-Calimero cells untreated (UTR) and transfected with AVA1-SSO. B: Single-cell plot of the relative protein PSI, pPSI(EGFP), at different AVA1a-SSO concentrations; C: Correlation between pPSI (EGFP) and taqman PSI, tPSI (FS), upon transfection of various concentrations of AVA1a-SSO at 24 or 72 hours compared to AVA1-sgRNA (CRISPR/Cas modality) from Figure 2C. D: Dose-response curves for full 2'-MOE and full 2'-OMe AVA1a-SSO; the graph shows both the switching efficiency (pPSI (EGFP), left Y-axis) and the absolute cell count (right Y-axis) upon transfection of the same SSO sequence (AVA1a) with the two different chemistries. Data was obtained in B16F10-Calimero cells at 24h after transfected with various concentrations of AVA1-SSO. E: Relative luminescence at 24h in three cell lines transfected with 100nM of AVA2-SSO (labelled as 2087). NC2 is a non-targeting SSO control. F: (Top panel) Flow cytometry profiles of CALIMERO-transfected CT26 cells. EGFP (left) and tRFP (right) fluorescence was measured after 24 and 72 hours for untreated (UTR) and 25nM of AVA1 transfected cells. (Bottom panel) pPSI of FS exon calculated for each individual cell from top panel. G: (Left panel) Untreated-normalized (UTR) luminescence of CALIMERO-expressing cell lines, B16F10, CT26 and HepG2, treated with 100nM of scramble (SCR) and AVA2 SSO for 48 hours. (Right panel) Linear correlation between Cyclophilin A-normalized TaqMan probe 2-3B, which measures abundance of Exon 3B-containing CALIMERO mature

transcripts, and luminescence. CALIMERO-expressing HeLa cells were transfected with 50nM of each of the ten AVA2 candidates separately, and 15,000 cells were collected for luciferin incubation at 24 hours post- transfection. **H:** (Left panel) *In vivo* luciferase assay of wild type (WT) and CALIMERO expressing mice, injected subcutaneously with PBS (UTR) and 25 mg/kg of AVA2 (AVA2) for 48 hours. (Right panel) Untreated-normalized luminescence of CALIMERO expressing mice injected subcutaneously with 25 mg/kg of AVA2 on day 0 and day 28 over a period of 35 days. **I:** Relative Fluorescence Units (RFU) of the peritoneal mouse area for wild-type C57BL/6 mice, and CALIMERO heterozygous and homozygous mice after luciferin injection. ROI of the same size were drawn over the shaved region of the mice. The luciferase signal was measured as average radiance (photons/s/sr/cm²). **J:** Agarose gel of CALIMERO amplicons obtained from CT26 cells transfected with either a scrambled or AVA1 SSO. Cyclophilin A housekeeper was used as a loading control. **K:** Western blot of protein lysates obtained from CALIMERO-expressing CT26 cells transfected with 100nM of a scrambled, AVA1 SSO, or AVA2 SSO, probed for FLAG-tag antibody. Cells were collected at 48h post-transfection. **L:** Flow cytometry of CALIMERO-expressing HepG2 cells incubated for 48h with 1 μ M of AVA1 (2'OMe, central panel) and in presence of 9mM calcium-enriched media (CEM, CaCl₂, right panel). **M:** Time-course of tPSI (FS), calculated from TaqMan qPCR data, from two cell lines, CT26 and HepG2, after a single 10nM transfection of AVA1 with RNAiMAX. **N:** Agarose gel and subsequent Sanger sequencing of CALIMERO amplicons obtained from CT26 cells transfected with a scrambled SSO, AVA2, or AVA1 and AVA2. Cyclophilin A housekeeper was used as a loading control. **O:** Sanger sequencing of the splicing products after transfection with AVA1, AVA2, or AVA1 and AVA2 SSOs.

[0031] **Fig. 5** shows the compound tools for siRNA modality. **A:** tRFP intensity of wild-type cells (WT) and cells transfected with the CALIMERO plasmid. A basally measurable tRFP signal (MFI) can be detected in CALIMERO-transfected cells; **B:** Flow-cytometry data of cells untreated or transfected with AVA1-siRNA1 or AVA1-siRNA2 targeting FS exon at 48h post transfection. A clear EGFP reduction is observed in the AVA1-siRNA1 or AVA1-siRNA2 treated samples, while tRFP protein level remained stable across; **C:** Histograms of EGFP and tRFP intensity of B16F10-CALIMERO cells transfected with Scramble siRNA, AVA1-siRNA1 and AVA1-siRNA2. EGFP signal is decreasing while tRFP remains constant. **D:** qPCR results from the cDNA of the cells in panel C on total CALIMERO gene expression normalized for Cyclophilin A housekeeping gene. Every dot represents a biological replicate in two different cell lines; **E:** pPSI (EGFP) calculated for the same set of experiments correlates with the qPCR data from panel D.

[0032] Fig. 6 shows the compound tools for GAPmer modality. **A:** Flow cytometry data of CT26 cell line at 48h post transfection with AVA1-EX2-GAPmer and AVA2-EX3-GAPmer. AVA1-EX2-GAPmer degraded EGFP (X axis) only whereas AVA2-EX3-GAPmer degraded both EGFP and tRFP (Y axis); **B:** Normalized EGFP Median Fluorescence Intensity and normalized cell count of CT26 cells transfected with AVA1-EX2-GAPmer, AVA2-EX3-GAPmer or Untreated (UTR) at three different concentrations. **C:** CALIMERO-expressing B16F10 cells were transfected with AVA1-EX2-GAPmer at various concentrations and analyzed through flow cytometry at 48 hours post-transfection. The three nucleotides at both 5' and 3' wings of GAPmer1 were modified with 2'OMe and all backbone are PS. **D:** Single-cell pPSI (FS) of the samples plotted in panel C. Each dot represents a single cell.

[0033] Fig. 7 shows the compound tools for snRNP modality. Normalised tRFP MFI signal after transfection of cells with the pC4 U7 AAV plasmids carrying a scrambled sequence or a sequence targeting the acceptor splice site and the ESE of CALIMERO exon 2 (FS exon).

DETAILED DESCRIPTION OF THE INVENTION

[0034] In one embodiment, the CALIMERO platform is an internally normalized and unlabelled reporter gene construct and a tool compound library for the characterization of functional cellular uptake of nucleic acid therapeutics (NATs). The terms "CALIMERO" and "NATURA (Nucleic Acids Therapeutics Unlabelled Reporter Assay)" are used interchangeably.

[0035] The CALIMERO (Conjugated Antisense oligonucleotide Library Mouse switchER mOdel) platform consists of two essential components.

[0036] (1) The CALIMERO gene – an artificial reporter gene construct whose expressed products are used as molecular readouts to detect and quantify functional cellular uptake of nucleic acids and their internal chemical modifications and/or covalently-linked conjugates or other delivery reagents, including the quantitation of the on-target efficiency of the nucleic acids as a result of the internal chemical modifications and delivery systems. The CALIMERO gene, when expressed in a cell line, organoid, or into a transgenic animal model, allows quantitative characterization of functional uptake of NATs rapidly and with single-cell sensitivity. The term "CALIMERO mouse" refers to a mouse model bearing the CALIMERO gene. The gene is designed to express 4 spliced-isoforms, each translating to a distinct protein. The 4-protein system allows, for the first-time, to:

A. Explicitly account and normalize for cell-to-cell variability in CALIMERO expression, without the need to rely on housekeeping genes that are also subjected to cell-to-cell variability. It thus makes accurate quantitative characterization a reality.

- B. Detect and characterize functional uptake via fluorescence, luminescence or immuno-histochemistry methods which provide the flexibility and throughput for investigation at the cellular, tissue and organs (animal) level.

[0037] (2) A library of compound tools that modulate the CALIMERO gene via specific NAT modalities. The compound tools are essential to use the CALIMERO gene, and thus, both the gene and the tool library are indispensable. As the CALIMERO gene is purposefully designed to be amenable for modulating the relative abundance of its spliced-isoforms, either by means of splice modulation or direct suppression of specific splicing isoforms, the platform is precisely generalized for all clinically approved and most promising NAT modalities, listed as follows.

5

10 A. Steric blocker antisense oligonucleotide – a short single-stranded RNA that is typically modified chemically at every ribose sugar and every backbone, and/or bases. They exert their on-target activity by binding to nascent and/or mature mRNA (coding and non-coding) and effect steric blockage against specific RNA-binding proteins and/or ribonucleoproteins. As such, they do not require any intracellular biological and

15 biochemical pre-processing to be functional. Potential applications include exon splicing modulation (a.k.a. SSOs or splice-switching oligonucleotides) to induce isoform-switching or transcript degradation via nonsense-mediated decay pathway, inhibition and inducing ADAR-mediated RNA editing, and inhibition of protein translation. The compound tools for this modality were designed as SSOs to modulate splicing of the CALIMERO gene.

20

B. GAPmer – a short single-stranded DNA with typically 3 to 5 RNA bases flanked on both ends. When bound to a target mRNA, the mechanistic action of a GAPmer is by activating the RNaseH enzyme, an endoribonuclease that is recruited to a DNA-RNA duplex and cleaves the mRNA. The DNA ribose sugar is not amenable for chemical

25 modification as doing so abrogates the recruitment/activation of RNaseH, resulting in the loss of a GAPmer's potency. The GAPmer compound tools were designed to target specific spliced-isoforms expressed by the CALIMERO gene.

