

NON-ACTIVE LIPID NANOPARTICLES WITH NON-VIRAL, CAPSID FREE DNA**RELATED APPLICATIONS**

[0001] This application claims priority to U.S. Provisional Application No. 62/814,460, filed on March 6, 2019 and U.S. Provisional Application No. 62/857,557, filed on June 5, 2019, the contents of each of which are hereby incorporated by reference in their entireties.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing, as well as sequences in Tables 1-11 herein, and each are hereby incorporated by reference in its entirety. Said ASCII copy, created on March 5, 2020, is named 131698-05020_SL.txt and is 325,465 bytes in size.

TECHNICAL FIELD

[0003] The present invention is directed to compositions and methods for delivery of non-viral, capsid-free DNA vectors to the cytosol of a target cell.

BACKGROUND

[0004] Gene therapy aims to improve clinical outcomes for patients suffering from either genetic mutations or acquired diseases caused by an aberration in the gene expression profile. Gene therapy includes the treatment or prevention of medical conditions resulting from defective genes or abnormal regulation or expression, *e.g.* underexpression or overexpression, that can result in a disorder, disease, malignancy, *etc.* For example, a disease or disorder caused by a defective gene might be treated, prevented or ameliorated by delivery of a corrective genetic material to a patient, or might be treated, prevented or ameliorated by altering or silencing a defective gene, *e.g.*, with a corrective genetic material to a patient resulting in the therapeutic expression of the genetic material within the patient.

[0005] The basis of gene therapy is to supply a transcription cassette with an active gene product (sometimes referred to as a transgene), *e.g.*, that can result in a positive gain-of-function effect, a negative loss-of-function effect, or another outcome. Such outcomes can be attributed to expression of a therapeutic protein, *e.g.*, an antibody, functional enzyme, or fusion protein. Gene therapy can also be used to treat a disease or malignancy caused by other factors. Human monogenic disorders can be treated by the delivery and expression of a normal gene to the target cells. Delivery and expression of a corrective gene in the patient's target cells can be carried out *via* numerous methods, including the use of engineered viruses and viral gene delivery vectors. Among the many virus-derived vectors available (*e.g.*, recombinant retrovirus, recombinant lentivirus, recombinant

adenovirus, and the like), recombinant adeno-associated virus (rAAV) is gaining popularity as a versatile vector in gene therapy.

[0006] Adeno-associated viruses (AAV) belong to the Parvoviridae family and more specifically constitute the Dependoparvovirus genus. Vectors derived from AAV (*i.e.*, recombinant AAV (rAAV) or AAV vectors) are attractive for delivering genetic material because (i) they are able to infect (transduce) a wide variety of non-dividing and dividing cell types including myocytes and neurons; (ii) they are devoid of the virus structural genes, thereby diminishing the host cell responses to virus infection, *e.g.*, interferon-mediated responses; (iii) wild-type viruses are considered non-pathogenic in humans; (iv) in contrast to wild type AAV, which are capable of integrating into the host cell genome, replication-deficient AAV vectors lack the rep gene and generally persist as episomes, thus limiting the risk of insertional mutagenesis or genotoxicity; and (v) in comparison to other vector systems, AAV vectors are generally considered to be relatively poor immunogens and therefore do not trigger a significant immune response (see ii), thus gaining persistence of the vector DNA and potentially, long-term expression of the therapeutic transgenes.

[0007] However, there are several major deficiencies in using AAV particles as a gene delivery vector. One major drawback associated with rAAV is its limited viral packaging capacity of about 4.5 kb of heterologous DNA (Dong *et al.*, 1996; Athanasopoulos *et al.*, 2004; Lai *et al.*, 2010), and as a result, use of AAV vectors has been limited to less than 150,000 Da protein coding capacity. The second drawback is that as a result of the prevalence of wild-type AAV infection in the population, candidates for rAAV gene therapy have to be screened for the presence of neutralizing antibodies that eliminate the vector from the patient. A third drawback is related to the capsid immunogenicity that prevents re-administration to patients that were not excluded from an initial treatment. The immune system in the patient can respond to the vector which effectively acts as a “booster” shot to stimulate the immune system generating high titer anti-AAV antibodies that preclude future treatments. Some recent reports indicate concerns with immunogenicity in high dose situations. Another notable drawback is that the onset of AAV-mediated gene expression is relatively slow, given that single-stranded AAV DNA must be converted to double-stranded DNA prior to heterologous gene expression.

[0008] Additionally, conventional AAV virions with capsids are produced by introducing a plasmid or plasmids containing the AAV genome, rep genes, and cap genes (Grimm *et al.*, 1998). However, such encapsidated AAV virus vectors were found to inefficiently transduce certain cell and tissue types and the capsids also induce an immune response. Accordingly, use of adeno-associated virus (AAV) vectors for gene therapy is limited due to the single administration to patients (owing to the patient immune response), the limited range of transgene genetic material suitable for delivery in AAV vectors due to minimal viral packaging capacity (about 4.5kb), and slow AAV-mediated gene expression.

[0009] Recently, non-viral, capsid-free DNA vectors with covalently-closed ends (ceDNA) that contain transgenes flanked by AAV 2 ITRs were reported. However, delivery of these DNA vectors to cells, *in vitro* and *in vivo*, remains challenging.

[0010] Although conceptually elegant, the prospect of using nucleic-acid molecules for gene therapy for treating human diseases remains uncertain. The main cause of this uncertainty is the apparent adverse events relating to host's innate immune response to nucleic acid therapeutics and, thus, the way in which these materials modulate expression of their intended targets in the context of the immune response. The current state of the art surrounding the creation, function, behavior and optimization of nucleic acid molecules that may be adopted for clinical applications has a particular focus on: (1) antisense oligonucleotides and duplex RNAs that directly regulate translation and gene expression; (2) transcriptional gene silencing RNAs that result in long-term epigenetic modifications; (3) antisense oligonucleotides that interact with and alter gene splicing patterns; (4) creation of synthetic or viral vectors that mimic physiological functionalities of naturally occurring AAV or lentiviral genome; and (5) the *in vivo* delivery of therapeutic oligonucleotides. However, despite the advances made in the development of nucleic acid therapeutics that are evident in recent clinical achievements, the field of gene therapy is still severely limited by unwanted adverse events in recipients triggered by the therapeutic nucleic acids, themselves.

[0011] Accordingly, there remains a need in the art for methods and formulations that address these challenges.

SUMMARY

[0012] In recent years it has emerged that foreign nucleic acids, *e.g.*, DNA potently stimulates the innate immune response, particularly type 1 interferon (IFN) production. This occurs through a pathway dependent upon DNA sensor cyclic guanosine monophosphate-adenosine monophosphate synthase (cGAS) and the downstream adaptor protein stimulation of IFN genes (STING). See, for example Thomsen *et al.*, Hepatology (2016), 64(3):746-759. Inventors have discovered, *inter alia*, that innate immune response to foreign DNA, such as ceDNA, delivered by lipid nanoparticle formulations can be attenuated, reduced or inhibited by sequestering cytosolic release of foreign DNA, such as ceDNA, to desired cells. Without wishing to be bound by a theory, the method limits expression of foreign DNA, such as ceDNA, to the desired cells. This can reduce or inhibit the innate immune response. For example, when the foreign DNA, such as ceDNA is delivered to cells that lack or do not express a functional innate DNA-sensing pathway, for example, a cell that lacks or does not express functional cGAS and/or STING, no innate immune response is produced. When the foreign DNA, such as ceDNA is delivered to cells that do express a functional innate DNA-sensing pathway, the innate immune response is limited to the response produced by those cells, avoiding the innate immune response in other cells where ceDNA expression is not desired.

[0013] Thus, in one aspect, provided herein is the use of a capsid free, non-viral vector in the manufacture of a medicament for delivering the capsid free, non-viral vector to the cytosol of a target cell within a subject, wherein the capsid free, non-viral vector is encapsulated in a lipid nanoparticle (LNP), wherein the LNP lacks fusogenic activity; and wherein the capsid free, non-viral vector and an endosomolytic agent are to be co-administered. In another aspect, provided herein is the use of a capsid free, non-viral vector in the manufacture of a medicament for delivering the capsid free, non-viral vector to the cytosol of a target cell within a subject, wherein the capsid free, non-viral vector is encapsulated in a lipid nanoparticle (LNP), wherein the LNP lacks fusogenic activity; wherein the capsid free, non-viral vector and an endosomolytic agent are to be co-administered, and wherein at least one of the lipid nanoparticle and the endosomolytic agent includes a first targeting ligand that binds to a molecule on the surface of the target cell. In another example, provided herein are methods for delivering a capsid free, non-viral vector (ceDNA) to the cytosol of a target cell within a subject. Generally, the method comprises co-administering to the subject: (a) a capsid free, non-viral vector encapsulated in a non-fusogenic lipid nanoparticle (LNP); and (b) an endosomolytic agent. According to some examples, the capsid free, non-viral vector when digested with a restriction enzyme having a single recognition site on the DNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA when analyzed on a non-denaturing gel. According to some examples, the endosomolytic agent targets the target cell. According to some examples, the capsid free, non-viral vector is translocated to nucleus of the cell after administration. According to some examples, the LNP releases less than 10% of the ceDNA comprised therein at endosomal pH. According to some examples, the LNP does not induce an immune response when administered without the endosomolytic agent. According to some examples, the target cell is a cell that lacks or does not express a functional innate DNA-sensing pathway, or which has reduced innate DNA-sensing pathway activity. According to some examples, the target cell is a cell that lacks or does not express functional cGAS and/or STING, or which has reduced cGAS and/or STING activity. According to some examples, the target cell is a hepatocyte. According to some examples, the endosomolytic agent is a membrane-destabilizing polymer. According to some examples, the membrane-destabilizing polymer is a copolymer, a peptide, a membrane-destabilizing toxin or a derivative thereof, or a viral fusogenic peptide or derivative thereof. According to some examples, the endosomolytic agent is a pH-sensitive polymer. According to some examples, the endosomolytic agent is a polyanionic peptide, polycationic peptide, amphipathic peptide, hydrophobic peptide or a peptidomimetic.

[0014] Without wishing to be bound by a theory, the lipid nanoparticle functions to encapsulate the capsid free, non-viral vector, preventing its interaction with various components of the systemic circulation. Thus, the LNP acts to shield the encapsulated vector from degradation, clearance, or unwanted interactions. Generally, the lipid nanoparticle is non-fusogenic. In other words, the lipid nanoparticle does not have, or has very little, fusogenic activity that would enable it

to fuse with and consequently destabilize a membrane. As described herein, a non-fusogenic lipid nanoparticle refers to a nanoparticle that does not or substantially does not fuse with a membrane or, if it does fuse with a membrane, does not destabilize the membrane. For example, the lipid nanoparticle does not comprise a component having fusogenic activity. Further, the lipid nanoparticle does not have, or has very little, fusogenic activity at any pH, such low pH (*e.g.*, about pH 6.5 or lower), neutral pH (*e.g.*, about pH 7-8, *e.g.*, pH 7, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9 or 8) or high pH (*e.g.*, about pH 8.5 or higher.)

[0015] In some examples, the fusogenic activity of the lipid nanoparticle differs by less than 10%, *e.g.*, less than 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% or 1% at a low pH vs neutral pH as measured by a membrane-impermeable fluorescent dye exclusion assay. In some examples, the fusogenic activity of the lipid nanoparticle is substantially the same, *e.g.*, differs by less than 0.5%, 0.25%, 0.1% or an undetectable amount at a low pH vs neutral pH as measured by a membrane-impermeable fluorescent dye exclusion assay, *e.g.*, the assay described in the Examples section.

[0016] Various methods are known in the art for determining fusogenic activity or lack thereof. For example, fusogenic activity or lack thereof can be determined in an *in vitro* cell assay, such as the red blood cell hemolysis assay or a liposomal leakage assay. A two-step assay can also be performed, wherein a first assay evaluates the fusogenic activity of the lipid nanoparticle constituents alone, and a second assay evaluates the fusogenic activity of the assembled nanoparticle.

[0017] Lipid nanoparticles are typically used as carriers for nucleic acid delivery in the context of pharmaceutical development. Generally, lipid nanoparticle compositions for such delivery are composed of synthetic ionizable or cationic lipids. In addition to the ionizable lipid, the lipid nanoparticles can comprise one or more phospholipids, especially compounds having a phosphatidylcholine group, cholesterol and a polyethylene glycol (PEG) lipid. However, these compositions may also include other lipids. The sum composition of lipids typically dictates the surface characteristics and thus the protein (opsonization) content in biological systems; thus, driving biodistribution and cell uptake properties. The lipid nanoparticles of present disclosure differ from the lipid nanoparticles typically used as carriers for nucleic acid delivery in the art. The focus in the art is on lipid nanoparticles that can fuse with and destabilize cell membranes so that any nucleic acid encapsulated in the lipid nanoparticle can be released into the cell. Stated in another way, taken as a whole, the prior art teaches against using non-fusogenic lipid nanoparticles for delivering nucleic acids in to cells. In contrast, the lipid nanoparticles of the disclosure lack fusogenic activity. In other words, the lipid nanoparticles of the disclosure go against the common knowledge in the art suggesting to use lipid nanoparticles having fusogenic activity.

[0018] Generally, the non-fusogenic lipid nanoparticle is an inactive lipid nanoparticle. An “inactive lipid nanoparticle” means a lipid nanoparticle that does not release encapsulated ceDNA. For example, the inactive lipid nanoparticle releases less than 10%, *e.g.*, less than 5%, 4%, 3%, 2% or 1% of the encapsulated ceDNA at an acidic pH, *e.g.*, pH 6, as measured by a membrane-impermeable

fluorescent dye exclusion assay, *e.g.*, the assay described in the Examples section. In some examples, the nanoparticle releases substantially no ceDNA, *e.g.*, less than 0.5%, 0.25%, 0.1% or an undetectable amount, of the encapsulated ceDNA at an acidic pH, *e.g.*, pH 6, as measured by a membrane-impermeable fluorescent dye exclusion assay, *e.g.*, the assay described in the Examples section.

[0019] In some examples, the inactive lipid nanoparticle releases less than 10%, *e.g.*, less than 5%, 4%, 3%, 2% or 1% of the encapsulated ceDNA at endosomal pH. For example, the nanoparticle releases substantially no ceDNA, *e.g.*, less than 0.5%, 0.25%, 0.1% or an undetectable amount, of the encapsulated ceDNA at endosomal pH.

[0020] In some examples, the nanoparticle releases less than 10%, *e.g.*, less than 5%, 4%, 3%, 2% or 1% of the encapsulated ceDNA into the cytoplasm when the nanoparticle is administered alone relative to when the lipid nanoparticle is co-administered with an endosomolytic agent. In some further examples of this, the nanoparticle releases substantially no ceDNA, *e.g.*, less than 0.5%, 0.25%, 0.1% or an undetectable amount, of the encapsulated ceDNA into the cytoplasm when the nanoparticle is administered alone relative to when the lipid nanoparticle is co-administered with an endosomolytic agent.

[0021] Without wishing to be bound by a theory, the endosomolytic agent targets the target cell. For example, the endosomolytic agent preferentially or specifically binds to and/or is taken-up by the target cell relative to a non-target cell. For example, uptake of the endosomolytic agent by the target cell is at least 1-fold, 10-folds, 25-folds, 50-fold, 75-folds, 100-folds, 250-folds, 500-folds, 750-folds, 1000-folds or higher than uptake by a non-target cell. In some examples, the endosomolytic agent is preferentially or specifically taken up by a cell that lacks or does not express a functional innate DNA-sensing pathway. For example, the endosomolytic agent is preferentially or specifically taken up by a cell that lacks or does not express functional cGAS and/or STING.