C. siRNA – a short double-stranded RNA that requires processing by the RNAi (RNA interference) pathway whereby the antisense strand is cleaved from the double-strand

30 RNA and loaded onto the RISC complex, which thereafter guides the binding to and cleavage of its target mRNA. Both the ribose sugars and backbones, but not the bases, are amenable for modification with a limited type of chemistries and to a limited extent. As the antisense strand is the active component, it is less amenable than the sense strand for modifications. The siRNA compound tools were designed to target a specific

35 spliced-isoform expressed by the CALIMERO gene.

D. CRISPR/Cas-sgRNA – single-stranded guide RNA (sgRNA) ligated to the Cas9 protein. The gRNA compound tools were meticulously designed to disrupt specific NATURA splice sites through CRISPR-Cas9-mediated genomic introduction of indels or substitutions surrounding their consensus sequence, while avoiding potential off-targets in both mouse and human genomes.

E. snRNP-AAV – small nuclear ribonucleoproteins. snRNPs are transcribed as snRNAs (small nuclear RNAs) that undergo subsequent processing that involves binding to RNA-binding proteins to form ribonucleoproteins. snRNPs are expressed ubiquitously in every cell with many of them playing important roles in splicing and processing of nascent mRNAs. In particular, the mRNA target binding sequence of both U1 and U7 snRNAs can be engineered to guide its binding site. Engineered snRNAs are coded in DNA plasmids that are packaged in associated adenovirus (AAV) particles as delivery vehicles. While there is no therapeutics approved or in advanced clinical studies, the objective to consider this modality is to provide an expedient avenue for studying AAVs as functional delivery vehicles, bypassing the use of a reporter gene as the cargo; cargo size of engineered U1 or U7 is substantially smaller than typical reporter genes.

[0038] Based on the NAT modality of interest, the respective compound tools shall be formulated with specific delivery agents (such as by, but not limited to, covalent linkages to a cell-targeting molecule, non-covalent complexation, encapsulation) or shall be chemically modified internally, while keeping invariant the sequence composition. If a formulation can indeed result in uptake by specific cells and retains the on-target potency to modulate the CALIMERO gene, the functional uptake can be visualized and characterised, both qualitatively and quantitatively, by fluorescence (with confocal microscopy or flow cytometry), luminescence (for in vivo, real time measurements), and immuno-histochemistry (tissue staining). By eliminating the need to perform laborious RNA extraction and analyses as readouts that are associated with intrinsic variability, the CALIMERO platform empowers rapid and precise readouts that scales with the number of delivery formulations/agents and internal chemical modifications tested.

[0039] The CALIMERO gene and its library of compound tools are easy to use. The CALIMERO gene is in a PiggyBac backbone plasmid for ease of stable integrating in the genome upon which CALIMERO-expressing cells can be selected and/or sorted quickly, resulting in stable expression. Its sensitivity in splicing switches/relative abundance is extremely high and reproducible. In the present study, a transgenic mouse (CALIMERO: Conjugated AON Library Mouse switcher mOdel) bearing the CALIMERO gene was created.

[0040] In one aspect, the present invention provides a DNA construct comprising: (a) a starter exon encoding a start codon; (b) a first frame-switch exon located downstream of the starter exon, the first frame-switch exon comprising a first acceptor splice site, wherein inclusion or exclusion of the first frame-switch exon during alternative splicing determines the downstream reading frame of the mature mRNA transcript; and (c) a first specific exon located downstream of the first frame-switch exon, the first specific exon comprising a second acceptor splice site, a first reporter gene and a second reporter gene, wherein the first reporter gene is upstream or downstream of the second reporter gene and wherein the first reporter gene is in a different reading frame from the second reporter gene, wherein the exclusion of the first frame-switch exon causes the first reporter gene to be in-frame with the start codon and wherein the inclusion of the first frame-switch exon causes the second reporter gene to be in-frame with the start codon.

[0041] In various embodiments, the terms "DNA construct" and "CALIMERO gene" are used interchangeably.

[0042] As used herein, the term "exon" refers to a segment of a nucleic acid sequence that is transcribed into pre-mRNA and that is present in mature mRNA after splicing. The term "specific exon" is meant to refer to an exon that comprises one or more relay nodes (RNs), each with its own acceptor site, within which contains one or more protein-coding genes (e.g. reporter genes) that could be interspaced with a STOP codon. The term "starter exon" refers to an exon that comprises a start codon. The term "frame-switch exon" refers to an exon whose inclusion or exclusion changes the downstream reading frame (i.e. reading frame downstream of the frame-switch exon) of the resulting mature mRNA. The function of the frame-switch exon is to modulate the reading frame of the specific exon(s) and the reporter genes therein.

[0043] Multiple reporter genes can be located in one specific exon. The term "specific exon" is meant to refer to an exon that comprises one or more reporter genes. In one example, where there are only 2 reporter genes in one specific exon, the two are in different reading frames (in one RN). In another example, where there are 4 reporter genes in one specific exon, it depends on whether they are all placed into one exon (Exons 3A and 3B as shown in Figure 1) or split into two exons. As shown in Figure 1, as a single exon, tRFP and His-tag are in-frame while EGFP and Luc are in-frame. The acceptor splice site at Exon 3A (RN1) is used over the acceptor splice site at Exon 3B (RN2) in the absence of SSOs. On the other hand, if the 4 reporter genes are split into 2 exons, the reporter genes within an exon are in different reading frames.

[0044] In various embodiments, the first frame-switch exon is designed to be spliced into the majority but not all mature mRNA transcripts. This allows NAT modalities that do not

modulate splicing to target mature mRNA transcripts containing the first frame-switch exon (major transcript), while keeping the abundance of the mature mRNA transcripts without the first frame-switch exon constant (minor transcript). This facilitates internal normalization against the latter expressed product.

- 5 [0045] In various embodiments, the first and second FS exons are designed to be predominantly spliced into the mature mRNA transcript. This may be achieved by engineering the acceptor splice sites of the first and second FS exons to be strong so that the acceptor splice sites can promote the inclusion of the FS exon(s) while allowing the FS exon(s) to be spliced out in some transcripts.
- 10 [0046] In one embodiment, the DNA construct as described herein further comprises a second specific exon located downstream of the first specific exon.
- [0047] In one embodiment, the DNA construct as described herein further comprises a second frame-switch exon located downstream of the first frame-switch exon and upstream of the first specific exon, the second frame-switch exon comprising a third acceptor splice site,
- 15 wherein the second frame-switch exon is in a different reading frame from the first frame-switch exon and wherein the combination of inclusion or exclusion of the first frame-switch exon and the second frame-switch exon during alternative splicing determines the downstream reading frame of the mature mRNA transcript.
- [0048] In one embodiment, the DNA construct as described herein further comprises a
- 20 third reporter gene.
- [0049] In one embodiment, the third reporter gene is located in the first specific exon.
- [0050] In another embodiment, the third reporter gene is located in the second specific exon.
- [0051] In one embodiment, the third reporter gene is in the same reading frame as the first
- 25 or second reporter genes.
- [0052] In another embodiment, the third reporter gene is in a different reading frame from the first and second reporter genes.
- [0053] In one embodiment, the DNA construct as described herein further comprises a fourth reporter gene.
- 30 [0054] In one embodiment, the fourth reporter gene is located in the first specific exon.
- [0055] In another embodiment, the fourth reporter gene is located in the second specific exon.
- [0056] In one embodiment, the DNA construct as described herein further comprises a fifth reporter gene and a sixth reporter gene.

- [0057] In one embodiment, the fifth reporter gene and the sixth reporter gene are located in the second specific exon.
- [0058] In various embodiments, the reporter genes are selected from the group consisting of turboRFP, EGFP, luciferase and His-tag.
- 5 [0059] In one embodiment, each reporter gene comprises a stop codon.
- [0060] The advantage of each reporter gene having its own stop codon is that the resultant mRNA transcript would have a defined stop codon to prevent nonsense-mediated decay, non-stop decay and/or additional amino acids that are otherwise translated interfering with the folding of the desired reporter protein.
- 10 [0061] In one embodiment, the first frame-switch exon is predominantly spliced into the mature mRNA transcript.
- [0062] In one embodiment, the second frame-switch exon is predominantly spliced into the mature mRNA transcript.
- [0063] In one embodiment, the DNA construct as described herein comprises the
- 15 sequence of SEQ ID NO: 67.
- [0064] In another aspect, the present disclosure provides a method of measuring the efficacy of a nucleic acid delivery into a cell by a delivery agent, the method comprising: (a) using the delivery agent to introduce the nucleic acid into the cell, wherein the cell expresses the DNA construct of any one of claims 1 to 18 and wherein the nucleic acid is configured to
- 20 modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct; and (b) quantifying the functional uptake of the nucleic acid into the cell by measuring the relative expression of two of the reporter genes, wherein the functional uptake of the nucleic acid is indicative of the efficacy of the delivery agent.
- 25 [0065] The method of measuring the efficacy of a nucleic acid delivery into a cell by a delivery agent can be used to screen existing delivery agents and also expedite the development of new delivery agents for the NAT modality of compound tool used in the method. For delivery agents that are composed of multiple ingredients, the system of CALIMERO gene and compound tools will allow for the screening of compositions that lead to
- 30 optimal functional delivery specific to each nucleic acid modality.
- [0066] In one embodiment, step (b) of the method as described herein comprises quantifying the functional uptake of the nucleic acid into the cell by measuring the relative expression of the first reporter gene and the second reporter gene.
- [0067] In various embodiments, the efficacy of the delivery agent may be measured by the
- 35 relative expression of the first reporter gene to the second reporter gene. While the efficacy of

- the delivery agent may also be measured by the expression of one of the reporter genes, the advantage of using the relative expression as a readout is that the expression variability of the DNA construct between cell-to-cell, cell types, tissues, organs and between mice is corrected for. As a result, the readout is less variable and hence less mice will be needed. An additional
- 5 advantage from the internally normalized readout is that results between different delivery modalities or formulations or internal chemical modifications performed on different experiments can be compared in quantitative terms.
- [0068] In the method of measuring the efficacy of nucleic acid delivery into a cell by a delivery agent, a high functional uptake of the nucleic acid as quantified by measuring the
- 10 expression of the first reporter gene relative to the expression of the second reporter gene also implies a non-interference on the nucleic acid activity by the delivery agent.
- [0069] In one embodiment, the nucleic acid is selected from the group consisting of a steric blocker antisense oligonucleotide, a RNaseH-activating GAPmer, an siRNA, a CRISPR/Cas-sgRNA and a snRNA.
- 15 [0070] In one embodiment, the steric blocker antisense oligonucleotide is a splice-switching oligonucleotide (SSO).
- [0071] In one embodiment, the delivery agent is selected from the group consisting of an intra-molecular chemical modification, an inter-molecular delivery agent, an encapsulation delivery agent, a complexation delivery agent, a viral-based delivery agent, a cellular delivery
- 20 agent and a sub-cellular delivery agent.
- [0072] In various embodiments, the expression of the reporter genes is measured by fluorescence, luminescence, immunohistochemistry or quantitative PCR.
- [0073] Where the delivery agent used is the same, the CALIMERO gene can also be used to measure the potency or activity of a nucleic acid (static sequence) but is intramolecularly
- 25 modified with different chemistr(ies) at different ribose sugar, backbone and/or nucleobase.
- [0074] For each NAT modality (except for snRNA) in the associated nucleic acid compound toolbox, examples of multiple compounds, each with different nucleobase sequences, have been provided in the present disclosure. This enables the investigation of sequence dependency on functional delivery and potency, if desired.
- 30 [0075] In another aspect, the present disclosure provides a method for measuring the effect of chemical modification on the potency or activity of a nucleic acid, the method comprising: (a) introducing a chemically-modified nucleic acid into a cell, wherein the cell expresses the DNA construct of any one of claims 1 to 18 and wherein the chemically-modified nucleic acid is configured to modify the DNA construct or to modulate the expression or
- 35 processing of a pre-mRNA or mature mRNA expressed from the DNA construct; and (b)

quantifying the potency or activity of the chemically-modified nucleic acid by measuring the relative expression of two of the reporter genes.

[0076] As used herein, the term "potency" is meant to include the concentration needed for a drug (such as a nucleic acid) to achieve a given level of efficiency. As used herein, the
5 term "activity" is meant to include the efficiency of a drug (such as a nucleic acid) at a given concentration. Potency is a function of activity. In one example, the activity of a nucleic acid can be quantified by measuring the numerical change in the relative expression of two of the reporter genes. For example, the activity of the nucleic acid can be quantified by measuring the numerical change in the relative expression of the first reporter gene to the second reporter
10 gene.

[0077] By "modify the DNA construct", it is meant to include the modifying of the DNA construct by CRISPR/Cas-sgRNA, such as by disrupting specific NATURA splice sites through CRISPR-Cas9-mediated genomic introduction of indels (insertions and/or deletions) or substitutions surrounding their consensus sequence. By "modulate the expression or
15 processing of a pre-mRNA or mature mRNA expressed from the DNA construct", it is meant to include the modulation of the splicing of the pre-mRNA (such as by splice-switching oligonucleotides and snRNA), the abundance of specific spliced-isoforms (such as by GAPmer and siRNA) expressed from the DNA construct, or the modulation of the processing of the mature mRNA.

[0078] In one embodiment, step (b) of the method as described herein comprises quantifying the potency or activity of the chemically-modified nucleic acid by measuring the relative expression of the first reporter gene and the second reporter gene.

[0079] In one example, in the absence of the chemically-modified nucleic acid (i.e. the compound tool), the relative expression of the first reporter gene to the second reporter gene
25 is around 0.9. A highly potent chemically-modified nucleic acid (i.e. a highly potent NAT with specific chemical modification) can lower the score to less than 0.2. It would be generally understood by a person skilled in the art that the score would vary according to the dose since efficiency is a function of the dose.

[0080] In one aspect, the present disclosure provides a transgenic animal that expresses
30 the DNA construct as described herein.

[0081] In one embodiment, the animal is a rodent. In some examples, the rodent is a mouse or a rat.

[0082] In another aspect, the present disclosure provides a nucleic acid configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or
35 mature mRNA expressed from the DNA construct of any one of claims 1 to 18, wherein the

nucleic acid is selected from the group consisting of a steric blocker antisense oligonucleotide, a RNaseH-activating GAPmer, an siRNA, a CRISPR/Cas-sgRNA and a snRNA.

[0083] The term “nucleic acid” as used herein may be used interchangeably with “nucleic acid therapeutic (NAT) modality” or “compound tool”.

5 [0084] In one embodiment, the steric blocker antisense oligonucleotide is a splice-switching oligonucleotide (SSO).

[0085] In one embodiment, the steric blocker antisense oligonucleotide comprises a sequence selected from the group consisting of SEQ ID NOs 1 to 57.

10 [0086] In one embodiment, the RNaseH-activating GAPmer comprises a sequence selected from the group consisting of SEQ ID NOs 58 and 59.

[0087] In one embodiment, the siRNA comprises a sequence selected from the group consisting of SEQ ID NOs 60 to 63.

[0088] In one embodiment, the CRISPR/Cas-sgRNA comprises a sequence selected from the group consisting of SEQ ID NOs 64 and 65.

15 [0089] In one embodiment, the snRNA comprises the sequence of SEQ ID NO: 66.

[0090] In various embodiments, snRNA may be delivered via viral vectors such as AAVs and lentiviruses. With knowledge of the endogenous modifications during the processing of the snRNA, non-viral delivery methods can also be used. The term “snRNPs” refer to snRNA (small nuclear RNA) complexed with RNA-binding proteins.

20 [0091] In one aspect, the present disclosure provides a system for characterizing the functional cellular uptake of a nucleic acid, the system comprising: (a) a DNA construct comprising at least two reporter genes; and (b) a nucleic acid configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct such that one of the at least two reporter genes is
25 preferentially expressed.

[0092] The term “preferentially expressed” in the context of a reporter gene is meant to include the scenario where the reporter gene is expressed to a greater degree than another reporter gene, and the scenario where the reporter gene is exclusively expressed and the other reporter gene is not expressed. Similarly, the term “preferentially used” in the context of
30 an acceptor splice site is meant to include the scenario where the acceptor splice site is used to a greater degree (i.e. with greater likelihood) than another acceptor splice site, and the scenario where the acceptor splice site is exclusively used and the other acceptor splice site is not used.

[0093] In the context of the nucleic acid configured to modify the DNA construct or to
35 modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the

- DNA construct as described herein, the phrase "one of the at least two reporter genes is preferentially expressed" is meant to include a scenario where the nucleic acid, such as a steric blocker antisense oligonucleotide, causes a switch in expression from a first reporter gene to a second reporter gene, or from the second reporter gene to the first reporter gene.
- 5 This phrase also includes the scenario where the nucleic acid causes a decrease in the expression of the first reporter gene or of the second reporter gene, such as where the nucleic acid is a siRNA or GAPmer and where the siRNA or GAPmer degrades the mature mRNA of one of the reporter genes. In one example, two siRNAs that target FS exon were designed to each degrade the dominantly expressed "EGFP isoform". The "tRFP isoform", which does not
- 10 contain the FS exon and is thus not targeted by the siRNA, is used as the normalization to account for variability in gene expression. In another example, the sequence of AVA1, which targets the FS exon, was converted to a gapmer. With the same principle of readout as the example on siRNA modality above, transfection of the gapmer results in EGFP transcript degradation that reduces EGFP protein abundance but not tRFP.
- 15 [0094] In one embodiment, the DNA construct comprises the DNA construct as described herein.
- [0095] In one embodiment, the nucleic acid comprises the nucleic acid as described herein.

20 EXAMPLES

- [0096] Non-limiting examples of the invention and comparative examples will be further described in greater detail by reference to specific Examples, which should not be construed as in any way limiting the scope of the invention.
- [0097] **(A) CALIMERO gene**
- 25 [0098] The underlying idea is to provide a gene structure that expresses N spliced-isoforms (from a single gene construct), for which each isoform translates to a distinct protein. While there is no theoretical limit on N and on the corresponding choice of proteins, N = 4 is the minimally sufficient system that allows for internal normalization, amenability to multiple NAT modalities and compositions, and visualization of functional NAT delivery from in vitro cell
- 30 culture to in vivo animal models. In our context, the four proteins were selected for their widely proven applications in confocal microscopy and flow cytometry (turboRFP and EGFP), in vivo imaging (firefly luciferase) and immunohistochemistry staining (polyhistidine-tag, a.k.a. His-tag, and Flag-tag), as the means for functional delivery readout. Although the most intuitive and direct design is to construct the entire coding sequence of each protein as individual exons
- 35 that are mutually exclusive, i.e. only one exon can be spliced in on the mature mRNA, it is

extremely difficult or near impossible to design four mutually exclusive exons with similar propensity for splicing; at least no four-mutually-exclusive exon exists in nature. And as a result, it is not scalable with N.