[0022] In various examples, at least one of the endosomolytic agent and the lipid nanoparticle includes a first targeting ligand that specifically binds to a molecule on surface of the target cell. In some preferred examples, the endosomolytic agent includes the first targeting ligand.

[0023] Without limitations, the lipid nanoparticle and endosomolytic agent can be administered separately or within a single composition. When the lipid nanoparticle and endosomolytic agent are administered separately, they can be administered in any order. For example, the nanoparticle can be administered prior to administering the endosomolytic agent or the nanoparticle can be administered after administering the endosomolytic agent.

[0024] In some examples, the endosomolytic agent and the lipid nanoparticle are formulated into separate compositions for administering. When the endosomolytic agent and the lipid nanoparticle are formulated into separate compositions the two compositions can either be simultaneously administered or they can be sequentially or subsequently administered.

[0025] In some examples, at least one of the lipid nanoparticle and the endosomolytic agent is administered in a repeat dosage regime (*e.g.*, a weekly or bi-weekly repeated administration protocol). In some other examples, both the lipid nanoparticle and the endosomolytic agent are administered in a repeat dosage regime (*e.g.*, a weekly or bi-weekly repeated administration protocol).

[0026] In another aspect, also provided herein is a composition comprising: (a) a capsid free, non-viral vector encapsulated in a lipid nanoparticle (LNP), wherein the LNP lacks fusogenic activity; and (b) an endosomolytic agent. According to some examples, the capsid free, non-viral vector when digested with a restriction enzyme having a single recognition site on the DNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA when analyzed on a non-denaturing gel. According to some examples, the endosomolytic agent targets the target cell. According to some examples, the capsid free, non-viral vector is translocated to nucleus of the cell after administration. According to some examples, the LNP releases less than 10% of the ceDNA comprised therein at endosomal pH. According to some examples, the LNP does not induce an immune response when administered without the endosomolytic agent. According to some examples, the target cell is a cell that lacks or does not express a functional innate DNA-sensing pathway, or which has reduced innate DNA-sensing pathway activity. According to some examples, the target cell is a cell that lacks or does not express functional cGAS and/or STING, or which has reduced cGAS and/or STING activity. According to some examples, the target cell is a hepatocyte. According to some examples, the endosomolytic agent is a membrane-destabilizing polymer. According to some examples, the membrane-destabilizing polymer is a copolymer, a peptide, a membrane-destabilizing toxin or a derivative thereof, or a viral fusogenic peptide or derivative thereof. According to some examples, the endosomolytic agent is a pH-sensitive polymer. According to some examples, the endosomolytic agent is a polyanionic peptide, polycationic peptide, amphipathic peptide, hydrophobic peptide or a peptidomimetic. According to some examples, the endosomolytic agent is in form of a nanoparticle. According to some examples, the nanoparticle further comprises a cationic lipid, a non-cationic lipid, a sterol or a derivative thereof, a conjugated lipid, or any combination thereof. According to some examples, the lipid nanoparticle comprises the endosomolytic agent. According to some examples, the endosomolytic agent is preferentially or specifically taken up by the target cell relative to a non-target cell. According to some examples, the endosomolytic agent is preferentially or specifically taken up by a cell that lacks or does not express a functional innate DNA-sensing pathway. According to some examples, the endosomolytic agent is preferentially or specifically taken up by a cell that lacks or does not express functional cGAS and/or STING. According to some examples, at least one of the endosomolytic agent and the lipid nanoparticle includes a first targeting ligand. According to some examples, the targeting ligand binds to a cell surface molecule on a cell that lacks or does not express a functional innate DNA-sensing pathway or wherein the targeting ligand binds to a cell surface molecule on a cell that lacks or does not express functional cGAS and/or STING. According to some

examples, the endosomolytic agent includes the first targeting ligand. According to some examples, one of the lipid nanoparticle and endosomolytic agent includes the first targeting ligand, and the other of the lipid nanoparticle and endosomolytic agent includes a second targeting ligand. According to some examples, the first and the second targeting ligand recognize and bind to same cell surface molecule. According to some examples, the first and the second targeting ligand are the same or wherein the first and the second targeting ligand are different. According to some examples, the first and the second targeting ligand recognize and bind to same cell surface molecule. According to some examples, the ligand binds to a cell surface molecule selected from the group consisting of a transferrin receptor type 1, transferrin receptor type 2, the EGF receptor, HER2/Neu, a VEGF receptor, a PDGF receptor, an integrin, an NGF receptor, CD2, CD3, CD4, CD8, CD19, CD20, CD22, CD33, CD43, CD38, CD56, CD69, the asialoglycoprotein receptor (ASGPR), GalNAc receptor, prostate-specific membrane antigen (PSMA), a folate receptor, and a sigma receptor. According to some examples, the cell surface molecule is asialoglycoprotein receptor (ASGPR) or GalNAc receptor. According to some examples, the targeting ligand is a monovalent or multivalent D-galactose or N-acetyl-D-galactose. According to some examples, the composition further comprises an additional compound. According to some examples, said additional compound is encompassed in a lipid nanoparticle, and wherein said lipid nanoparticle is different from the nanoparticle comprising the ceDNA. According to some examples, said additional compound is encompassed in the lipid nanoparticle comprising the ceDNA. According to some examples, said additional compound and the endosomolytic agent are comprised in a nanoparticle. According to some examples, said additional compound is a therapeutic agent. According to some examples, said additional compound is an immune modulating agent. According to some examples, the immune modulating agent is an immunosuppressant. According to some examples, the immune modulating agent is selected from the group consisting of like cGAS inhibitors, TLR9 antagonists, Caspase-1 inhibitors, and any combination thereof. According to some examples, said additional compound is a second capsid free, non-viral vector, wherein the first and second capsid free, non-viral vectors are different. According to some examples, the capsid free, non-viral vector is a close-ended DNA (ceDNA) vector comprising at least one heterologous nucleotide sequence between flanking inverted terminal repeats (ITRs), wherein at least one heterologous nucleotide sequence encodes at least one transgene or therapeutic protein of interest. In another aspect, provided herein is a use of the composition the disclosure in the manufacture of a medicament for expressing a desired transgene or therapeutic protein in a cell *in vivo* comprising contacting the cell with the composition of the disclosure. In another aspect, provided herein is a cell containing a composition of the disclosure. According to some examples, the least one heterologous nucleotide sequence that encodes at least one transgene or therapeutic protein is a nucleic acid RNAi agent. In another aspect, the present disclosure provides an *in vitro* method of expressing a desired transgene or therapeutic protein in a cell, comprising contacting the cell with the composition as disclosed herein.

[0027] Other aspects and examples of the disclosure as described herein.

DESCRIPTION OF THE DRAWINGS

[0028] Embodiments of the present disclosure, briefly summarized above and discussed in greater detail below, can be understood by reference to the illustrative embodiments of the disclosure depicted in the appended drawings. However, the appended drawings illustrate only typical embodiments of the disclosure and are therefore not to be considered limiting of scope, for the disclosure may admit to other equally effective embodiments.

[0029] **FIG. 1A** illustrates an exemplary structure of a ceDNA vector for expression of a transgene as disclosed herein, comprising asymmetric ITRs. In this embodiment, the exemplary ceDNA vector comprises an expression cassette containing CAG promoter, WPRE, and BGHpA. An open reading frame (ORF) encoding a transgene can be inserted into the cloning site (R3/R4) between the CAG promoter and WPRE. The expression cassette is flanked by two inverted terminal repeats

(ITRs) – the wild-type AAV2 ITR on the upstream (5'-end) and the modified ITR on the downstream (3'-end) of the expression cassette, therefore the two ITRs flanking the expression cassette are asymmetric with respect to each other.

[0030] **FIG. 1B** illustrates an exemplary structure of a ceDNA vector for expression a transgene as disclosed herein comprising asymmetric ITRs with an expression cassette containing CAG promoter, WPRE, and BGHpA. An open reading frame (ORF) encoding the transgene can be inserted into the cloning site between CAG promoter and WPRE. The expression cassette is flanked by two inverted terminal repeats (ITRs) – a modified ITR on the upstream (5'-end) and a wild-type ITR on the downstream (3'-end) of the expression cassette.

[0031] **FIG. 1C** illustrates an exemplary structure of a ceDNA vector for expression of a transgene as disclosed herein comprising asymmetric ITRs, with an expression cassette containing an enhancer/promoter, the transgene, a post transcriptional element (WPRE), and a polyA signal. An open reading frame (ORF) allows insertion of transgene encoding a protein of interest, or therapeutic nucleic acid into the cloning site between CAG promoter and WPRE. The expression cassette is flanked by two inverted terminal repeats (ITRs) that are asymmetrical with respect to each other; a modified ITR on the upstream (5'-end) and a modified ITR on the downstream (3'-end) of the expression cassette, where the 5' ITR and the 3' ITR are both modified ITRs but have different modifications (*i.e.*, they do not have the same modifications).

[0032] **FIG. 1D** illustrates an exemplary structure of a ceDNA vector for expression of a transgene as disclosed herein, comprising symmetric modified ITRs, or substantially symmetrical modified ITRs as defined herein, with an expression cassette containing CAG promoter, WPRE, and BGHpA. An open reading frame (ORF) encoding the transgene is inserted into the cloning site between CAG promoter and WPRE. The expression cassette is flanked by two modified inverted terminal repeats (ITRs), where the 5' modified ITR and the 3' modified ITR are symmetrical or substantially symmetrical.

[0033] **FIG. 1E** illustrates an exemplary structure of a ceDNA vector for expression of a transgene as disclosed herein comprising symmetric modified ITRs, or substantially symmetrical modified ITRs as defined herein, with an expression cassette containing an enhancer/promoter, a transgene, a post transcriptional element (WPRE), and a polyA signal. An open reading frame (ORF) allows insertion of a transgene into the cloning site between CAG promoter and WPRE. The expression cassette is flanked by two modified inverted terminal repeats (ITRs), where the 5' modified ITR and the 3' modified ITR are symmetrical or substantially symmetrical.

[0034] **FIG. 1F** illustrates an exemplary structure of a ceDNA vector for expression of a transgene as disclosed herein, comprising symmetric WT-ITRs, or substantially symmetrical WT-ITRs as defined herein, with an expression cassette containing CAG promoter, WPRE, and BGHpA. An open reading frame (ORF) encoding a transgene is inserted into the cloning site between CAG

promoter and WPRE. The expression cassette is flanked by two wild type inverted terminal repeats (WT-ITRs), where the 5' WT-ITR and the 3' WT ITR are symmetrical or substantially symmetrical.

[0035] FIG. 1G illustrates an exemplary structure of a ceDNA vector for expression of a transgene as disclosed herein, comprising symmetric modified ITRs, or substantially symmetrical modified ITRs as defined herein, with an expression cassette containing an enhancer/promoter, a transgene, a post transcriptional element (WPRE), and a polyA signal. An open reading frame (ORF) allows insertion of a transgene into the cloning site between CAG promoter and WPRE. The expression cassette is flanked by two wild type inverted terminal repeats (WT-ITRs), where the 5' WT-ITR and the 3' WT ITR are symmetrical or substantially symmetrical.

[0036] FIG. 2A provides the T-shaped stem-loop structure of a wild-type left ITR of AAV2 (SEQ ID NO: 52) with identification of A-A' arm, B-B' arm, C-C' arm, two Rep binding sites (RBE and RBE') and also shows the terminal resolution site (*trs*). The RBE contains a series of 4 duplex tetramers that are believed to interact with either Rep 78 or Rep 68. In addition, the RBE' is also believed to interact with Rep complex assembled on the wild-type ITR or mutated ITR in the construct. The D and D' regions contain transcription factor binding sites and other conserved structure. **FIG. 2B** shows proposed Rep-catalyzed nicking and ligating activities in a wild-type left ITR (SEQ ID NO: 53), including the T-shaped stem-loop structure of the wild-type left ITR of AAV2 with identification of A-A' arm, B-B' arm, C-C' arm, two Rep Binding sites (RBE and RBE') and also shows the terminal resolution site (*trs*), and the D and D' region comprising several transcription factor binding sites and other conserved structure.

[0037] FIG. 3A provides the primary structure (polynucleotide sequence) (left) and the secondary structure (right) of the RBE-containing portions of the A-A' arm, and the C-C' and B-B' arm of the wild type left AAV2 ITR (SEQ ID NO: 54). **FIG. 3B** shows an exemplary mutated ITR (also referred to as a modified ITR) sequence for the left ITR. Shown is the primary structure (left) and the predicted secondary structure (right) of the RBE portion of the A-A' arm, the C arm and B-B' arm of an exemplary mutated left ITR (ITR-1, left) (SEQ ID NO: 113). **FIG. 3C** shows the primary structure (left) and the secondary structure (right) of the RBE-containing portion of the A-A' loop, and the B-B' and C-C' arms of wild type right AAV2 ITR (SEQ ID NO: 55). **FIG. 3D** shows an exemplary right modified ITR. Shown is the primary structure (left) and the predicted secondary structure (right) of the RBE containing portion of the A-A' arm, the B-B' and the C arm of an exemplary mutant right ITR (ITR-1, right) (SEQ ID NO: 114). Any combination of left and right ITR (*e.g.*, AAV2 ITRs or other viral serotype or synthetic ITRs) can be used as taught herein. Each of **FIGS. 3A-3D** polynucleotide sequences refer to the sequence used in the plasmid or bacmid/baculovirus genome used to produce the ceDNA as described herein. Also included in each of **FIGS. 3A-3D** are corresponding ceDNA secondary structures inferred from the ceDNA vector configurations in the plasmid or bacmid/baculovirus genome and the predicted Gibbs free energy values.

[0038] FIG. 4A is a schematic illustrating an upstream process for making baculovirus infected insect cells (BIICs) that are useful in the production of a ceDNA vector for expression of a transgene as disclosed herein in the process described in the schematic in FIG. 4B. FIG. 4B is a schematic of an exemplary method of ceDNA production and FIG. 4C illustrates a biochemical method and process to confirm ceDNA vector production. FIG. 4D and FIG. 4E are schematic illustrations describing a process for identifying the presence of ceDNA in DNA harvested from cell pellets obtained during the ceDNA production processes in FIG. 4B. FIG. 4D shows schematic expected bands for an exemplary ceDNA either left uncut or digested with a restriction endonuclease and then subjected to electrophoresis on either a native gel or a denaturing gel. The leftmost schematic is a native gel, and shows multiple bands suggesting that in its duplex and uncut form ceDNA exists in at least monomeric and dimeric states, visible as a faster-migrating smaller monomer and a slower-migrating dimer that is twice the size of the monomer. The schematic second from the left shows that when ceDNA is cut with a restriction endonuclease, the original bands are gone and faster-migrating (*e.g.*, smaller) bands appear, corresponding to the expected fragment sizes remaining after the cleavage. Under denaturing conditions, the original duplex DNA is single-stranded and migrates as a species twice as large as observed on native gel because the complementary strands are covalently linked. Thus in the second schematic from the right, the digested ceDNA shows a similar banding distribution to that observed on native gel, but the bands migrate as fragments twice the size of their native gel counterparts. The rightmost schematic shows that uncut ceDNA under denaturing conditions migrates as a single-stranded open circle, and thus the observed bands are twice the size of those observed under native conditions where the circle is not open. In this figure “kb” is used to indicate relative size of nucleotide molecules based, depending on context, on either nucleotide chain length (*e.g.*, for the single stranded molecules observed in denaturing conditions) or number of basepairs (*e.g.*, for the double-stranded molecules observed in native conditions). FIG. 4E shows DNA having a non-continuous structure. The ceDNA can be cut by a restriction endonuclease, having a single recognition site on the ceDNA vector, and generate two DNA fragments with different sizes (1kb and 2kb) in both neutral and denaturing conditions. FIG. 4E also shows a ceDNA having a linear and continuous structure. The ceDNA vector can be cut by the restriction endonuclease, and generate two DNA fragments that migrate as 1kb and 2kb in neutral conditions, but in denaturing conditions, the stands remain connected and produce single strands that migrate as 2kb and 4kb.