[0099] To overcome the technical hurdle, a novel concept of modular “multiplex switch-relay (MSR)” was invented and implemented. The MSR can be implemented as an either binary or ternary system. A binary MSR module consists of a frame-switch (FS) exon, whose inclusion or exclusion determines the codon reading frame of the mature transcript, and at least one downstream exon that serves as a relay nodule (RN). The RN exon composes of an acceptor splice site along with the juxtaposition of two protein coding sequences at different reading frames, each retaining its STOP codon. By pairing proteins into a RN exon, as compared to the use of two individual exons, similar propensities for splicing are enforced. Subsequently, the functional translation of either one of the two proteins can be controlled by compound tools designed to modulate FS exon splicing or abundance, depending on the NAT modality. Since the functional delivery readout is always the expression ratio of the two proteins, rather than the absolute expression level of each, the MSR module achieves internal normalization. To scale with N as discussed above, one can append more RN exons. In fact, the number of RN exons required equals $N / 2$.

[00100] The second and alternative way to scale with N is through a ternary MSR module. Here, it consists of two independent FS exons, whose combinations of respective inclusion or exclusion determine the codon reading frame of the mature transcript, and at least one RN exon that juxtaposes three protein coding sequences at different reading frames. The advantage of similar splicing propensities of the three proteins are conserved. Analogously, the functional translation of one of the three proteins can be controlled by compound tools designed to modulate one or both FS exons splicing or abundance, depending on the NAT modality. Internal normalization is also achieved via the readouts of functional delivery in terms of expression ratio among the three proteins, or between any of the two proteins. The ternary system is well-suited for N that is a multiple of 3, as the number of RN exons required is $N / 3$.

[00101] Lastly, one can combine multiple binary (for instance, using multiple RNs) and/or ternary MSR modules in a single gene construct. Although the application is unclear now, the modular design nonetheless provides the basic building block to construct more complex systems.

[00102] To demonstrate that the proposed MSR concept is feasible and can be implemented in practice, a $N = 4$ binary MSR module with two RNs, as described in the following, and the corresponding compound tool library, as described in (B), was designed

and validated. Without any loss of generality, one embodiment of the CALIMERO gene is depicted in Figure 1, with the following exons and modulation logic tabulated in Table 1.

- Exon 1 codes for the ATG start codon, Flag-tag, and a nuclear localization signal (NLS).
- 5 • Exon 2 is the FS exon, and is designed to be predominately spliced into the mature mRNA, allowing for the development of compound tools from NATs that do not modulate splicing (refer to (B)).
- Exon 3A, contained in exon 3, is the first RN, and is composed of the coding sequences of the fluorescent proteins turboRFP and EGFP in tandem. The codon reading
10 frame of EGFP sequence is designed to be inframe with the FS exon, so that the protein is expressed only when FS exon is spliced-in. On the contrary, the turboRFP sequence is inframe when FS exon is spliced-out. The relative protein expressions of turboRFP and EGFP is the internally normalized readout for functional delivery.
- Exon 3B, contained in exon 3, is the second RN, and besides its own acceptor
15 splice site, is composed of the coding sequences of His-tag and firefly luciferase in tandem. Here, the codon reading frame of luciferase sequence is inframe with the FS exon while the His-tag sequence is inframe when FS exon is spliced-out. Exon 3A is designed to be preferentially spliced into the mature mRNA over exon 3B.

State	FS	RN1	Expressed protein	Readout method	Switching compound tools
0	1	1	EGFP	Confocal microscopy / flow cytometry	-
1	0	1	tRFP		AVA1
2	1	0	Luciferase (RN2)	<i>in vivo</i> imaging	AVA2
3	0	0	His-Tag (RN2)	IHC staining	AVA1 + AVA2

[00103] **Table 1.** Logic table manifested by the CALIMERO gene as depicted in Figure 1

20 [00104] **(B) Compound tool library**

[00105] Compound tool library from the five NAT modalities described above were designed and validated to implement the logic in Table 1. For easy reference, the compound tools for all NAT modality shall be referenced to as listed in Table 1 last column, in accordance to the logic state that they induce as enumerated in Table 1 first column. FS = 1 indicates the FS exon is spliced in the mature transcript while FS = 0 indicates otherwise. Analogously, RN1 =
25 1 indicates that the proximal acceptor splice site of exon 3 is used resulting in exon 3A

inclusion in the mature transcript while RN1 = 0 indicates that the alternate acceptor splice site of exon 3 is used resulting in exon 3B inclusion.

[00106] Splice-switching oligonucleotides (SSOs), CRISPR-Cas9 and snRNP modulate the splice switching of CALIMERO, whereas siRNAs and GAPmers degrade specific spliced-isoforms of the CALIMERO gene for an internally-normalized quantification.

[00107] CRISPR/Cas-sgRNA Modality

[00108] AVA1-sgRNA was designed and validated to target the donor splice site of the FS exon (exon 2). This will result in exon 2 exclusion from the mature transcript of CALIMERO, thereby switching the logic state from 0 to 1, as demonstrated in Figure 2A using RNAiMAX and lentiCRISPRv2 viral particles as delivery carriers, and using other carriers (Figure 2B). Our data showed that the extent of State 0 to 1 switch by the AVA1-sgRNA as quantitated at the DNA, mRNA and protein are perfectly linear (Figure 2C). Hence, functional delivery can be visualized and characterized at the protein level, bypassing the tedious and non-scalable mRNA extraction and analyses.

[00109] Next, AVA2-sgRNA was designed and validated to target the (proximal) acceptor splice site of RN1 exon (exon 3A), thereby switching State 0 to 2 (Table 1). This will induce luciferase expression (Figure 2D) and suppresses both tRFP and EGFP expressions (Figure 2E).

[00110] With the compound tools for the CRISPR/Cas-sgRNA modality validated, it was demonstrated how the system can quantitatively characterize and compare the efficiency of functional delivery among transient and permanent delivery agents using the AVA1-sgRNA tool. Figures 3A showed the flow cytometry readout, with Figure 3B showing the corresponding proportion of cells on which functional delivery was achieved by RNAiMAX, TaraTru, QiagenHi, and Lipofectamine2000 transient agents, and lentivirus (permanent). The protein readout was corroborated by Taqman PCR at the RNA level; Figures 3C and 3D depicted the PCR primers' location and the PSI of the FS exon, respectively.

[00111] Steric Blocker Antisense Oligonucleotide Modality

[00112] SSO (splice-switching oligonucleotide) were rationally designed to induce exclusion of FS (AVA1-SSO) or RN1 (AVA2-SSO) exons. Three groups of AVA1-SSO were designed namely, AVA1a-SSOs that target exonic splicing enhancers (ESEs), AVA1b-SSOs that block FS's acceptor splice site, and AVA1c-SSOs that block FS's donor splice site. The expression switch from EGFP (State 0) to tRFP (State 1) induced by AVA1a-SSO was validated by both confocal high-content protein imaging and taqman RNA PCR (Figures 4A, 4B and 4C), by flow cytometry (Figures 4F and 4L), by PCR and Sanger sequencing (Figures 4J, 4N and 4O), by western blot (Figure 4K); similar results were obtained for AVA1b1 and b2-SSOs and AVA1c2-

SSOs (Figure 4D). The tPSI time-course induced by AVA1a-SSO calculated from TaqMan qPCR was also determined (Figure 4M). On the other hand, AVA2-SSOs were designed to block RN1's acceptor splice site so as to induce the usage of RN2's acceptor splice site, which switches the expression from EGFP (State 0) to luciferase (State 2) (Figures 4E, 4G, 4N and 4O) in vitro, and in mice model expressing the CALIMERO gene (Figures 4H and 4I). Lastly, the simultaneous use of AVA1 and AVA2 induced the usage of Exon 3B acceptor splice site (State 3) (Figure 4O).

[00113] siRNA Modality

[00114] Given that the FS exon is designed to be predominately but not entirely spliced into the mature mRNA transcripts, a low basal expression level of tRFP (State 1) in CALIMERO-transfected cells is consistently detected, as shown in Figure 5A. This feature allows the system to be amenable to NATs that do not modulate splicing, such as both the siRNA and GAPmer modalities. By designing the respective compound tools that specifically degrade "State 0" CALIMERO transcripts (that is, EGFP expressing), which is achieved by targeting the FS exon, the constant basal tRFP protein level can be used as internal normalization for the readout of functional delivery. Advantageously, by using CALIMERO's internal normalization (EGFP/tRFP), false positive hits in down-screenings caused by unspecific gene expression downregulation can be avoided.

[00115] Two siRNAs, AVA1-siRNA1 and AVA1-siRNA2, were designed and validated to reduce EGFP but not tRFP protein levels (Figure 5B). It was further confirmed that the observations were the result of the on-target specificity of both siRNA compound tools, and not due to transcriptome-wide off-target events or shutdown, with a scrambled siRNA that did not result in changes to CALIMERO gene expression, and EGFP and tRFP protein levels (Figures 5C, 5D and 5E).

[00116] GAPmer Modality

[00117] With the same principle as the siRNA compound tools, AVA1-GAPmer were rationally designed and validated to reduce EGFP but not tRFP protein levels (Figures 6A, 6B, 6C and 6D). In addition, a GAPmer, AVA2-EX3-GAPmer, targeting RN1 (exon 3A) was designed and validated, which in this context leads to the degradation of all CALIMERO transcripts, resulting in the reduction of both EGFP and tRFP protein levels (Figures 6A and 6B).

[00118] snRNP Modality

[00119] AVA1-U7s were engineered from the endogenous human U7-snRNA to target either the acceptor splice site of or exonic splicing enhancer (ESE) within FS exon, with the purpose to induce FS exon exclusion from the mature transcript of CALIMERO, thereby

switching the logic state from 0 to 1. Figure 7 depicted the increase in normalized tRFP signal upon transfection with the AVA1-U7s (labelled as pC4.4).

METHODS:

5 Cell lines

The mouse and human cell lines were propagated in accordance with ATCC guidelines.

To generate stable NATURA-expressing cell lines, 3×10^5 cells were seeded in 6-well plates containing 1 mL of complete growth medium the day prior to transfection. Transfection was carried out using 4.5 µg of NATURA plasmid and 0.5 µg of PRP-mCherry-CAGhybase plasmid in conjunction with the Lipofectamine 3000 reagent after 16-20 hours of incubation. The following day, the medium was replaced with complete growth medium supplemented with Blasticidine antibiotic (1:1000, Sigma-Aldrich, cat- 15205) and the cells were grown under selection for 14 days. Subsequently, the cells were sorted based on EGFP expression.

15 FACS analysis

Cells were trypsinized and washed in PBS. Collected cells were resuspended in ice-cold 5% FBS/PBS and analyzed within 1hr using the Becton & Dickinson LSRII or Penton-Novocyte Agilent L5 analyser. The HTS module was used for analysing 96-well plates.

The laser settings for the EGFP and tRFP channels were kept consistent throughout the experiments and across different cell lines. For experiments requiring absolute count, events/sec were recorded while keeping the volume of sample constant. The pPSI was calculated with the following formula, where MFI stands for Mean Fluorescence Intensity:

$$pPSI = \frac{MFI_{EGFP}}{MFI_{EGFP} + MFI_{tRFP}}$$

Luciferase assay

25 Live cells were stained by adding 0.4% trypan blue (1:1, #15250061; Gibco) and counted using Countess II FL automated cell counter (Thermo Fisher Scientific). Luciferase assay (#E1960; Promega) was performed by lysing 1×10^5 cells in a flat white-96 well plate with 22 µl of 1x passive lysis buffer at room temperature for 15 mins. Subsequently, the sample was incubated with 100 µl of luciferase assay substrate for 5 mins in humidified incubator.

30 Readings were taken with Tecan multimode spark 10M plate reader.

Taqman assay

The cDNA obtained as described in the "Agarose Gel and PCR" paragraph was analysed by Taqman Assay (#4304437; Thermo Fisher Scientific) using primers and probes specified in (table#).

Each reaction mix contained 10 µl of 2x Taqman master mix, 1 µl of NATURA primer-probe mix (2.5 µM probes and 10 µM primers), 1 µl of Cyclophilin A primer-probe mix (2.5 µM probes and 10 µM primers) and 8 µl of cDNA (3.33 ng/µl). The assays were run on Biorad CFX96 Real time system using cycling conditions- 2 mins at 95oC followed by 40 cycles of 5
 5 secs at 95oC and 40 secs at 60oC with no dissociation curve. The data was analysed using Graphpad prism 9.

[00120] Key Technical Features:

- 10 A. Design of CALIMERO gene structure that meets these technical requirements: internal normalization, amenability to multiple NAT modalities, visualization of functional NAT delivery with multiple optical microscopy and flow cytometry platforms, and scalable to express N spliced-isoforms (from one gene construct).
- B. An embodiment of CALIMERO gene with N = 4, with the full DNA sequence, to demonstrate it is practical to implement the proposed gene structure.
- 15 C. Compound tool library, with full composition of each tool, to modulate the CALIMERO gene to facilitate functional delivery readout for the following NAT modalities: Steric blocker antisense oligonucleotide, GAPmer, siRNA, CRISPR/Cas-sgRNA, and snRNP-AAV.

[00121] Advantageously, the present disclosure provides a model/platform that is
 20 unlabelled-based, internally-normalized and scalable for screening and quantitative characterization of delivery agents resulting in functional delivery of multiple NAT modalities with a single cell-resolution, without the need for tedious RNA extraction and analyses, in vitro cell culture and in vivo animal models.

[00122] Applications of the technology:

- 25 1. Unbiased qualitative and quantitative high-throughput discovery, validation, characterization and/or comparative studies of delivery agents/methods/technologies for NATs; in vitro and in vivo.
2. Unbiased qualitative and quantitative high-throughput in vivo biodistribution and pharmacokinetics (PK) characterization of delivery agents/methods/technologies for
 30 NATs.
3. Unbiased quantitative high-throughput measurement of on-target potency; in vitro and in vivo.

[00123] A comparison of the CALIMERO platform against existing methods for screening NAT delivery is shown in Table 2 below.

	Labelled methods	Unlabeled methods		CALIMERO platform
	<i>The NAT molecule is conjugated to a fluorophore (e.g.: FAM)</i>	<i>The NAT molecule is targeting a known gene and the outcome is the quantification of the RNA</i>	<i>The NAT molecule is targeting a known reporter gene (e.g.: GFP) and the outcome is the quantification of the protein</i>	
NAT classes	siRNA, GAPmer, SSO, CRISPR-Cas	siRNA, GAPmer, CRISPR		siRNA, GAPmer, SSO, CRISPR-Cas
True-positive rate	Cannot discern if NAT is functionally uptake, or trapped at cell membrane or sequestered at intracellular compartments	High	High	High
Native NAT properties	The labelling alters the biochemical properties/size of the NAT	✓	✓	✓
Internally normalized	Not possible	~ <i>Housekeeper gene</i>	Toxicity/transcription inhibition cause false positives (<i>down screening</i>)	✓
Toxicity quantification	Separate measurement (cell count). It is biased in the presence of transcription/translation arrest response.			✓
Single-cell resolution	✓	Macroscopic resolution (tissue level)	✓	✓
High throughput	✓	Requires RNA extraction/RT	✓	✓ 10 samples/min
Rapid outcome	✓	Requires multiple steps for RNA extraction/RT/qPCR	✓	
Inexpensive setup	✓		✓	✓
Kinetic measurements	Endpoint only			✓

Table 2. Comparison of the CALIMERO platform against existing methods for screening NAT delivery

[00124] CALIMERO is different from existing platforms in the following manner:

- Is both an *in vitro* (high throughput) and *in vivo* (CALIMERO mouse, mid throughput) system
- It is internally normalized (important to eliminate false positive resulting from cellular toxicity of NATs or the delivery agents that cause transcription/translation shutdown)

- It's fast and cheap to screen NAT *in vitro* at 96 samples in 10 minutes
- It has a single cell resolution both *in vitro* and *in vivo*
- It allows functional, unlabeled screenings
- *In vivo* NAT real time kinetic measurements

5 • It allows concurrent cell viability/fitness and functional uptake readouts

[00125] Examples of gene structures for the DNA construct are shown in Table 3 below.

Scenario	Arrangement of reporter genes	Logic: Binary (1 FS exon) or ternary (2 FS exons) system?	Reading frame	Acceptor splice site	Remarks
Scenario A Number of reporter genes: 3	1 st , 2 nd and 3 rd reporter genes in 1 st specific exon	Binary	3 rd reporter gene is in same reading frame as 1 st or 2 nd reporter gene. If 3 rd reporter gene is in frame with 1 st reporter gene, a stop codon is located after the 1 st reporter gene. If 3 rd reporter gene is in frame with 2 nd reporter gene, a stop codon is located after 2 nd reporter gene.	3 rd reporter gene has own acceptor SS. But the upstream acceptor SS is preferentially used.	-
Scenario B Number of reporter genes: 3	1 st , 2 nd and 3 rd reporter genes in 1 st specific exon	Ternary	1 st , 2 nd , and 3 rd reporter genes are in different reading frames. 1 st FS exon and 2 nd FS exon are in different reading frames.	3 rd reporter gene does not have own acceptor SS.	-
Scenario C Number of 3 rd reporter gene in 2 nd	1 st and 2 nd reporter genes in 1 st specific exon. 3 rd reporter gene in 2 nd	Binary	3 rd reporter gene is in same reading frame as 1 st or 2 nd reporter gene.	3 rd reporter gene has own acceptor SS. Preferably, the upstream acceptor SS	Alternatively, the 1 st specific exon can contain only the 1 st reporter gene and the 2 nd specific exon contains the 2 nd and 3 rd reporter genes. In such a