[0039] FIG. 5 is an exemplary picture of a denaturing gel running examples of ceDNA vectors with (+) or without (-) digestion with endonucleases (EcoRI for ceDNA construct 1 and 2; BamHI for ceDNA construct 3 and 4; SpeI for ceDNA construct 5 and 6; and XhoI for ceDNA construct 7 and 8) Constructs 1-8 are described in Example 1 of International Application PCT PCT/US18/49996, which is incorporated herein in its entirety by reference. Sizes of bands highlighted with an asterisk were determined and provided on the bottom of the picture.

[0040] **FIG. 6** depicts the results of the experiments described in Example 7 and specifically shows the IVIS images obtained from mice treated with LNP-polyC control (mouse furthest to the left) and four mice treated with LNP-ceDNA-Luciferase (all but the mouse furthest to the left). The four ceDNA-treated mice show significant fluorescence in the liver-containing region of the mouse.

[0041] **FIG. 7** depicts the results of the experiment described in Example 8. The dark specks indicate the presence of the protein resulting from the expressed ceDNA transgene and demonstrate association of the administered LNP-ceDNA with hepatocytes.

[0042] **FIGS. 8A and 8B** depict the results of the ocular studies set forth in Example 9. **FIG. 8A** shows representative IVIS images from JetPEI®-ceDNA-Luciferase-injected rat eyes (upper left) versus uninjected eye in the same rat (upper right) or plasmid-Luciferase DNA-injected rat eye (lower left) and the uninjected eye in that same rat (lower right). **FIG. 8B** shows a graph of the average radiance observed in treated eyes or the corresponding untreated eyes in each of the treatment groups. The ceDNA-treated rats demonstrated prolonged significant fluorescence (and hence luciferase transgene expression) over 99 days, in sharp contrast to rats treated with plasmid-luciferase where minimal relative fluorescence (and hence luciferase transgene expression) was observed.

[0043] **FIGS. 9A and 9B** depict the results of the ceDNA persistence and redosing study in Rag2 mice described in Example 10. **FIG. 9A** shows a graph of total flux over time observed in LNP-ceDNA-Luc-treated wild-type c57bl/6 mice or Rag2 mice. **FIG. 9B** provides a graph showing the impact of redose on expression levels of the luciferase transgene in Rag2 mice, with resulting increased stable expression observed after redose (arrow indicates time of redose administration).

[0044] **FIG. 10** provides data from the ceDNA luciferase expression study in treated mice described in Example 11, showing total flux in each group of mice over the duration of the study. High levels of unmethylated CpG correlated with lower total flux observed in the mice over time, while use of a liver-specific promoter correlated with durable, stable expression of the transgene from the ceDNA vector over at least 77 days.

[0045] **FIG. 11** is a bar graph showing particle size of exemplary LNPs.

[0046] **FIG. 12** is a bar graph showing zeta potential of exemplary LNPs.

[0047] **FIG. 13** is a bar graph showing encapsulation efficiency of some exemplary LNPs.

[0048] **FIG. 14** is a bar graph showing ceDNA release from exemplary LNPs when incubated with anionic liposomes at pH 7.4 and pH 6.0.

[0049] **FIGs. 15A, 15B, and 15C** show the effect of exemplary LNPs on bodyweight (**FIG. 15A**) and liver enzymes: ALT (**FIG. 15B**) and AST (**FIG. 15C**).

[0050] **FIG. 16** shows inactive LNPs encapsulating ceDNA attenuate or restrict cytokine stimulation.

DETAILED DESCRIPTION

[0051] Nucleic acid transfer vectors and therapeutic agents are promising therapeutics for a variety of applications, such as gene expression and modulation thereof. Viral transfer vectors may comprise transgenes that encode proteins or nucleic acids. Examples of such include AAV vectors, microRNA (miRNA), small interfering RNA (siRNA), as well as antisense oligonucleotides that bind mutation sites in messenger RNA (such as small nuclear RNA (snRNA)). Unfortunately, the promise of these therapeutics has not yet been realized, in large part due to cellular and humoral immune responses directed against the viral transfer vector. These immune responses include antibody, B cell and T cell responses, and are often specific to viral antigens of the viral transfer vector, such as viral capsid or coat proteins or peptides thereof.

[0052] Currently, many potential patients harbor some level of pre-existing immunity against the viruses on which viral transfer vectors are based. In fact, antibodies against viral nucleic acids (both DNA and RNA) or protein are highly prevalent in the human population. In addition, even if the level of pre-existing immunity is low, for example, due to the low immunogenicity of the viral transfer vector, such low levels may still prevent successful transduction (*e.g.*, Jeune, *et al.*, Human Gene Therapy Methods, 24:59-67 (2013)). Thus, even low levels of pre-existing immunity may hinder the use of a specific viral transfer vector in a patient, and may require a clinician to choose a viral transfer vector based on a virus of a different serotype that may not be as efficacious, or even opt out for a different type of therapy altogether if another viral transfer vector therapy is not available.

[0053] Additionally, viral vectors, such as adeno-associated vectors, can be highly immunogenic and elicit humoral and cell-mediated immunity that can compromise efficacy, particularly with respect to re-administration. In fact, cellular and humoral immune responses against a viral transfer vector can develop after a single administration of the viral transfer vector. After viral transfer vector administration, neutralizing antibody titers can increase and remain high for several years, and can reduce the effectiveness of re-administration of the viral transfer vector. Indeed, repeated administration of a viral transfer vector generally results in enhanced, undesired immune responses. In addition, viral transfer vector-specific CD8⁺ T cells may arise and eliminate transduced cells expressing a desired transgene product, for example, on re-exposure to a viral antigen like viral nucleic acid or capsid protein. For example, it has been shown that AAV nucleic acids or capsid antigens can trigger immune-mediated destruction of hepatocytes transduced with an AAV viral transfer vector. For many therapeutic applications, it is thought that multiple rounds of administration of viral transfer vectors are needed for long-term benefits. The ability to do so, however, would be severely limited, particularly if re-administration is needed, without the methods and compositions provided herein.

[0054] Methods and compositions are provided that offer solutions to the aforementioned obstacles to effective use of variety of nucleic acid therapeutics, including viral or non-viral (synthetic) transfer vectors, and other nucleic acid therapeutics for treatment.

[0055] The present disclosure relates, *inter alia*, to formulations and methods for delivery of capsid free, non-viral closed-ended vectors (ceDNA vectors) to the cytosol of a target cell within a subject. Generally, the capsid free, non-viral vector is formulated in a lipid nanoparticle and either an endosomolytic agent is added to the formulation (a co-formulation for co-injection of lipid nanoparticle and endosomolytic agent) or the lipid nanoparticle and the endosomolytic agents are used separately via separate (*e.g.*, co or sequential) administration to a subject.

[0056] Without wishing to be bound by a theory, the lipid nanoparticle may or may not participate in lysis of endosomes. In some embodiments, the lipid nanoparticle does not participate in lysis of endosomes. In other words, the lipid nanoparticle does not have, or has very little, endosomolytic activity. For example, the lipid nanoparticle does not comprise a component having endosomolytic activity. Various methods are known in the art for determining endosomolytic activity. For example, endosomolytic activity can be determined in an *in vitro* cell assay, such as the red blood cell hemolysis assay or a liposomal leakage assay. Such an assay can comprise: contacting blood cells with lipid nanoparticles (or constituents of the lipid nanoparticles), wherein the pH of the medium in which the contact occurs is controlled; determining whether the lipid nanoparticles (or constituents of the lipid nanoparticles) induce differential lysis of blood cells at a low pH (*e.g.*, about pH 5-6) versus neutral pH (*e.g.*, about pH 7-8). A two-step assay can also be performed, wherein a first assay evaluates the ability of the lipid nanoparticle constituents alone to respond to changes in pH, and a second assay evaluates the ability of the assembled nanoparticle to respond to changes in pH.

[0057] In some embodiments, the lipid nanoparticle does not induce an immune response when administered to the subject. For example, the lipid nanoparticle induces very little or no (*e.g.*, less than 10%, 5%, or 2.5%) liver enzyme activity and/or inflammatory cytokines relative to a control, *e.g.*, a buffer.

[0058] Without wishing to be bound by a theory, the endosomolytic agent promotes lysis of the endosomal/lysosomal compartments and/or translocation across a cellular membrane and release of contents of endosomal/lysosomal compartments into the cytoplasm of the cell. It is believed that the endosomolytic agent functions as an agent to elicit or enhance the delivery of the ceDNA into the cytosol of target cells, possibly by improving endosomal escape of the lipid nanoparticle from the endosome. For example, the lipid nanoparticle and the endosomolytic agent may co-localize to an intracellular vesicle, *e.g.*, an endosome/lysosome within the target cell, where the endosomolytic agent can release of the ceDNA by lysis of the endosomal/lysosomal membrane. In some embodiments, the endosomolytic agent assumes its active conformation at endosomal pH, *e.g.*, pH 5-6.5. The “active” conformation is that conformation in which the endosomolytic agent promotes lysis of the endosomal/lysosomal compartments and/or translocation of contents of endosomal/lysosomal compartments into the cytoplasm of the cell.

[0059] In some embodiments, the membrane active functionality of the endosomolytic agent is masked. When the endosomolytic agent reaches the endosome, the membrane active functionality is unmasked and the agent becomes active. The unmasking may be carried out more readily under the conditions found in the endosome than outside the endosome. For example, the membrane active functionality can be masked with a molecule through a cleavable linker that undergoes cleavage in the endosome. Without wishing to be bound by theory, it is envisioned that upon entry into the endosome, such a linkage will be cleaved and the masking agent released from the endosomolytic agent.

[0060] Endosomolytic agents include, but are not limited to, imidazoles, poly or oligoimidazoles, polyethylene imidazoles (PEIs), peptides, fusogenic peptides, polycarboxylates, polycations, masked oligo or poly cations or anions, acetals, polyacetals, ketals/polyketals, orthoesters, polymers with masked or unmasked cationic or anionic charges, dendrimers with masked or unmasked cationic or anionic charges.

[0061] In some embodiments, the endosomolytic agent is a membrane-destabilizing polymer. A variety of membrane-destabilizing polymers are generally known in the art and may be used in accordance with the present methods described herein. Known types of membrane-destabilizing polymers include, for example, copolymers such as amphipathic copolymers, polycationic or amphipathic peptides, membrane active toxins, and viral fusogenic peptides. Exemplary membrane-destabilizing polymers are described, for example, in International PCT Application Publication Nos. WO2009/140427, WO2009/140429, and WO2016/118697, contents of each of which are incorporated herein by reference in their entireties.

[0062] The membrane-destabilizing polymer can be a copolymer, a synthetic peptide, a membrane-destabilizing toxin or derivative thereof, or a viral fusogenic peptide or derivative thereof. In some embodiments, the membrane-destabilizing polymer is a pH-sensitive polymer, for example, a pH-sensitive copolymer. The copolymer may be a block copolymer such as, for example, a diblock copolymer. In some embodiments, the block copolymer includes a hydrophobic, membrane-destabilizing block and a hydrophilic block.

[0063] In some such embodiments, the hydrophilic block is polymerized from both hydrophilic monomers and hydrophobic monomers such that there are more hydrophilic monomeric residues than hydrophobic monomeric residues in the hydrophilic block. The hydrophilic block may be cleavably linked to the hydrophobic block, such as through a disulfide bond or a pH-sensitive bond. In some embodiments, the hydrophilic block includes monomeric residues linked to a pendant shielding moiety such as, *e.g.*, a polyethylene glycol (PEG) moiety. The shielding moiety may be cleavably linked to the hydrophilic block, such as through a disulfide bond or a pH-sensitive bond. Particularly suitable pH-sensitive bonds (for linkage of the hydrophilic and hydrophobic blocks or linkage of the shielding moiety to the hydrophilic block) include hydrazone, acetal, ketal, imine, orthoester, carbonate, and maleamic acid linkages.

[0064] The pH-sensitive polymer may include monomeric residues having a carboxylic acid functional group, monomeric residues having an amine functional group, and/or monomeric residues having a hydrophobic functional group. In some variations, the pH-sensitive polymer includes monomeric residues derived from polymerization of a (C₂-C₈) alkylacrylic acid (*e.g.*, propylacrylic acid); monomeric residues derived from polymerization of a (C₂-C₈) alkyl-ethacrylate, a (C₂-C₈) alkyl-methacrylate, or a (C₂-C₈) alkyl-acrylate; and/or monomeric residues derived from polymerization of (N,N-di(C₁-C₆)alkyl-amino (C₁-C₆)alkyl-ethacrylate, (N,N-di(C₁-C₆)alkyl-amino (C₁-C₆)alkyl-methacrylate, or (N,N-di(C₁-C₆)alkyl-amino (C₁-C₆)alkyl-acrylate. In a specific variation, the pH-sensitive polymer includes a random copolymer chain having monomeric residues derived from polymerization of propyl acrylic acid, N,N-dimethylaminoethylmethacrylate, and butyl methacrylate. In some embodiments, the pH-sensitive polymer is a block copolymer comprising the random copolymer chain as a membrane disrupting polymer block, and further including one or more additional blocks.

[0065] In some embodiments, the pH-sensitive membrane-destabilizing polymer is a diblock copolymer having a hydrophilic random copolymer block and a hydrophobic random copolymer block, where (i) the hydrophilic block is an amphiphilic block comprising both hydrophilic monomeric residues and hydrophobic monomeric residues, where the number of hydrophilic monomeric residues in the hydrophilic block is greater than the number of hydrophobic monomeric residues, (ii) the hydrophobic block is an amphiphilic, membrane-destabilizing block comprising both hydrophobic monomeric residues and hydrophilic monomeric residues and having an overall hydrophobic character at a pH of about 7.4; and (iii) each of the hydrophilic monomeric residues of the hydrophilic and hydrophobic blocks is independently selected from the group consisting of monomeric residues that are ionic at a pH of about 7.4, monomeric residues that are neutral at a pH of about 7.4, and monomeric residues that are zwitterionic at a pH of about 7.4.

[0066] In yet some other embodiments, the pH-sensitive polymer is covalently linked to a membrane-destabilizing peptide. In some such embodiments, the pH-sensitive polymer includes a plurality of pendant linking groups, and a plurality of membrane-destabilizing peptides are linked to the pH-sensitive polymer via the plurality of pendant linking groups. Exemplary pH-sensitive polymers include the random block copolymers of Formula I, II, V, Ia, Va, Vb, Vc, Vd, Ve, Vf, Vg, Vh, Vi, Vj, Vk, Vl or Vm as described in WO2016/118697, the content of which is incorporated herein by reference in its entirety.

[0067] In some embodiments, the pH-sensitive polymer is selected from the group consisting polymers P67-P124 as described in WO2016/118697, incorporated by reference in its entirety herein, and shown below.