reporter genes: 3	specific exon (i.e. there is an intron upstream of 3 rd reporter gene)			(at the 1 st specific exon) is preferentially used.	scenario, the 2 nd and 3 rd reporter genes are not in frame, but one of these would be in frame with the 1 st reporter gene, and the 2 nd reporter gene has its own acceptor SS instead.
Scenario D Number of reporter genes: 3	1 st and 2 nd reporter genes in 1 st specific exon. 3 rd reporter gene in 2 nd specific exon	Ternary	3 rd reporter gene is in different reading frame from 1 st and 2 nd reporter gene. 1 st FS exon and 2 nd FS exon are in different reading frames.	3 rd reporter gene has own acceptor SS.	Alternatively, the 1 st specific exon can contain only the 1 st reporter gene and the 2 nd specific contains the 2 nd and 3 rd reporter genes. In such a scenario, the 2 nd reporter gene has its own acceptor SS instead.
Scenario E Number of reporter genes: 4	1 st , 2 nd , 3 rd and 4 th reporter genes in 1 st specific exon (as shown in Figure 1)	Binary	3 rd reporter gene is in-frame with 1 st reporter gene, 4 th reporter gene is in-frame with 2 nd reporter gene, or vice versa.	3 rd reporter gene has own acceptor SS. But the upstream acceptor SS is preferentially used.	-
Scenario F Number of reporter genes: 4	1 st , 2 nd , 3 rd and 4 th reporter genes in 1 st specific exon	Ternary	1 st , 2 nd , and 3 rd reporter genes are in different reading frames. 1 st FS exon and 2 nd FS exon are in different reading frames. There is a stop codon after each reporter gene.	4 th reporter gene has own acceptor SS.	Alternatively, the 2 nd reporter gene can have its own acceptor SS. In such a scenario, the 2 nd , 3 rd and 4 th reporter genes are each in different reading frames.
Scenario G Number of reporter genes: 4	1 st and 2 nd reporter genes in 1 st specific exon. 3 rd and 4 th reporter genes in 2 nd specific exon ("Exon 4")	Binary	3 rd reporter gene is in-frame with 1 st reporter gene. 4 th reporter gene is in-frame with 2 nd reporter gene.	3 rd reporter gene has own acceptor SS. Preferably, the upstream acceptor SS (at the 1 st specific exon) is preferentially used.	Alternatively, the 1 st specific exon can contain only the 1 st reporter gene, and the 2 nd specific exon can contain the 2 nd to 4 th reporter genes. In such a scenario, the 2 nd reporter gene has its own acceptor SS. Within the 2 nd specific exon, the

					alternative acceptor SS can be at the 3rd or 4th reporter gene. When the alternative acceptor SS is at the 3rd reporter gene, its reading frame is different from the 4th reporter gene. When the alternative acceptor SS is at the 4th reporter gene, the 2nd and 3rd reporter genes are in different reading frames. In another alternative scenario, the 1st specific exon can contain the 1st to 3rd reporter genes, and the 2nd specific exon contains the 4th reporter gene. In such a scenario, the 4th reporter gene has its own acceptor SS. Within the 1st specific exon, the alternative acceptor SS can be at the 2nd or 3rd reporter gene. When the alternative acceptor SS is at the 2nd reporter gene, its reading frame is different from the 3rd reporter gene. When the alternative acceptor SS is at the 3rd reporter gene, the 1st and 2nd reporter genes are in different reading frames.
Scenario H Number of	1 st and 2 nd reporter genes in 1 st specific exon. 3 rd and 4 th	Ternary	3 rd reporter gene is in a different reading frame from 1 st and 2 nd reporter genes. 4 th reporter gene is in a	3 rd reporter gene has own acceptor SS. Preferably, the upstream	Alternatively, the 1 st specific exon can contain only the 1 st reporter gene, and the 2 nd specific exon can contain the 2 nd to 4 th

reporter genes: 4	reporter genes in 2 nd specific exon ("Exon 4")		different reading frame from 3 rd reporter gene. If the upstream SS (at the 1 st specific exon) is preferentially used, the 3 rd reporter need not be in a different frame as the 1 st and 2 nd reporter genes, so long as the 3 rd reporter gene is in a different reading frame from the 4 th reporter gene. 1 st FS exon and 2 nd FS exon are in different reading frames.	acceptor SS (at the 1 st specific exon) is preferentially used.	reporter genes. In such a scenario, the 2nd reporter gene has its own acceptor SS. If the three reporter genes in the 2nd specific exon are not in frame, there is no need for an alternative acceptor SS within the exon. If any two of the reporter genes are in frame, they must be separated by an alternative acceptor SS within the 2nd specific exon. In another alternative scenario, the 1st specific exon can contain the 1st to 3rd reporter genes, and the 2nd specific exon contains the 4th reporter gene. In such a scenario, the 4th reporter gene has its own acceptor SS. If the three reporter genes in the 1st specific exon are not in frame, there is no need for an alternative acceptor SS within the exon. If any two of the reporter genes are in frame, they must be separated by an alternative acceptor SS within the 1st specific exon.
Scenario I	1 st , 2 nd and 3 rd reporter genes in 1 st specific exon.	Ternary	1 st , 2 nd , and 3 rd reporter genes are in different reading frames. 4 th , 5 th and 6 th reporter genes	4 th reporter gene has own acceptor SS. Preferably, the	Alternatively, a scenario with a total of 5 reporter genes split into 2 specific exons is also possible

Number of reporter genes: 6	4 th to 6 th reporter genes in 2 nd specific exon.		are in different reading frames. 1 st FS exon and 2 nd FS exon are in different reading frames.	upstream acceptor SS (at the 1 st specific exon) is preferentially used.	with either binary or ternary logic.
-----------------------------	---	--	---	---	--------------------------------------

Table 3. Various embodiments of DNA construct

[00126] In one embodiment, the invention relates to a DNA construct (such as the CALIMERO gene) and the associated nucleic acid compound tool that can be used to characterise functional delivery efficacy of nucleic acid therapeutics (NATs) into cells and measuring the effect of chemical modification on the potency or activity of nucleic acid therapeutics (NATs) in cells.

[00127] Examples of the nucleic acid as described herein are provided in Table 4 below.

SEQ ID NO	Type of nucleic acid therapeutic modality	Name of sequence	Sequence (5' to 3')
1	SSO (AVA1a)	1992	AUG GCA GUC CAG AGC UUU CAU UCU GU
2	SSO (AVA1a)	1993	GGA UGG CAG UCC AGA GCU UUC AUU
3	SSO (AVA1a)	1994	CAG AGG CAC GGA UGG CAG UC
4	SSO (AVA1a)	1995	GAU GGC AGU CCA GAG CUU UCA UUC UG
5	SSO (AVA1a)	1996	GAU GGC AGU CCA GAG CUU UCA UUC U
6	SSO (AVA1a)	1997	GCU UUC AUU CUG UAC AGG UAG UUA UAG CA
7	SSO (AVA1a)	2076	AGU CCA GAG CUU UCA UUC UGU
8	SSO (AVA1a)	2077	UGG CAG UCC AGA GCU UUC AUU CUG U
9	SSO (AVA1b)	2052	CAU UCU GUA CAG GUA GUU AUA GCA ACC UAU UGA
10	SSO (AVA1b)	2053	UCU GUA CAG GUA GUU AUA GCA ACC UAU UGA
11	SSO (AVA1b)	2054	UCA UUC UGU ACA GGU AGU UAU AGC AAC CUA
12	SSO (AVA1c)	2055	AGG UAG CUC UUA CCA GAG GCA CG
13	SSO (AVA1c)	2056	GGU AGC UCU UAC CAG AGG CAC
14	SSO (AVA1c)	2057	ACC CUU UCA GGU AGC UCU UAC CAG
15	SSO (AVA1c)	2058	ACC CUU UCA GGU AGC UCU UAC CA
16	SSO (AVA1a)	2045	GUC CAG AGC UUU CAU UCU GU
17	SSO (AVA1a)	2099	UGG CAG UCC AGA GCU UUC
18	SSO (AVA1b)	2100	AGU UAU AGC AAC CUA UUG A
19	SSO (AVA1b)	2101	AGU UAU AGC AAC CUA

20	SSO (AVA1c)	2102	AGG UAG CUC UUA CCA G
21	SSO (AVA1c)	2103	CUC UUA CCA GAG GCA CG
22	SSO (AVA1a)	2186	CCA GAG CUU UCA UUC UGU
23	SSO (AVA1a)	2187	GUC CAG AGC UUU CAU UC
24	SSO (AVA1b)	1948	AGU UGU AGC AAC CUG UUG AGG AAA AGA CUG GU
25	SSO (AVA1a)	1949	CAU GGC AUC CAG AGC UUU CAU CCU AU
26	SSO (AVA2)	2068	GGG UGU GGU UGA UGA AGG CUU UGC
27	SSO (AVA2)	2069	GGG UUU CUU GGA UCU GUA UGU GGU CU
28	SSO (AVA2)	2070	GAA GAA GUC GUG CUG CUU CAU GUG GU
29	SSO (AVA2)	2071	AUU CUU UCC AGU CUG UGG UCC ACG AA
30	SSO (AVA2)	2072	UUU GUC GGC CUC CUU GAU UCU UUC C
31	SSO (AVA2)	2073	UGC UGC UCG ACG UAG GUC UCU
32	SSO (AVA2)	2074	UGA AGU UCA CCU UGA UGC CGU UCU UCU
33	SSO (AVA2)	2075	GAU GUU GCC GUC CUC CUU GAA GUC
34	SSO (AVA2)	2087	ACA UGA CAC CGG CUG AAA AGA AAA GAG CAC AAU
35	SSO (AVA2)	2088	ACA CCG GCU GAA AAG AAA AGA GCA CAA UGA UAC C
36	SSO (AVA2)	2089	UCG GAC AUG ACA CCG GCU
37	SSO (AVA2)	2120	CUC GGA CAU GAC ACC GGC
38	SSO (AVA2)	2121	CAG CUC GGA CAU GAC ACC GG
39	SSO (AVA2)	2122	UCA GCU CGG ACA UGA CAC CG
40	SSO (AVA2)	2123	UUG AUC AGC UCG GAC AUG ACA CC
41	SSO (AVA2)	2124	CCU UGA UCA GCU CGG ACA UGA CA
42	SSO (AVA2)	2125	CUC CUU GAU CAG CUC GGA CAU GAC
43	SSO (AVA2)	2126	UCU CCU UGA UCA GCU CGG ACA UGA
44	SSO (AVA2)	2127	AUG UUC UCC UUG AUC AGC UCG GAC AUG
45	SSO (AVA2)	2167	ACA UGA CAC CGG CUG AAA AGA
46	SSO (AVA2)	2168	UGA UCA GCU CGG ACA UGA C
47	SSO (AVA2)	2169	AUG UUC UCC UUG AUC AGC UC
48	SSO (AVA2)	1950	GAU GAC GUU CUU GGA GGA GCG C
49	SSO (AVA2)	1951	UCU CGA ACU CGU GGC CGU UCA
50	SSO (AVA2)	1952	GGG AAG GAC AGC UUC UUG UAG UCG G
51	SSO (AVA2)	1953	GCC GCC GUC CUC GAA GUU C
52	SSO (AVA2)	1954	GUC UUC UUC UGC AUU ACG GGG C
53	SSO (AVA2)	1955	UCA CCU UCA GCU UCA CGG UGU UG
54	SSO (AVA2)	1956	CCG CCG UCC UUC AGC UUC AG
55	SSO (AVA2)	1957	GAA GAA GUC GUG CUG CUU CAU GUG GU