- A. P67: NAG-PEG12-[PEGMA(300, 79.1%)-BPAM(10.0%)-PDSMA(10.9%)] 3.56 KDa-b-[DMAEMA(34.7%)-BMA(54.7%)-PAA(10.5%)]4.71 KDa;

- B. P68: NAG-PEG12-[PEGMA(300; 89.8%)-PhEMA(10.2%)]3.23KDa-b-[DMAEMA(33%)-BMA(57%)-PAA(10%)]6.0 KDa;
- C. P69: NAG-PEG12-[PEGMA(300;78.7%)-PhEMA(21.3%)]3.25KDa-b-[DMAEMA(32.9%)-BMA(54.8%)-PAA(12.3%)]5.4KDa;
- D. P70: NAG-PEG12-[PEGMA(300,88.6)-PhEMA(11.4%)]3.02KDa-b-[DMAEMA(36.8%)-BMA(56.3%)-PAA(6.9%)]4.39KDa
- E. P71: NAG-PEG12-[PEGMA(300, 69.5%)-BPAM (19.2%)-PDSMA (11.3%)] 3.59 KDa-b-[DMAEMA (35.2%)-BMA (53.9%)-PAA (10.9%)]5.27 Kda;
- F. P72: NAG-PEG12-[PEGMA(300, 80.3%)-ImMA(19.7)]13.7KDa-b-[DMAEMA(35.9%)-BMA(53.9%)-PAA(10.2%)]4.7KDa;
- G. P73: NAG-PEG12-[PEGMA(300, 73.1%)-BMA(14.4%)-PhEMA(12.5%)]3.8KDa-b-[DMAEMA(37.6%)-BMA(52.3%)-PAA(10.1%)]4.2KDa
- H. P74: NAG-PEG12-[PEGMA(300, 80.3%)-BMA(23.3%)]3.8KDa-b-[DMAEMA(38.2%)-BMA(51.5%)-PAA(10.3%)]3.5KDa;
- I. P75: NAG-PEG12-[PEGMA(300, 75.8%)-isoA-MA(11.8%)-PhEMA(12.4%)]3.3KDa-b-[DMAEMA(39.3%)-BMA(51.6%)-PAA(9%)]4.95KDa;
- J. P76: NAG-PEG12-[PEGMA(300, 74.9%)-isoA-MA(25.1%)]2.9KDa-b-[DMAEMA(38%)-BMA(53%)-PAA(9.1%)]5.2KDa ;
- K. P77: NAG-PEG12-[PEGMA (300, 86%)-CyHexMA (14%)]2.98 KDa-b-[DMAEMA (36.2%)-BMA (51.7%)-PAA (12.2%)]4.66 KDa ;
- L. P78: NAG-PEG12-[PEGMA(300, 72.5%)-BPAM(27.5%)]3.8KDa-b-[DMAEMA(25.6%)-BMA(64.8%)-PAA(9.6%)]5.5KDa;
- M. P79: NAG-PEG12-[PEGMA (300, 69.9%)-HMA(30.1%)] 2.93 KDa-b-[DMAEMA (34.4%)-BMA (53.6%)-PAA (12%)]4.43 Kda;
- N. P80: NAG-PEG12-[PEGMA (300, 85.4%)-EHMA(14.6%)] 3.36 KDa-b-[DMAEMA (36.5%)-BMA (53.7%)-PAA (9.7%)]4.18 KDa ;
- O. P81: NAG-PEG12-[PEGMA(300, 72%)-F1-BMA(28%)]3.75KDa-b-[DMAEMA(30.7%)-BMA(56.7%)-PAA(12.6%)]5.7KDa ;
- P. P82: NAG-PEG12-[PEGMA(300, 71.9%)-F1-BMA(28.1%)]3.55Kda-b-[DMAEMA(29.9%)-BMA(57.6%)-PAA(12.4%)]5.3KDa;
- Q. P83: NAG-PEG12-[PEGMA(300, 78.9%)-F-CyHexMA(21.1%)]4.56KDa-b-[DMAEMA(33.2%)-BMA(55.4%)-PAA(11.4%)]5.3KDa;
- R. P84: NAG-PEG12-[PEGMA(300, 77.9%)-F-HPenMA(22.1%)]3.26Kda-b-[DMAEMA(30.9%)-BMA(57.4%)-PAA(11.6%)]6.5KDa;
- S. P85: NAG-PEG12-[PEGMA(300, 79%)-BMA(21%)]2.9KDa-b-[DMAEMA(29.3%)-BMA(26.6%)- F1-BMA(34.6%)-PAA(9.5%)]5.8KDa;

- T. P86: NAG-PEG12-[PEGMA (300, 78.1%)-C12MA(21.9%)]3.67 KDa-b-[DMAEMA (32.1%)-BMA (53.7%)-PAA (14.2%)]4.7 KDa;
- U. P87: NAG-PEG12-[PEGMA (300, 69.7%)-EHMA (30.3%)]3.9 KDa-b-[DMAEMA (31.1%)-BMA(56.7%)-PAA (12.1%)]5.1 KDa;
- V. P88: NAG-PEG12-[PEGMA (300, 76%)-5-NMA (24%)] 3.0 KDa-b-[DMAEMA (34.4%)-BMA (54%)-PAA (11.6%)]5.6 KDa;
- W. P89: NAG-PEG12-[PEGMA (300,73.8%)-BMA (26.2%)]3.5 KDa-b-[DMAEMA (30.7%)-BMA(58.9%)-PAA (10.4%)]4.9 KDa;
- X. P90: NAG-PEG12-[PEGMA (300, 72.6%)-HMA (27.4%)]3.58 KDa-b-[DMAEMA (30.6%)-BMA(56.2%)-PAA (13.3%)]5.6 KDa;
- Y. P91: CH3O-PEG12-[PEGMA (300, 92.8%)-PDSMA (7.2%)]3.6 KDa-b-[DMAEMA (34.2%)-BMA(54.7%)-PAA (11%)]6.5 KDa ;
- Z. P92: NAG-PEG12-[PEGMA (300, 83.2%)-AEOMA (16.8%)]3.0 KDa-b-[DMAEMA (36.2%)-BMA(52.2%)-PAA (11.6%)]5.6;
- AA. P93: NAG-PEG12-[PEGMA (300, 77.6%)-CyHexMA (22.4%)]2.64 KDa-b-[DMAEMA (32.1%)-BMA(43.1%)-PAA (12.6%)-CyHexMA(12.3%)]4.67 KDa;
- BB. P94: NAG-PEG12-[PEGMA (300, 72.2%)-B-F1-HMA (27.8%)]4.2 KDa-b-[DMAEMA (35.7%)-BMA(54.4%)-PAA (9.9%)]4.7 KDa;
- CC. P95: NAG-PEG12-[PEGMA(300, 71.2%)-F1-BMA(28.8%)]3.55KDa-b-[DMAEMA(34.2%)-BMA(57.9%)-PAA(7.9%)]4.9KDa;
- DD. P96: NAG-PEG12-[PEGMA(300, 72.6%)-F1-BMA(27.4%)]3.55KDa-b-[DMAEMA(30.7%)-BMA(56.1%)-PAA(13.2%)]4.9KDa;
- EE. P97: NAG-PEG12-[PEGMA(300, 70.0%)-F1-BMA(30.0%)]3.55KDa-b-[DMAEMA(31.3%)-BMA(60.7%)-PAA(8.0%)]5.1KDa;
- FF. P98: NAG-PEG12-[PEGMA(300, 75 %) -2-Bul-OMA(25 %14.26KDa-b-[DMAEMA(32.1%)-BMA(55.7%)-PAA(12.2%)]5.69KD;
- GG. P99: NAG-PEG12-[PEGMA(300, 73.3%)-5-NMA(26.7%)]4.051(1)a-b-[DMAEMA(31.5%)-BMA(55.2%)-PAA(13.3%)]5.20KD;
- HH. P100: NAG-PEG12-[PEGMA(300, 74.1%)-F1-BMA(25.9%)]3.79KDa-b-[DMAEMA(29.9%)-BMA(56.2%)-PAA(13.9%)]5.44KDa;
- II. P101: NAG-PEG12-[PEGMA(300, 72.2%)-B-F1-0MA(27.8%)]4.2KDa-b-[DMAEMA(35.7%)-BMA(54.4%)-PAA(9.9%)]5.6KD;
- JJ. P102: NAG-PEG12-[PEGMA(300, 71.9%)-F1-BMA(28.1%)]3.55KDa-b-[DMAEMA(27.3%)-BMA(60.9%)-PAA(11.9%)]4.55KDa;
- KK. P103: NAG-PEG12-[PEGMA(300, 70.3%)-F1-BMA(29.7%)]3.6KDa-b-[DMAEMA(32.2%)-BMA(57.6%)-PAA(10.2%)]5KDa;

- LL. P104: NAG-PEG12-[PEGMA(300, 68%)-F1-BMA(32%)]*3.7KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5.3KDa;
- MM. P106: NAG-PEG12-[PEGMA(300, 74%)-HMA(26%)]4.1KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5 KDa;
- NN. P107: NAG-PEG12-[PEGMA(300, 74%)-HMA(26%)]*4.1KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*4.2KDa;
- OO. P108: NAG-PEG12-[PEGMA(300, 80%)-HMA(20%)]*4.96Kda-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5.5KDa;
- PP. P109: NAG-PEG12-[PEGMA(300, 80%)-HMA(20%)]*4.96Kda-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*6.5KDa;
- QQ. P110: NAG-PEG12-[PEGMA(300, 77.7%)-EHMA(22.3%)]4.37KDa-b-[DMAEMA(31 %)-BMA(56%)-PAA(13%)]*6KDa ;
- RR. P111: NAG-PEG12-[PEGMA(300, 77%)-F1-BMA(23%)]5.80KDa-b-[DMAEMA(27.3%)-BMA(60.9%)-PAA(11.9%)]*5.74KDa;
- SS. P111: NAG-PEG12-[PEGMA(300, 77%)-F1-BMA(23%)]5.80KDa-b-[DMAEMA(27.3%)-BMA(60.9%)-PAA(11.9%)]*5.74KDa;
- TT. P113: NAG-PEG12-[PEGMA(300, 84.9%)-Chol-MA(15.1%)]*3.5KDa-b-[DMAEMA(31 %)-BMA(56%)-PAA(13%)]*4.67KDa;
- UU. P114: NAG-PEG12-[PEGMA(300, 67%)-HMA(33%)]*5.7KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*6.15KDa;
- VV. P115: NAG-PEG12-[PEGMA(300, 67%)-HMA(33%)]*5.7KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*6 KDa;
- WW. P116: NAG-PEG12-[PEGMA(300, 73%)-F1-BMA(27%)]* 6.3KDa-b-[DMAEMA(27.3%)-BMA(60.9%)-PAA(11.9%)]* 5.9KDa;
- XX. P117: NAG-PEG12-[PEGMA(300, 72%)-PF-BMA(28%)]*+3.7KDa-b-[DMAEMA(27.3%)-BMA(60.9%)-PAA(11.9%)]*5.0KDa;
- YY. P118: NAG-PEG12-[PEGMA(300, 70%)-HMA(30%)]*5.2KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5.7KDa;
- ZZ. P119: NAG-PEG12-[PEGMA(300, 70%)-HMA(30%)]*5.2KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5 KDa;
- AAA. P120: NAG-PEG12-[PEGMA(300, 75%)-CyHexMA(25%)]*4KDa-b-[DMAEMA(31 %)-BMA(56%)-PAA(13%)]*5.2KDa;
- BBB. P121: NAG-PEG12-[PEGMA(300, 75%)-Me-CyHexMA(25%)]*4.3KDa-b-[DMAEMA(31 %)-BMA(56%)-PAA(13%)]*5.1KDa;
- CCC. P122: NAG-PEG12-[PEGMA(300, 73%)-F1-BMA(27%)]* 6.3KDa-b-[DMAEMA(27.3%)-BMA(60.9%)-PAA(11.9%)]* 6.9KDa;

- DDD. P123: NAG-PEG12-[PEGMA(300, 79%)-Bul-O-MA(21%)]*4.88KDa-b-[DMAEMA(31 %)-BMA(56%)-PAA(13%)]*4.6KDa;
- EEE. P124: NAG-PEG12-[PEGMA(300, 74%)-HMA(26%)]*4.15KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5 KDa; and
- FFF. P125: NAG-PEG12-[PEGMA(300, 74%)-HMA(26%)]*4.15KDa-b-[DMAEMA(31%)-BMA(56%)-PAA(13%)]*5 KDa.

[0068] Additional exemplary synthetic polymers with endosomolytic activity are described, for example, in United States Patent Application Publications Nos. 2009/0048410; 20090023890; 2008/0287630; 20080287628; 2008/0281044; 2008/0281041; 2008/0269450; 2007/0105804; 20070036865; and 2004/0198687, the contents of all which are hereby incorporated by reference in their entireties.

[0069] In some embodiments, the endosomolytic agent can be a polyanionic peptide, polycationic peptide, amphipathic peptide, hydrophobic peptide or a peptidomimetic which shows pH-dependent membrane activity and/or fusogenicity. In some embodiments, the endosomolytic agent is a cell-permeation or Cell Penetrating Peptide (CPP). Exemplary primary sequences of endosomolytic peptides include the following amino acid sequences:

AALEALAEALAEALAEALAEAAAAGGC (SEQ ID NO: 530);
 AALAEALAEALAEALAEALAEALAAAAGGC (SEQ ID NO: 531);
 ALEALAEALAEALAEA (SEQ ID NO: 532);
 GLFEAIEGFIENGWEGMIWDYG (SEQ ID NO: 533);
 GLFGAIAGFIENGWEGMIDGWYG (SEQ ID NO: 534);
 GLFEAIEGFIENGWEGMIDGWYGC (SEQ ID NO: 535);
 GLFEAIEGFIENGWEGMIDGWYGC (SEQ ID NO: 536);
 GLFEAIEGFIENGWEGMIDGGC (SEQ ID NO: 537);
 GLFEAIEGFIENGWEGMIDGGC (SEQ ID NO: 538);
 CGLFGEIEELIEEGLNLDWNG (SEQ ID NO: 539);
 GLFGALAEALAEALAEHLAEALAEALAEALAAGGSC (SEQ ID NO: 540);
 GLFEAIEGFIENGWEGLAELAEALAEALAEALAAGGSC (SEQ ID NO: 541);
 GLFEAIEGFIENGWEGnIDGK (n = norleucine) (SEQ ID NO: 542);
 GLFEAIEGFIENGWEGnIDG (n = norleucine) (SEQ ID NO: 543);
 GLFEALLELLESLWELLLEA (SEQ ID NO: 544);
 GLFKALLKLLKSLWKLLKA (SEQ ID NO: 545);
 GLFRALLRLLRSLWRLLRA (SEQ ID NO: 546);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 547);
 GLFFEIAIEFIEGGWEGGLIEGC (SEQ ID NO: 548);
 GIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 549);

H5WYG (SEQ ID NO: 550);
CHK₆HC (SEQ ID NO: 551);
RQIKIWFQNRRMKWKK (SEQ ID NO: 552);
GRKKRRQRRRPPQC (SEQ ID NO: 553);
GALFLGWLGAAGSTM (SEQ ID NO: 554);
GAWSQPKKKRKV (SEQ ID NO: 555);
LLIILRRRIRKQAHASK (SEQ ID NO: 556);
GWTLSAGYLLKINLKALAALAKKIL (SEQ ID NO: 557);
KLALKLALKALKAALKLA (SEQ ID NO: 558);
RRRRRRRRR (SEQ ID NO: 559);
KFFKFFKFFK (SEQ ID NO: 560);
LLGDFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTE (SEQ ID NO: 561);
SWLSKTAKKLENSAKKRISGIAIAIQGGPR (SEQ ID NO: 562);
ACYCRIPACIAGERRYGTCTIYQGRWAFCC (SEQ ID NO: 563);
DHYNCSVSSGGQCLYSACPIFTKIQTCTYRGKAKCCK ((SEQ ID NO: 564);
RKCRIVVIRVCRRRRPRPPYLPRPRPPFFPPRLPPRIPPGFPPRFPPRFPGKR (SEQ ID NO: 565);
ILPWKWPWWPWRR (566);
WEAALAEALAEALAEHLAEALAEALAEALAA (SEQ ID NO: 567);
CAEALAEALAEALAEALA (SEQ ID NO: 568);
GIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 569);
CGIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 570);
FIIDIIAFLLMGGFIVYVKNL (SEQ ID NO: 571);
CAAFIIDHAFLLMGGFIVYVKNL (SEQ ID NO: 572);
CARGWEVLKYWWNLLQY (SEQ ID NO: 573);
MVKSKIGSWILVLFVAMWSDVGLCKKRPKP (SEQ ID NO: 574);
KLALKLALKALKAALKLA (SEQ ID NO: 575);
YARAAARQARA (SEQ ID NO: 576);
GDCLPHLKLCKENKDCCSKKCKRRGTNIE (SEQ ID NO: 577);
RRLSYSRRRF (SEQ ID NO: 578);
RGGRLSYSRRRFSTSTGR (SEQ ID NO: 579);
IAWVKAFIRKLKRGPLG (SEQ ID NO: 580);
YTAIAWVKAFIRKLK (SEQ ID NO: 581);
GLWRALWRLRLWRLWRA (SEQ ID NO: 582);
KWFETWFTEWPKKRK (SEQ ID NO: 583);
KETWWETWTEWSQPKKKRKV (SEQ ID NO: 584);
AGYLLGK(eNH_a)INLKALAALAKKIL (SEQ ID NO: 585);

AGYLLGKINLKALAALAKKIL (SEQ ID NO: 586);
 RQIKIVVFQNRMRKWK (SEQ ID NO: 587);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 588);
 LLILRRRIRKQAHASK (SEQ ID NO: 589);
 YTIVVMPENPRPGTPCDIFTNSRGKRASNG (SEQ ID NO: 590);
 AAVALLPAVLLALLAK (SEQ ID NO: 591);
 GWTLNSAGYLLGKINLKALAALAKKIL (SEQ ID NO: 592);
 GRKKRRQRRPPQ (SEQ ID NO: 593);
 KMTRAQRRAAARRNRRTAR (SEQ ID NO: 594);
 KKRKAPKKRKFA-KFHTFPQTAIGVGAP (SEQ ID NOS 595 and 600, respectively);
 MVTVLFRRLRIRASGPPRV (SEQ ID NO: 596);
 LIRLSHLIHIVVFQNRRLKWK (SEQ ID NO: 597);
 GALFLGFLGAAGSTMGAWSQPKKKRKV (SEQ ID NO: 598); and
 GALFLAFLAALSLMGLWSQPKKKRKV (SEQ ID NO: 599).

[0070] Additional exemplary peptide sequences for the endosomolytic agent include, SEQ ID NOS: 1-1802, as described in Table 2 of WO 2015/069586, content of which is incorporated herein by reference in its entirety.

[0071] In some embodiments, the endosomolytic agent is a peptide of Formula (P)_c-(L)_d-(G)_e, where P is an endosomolytic peptide; G is a linker; G is a targeting ligand; each of c and e is 1, 2, 3, 4, 5, or 6; and d is 0, 1, 2, 3, 4, 5 or 6. In some embodiments, the endosomolytic agent is of Formula (P)_c-(L)_d-(G)_e, as described in Table 2 of WO 2015/069586, incorporated by reference in its entirety herein.

[0072] Lipids having membrane activity are also amenable to the present invention as endosomolytic agents. Such lipids are also described as fusogenic lipids in the art. Without wishing to be bound by a theory, these fusogenic lipids are thought to fuse with and consequently destabilize a membrane. Fusogenic lipids usually have small head groups and unsaturated acyl chains. Exemplary fusogenic lipids include, but are not limited to, 1,2-dioleoyl-sn-3-phosphoethanolamine (DOPE), phosphatidylethanolamine (POPE), palmitoyloleoylphosphatidylcholine (POPC), (6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-ol (Di-Lin), N-methyl(2,2-di((9Z,12Z)-octadeca-9,12-dienyl)-1,3-dioxolan-4-yl)methanamine (DLin-k-DMA) and N-methyl-2-(2,2-di((9Z,12Z)-octadeca-9,12-dienyl)-1,3-dioxolan-4-yl)ethanamine (XTC).

[0073] Without wishing to be bound by a theory, many endosomolytic agents can self-assemble into particles. Accordingly, in some embodiments, the endosomolytic agent is in form of a particle, *e.g.*, a nanoparticle.

[0074] For example, the endosomolytic agent can be nanoparticle having a mean diameter from about 30 nm to about 150 nm, more typically from about 50 nm to about 150 nm, more typically about 60 nm to about 130 nm, more typically about 70 nm to about 110 nm, most typically about 85

nm to about 105nm, and preferably about 100 nm. Nanoparticle particle size can be determined by quasi-elastic light scattering using, for example, a Malvern Zetasizer Nano ZS (Malvern, UK) system.

[0075] In some embodiments, the endosomolytic agent is in form of particle, wherein the particle further comprises one or more of a ionizable lipid, a non-cationic lipid, a sterol or derivative thereof, and a conjugate lipid, *e.g.*, PEG-lipid. Exemplary non-cationic lipids, sterols and derivative thereof, and conjugate lipids are described below and exemplary ionizable lipids are described in **Table 1A**.

[0076] In some embodiments, the endosomolytic agent is comprised in the lipid nanoparticle that encapsulates the ceDNA.

[0077] In some embodiments, at least one of the lipid nanoparticle and the endosomolytic agent includes a first targeting ligand that binds to a molecule on the surface of the target cell. For example, the endosomolytic agent, the lipid nanoparticle, or both the endosomolytic agent and lipid nanoparticle can include the first targeting ligand. In some embodiments, the endosomolytic agent includes the first targeting ligand and the nanoparticle does not include a targeting ligand. Generally, a “targeting ligand” refers to a moiety that confers some degree of target specificity to one or more cells, tissues, or organs, such as in a subject or organism and thus the ability to target such cells, tissues, or organs with a compound or composition of interest, *e.g.*, endosomolytic agent and/or lipid nanoparticle.

[0078] In some embodiments, one of the lipid nanoparticle and endosomolytic agent includes the first targeting ligand, and the other of the lipid nanoparticle and endosomolytic agent includes a second targeting ligand. The first and second targeting ligands can be the same or different. Further, the second targeting ligand can bind to the same cell surface molecule recognized by the first targeting ligand or the second targeting ligand can bind to a cell surface molecule that is different from the one recognized by the first targeting ligand.

[0079] In some embodiments, the first and/or second targeting ligand specifically binds to a molecule on the surface of the target cell. By specifically binds is meant that the targeting ligand binds to the molecule on surface of the target cell with at least 2-fold greater affinity relative to molecules on the surface of a non-target cell, *e.g.*, at least 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 25-fold, 50-fold, or 100-fold greater affinity.

[0080] Exemplary cell surface molecules to which the first and/or second targeting ligand can bind include, but are not limited to transferrin receptor type 1, transferrin receptor type 2, the EGF receptor, HER2/Neu, a VEGF receptor, a PDGF receptor, an integrin, an NGF receptor, CD2, CD3, CD4, CD8, CD19, CD20, CD22, CD33, CD43, CD38, CD56, CD69, the asialoglycoprotein receptor (ASGPR), GalNAc receptor, prostate-specific membrane antigen (PSMA), a folate receptor, and a sigma receptor.

[0081] In some embodiments, the first and/or the second targeting ligand bind to a molecule on surface of a hepatocyte. Exemplary receptors on hepatocytes include, but are not limited to, the

asialoglycoprotein receptor (ASGPR) and GalNAc receptor. Accordingly, in some embodiments, the first and/or the second targeting ligand binds to ASGPR or GalNAc receptor.

[0082] Without limitations, the targeting ligand, *e.g.*, the first and/or second targeting ligand can be selected from small molecules, proteins (*e.g.*, an antibody or antigen binding fragment thereof, a peptide aptamer, or a protein derived from a natural ligand the cell surface molecule), a peptide (such as, an integrin-binding peptide, a LOX-1-binding peptide, or epidermal growth factor (EGF) peptide, a neurotensin peptide, an NL4 peptide, or a YIGSR laminin peptide (SEQ ID NO: 601)), and nucleic acid aptamer.

[0083] In some embodiments, the targeting ligand is a sugar (*e.g.*, lactose, galactose, N-acetyl galactosamine (NAG, also referred to as GalNAc), mannose, and mannose-6-phosphate (M6P)), a vitamin (*e.g.*, folate), a bisphosphonate, or an analogue thereof. Carbohydrates and carbohydrate clusters with multiple carbohydrate motifs represent an important class of targeting ligands, which allow the targeting of drugs to a wide variety of tissues and cell types. Accordingly, in some embodiments, the first and/or second targeting ligand is a carbohydrate or a carbohydrate cluster. Exemplary, carbohydrate based targeting ligands include, but are not limited to, D-galactose, multivalent galactose, N-acetyl-D-galactose (GalNAc), multivalent GalNAc, *e.g.* GalNAc2 and GalNAc3; D-mannose, multivalent mannose, multivalent lactose, N-acetyl-galactosamine, N-acetyl-gulucosamine, multivalent fucose, glycosylated polyaminoacids and lectins. The term multivalent indicates that more than one monosaccharide unit is present. Such monosaccharide subunits may be linked to each other through glycosidic linkages or linked to a scaffold molecule.

[0084] In some embodiments, the first and/or second targeting ligand can be selected from the group consisting of selected from the group consisting of an antibody, a ligand-binding portion of a receptor, a ligand for a receptor, an aptamer, D-galactose, N-acetyl- D-galactose (GalNAc), multivalent N-acetyl-D-galactose, D-mannose, cholesterol, a fatty acid, a lipoprotein, folate, thyrotropin, melanotropin, surfactant protein A, mucin, carbohydrate, multivalent lactose, multivalent galactose, N-acetyl-galactosamine, N-acetyl-glucosamine, multivalent mannose, multivalent fructose, glycosylated polyaminoacids, transferin, bisphosphonate, polyglutamate, polyaspartate, a lipophilic moiety that enhances plasma protein binding, a steroid, bile acid, vitamin B12, biotin, an RGD peptide, an RGD peptide mimic, ibuprofen, naproxen, aspirin, folate, and analogs and derivatives thereof.

[0085] In some embodiments, the first and/or second targeting ligand is selected from the group consisting of D-galactose, N-acetyl-D-galactose (GalNAc), GalNAc2, and GalNAc3, cholesterol, folate, and analogs and derivatives thereof. In some preferred embodiments, the first and/or second ligand is monovalent or multivalent, *e.g.*, trivalent GalNAc (GalNAc3).

[0086] Additional exemplary targeting ligand are described, for example, in Table 4 of WO2015/069586, the content of which is incorporated herein by reference in its entirety. Exemplary

GalNAc ligands are also described in WO2009126933, the content of which is incorporated herein by reference in its entirety.

[0087] A targeting ligand, *e.g.*, the first and/or second targeting ligand, can be linked to the endosomolytic agent or the lipid nanoparticle via a linker. This linker may be cleavable or non-cleavable, depending on the application. In some embodiments, a cleavable linker may be used to release the nucleic acid after transport from the endosome to the cytoplasm. The intended nature of the conjugation or coupling interaction, or the desired biological effect, will determine the choice of linker group. In some embodiments, the linker is a linker as described in Table 3 of WO 2015/069586, the content of which is incorporated herein by reference in its entirety. In some embodiments, the linker is a linker described in WO2016/118697, content of which is incorporated herein by reference in its entirety.

[0088] In some embodiments, the target cell is selected from a secretory cell, a chondrocyte, an epithelial cell, a nerve cell, a muscle cell, a blood cell, an endothelial cell, a pericyte, a fibroblast, a glial cell, and a dendritic cell. Other suitable target cells include cancer cells, immune cells, bacterially-infected cells, virally-infected cells, or cells having an abnormal metabolic activity.

[0089] As discussed above, foreign nucleic acids, *e.g.*, DNA potently stimulates the innate immune response and this occurs through a pathway dependent upon DNA sensor Cgas STING. Accordingly, in some embodiments, the target cell is a cell that lacks or does not express a functional innate DNA-sensing pathway, for example, a cell that lacks or does not express functional cGAS and/or STING. Thomsen *et al.* also demonstrated that hepatocytes lack a functional innate DNA-sensing pathway since they do not express STING. Thus, in some embodiments, the target cell is a hepatocyte.

[0090] Generally, the lipid particles are prepared at a total lipid to ceDNA (mass or weight) ratio of from about 10:1 to 30:1. In some embodiments, the lipid to ceDNA ratio (mass/mass ratio; w/w ratio) can be in the range of from about 1:1 to about 25:1, from about 10:1 to about 14:1, from about 3:1 to about 15:1, from about 4:1 to about 10:1, from about 5:1 to about 9:1, or about 6:1 to about 9:1. The amounts of lipids and ceDNA can be adjusted to provide a desired N/P ratio, for example, N/P ratio of 3, 4, 5, 6, 7, 8, 9, 10 or higher. Generally, the lipid particle formulation's overall lipid content can range from about 5 mg/ml to about 30 mg/mL.

[0091] In some embodiments, the lipid nanoparticle comprises an agent for condensing and/or encapsulating nucleic acid cargo, such as ceDNA. Such an agent is also referred to as a condensing or encapsulating agent herein. Without limitations, any compound known in the art for condensing and/or encapsulating nucleic acids can be used as long as it is non-fusogenic. In other words, an agent capable of condensing and/or encapsulating the nucleic acid cargo, such as ceDNA, but having little or no fusogenic activity. Without wishing to be bound by a theory, a condensing agent may have some fusogenic activity when not condensing/encapsulating a nucleic acid, such as ceDNA, but a

nucleic acid encapsulating lipid nanoparticle formed with said condensing agent can be non-fusogenic.

[0092] Generally, a condensing agent employed in the nanoparticles of the invention does not have, or has very little, fusogenic activity at any pH. For example, the fusogenic activity of the condensing agent differs by less than 10%, *e.g.*, less than 5%, 4%, 3%, 2% or 1% at a low pH vs neutral pH as measured by a membrane-impermeable fluorescent dye exclusion assay. In some embodiments, the fusogenic activity of the condensing agent is substantially the same, *e.g.*, differs by less than 0.5%, 0.25%, 0.1% or an undetectable amount at a low pH vs neutral pH as measured by a membrane-impermeable fluorescent dye exclusion assay, *e.g.*, the assay described in the Examples section.