56	SSO (AVA2)	1958	GUC UUG UAG UUG CCG UCG UCC UUG
57	SSO (AVA2)	1959	CUU GAU GCC GUU CUU CUG CUU GUC G
58	GAPmer	1992	AUG GCA GTC CAG AGC TTT CAT TCU GU
59	GAPmer	58	G CAG ATG AAC TTC AGG GTC AGC UUG
60	siRNA	si1 (sense)	AUAACUACCUGUACAGAAUUU
61	siRNA	si1 (antisense)	AUUCUGUACAGGUAGUUAUUU
62	siRNA	si2 (sense)	GAAUGAAAGCUCUGGACUGUU
63	siRNA	si2 (antisense)	CAGUCCAGAGCUUUAUUCUU
64	sgRNA	sg1	UUUCAGGUAGCUCUUACCAG
65	sgRNA	sg2	UGUGCUCUUUUCUUUUCAGC
66	snRNA	pC4 U7	UGGCAGUCCAGAGCUUUAUUCUGUAGUUAUAGCAA CCUAUUGA

Table 4. Examples of nucleic acids

[00128] An embodiment of the DNA construct as described herein, i.e. the CALIMERO gene sequence (SEQ ID NO: 67), is shown as follows:

[00129] GGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGA
5 GAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGG
TAAACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGA
ACCGTATATAAGTGCAAGTAGTCGCCGTGAACGTTCTTTTCGCAACGGGTTTGCCGCCA
GAACACAGGTAAGTGCCGTGTGTGGTTCGCGCGGGCCTGGCCTCTTACGGGTTATGG
10 CCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGC
TTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGC
CTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGT
GGCACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAATTTTGTAT
GACCTGCTGCGACGCTTTTTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTG
CACACTGGTATTTTCGGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCAG
15 CGCACATGTTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAGAATCGGACGGGG
GTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGTCTCGCGCCGCCGTGTATCGC
CCCGCCCTGGGCGGCAAGGCTGGCCCGGTGCGCACCAAGTTGCGTGAGCGGAAAGAT
GGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGA
GAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTCCGTCCTCAGCCGTC
20 GCTTCATGTGACTCCACGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTCGA
GCTTTTGGAGTACGTCGTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCC
CCACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCC

TTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTT
CAAAGTTTTTTTCTTCATTTTCAGGTGTCGTGACAAGTTTGTACAAAAAGCAGGCTGGA
TCCACCATGGATTACAAGGATGACGATGACAAGGGGGTACCTGCCCCAAAAAAAAC
GCAAAGTGGAGGACCCAGTACCAGGATCTAGAGGTAAGTGGCAAGCTTTTGGGAGGA
5 GGCCTCAGTTAAGAGGCCACAAAGTTTGAATGCTTAGTTCTTCAGAGTTGTCAGTACG
TGCTAGCAAAGCCTCAGTTATTGCCAGCTACTGAATGTATGCCAGATGCGGGGAGAC
AGTTTCTTACCATGTTGGAGTTTGGCGAATCTTGAGAGTTATGCCTGAGACTGATCTTTT
TTTCAATAGGTTGCTATACTACCTGTACAGAATGAAAGCTCTGGACTGCCATCCGTGC
CTCTGGTAAGAGCTACCTGAAAGGTCCTTTCTAGGCTCTCTGGGATTTACATGGGCAC
10 AGAGTGGTAAACCTATAATGTATGTTAGGATTTGTCCTTCAGCACTTCTCTGGATATTA
AGAATTACCTAGAGTTTTGTAATCACTGAATTATTTAGCCAAAGGAAATTCTGTTTTTCT
CCAAAGAATCTAGGGAGTTTTGTTCTTCTTCTTATTATTTTTTTTTTAAGATTTAGTTATTT
ATTTATTTGAAAGGCAGAGAGAGAGAAAGGTCTTCCATTCACTGGTTCACTCTCCAGAT
GGCCGCAACAGCCAGAGCTGCACCGACCTGAAGCCAGGAGCTAGGAGTTTCTTCCAG
15 GTCTCCCATGTGGGTGCAGGGGCCCAAGCACTTGGACCATCATCTTCTGCTTTCCAG
GCCATAGCAGAGAACTGGATCGGAAGTGGAGCAGCCGGGACTTGAAGTGGTGTCCATA
TGGAATGCTGGCACTGCAGACCAGGGCTTTAACCCGCTGTGCCACAGCGCTGGCCCC
TGTTCTTTTCTGAGATTGTTAGTTTCATGATTGGGTAAAGGTAATTTGTTGTGGGGTTT
GGGTTTTTTTGTGTTTGTGTTTGTGTTTTTTTTTTAATTTTATTTGAAAGAGTTACACA
20 GAGAGCAGTCTTCCATCCACTGGTTCAGTCTCCAGTTGGCCACAATGGCCAGAGCTGG
GCCGATCTGAAGCCAGGAACCACTAGCTTCTTCCAGGTCTCCCATGCGAGTGCAAGGG
GCCAAGGGCTTGGGCCATCTTCTGCTGCTTTCCAGGCCACAGCAGAGAGCTGGATCA
GAAGTGGAGCAGGCAGTGACTTTACCCGCTATGCCATAGCACAGACCCTGGCAATTTG
TTCTTGACAAAAGGCTAGTGGTGTCTGGTTTTTCCACTAAGCCTAAGATAGATGGTTCT
25 TGTTGGATTTCTTCATTCTGTTACCACTTATCATATATGTTCTATGTCTACCAAAACATAG
TTTATATCTAAAGCCAGGAAGGTTAAGTAGGGTGGTAAAAGGATGTTGACGCTATCCG
TTATGCTTGGCATAGCTTACAGTTGTGCTTCTACCGCTAAGGCCATTTAGGTACAGCCG
CATACCTTTGGTTTTCTTAATTCCATTAGAATTTCACTCCAATTCAGTCAATTTAGTAC
CTATTTCTGAGTTTGGGAAAGTAAGTAAAGAAATGGTAGAAGAGGTATCATTGTGCTC
30 TTTTCTTTTTCAGCCGGTGTCTGTCGAGCTGATCAAGGAGAACATGCACATGAAGCTG
TACATGGAGGGCACCGTGAACAACCACTTCAAGTGCACATCCGAGGGCGAAGGCA
AGCCCTACGAGGGCACCCAGACCATGAAGATCAAGGTGGTTCGAGGGCGGCCCTCTCC
CCTTCGCCTTCGACATCCTGGCTACCAGCTTCATGTACGGCAGCAAAGCCTTCATCAAC
CACACCCAGGGCATCCCCGACTTCTTCAAGCAGTCTTCCCcGAGGGCTTCACATGGGA
35 GAGAATCACACATACGAAGACGGGGGCGTGCTGACCGCTACCCAGGACACCAGCTT