[0093] In some embodiments, the condensing agent is a lipid, for example a non-fusogenic cationic lipid. By a “non-fusogenic cationic lipid” is meant a cationic lipid that can condense and/or encapsulate the nucleic acid cargo, such as ceDNA, but does not have, or has very little, fusogenic activity. In some embodiments, the condensing agent is a cationic lipid described in PCT and US patent publication listed in **Table 1A** and determined to be non-fusogenic, as measured, for example, by a membrane-impermeable fluorescent dye exclusion assay, *e.g.*, the assay described in the Examples section herein. Contents of all of the PCT and US patent publication listed in **Table 1A** are incorporated herein by reference in their entireties.

Table 1A. Exemplary cationic lipids

PCT Publication	US Publication
WO2015/095340	US2016/0311759
WO2015/199952	US2015/0376115
WO2018/011633	US2016/0151284
WO2017/049245	US2017/0210697
WO2015/061467	US2015/0140070
WO2012/040184	US2013/0178541
WO2012/000104	US2013/0303587
WO2015/074085	US2015/0141678
WO2016/081029	US2015/0239926
WO2017/004143	US2016/0376224
WO2017/075531	US2017/0119904
WO2017/117528	
WO2011/022460	US2012/0149894
WO2013/148541	US2015/0057373
WO2013/116126	
WO2011/153120	US2013/0090372
WO2012/044638	US2013/0274523
WO2012/054365	US2013/0274504
WO2011/090965	US2013/0274504
WO2013/016058	
WO2012/162210	
WO2008/042973	US2009/0023673
WO2010/129709	US2012/0128760
WO2010/144740	US201/003241240

WO2012/099755	US2014/0200257
WO2013/049328	US2015/0203446
WO2013/086322	US2018/0005363
WO2013/086373	US2014/0308304
WO2011/071860	US2013/0338210
WO2009/132131	
WO2010/048536	
WO2010/088537	US2012/0101148
WO2010/054401	US2012/0027796
WO2010/054406	
WO2010/054405	
WO2010/054384	US2012/0058144
WO2012/016184	US2013/0323269
WO2009/086558	US2011/0117125
WO2010/042877	US2011/0256175
WO2011/000106	US2012/0202871
WO2011/000107	US2011/0076335
WO2005/120152	US2006/0083780
WO2011/141705	US2013/0123338
WO2013/126803	US2015/0064242
WO2006/07712	US2006/0051405
WO2011/038160	US2013/0065939
WO2005/121348	US2006/0008910
WO2011/066651	US2003/0022649
WO2009/127060	US2010/0130588
WO2011/141704	US2013/0116307
WO2006/069782	US2010/0062967
WO2012/031043	US2013/0202684
WO2013/006825	US2014/0141070
WO2013/033563	US2014/0255472
WO2013/089151	US2014/0039032
WO2017/099823	US2018/0028664
WO2015/095346	US2016/0317458
WO2013/086354	US2013/0195920

[0094] In some embodiments, the condensing agent, *e.g.* a cationic lipid, is selected from the group consisting of N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA); N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTAP); 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (DOEPC); 1,2-dilauroyl-sn-glycero-3-ethylphosphocholine (DLEPC); 1,2-dimyristoyl-sn-glycero-3-ethylphosphocholine (DMEPC); 1,2-dimyristoleoyl-sn-glycero-3-ethylphosphocholine (14:1), N1-[2-((1S)-1-[(3-aminopropyl)amino]-4-[di(3-amino-propyl)aminobutyl]carboxamido)ethyl]-3,4-di[oleyloxy]-benzamide(MVL5); Dioctadecylamido-glycylspermine (DOGS); 3b-[N-(N',N'-dimethylaminoethyl)carbonyl] cholesterol (DC-Chol); Dioctadecyldimethylammonium Bromide (DDAB); a Saint lipid (*e.g.*, SAINT-2, N-methyl-4-(dioleoyl)methylpyridinium); 1,2-dimyristyloxypropyl-3-dimethylhydroxyethylammonium bromide (DMRIE); 1,2-dioleoyl-3-dimethyl-hydroxyethyl ammonium bromide (DORIE); 1,2-dioleoyloxypropyl-3-dimethylhydroxyethyl ammonium chloride (DORI); Di-alkylated Amino Acid

(DILA2) (*e.g.*, C18 :1 -norArg -C16); Dioleyldimethylammonium chloride (DODAC); 1-palmitoyl-2-oleoyl-sn-glycero-3 -ethylphosphocholine (POEPC); and 1,2 -dimyristoleoyl-sn-glycero-3-ethylphosphocholine (MOEPC). In some variations, the condensing agent, *e.g.* a cationic lipid, is a lipid such as, *e.g.*, Dioctadecyldimethylammonium bromide (DDAB), 1,2-dilinoleyloxy-3-dimethylaminopropane (DLinDMA), 2,2-dilinoley1-4-(2dimethylaminoethyl)-[1,31-dioxolane (DLin-KC2-DMA), heptatriaconta-6,9,28,31-tetraen-19-yl-4-(dimethylamino)butanoate (DLin-MC3-DMA), 1,2-Dioleoyloxy-3-dimethylaminopropane (DODAP), 1,2-Dioleoyloxy-3-dimethylaminopropane (DODMA), Morpholinocholesterol (Mo-CHOL), (R)-5-(dimethylamino)pentane-1,2-diyl dioleate hydrochloride (DODAPen-C1), (R)-5-guanidinopentane-1,2-diyl dioleate hydrochloride (DOPen-G), (R)-N,N,N-trimethyl-4,5-bis(oleoyloxy)pentan-1-aminium chloride(DOTAPen).

[0095] In some embodiments, the condensing lipid is DOTAP.

[0096] Without limitations, the condensing agent, *e.g.* a cationic lipid, can comprise 10-90% (mol) of the total lipid present in the lipid nanoparticle. For example, condensing agent, *e.g.* a cationic lipid, molar content can be 20-90% (mol), 20-70% (mol), 30-60% (mol) or 40-50% (mol) of the total lipid present in the lipid nanoparticle. In some embodiments, condensing agent, *e.g.* a cationic lipid, comprises from about 50 mol % to about 90 mol % of the total lipid present in the lipid nanoparticle.

[0097] In some aspects, the lipid nanoparticle can further comprise a non-cationic lipid. Non-ionic lipids include amphipathic lipids, neutral lipids and anionic lipids. Accordingly, the non-cationic lipid can be a neutral uncharged, zwitterionic, or anionic lipid. In some embodiments, the non-cationic lipid is non-fusogenic, *i.e.*, a non-cationic lipid that does not or substantially does not fuse with a membrane or, if does fuse with a membrane, does not destabilize the membrane. Generally, a non-cationic lipid employed in the nanoparticles of the invention does not have, or has very little, fusogenic activity at any pH.

[0098] Exemplary non-cationic lipids include, but are not limited to, distearoyl-sn-glycero-phosphoethanolamine, distearoylphosphatidylcholine (DSPC), dioleoylphosphatidylcholine (DOPC), dipalmitoylphosphatidylcholine (DPPC), dioleoylphosphatidylglycerol (DOPG), dipalmitoylphosphatidylglycerol (DPPG), dioleoyl-phosphatidylethanolamine (DOPE), palmitoyl-oleoylphosphatidylcholine (POPC), palmitoyl-oleoylphosphatidylethanolamine (POPE), dioleoyl-phosphatidylethanolamine 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate (DOPE-mal), dipalmitoyl phosphatidyl ethanolamine (DPPE), dimyristoylphosphoethanolamine (DMPE), distearoyl-phosphatidyl-ethanolamine (DSPE), monomethyl-phosphatidylethanolamine (such as 16-O-monomethyl PE), dimethyl-phosphatidylethanolamine (such as 16-O-dimethyl PE), 18-1-trans PE, 1-stearoyl-2-oleoyl-phosphatidylethanolamine (SOPE), hydrogenated soy phosphatidylcholine (HSPC), egg phosphatidylcholine (EPC), dioleoylphosphatidylserine (DOPS), sphingomyelin (SM), dimyristoyl phosphatidylcholine (DMPC), dimyristoyl phosphatidylglycerol (DMPG), distearoylphosphatidylglycerol (DSPG), di-erucoylphosphatidylcholine (DEPC),

palmitoyl-oleoylphosphatidylglycerol (POPG), dielaidoyl-phosphatidylethanolamine (DEPE), 1,2-dilauroyl-sn-glycero-3-phosphoethanolamine (DLPE); 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (DPHyPE); lecithin, phosphatidylethanolamine, lysolecithin, lysophosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, sphingomyelin, egg sphingomyelin (ESM), cephalin, cardiolipin, phosphatidic acid, cerebrosides, dicetylphosphate, lysophosphatidylcholine, dilinoleoylphosphatidylcholine, or mixtures thereof. It is to be understood that other diacylphosphatidylcholine and diacylphosphatidylethanolamine phospholipids can also be used. The acyl groups in these lipids are preferably acyl groups derived from fatty acids having C₁₀-C₂₄ carbon chains, *e.g.*, lauroyl, myristoyl, palmitoyl, stearoyl, or oleoyl.

[0099] Other examples of non-cationic lipids suitable for use in the lipid nanoparticles include nonphosphorous lipids such as, *e.g.*, stearylamine, dodecylamine, hexadecylamine, acetyl palmitate, glycerolricinoleate, hexadecyl stearate, isopropyl myristate, amphoteric acrylic polymers, triethanolamine-lauryl sulfate, alkyl-aryl sulfate polyethoxylated fatty acid amides, dioctadecyldimethyl ammonium bromide, ceramide, sphingomyelin, and the like.

[00100] In some embodiments, the non-cationic lipid is a phospholipid. In some embodiments, the non-cationic lipid is selected from DSPC, DPPC, DMPC, DLPE, DMPE, DPHyPe, DOPC, POPC, DOPE, and SM. In some preferred embodiments, the non-cationic lipid is DSPC.

[00101] Exemplary non-cationic lipids are described in PCT and US patent applications listed in **Table 1B**, contents of all which are incorporated herein by reference in their entireties.

Table 1B. Non-cationic lipids

PCT Publication	US Publication
WO2017/099823	US2018/0028664

[00102] The non-cationic lipid can comprise 0-60% (mol) of the total lipid present in the lipid nanoparticle. For example, the non-cationic lipid content is 0-50% (mol), 5-20% (mol) or 10-15% (mol) of the total lipid present in the lipid nanoparticle. In various embodiments, the molar ratio of condensing agent, *e.g.* a cationic lipid, to the non-cationic lipid ranges from about 2:1 to about 8:1. In some other embodiments, the molar ratio of condensing agent, *e.g.* a cationic lipid, to the neutral lipid ranges from about 1:2 to about 1:8.

[00103] In some embodiments, the lipid nanoparticles do not comprise any phospholipids.

[00104] In some aspects, the lipid nanoparticle can further comprise a component, such as a sterol, to provide membrane integrity. Generally, the component used for providing membrane integrity is non-fusogenic, *i.e.*, a component that does not or substantially does not fuse with a membrane or, if does fuse with a membrane, does not destabilize the membrane. Generally, the component used in the nanoparticles of the invention for providing membrane integrity does not have, or has very little, fusogenic activity at any pH.

[00105] One exemplary sterol that can be used in the lipid nanoparticle is cholesterol and derivatives thereof. Non-limiting examples of cholesterol derivatives include polar analogues such as

5 α -cholestanol, 5 β -coprostanol, cholesteryl-(2'-hydroxy)-ethyl ether, cholesteryl-(4'-hydroxy)-butyl ether, and 6-ketocholestanol; non-polar analogues such as 5 α -cholestane, cholestenone, 5 α -cholestanone, 5 β -cholestanone, and cholesteryl decanoate; and mixtures thereof. In some embodiments, the cholesterol derivative is a polar analogue such as cholesteryl-(4'-hydroxy)-butyl ether. In some embodiments, cholesterol derivative is cholesteryl hemisuccinate (CHEMS).

[00106] Exemplary cholesterol derivatives are described in PCT and US patent applications listed in **Table 1C**, the contents of all which are incorporated herein by reference in their entireties.

Table 1C. Cholesterol derivatives

PCT Publication	US Publication
WO2009/127060	US2010/0130588

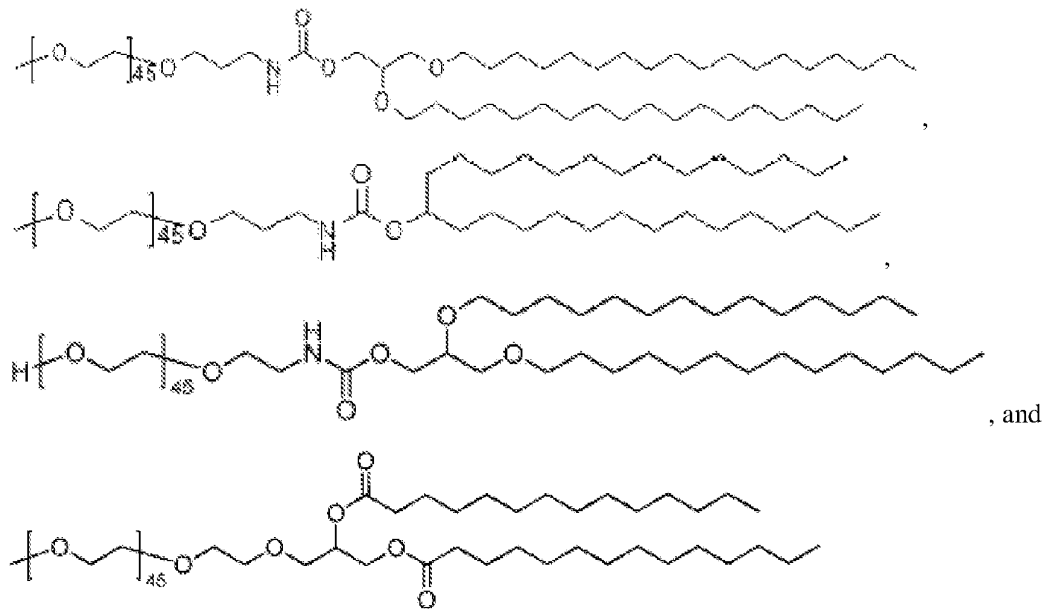
[00107] The component providing membrane integrity, such as a sterol, can comprise 0-50% (mol) of the total lipid present in the lipid nanoparticle. In some embodiments, such a component is 20-50% (mol) 30-40% (mol) of the total lipid content of the lipid nanoparticle.

[00108] In some aspects, the lipid nanoparticle can further comprise a polyethylene glycol (PEG) or a conjugated lipid molecule. Generally, these are used to inhibit aggregation of lipid nanoparticles and/or provide steric stabilization. Exemplary conjugated lipids include, but are not limited to, PEG-lipid conjugates, polyoxazoline (POZ)-lipid conjugates, polyamide-lipid conjugates (such as ATTA-lipid conjugates), cationic-polymer lipid (CPL) conjugates, and mixtures thereof. In some embodiments, conjugated lipid molecule is a PEG-lipid conjugate, for example, a (methoxy polyethylene glycol)-conjugated lipid. Generally, the conjugated lipid is non-fusogenic, *i.e.*, a conjugated lipid that does not or substantially does not fuse with a membrane or, if does fuse with a membrane, does not destabilize the membrane. Generally, a conjugated lipid in the nanoparticles of the invention for providing membrane integrity does not have, or has very little, fusogenic activity at any pH.

[00109] Exemplary PEG-lipid conjugates include, but are not limited to, PEG-diacylglycerol (DAG) (such as 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG-DMG)), PEG-dialkoxypentyl (DAA), PEG-phospholipid, PEG-ceramide (Cer), a pegylated phosphatidylethanolamine (PEG-PE), PEG succinate diacylglycerol (PEGS-DAG) (such as 4-O-(2',3'-di(tetradecanoyloxy)propyl)-1-O-(w-methoxy(polyethoxy)ethyl) butanedioate (PEG-S-DMG)), PEG dialkoxypentylcarbam, N-(carbonyl-methoxypolyethylene glycol 2000)-1,2-distearoyl-sn-glycero-3-phosphoethanolamine sodium salt, or a mixture thereof. Additional exemplary PEG-lipid conjugates are described, for example, in US5,885,613, US6,287,591, US2003/0077829, US2003/0077829, US2005/0175682, US2008/0020058, US2011/0117125, US2010/0130588, US2016/0376224, US2017/0119904, US2018/0028664, US2015/0376115 and US2016/0376224, the contents of all which are incorporated herein by reference in their entireties.

[00110] The PEG-DAA conjugate can be, for example, PEG-dilauryloxypropyl, PEG-dimyristyloxypropyl, PEG-dipalmitoyloxypropyl, or PEG-distearoyloxypropyl. The PEG-lipid can be

one or more of PEG-DMG, PEG-dilaurylglycerol, PEG-dipalmitoylglycerol, PEG-disterylgllycerol, PEG-dilaurylglycamide, PEG-dimyristylglycamide, PEG-dipalmitoylglycamide, PEG-disterylglycamide, PEG-cholesterol (1-[8'-(Cholest-5-en-3[beta]-oxy)carboxamido-3',6'-dioxaoctanyl] carbamoyl-[omega]-methyl-poly(ethylene glycol), PEG-DMB (3,4-Ditetradecoxybenzyl- [omega]-methyl-poly(ethylene glycol) ether), and 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000]. In some examples, the PEG-lipid can be selected from the group consisting of PEG-DMG, 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000],



[00111] In some embodiments, the PEG-lipid is selected from the group consisting N-(Carbonyl-methoxypolyethyleneglycol_n)-1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine (DMPE-PEG_n, where *n* is 350, 500, 750, 1000 or 2000), N-(Carbonyl-methoxypolyethyleneglycol_n)-1,2-distearoyl-sn-glycero-3-phosphoethanolamine (DSPE-PEG_n, where *n* is 350, 500, 750, 1000 or 2000), DSPE-polyglycelin-cyclohexyl-carboxylic acid, DSPE-polyglycelin-2-methylglutar-carboxylic acid, polyethylene glycol-dimyristolglycerol (PEG-DMG), polyethylene glycol-distearoyl glycerol (PEG-DSG), or N-octanoyl-sphingosine-1-{succinyl[methoxy(polyethylene glycol)2000]1 (C8 PEG2000 Ceramide). In some examples of DMPE-PEG_n, where *n* is 350, 500, 750, 1000 or 2000, the PEG-lipid is N-(Carbonyl-methoxypolyethyleneglycol 2000)-1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine (DMPE-PEG 2,000). In some examples of DSPE-PEG_n, where *n* is 350, 500, 750, 1000 or 2000, the PEG-lipid is N-(Carbonyl-methoxypolyethyleneglycol 2000)-1,2-distearoyl-sn-glycero-3-phosphoethanolamine (DSPE-PEG 2,000). In some preferred embodiments, the PEG-lipid is PEG-DMG.

[00112] In some embodiments, the conjugated lipid, *e.g.*, PEG-lipid, includes a targeting ligand, *e.g.*, first or second targeting ligand. For example, PEG-DMG conjugated with a GalNAc ligand.

[00113] Lipids conjugated with a molecule other than a PEG can also be used in place of PEG-lipid. For example, polyoxazoline (POZ)-lipid conjugates, polyamide-lipid conjugates (such as ATTA-lipid conjugates), and cationic-polymer lipid (CPL) conjugates can be used in place of or in addition to the PEG-lipid.

[00114] Exemplary conjugated lipids, *i.e.*, PEG-lipids, (POZ)-lipid conjugates, ATTA-lipid conjugates and cationic polymer-lipids are described in PCT and US patent applications listed in **Table 1D**, the contents of all which are incorporated herein by reference in their entireties.

Table 1D. Conjugated lipids

PCT Publication	US Publication
WO1996/010392	US5,885,613
WO1998/051278	US6,287,591
WO2002/087541	US2003/0077829
WO2005/026372	US2005/0175682
WO2008/147438	US2008/0020058
WO2009/086558	US2011/0117125
WO2012/000104	US2013/0303587
WO2017/117528	
WO2017/099823	US2018/0028664
WO2015/199952	US2015/0376115
WO2017/004143	US2016/0376224
WO2015/095346	US2016/0317458
	US6,320,017
	US6,586,559
WO2012/000104	US2013/0303587
WO2012/000104	US2013/0303587
WO2010/006282	US20110123453

[00115] The PEG or the conjugated lipid can comprise 0-20% (mol) of the total lipid present in the lipid nanoparticle. In some embodiments, PEG or the conjugated lipid content is 0.5-10%, 1-5%, 2-5%, 3-5%, or 3-4% (mol) of the total lipid present in the lipid nanoparticle.

[00116] Molar ratios of the condensing agent (*e.g.*, a cationic lipid), non-cationic-lipid, sterol, and PEG/conjugated lipid can be varied as needed. For example, the lipid particle can comprise 20-70% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 0-60% cholesterol by mole or by total weight of the composition, 0-30% non-cationic-lipid by mole or by total weight of the composition and 1-10% conjugated lipid by mole or by total weight of the composition. Preferably, the composition comprises 30-40% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 40-50% cholesterol by mole or by total weight of the composition, and 10-20% non-cationic-lipid by mole or by total weight of the composition. In some other embodiments, the composition is 50-75% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 20-40% cholesterol by mole or by total weight of the composition, and 5 to 10% non-cationic-lipid, by mole or by total weight of the composition and 1-10% conjugated lipid by mole or by total weight of the composition. The composition may contain 60-70%

condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 25-35% cholesterol by mole or by total weight of the composition, and 5-10% non-cationic-lipid by mole or by total weight of the composition. The composition may also contain up to 90% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition and 2 to 15% non-cationic lipid by mole or by total weight of the composition. The formulation may also be a lipid nanoparticle formulation, for example comprising 8-30% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 5-30% non-cationic lipid by mole or by total weight of the composition, and 0-20% cholesterol by mole or by total weight of the composition; 4-25% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 4-25% non-cationic lipid by mole or by total weight of the composition, 2 to 25% cholesterol by mole or by total weight of the composition, 10 to 35% conjugate lipid by mole or by total weight of the composition, and 5% cholesterol by mole or by total weight of the composition; or 2-30% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition, 2-30% non-cationic lipid by mole or by total weight of the composition, 1 to 15% cholesterol by mole or by total weight of the composition, 2 to 35% conjugate lipid by mole or by total weight of the composition, and 1-20% cholesterol by mole or by total weight of the composition; or even up to 90% condensing agent (*e.g.*, a cationic lipid) by mole or by total weight of the composition and 2-10% non-cationic lipids by mole or by total weight of the composition, or even 100% cationic lipid by mole or by total weight of the composition. In some embodiments, the lipid particle formulation comprises condensing agent (*e.g.*, a cationic lipid), phospholipid, cholesterol and a PEG-ylated lipid in a molar ratio of 50:10:38.5:1.5. In some embodiments, the lipid particle formulation comprises condensing agent (*e.g.*, a cationic lipid), phospholipid, cholesterol and a PEG-ylated lipid in a molar ratio of 50:10:37:3. In some other embodiments, the lipid particle formulation comprises condensing agent (*e.g.*, a cationic lipid), cholesterol and a PEG-ylated lipid in a molar ratio of 60:38.5:1.5. In some other embodiments, the lipid particle formulation comprises condensing agent (*e.g.*, a cationic lipid), cholesterol and a PEG-ylated lipid in a molar ratio of 58:39:3.

[00117] In some embodiments, the lipid particle comprises condensing agent (*e.g.*, a cationic lipid), non-cationic lipid (*e.g.* phospholipid), a sterol (*e.g.*, cholesterol) and a PEG-ylated lipid, where the molar ratio of lipids ranges from 20 to 70 mole percent for the condensing agent (*e.g.*, a cationic lipid), with a target of 40-60, the mole percent of non-cationic lipid ranges from 0 to 30, with a target of 0 to 15, the mole percent of sterol ranges from 20 to 70, with a target of 30 to 50, and the mole percent of PEG-ylated lipid ranges from 1 to 6, with a target of 2 to 5 or 2.5 to 4.

[00118] In some embodiments, the lipid particle comprises condensing agent (*e.g.*, a cationic lipid) / non-cationic-lipid / sterol / conjugated lipid at a molar ratio of 50:10:38.5:1.5. In some embodiments, the lipid particle comprises condensing agent (*e.g.*, a cationic lipid) / non-cationic-lipid / sterol / conjugated lipid at a molar ratio of 50:10:35:3. In some other embodiments, the lipid particle comprises condensing agent (*e.g.*, a cationic lipid) / non-cationic-lipid / sterol / conjugated lipid at a

molar ratio of 20:40:38.5:1.5. In some other embodiments, the lipid particle comprises condensing agent (*e.g.*, a cationic lipid) / non-cationic-lipid / sterol / conjugated lipid at a molar ratio of 20:40:37:3.

[00119] In other aspects, the disclosure provides for a lipid nanoparticle formulation comprising phospholipids, lecithin, phosphatidylcholine and phosphatidylethanolamine.

[00120] In some embodiments, the method comprises co-administering one or more additional compounds, *i.e.*, in addition to the lipid nanoparticle and the endosomolytic agent. These one or more additional compounds can be included in the lipid particle or a composition comprising the endosomolytic agent. Those compounds can be administered separately, or the additional compounds can be included one or both of the lipid nanoparticle and a composition comprising the endosomolytic agent. For example, the lipid nanoparticles can contain other compounds in addition to the ceDNA or at least a second ceDNA, different than the first. Without limitations, other additional compounds can be selected from the group consisting of small or large organic or inorganic molecules, monosaccharides, disaccharides, trisaccharides, oligosaccharides, polysaccharides, peptides, proteins, peptide analogs and derivatives thereof, peptidomimetics, nucleic acids, nucleic acid analogs and derivatives, an extract made from biological materials, or any combinations thereof.

[00121] In some embodiments, the additional compound can be a therapeutic agent. The therapeutic agent can be selected from any class suitable for the therapeutic objective. In other words, the therapeutic agent can be selected from any class suitable for the therapeutic objective. In other words, the therapeutic agent can be selected according to the treatment objective and biological action desired. For example, if the ceDNA within the LNP is useful for treating cancer, the additional compound can be an anti-cancer agent. In another example, if the LNP containing the ceDNA is useful for treating an infection, the additional compound can be an antimicrobial agent. In still another example, the additional compound can be a compound that inhibits an immune response, *e.g.*, an immunosuppressant. In some embodiments, different cocktails of different lipid nanoparticles containing different compounds, such as a ceDNA encoding a different protein, a different compound, such as a therapeutic, etc...

[00122] In some embodiments, the additional compound can be an immune modulating agent. For example, the additional compound can be an immunosuppressant. Exemplary immune modulating agents include, but are not limited to, cGAS inhibitors, TLR9 antagonists, and Caspase-1 inhibitors.

[00123] Also provided herein is a composition comprising a lipid nanoparticle and an endosomolytic agent.

[00124] Also provided herein is a pharmaceutical composition comprising the lipid nanoparticle, endosomolytic agent, and a pharmaceutically acceptable carrier or excipient.

Some Exemplary LNP Characteristics

[00125] Generally, the lipid nanoparticles have a mean diameter selected to provide an intended therapeutic effect. Accordingly, in some aspects, the lipid nanoparticle has a mean diameter from about 30 nm to about 150 nm, more typically from about 50 nm to about 150 nm, more typically about 60 nm to about 130 nm, more typically about 70 nm to about 110 nm, most typically about 85 nm to about 105 nm, and preferably about 100 nm. In some aspects, the disclosure provides for lipid particles that are larger in relative size to common nanoparticles and about 150 to 250 nm in size. Lipid nanoparticle particle size can be determined by quasi-elastic light scattering using, for example, a Malvern Zetasizer Nano ZS (Malvern, UK) system.

[00126] Depending on the intended use of the lipid particles, the proportions of the components can be varied and the delivery efficiency of a particular formulation can be measured using, for example, an endosomal release parameter (ERP) assay.

[00127] The ceDNA can be complexed with the lipid portion of the particle or encapsulated in the lipid position of the lipid nanoparticle. In some embodiments, the ceDNA can be fully encapsulated in the lipid position of the lipid nanoparticle, thereby protecting it from degradation by a nuclease, *e.g.*, in an aqueous solution. In some embodiments, the ceDNA in the lipid nanoparticle is not substantially degraded after exposure of the lipid nanoparticle to a nuclease at 37°C for at least about 20, 30, 45, or 60 minutes. In some embodiments, the ceDNA in the lipid nanoparticle is not substantially degraded after incubation of the particle in serum at 37°C for at least about 30, 45, or 60 minutes or at least about 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, or 36 hours.

[00128] In certain embodiments, the lipid nanoparticles are substantially non-toxic to mammals such as humans.

[00129] In some embodiments, lipid nanoparticles are solid core particles that possess at least one lipid bilayer. In other embodiments, the lipid nanoparticles have a non-bilayer structure, *i.e.*, a non-lamellar (*i.e.*, non-bilayer) morphology. Without limitations, the non-bilayer morphology can include, for example, three dimensional tubes, rods, cubic symmetries, *etc.* The non-lamellar morphology (*i.e.*, non-bilayer structure) of the lipid particles can be determined using analytical techniques known to and used by those of skill in the art. Such techniques include, but are not limited to, Cryo-Transmission Electron Microscopy (“Cryo-TEM”), Differential Scanning calorimetry (“DSC”), X-Ray Diffraction, *etc.* For example, the morphology of the lipid nanoparticles (lamellar vs. non-lamellar) can readily be assessed and characterized using, *e.g.*, Cryo-TEM analysis as described in US2010/0130588, content of which is incorporated herein by reference in its entirety.

[00130] In some further embodiments, the lipid nanoparticles having a non-lamellar morphology are electron dense.

[00131] In embodiments, the lipid nanoparticle is either unilamellar or multilamellar in structure. In some aspects, the disclosure provides for a lipid nanoparticle formulation that comprises multi-vesicular particles and/or foam-based particles.

[00132] The lipid nanoparticle may have positive or negative zeta potential. In various embodiments, the lipid nanoparticle has a positive zeta potential.

[00133] By controlling the composition and concentration of the lipid components, one can control the rate at which the lipid conjugate exchanges out of the lipid particle and, in turn, the rate at which the lipid nanoparticle becomes fusogenic. In addition, other variables including, *e.g.*, pH, temperature, or ionic strength, can be used to vary and/or control the rate at which the lipid nanoparticle becomes fusogenic. Other methods which can be used to control the rate at which the lipid nanoparticle becomes fusogenic will become apparent to those of skill in the art upon reading this disclosure. Also, by controlling the composition and concentration of the lipid conjugate, one can control the lipid particle size.

[00134] The pKa of formulated cationic lipids can be correlated with the effectiveness of the LNPs for delivery of nucleic acids (see Jayaraman *et al.*, *Angewandte Chemie, International Edition* (2012), 51(34), 8529-8533; Semple *et al.*, *Nature Biotechnology* 28, 172-176 (2010), both of which are incorporated by reference in their entireties). The preferred range of pKa is ~5 to ~7. The pKa of the cationic lipid can be determined in lipid nanoparticles using an assay based on fluorescence of 2-(p-toluidino)-6-naphthalene sulfonic acid (TNS).