CCAGAACGGCTGCATCATCTACAACGTCAAGATCAACGGGGTGAACCTCCCATCCAAC
GGCCCTGTGATGCAGAAGAAAACACGCGGCTGGGAGGCCAACACCGAGATGCTGTAC
CCCGC_cGACGGCGGCCTGAGAGGCCACAGCCAGATGGCCCTGAAGCTCGTGGGCGG
GGGCTACCTGCACTGCTCCTTCAAGACCACATACAGATCCAAGAAACCCGC_cAAGAACC
5 TCAAGATGCCCCGGCTTCCACTTCGTGGACCACAGACTGGAAGAATCAAGGAGGCCGA
CAAAGAGACCTACGTCGAGCAGCAGAGATGGCTGTGGCCAAGTACTGCGACCTCCC
CAGCAA_{ACT}GGGGCAGATAA_{gtc}CGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGT
GGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCAC
10 CACCGGAAGCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCC
ATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACA
AGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACA_{ACTA}
15 CAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAC
TTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCA
CAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGC
ACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTG
GAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA_aTAAGGGC
20 CCG_{gtc}atttaattaattcctgattgttctgttcattaatcattcagatagataaagattcttaggtcaaaaaaaaaaaaaaac
aggaaatcataccctacggttccatccacacaaaccta_{aaa}ataagcattggcactaag_{tact}cagggaaagaacactgttta
gatatgtcattgtatgtgtttttcctccctctagcactggaagaagTGTGcatcatcaccatcaccacTAAGCGGCC
GCATGGAAGACGCCAAAAACATAAAGAAAGGCCCGCGCCATTCTATCCGCTGGAAGA
TGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGTTTCTGGA
25 ACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTTCGA
AATGTCCGTTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAA
TCGTGATGTCAGTGAAAACCTCTTTCAATTCTTTATGCCGGTGTGGGCGCGTTATTTA
TCGGAGTTGCAGTTGCGCCCGCAACGACATTTATAATGAACGTGAATTGCTCAACAGT
ATGGGCATTTTCGACGCTACCGTGGTGTTCGTTTCCAAAAAGGGGTGCAAAAAATTTT
30 GAACGTGCAAAAAAGCTCCCAATCATCCAAAAAATTATTATCATGGATTCTAAAACGGA
TTACCAGGGATTTTCAGTCGATGTACACGTTTCGTACATCTCATCTACCTCCCGGTTTTAA
TGAATACGATTTTGTGCCAGAGTCCTTCGATAGGGACAAGACAATTGCACTGATCATGA
ACTCCTCTGGATCTACTGGTCTGCCTAAAGGTGTGCTCTGCCTCATAGAACTGCCTGC
GTGAGATTCTCGCATGCCAGAGATCCTATTTTTGGCAATCAAATCATTCCGGATACTGC
35 GATTTTAAGTGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACTCGGATATTT

GATATGTGGATTTCGAGTCGTCTTAATGTATAGATTTGAAGAAGAGCTGTTTCTGAGGA
 GCCTTCAGGATTACAAGATTCAAAGTGCGCTGCTGGTGCCAAACCCTATTCTCCTTCTTC
 GCCAAAAGCACTCTGATTGACAAATACGATTTATCTAATTTACACGAAATTGCTTCTGGT
 GCGCTCCCCTCTCTAAGGAAGTCGGGGAAGCGGTTGCCAAGAGGTTCCATCTGCCA
 5 GGTATCAGGCAAGGATATGGGCTCACTGAGACTACATCAGCTATTCTGATTACACCCGA
 GGGGGATGATAAACCGGGCGCGGTGCGTAAAGTTGTTCCATTTTTGAAGCGAAGGTT
 GTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAAAGAGGCGAACTGTGTGTGA
 GAGGTCTATGATTATGTCCGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGATT
 GACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGACGAACACT
 10 TCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGGCTCCC
 GCTGAATTGGAATCCATCTTGCTCCAACACCCCAACATCTTCGACGCAGGTGTGCAG
 GTCTTCCCGACGATGACGCCGGTGAACCTCCCGCCGCGTTGTTGTTTTGGAGCACGG
 AAAGACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAAGTAACAACCGCG
 AAAAAGTTGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAA
 15 AACTCGACGCAAGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGAT
 CGCCGTGTAAtgTCTGagacgGGATCCAGCCTGCTTTTTGTACAAGAAAGCTGGGTGCTA
 GCACCCAGCTTTCTTGACAAAGTGGTGATCCTCAGGTGCAGGCTGCCTATCAGAAGG
 TGGTGGCTGGTGTGGCCAATGCCCTGGCTCACAAATACCACTGAGATCTTTTCCCTCT
 GCCAAAAATTATGGGGACATCATGAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAG
 20 GAAATTTATTTTCATTGCAATAGTGTGTTGGAATTTTTGTGTCTCTCACTCGGAAGGAC
 ATATGGGAGGGCAAATCATTTAAACATCAGAATGAGTATTTGGTTTAGAGTTTGGCAAC
 ATATGCCCATATGCTGGCTGCCATGAACAAAGGTTGGCTATAAAGAGGTCATCAGTATA
 TGAAACAGCCCCCTGCTGTCCATTCCCTATTCCATAGAAAAGCCTTGACTTGAGGTTAG
 ATTTTTTTATATTTGTTTTGTGTTATTTTTCTTTAACATCCCTAAAATTTTCTTACAT
 25 GTTTTACTAGCCAGATTTTCTCCTCTCCTGACTACTCCCAGTCATAGCTGTCCCTCTT
 CTCTTATGGAGATCCCTCGACCTGCAGCCCAAGCTTC

[00130] While embodiments of the invention have been particularly shown and described
 with reference to specific embodiments, it should be understood by those skilled in the art that
 various changes in form and detail may be made therein without departing from the scope of
 30 the invention as defined by the appended claims. The scope of the invention is thus indicated
 by the appended claims and all changes which come within the meaning and range of
 equivalency of the claims are therefore intended to be embraced.

CLAIMS

1. A DNA construct comprising:
 - (a) a starter exon encoding a start codon;
 - 5 (b) a first frame-switch exon located downstream of the starter exon, the first frame-switch exon comprising a first acceptor splice site, wherein inclusion or exclusion of the first frame-switch exon during alternative splicing determines the downstream reading frame of the mature mRNA transcript; and
 - 10 (c) a first specific exon located downstream of the first frame-switch exon, the first specific exon comprising a second acceptor splice site, a first reporter gene and a second reporter gene, wherein the first reporter gene is upstream or downstream of the second reporter gene and wherein the first reporter gene is in a different reading frame from the second reporter gene,
wherein the exclusion of the first frame-switch exon causes the first reporter gene to
15 be in-frame with the start codon and wherein the inclusion of the first frame-switch exon causes the second reporter gene to be in-frame with the start codon.
- 20 2. The DNA construct of claim 1, further comprising a second specific exon located downstream of the first specific exon.
3. The DNA construct of claims 1 or 2, further comprising a second frame-switch exon located downstream of the first frame-switch exon and upstream of the first specific exon, the second frame-switch exon comprising a third acceptor splice site, wherein
25 the second frame-switch exon is in a different reading frame from the first frame-switch exon and wherein the combination of inclusion or exclusion of the first frame-switch exon and the second frame-switch exon during alternative splicing determines the downstream reading frame of the mature mRNA transcript.
- 30 4. The DNA construct of any one of claims 1 to 3, further comprising a third reporter gene.
5. The DNA construct of any one of claims 1 to 4, wherein the third reporter gene is in the same reading frame as the first or second reporter genes.
- 35 6. The DNA construct of any one of claims 1 to 4, wherein the third reporter gene is in a different reading frame from the first and second reporter genes.

7. The DNA construct of any one of claims 1 to 6, further comprising a fourth reporter gene.
- 5 8. The DNA construct of claim 1, comprising the sequence of SEQ ID NO:67.
9. A method of measuring the efficacy of a nucleic acid delivery into a cell by a delivery agent, the method comprising:
- 10 (a) using the delivery agent to introduce the nucleic acid into the cell, wherein the cell expresses the DNA construct of any one of claims 1 to 8 and wherein the nucleic acid is configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct; and
- 15 (b) quantifying the functional uptake of the nucleic acid into the cell by measuring the relative expression of two of the reporter genes, wherein the functional uptake of the nucleic acid is indicative of the efficacy of delivery of the nucleic acid by the delivery agent.
- 20 10. The method of claim 9, wherein step (b) comprises quantifying the functional uptake of the nucleic acid into the cell by measuring the relative expression of the first reporter gene and the second reporter gene.
- 25 11. The method of claims 9 or 10, wherein the nucleic acid is selected from the group consisting of a steric blocker antisense oligonucleotide, a RNaseH-activating GAPmer, an siRNA, a CRISPR/Cas-sgRNA and a snRNA.
12. The method of claim 11, wherein the steric blocker antisense oligonucleotide is a splice-switching oligonucleotide (SSO).
- 30 13. The method of any one of claims 9 to 12, wherein the delivery agent is selected from the group consisting of an intra-molecular chemical modification, an inter-molecular delivery agent, an encapsulation delivery agent, a complexation delivery agent, a viral-based delivery agent, a cellular delivery agent and a sub-cellular delivery agent.

14. The method of any one of claims 9 to 13, wherein the expression of the reporter genes is measured by fluorescence, luminescence, immunohistochemistry or quantitative PCR.
- 5 15. A method for measuring the effect of chemical modification on the potency or activity of a nucleic acid, the method comprising:
- (a) introducing a chemically-modified nucleic acid into a cell, wherein the cell expresses the DNA construct of any one of claims 1 to 8 and wherein the chemically-modified nucleic acid is configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct; and
- 10 (b) quantifying the potency or activity of the chemically-modified nucleic acid by measuring the relative expression of two of the reporter genes.
- 15 16. The method of claim 15, wherein step (b) comprises quantifying the potency or activity of the chemically-modified nucleic acid by measuring the relative expression of the first reporter gene and the second reporter gene.
17. A transgenic animal that expresses the DNA construct of any one of claims 1 to 8.
- 20 18. The transgenic animal of claim 17, wherein the animal is a rodent.
19. A system for characterizing the functional cellular uptake of a nucleic acid, the system comprising:
- 25 (a) a DNA construct comprising at least two reporter genes; and
- (b) a nucleic acid configured to modify the DNA construct or to modulate the expression or processing of a pre-mRNA or mature mRNA expressed from the DNA construct such that one of the at least two reporter genes is preferentially expressed.
- 30 20. The system of claim 19, wherein the DNA construct comprises the DNA construct of any one of claims 1 to 8.