[00135] Encapsulation of ceDNA in lipid particles can be determined by performing a membrane-impermeable fluorescent dye exclusion assay, which uses a dye that has enhanced fluorescence when associated with nucleic acid, for example, an Oligreen[®] assay or PicoGreen[®] assay. Generally, encapsulation is determined by adding the dye to the lipid particle formulation, measuring the resulting fluorescence, and comparing it to the fluorescence observed upon addition of a small amount of nonionic detergent. Detergent-mediated disruption of the lipid bilayer releases the encapsulated ceDNA, allowing it to interact with the membrane-impermeable dye. Encapsulation of ceDNA can be calculated as $E = (I_0 - I)/I_0$, where I and I_0 refers to the fluorescence intensities before and after the addition of detergent.

[00136] Lipid nanoparticles can form spontaneously upon mixing of ceDNA and the lipid(s). Depending on the desired particle size distribution, the resultant nanoparticle mixture can be extruded through a membrane (*e.g.*, 100 nm cut-off) using, for example, a thermobarrel extruder, such as Lipex Extruder (Northern Lipids, Inc). In some cases, the extrusion step can be omitted. Ethanol removal and simultaneous buffer exchange can be accomplished by, for example, dialysis or tangential flow filtration.

[00137] Generally, lipid nanoparticles can be formed by any method known in the art including. For example, the lipid nanoparticles can be prepared by the methods described, for example, in US2013/0037977, US2010/0015218, US2013/0156845, US2013/0164400, US2012/0225129, and

US2010/0130588, the content of each of which is incorporated herein by reference in its entirety. In some embodiments, lipid nanoparticles can be prepared using a continuous mixing method, a direct dilution process, or an in-line dilution process. The processes and apparatuses for preparing lipid nanoparticles using direct dilution and in-line dilution processes are described in US2007/0042031, the content of which is incorporated herein reference in its entirety. The processes and apparatuses for preparing lipid nanoparticles using step-wise dilution processes are described in US2004/0142025, the content of which is incorporated herein reference in its entirety.

[00138] In one non-limiting example, the lipid nanoparticles can be prepared by an impinging jet process. Generally, the particles are formed by mixing lipids dissolved in alcohol (*e.g.*, ethanol) with ceDNA dissolved in a buffer, *e.g.*, a citrate buffer, a sodium acetate buffer, a sodium acetate and magnesium chloride buffer, a malic acid buffer, a malic acid and sodium chloride buffer, or a sodium citrate and sodium chloride buffer. The mixing ratio of lipids to ceDNA can be about 45-55% lipid and about 65-45% ceDNA.

[00139] The lipid solution can contain a condensing agent (*e.g.*, a cationic lipid), a non-cationic lipid (*e.g.*, a phospholipid, such as DSPC), PEG or PEG conjugated molecule (*e.g.*, PEG-lipid), and a sterol (*e.g.*, cholesterol) at a total lipid concentration of 5-30 mg/mL, more likely 5-15 mg/mL, most likely 9-12 mg/mL in an alcohol, *e.g.*, in ethanol.

[00140] In the lipid solution, mol ratio of the lipids can range from about 25-98% for the cationic lipid, preferably about 35-65%; about 0-15% for the non-ionic lipid, preferably about 0-12%; about 0-15% for the PEG or PEG conjugated molecule, preferably about 1-6% or 2-5%; and about 0-75% for the sterol, preferably about 30-50%.

[00141] The ceDNA solution can comprise the ceDNA at a concentration range from 0.3 to 1.0 mg/mL, preferably 0.3-0.9 mg/mL in buffered solution, with pH in the range of 3.5-5.

[00142] For forming the LNPs, the two liquids are heated to a temperature in the range of about 15-40°C, preferably about 30-40°C, and then mixed, for example, in an impinging jet mixer, instantly forming the LNP. The mixing flow rate can range from 10-600 ml/min. The tube ID can have a range from 0.25 to 1.0 mm and a total flow rate from 10-600 mL/min. The combination of flow rate and tubing ID can have the effect of controlling the particle size of the LNPs between 30 and 200 nm. The solution can then be mixed with a buffered solution at a higher pH with a mixing ratio in the range of 1:1 to 1:3 vol:vol, preferably about 1:2 vol:vol. If needed this buffered solution can be at a temperature in the range of 15-40°C or 30-40°C. The mixed LNPs can then undergo an anion exchange filtration step. Prior to the anion exchange, the mixed LNPs can be incubated for a period of time, for example 30mins to 2 hours. The temperature during incubating can be in the range of 15-40°C or 30-40°C. After incubating the solution is filtered through a filter, such as a 0.8µm filter, containing an anion exchange separation step. This process can use tubing IDs ranging from 1 mm ID to 5 mm ID and a flow rate from 10 to 2000 mL/min.

[00143] After formation, the LNPs can be concentrated and diafiltered via an ultrafiltration process where the alcohol is removed and the buffer is exchanged for the final buffer solution, for example, phosphate buffered saline (PBS) at about pH 7, *e.g.*, about pH 6.9, about pH 7.0, about pH 7.1, about pH 7.2, about pH 7.3, or about pH 7.4.

[00144] The ultrafiltration process can use a tangential flow filtration format (TFF) using a membrane nominal molecular weight cutoff range from 30-500 KD. The membrane format is hollow fiber or flat sheet cassette. The TFF processes with the proper molecular weight cutoff can retain the LNP in the retentate and the filtrate or permeate contains the alcohol; citrate buffer and final buffer wastes. The TFF process is a multiple step process with an initial concentration to a ceDNA concentration of 1 -3 mg/mL. Following concentration, the LNPs solution is diafiltered against the final buffer for 10 -20 volumes to remove the alcohol and perform buffer exchange. The material can then be concentrated an additional 1-3 fold. The concentrated LNP solution can be sterile filtered.

I. Therapeutic nucleic acid vector in general

[00145] Nucleic acids are large, highly charged, rapidly degraded and cleared from the body, and offer generally poor pharmacological properties because they are recognized as a foreign matter to the body and become a target of an immune response (*e.g.*, innate immune response). Hence, certain nucleic acids, such as therapeutic nucleic acids or nucleic acids used for research purposes (*e.g.*, antisense oligonucleotide or viral vectors) can often trigger immune responses *in vivo*. The present disclosure provides pharmaceutical compositions and methods that may ameliorate, reduce or eliminate such immune responses and enhance efficacy of the nucleic acids by increasing expression levels through maximizing the durability of the nucleic acid in a reduced immune-responsive state in a subject recipient. This may also minimize any potential adverse events that may lead to an organ damage or other toxicity in the course of gene therapy.

[00146] The immunogenic / immunostimulatory nucleic acids can include both deoxyribonucleic acids and ribonucleic acids. For deoxyribonucleic acids (DNA), a particular sequence or motif has been shown to induce immune stimulation in mammals. These sequence or motifs include, but are not limited to, CpG motifs, pyrimidine-rich sequences, and palindrome sequences. CpG motifs in deoxyribonucleic acid are often recognized by the endosomal toll-like receptor 9 (TLR-9) which, in turn, triggers both the innate immune stimulatory pathway and the acquired immune stimulatory pathway.

[00147] In some embodiments, chemical modification of oligonucleotides for the purpose of altered and improved *in vivo* properties (delivery, stability, life-time, folding, target specificity), as well as their biological function and mechanism that directly correlate with therapeutic application, are described where appropriate.

[00148] Illustrative therapeutic nucleic acids of the present disclosure include, but are not limited to, minigenes, plasmids, minicircles, small interfering RNA (siRNA), microRNA (miRNA),

antisense oligonucleotides (ASO), ribozymes, closed ended double stranded DNA (*e.g.*, ceDNA, CELiD, linear covalently closed DNA (“ministring”), doggybone (dbDNA™), protelomere closed ended DNA, or dumbbell linear DNA), dicer-substrate dsRNA, small hairpin RNA (shRNA), asymmetrical interfering RNA (aiRNA), microRNA (miRNA), mRNA, tRNA, rRNA, and DNA viral vectors, viral RNA vector, and any combination thereof.

[00149] siRNA or miRNA that can downregulate the intracellular levels of specific proteins through a process called RNA interference (RNAi) are also contemplated by the present invention to be nucleic acid therapeutics. After siRNA or miRNA is introduced into the cytoplasm of a host cell, these double-stranded RNA constructs can bind to a protein called RISC. The sense strand of the siRNA or miRNA is removed by the RISC complex. The RISC complex, when combined with the complementary mRNA, cleaves the mRNA and release the cut strands. RNAi is by inducing specific destruction of mRNA that results in downregulation of a corresponding protein.

[00150] Antisense oligonucleotides (ASO) and ribozymes that inhibit mRNA translation into protein can be nucleic acid therapeutics. For antisense constructs, these single stranded deoxy nucleic acids have a complementary sequence to the sequence of the target protein mRNA, and Watson - capable of binding to the mRNA by Crick base pairing. This binding prevents translation of a target mRNA, and / or triggers RNaseH degradation of the mRNA transcript. As a result, the antisense oligonucleotide has increased specificity of action (*i.e.*, down-regulation of a specific disease-related protein).

[00151] In any of the methods provided herein, the therapeutic nucleic acid can be a therapeutic RNA. The therapeutic RNA can be an inhibitor of mRNA translation, agent of RNA interference (RNAi), catalytically active RNA molecule (ribozyme), transfer RNA (tRNA) or an RNA that binds an mRNA transcript (ASO), protein or other molecular ligand (aptamer). In any of the methods provided herein, the agent of RNAi can be a double-stranded RNA, single-stranded RNA, micro RNA, short interfering RNA, short hairpin RNA, or a triplex-forming oligonucleotide.

[00152] According to some embodiments, the therapeutic nucleic acid is a closed ended double stranded DNA, *e.g.*, a ceDNA. According to some embodiments, the expression and/or production of a therapeutic protein in a cell is from a non-viral DNA vector, *e.g.*, a ceDNA vector. A distinct advantage of ceDNA vectors for expression of a therapeutic protein over traditional AAV vectors, and even lentiviral vectors, is that there is no size constraint for the heterologous nucleic acid sequences encoding a desired protein. Thus, even a large therapeutic protein can be expressed from a single ceDNA vector. Thus, ceDNA vectors can be used to express a therapeutic protein in a subject in need thereof.

[00153] In some embodiments, the methods and compositions described herein relate to the use of a ceDNA vector with a non-fusogenic LNP and an endosomolytic agent where the ceDNA vector is, but is not limited to, a ceDNA vector comprising asymmetric ITRS as disclosed in International Patent Application PCT/US18/49996, filed on September 7, 2018 (see, *e.g.*, Examples 1-4),

incorporated by reference in its entirety herein; a ceDNA vector for gene editing as disclosed on the International Patent Application PCT/US18/64242 filed on December 6, 2018 (see, *e.g.*, Examples 1-7), incorporated by reference in its entirety herein, or a ceDNA vector for production of antibodies or fusion proteins, as disclosed in the International Patent Application PCT/US19/18016, filed on February 14, 2019, (*e.g.*, see Examples 1-4), incorporated by reference in its entirety herein, or a ceDNA vector for controlled transgene expression, as disclosed in International Patent Application PCT/US19/18927 filed on February 22, 2019, , incorporated by reference in its entirety herein. In some embodiments, it is also envisioned that the methods and compositions described herein can be used with a synthetically produced ceDNA vector, *e.g.*, a ceDNA vector produced in a cell free or insect-free system of ceDNA production, as disclosed in International Application PCT/US19/14122, filed on January 18, 2019, incorporated by reference in its entirety herein.

[00154] Embodiments of the invention are based on use of non-fusogenic LNP and an endosomolytic agents in methods and compositions for delivery of close ended linear duplexed (ceDNA) vectors, where the ceDNA vectors can express a desired transgene. In some embodiments, the transgene is a sequence encoding a therapeutic protein. The ceDNA vectors for expression of a desired transgene as described herein are not limited by size, thereby permitting, for example, expression of all of the components necessary for expression of a transgene from a single vector. The ceDNA vector for expression of a desired transgene is preferably duplex, *e.g.* self-complementary, over at least a portion of the molecule, such as the expression cassette (*e.g.* ceDNA is not a double stranded circular molecule). The ceDNA vector has covalently closed ends, and thus is resistant to exonuclease digestion (*e.g.* exonuclease I or exonuclease III), *e.g.* for over an hour at 37°C.

[00155] In general, a ceDNA vector for expression of a desired transgene is as disclosed herein, comprises in the 5' to 3' direction: a first adeno-associated virus (AAV) inverted terminal repeat (ITR), a nucleotide sequence of interest (for example an expression cassette as described herein) and a second AAV ITR. The ITR sequences selected from any of: (i) at least one WT ITR and at least one modified AAV inverted terminal repeat (mod-ITR) (*e.g.*, asymmetric modified ITRs); (ii) two modified ITRs where the mod-ITR pair have a different three-dimensional spatial organization with respect to each other (*e.g.*, asymmetric modified ITRs), or (iii) symmetrical or substantially symmetrical WT-WT ITR pair, where each WT-ITR has the same three-dimensional spatial organization, or (iv) symmetrical or substantially symmetrical modified ITR pair, where each mod-ITR has the same three-dimensional spatial organization.

[00156] Encompassed herein are methods and compositions comprising the ceDNA vector for transgene production, which may further include a delivery system, such as but not limited to, a liposome nanoparticle delivery system. Non-limiting exemplary liposome nanoparticle systems encompassed for use are disclosed herein. In some aspects, the disclosure provides for a lipid nanoparticle comprising ceDNA and an ionizable lipid. For example, a lipid nanoparticle formulation that is made and loaded with a ceDNA vector obtained by the process is disclosed in International

Application PCT/US2018/050042, filed on September 7, 2018, incorporated by reference in its entirety herein.

[00157] The ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein have no packaging constraints imposed by the limiting space within the viral capsid. ceDNA vectors represent a viable eukaryotically-produced alternative to prokaryote-produced plasmid DNA vectors, as opposed to encapsulated AAV genomes. This permits the insertion of control elements, *e.g.*, regulatory switches as disclosed herein, large transgenes, multiple transgenes etc.

[00158] FIG. 1A-1E show schematics of non-limiting, exemplary ceDNA vectors for expression of a desired transgene or therapeutic protein, or the corresponding sequence of ceDNA plasmids. ceDNA vectors for expression of a desired transgene or therapeutic protein are capsid-free and can be obtained from a plasmid encoding in this order: a first ITR, an expression cassette comprising a transgene and a second ITR. The expression cassette may include one or more regulatory sequences that allows and/or controls the expression of the transgene, *e.g.*, where the expression cassette can comprise one or more of, in this order: an enhancer/promoter, an ORF reporter (transgene), a post-transcription regulatory element (*e.g.*, WPRE), and a polyadenylation and termination signal (*e.g.*, BGH polyA).

[00159] The expression cassette can also comprise an internal ribosome entry site (IRES) and/or a 2A element. The cis-regulatory elements include, but are not limited to, a promoter, a riboswitch, an insulator, a mir-regulatable element, a post-transcriptional regulatory element, a tissue- and cell type-specific promoter and an enhancer. In some embodiments the ITR can act as the promoter for the transgene, *e.g.*, a desired transgene or therapeutic protein. In some embodiments, the ceDNA vector comprises additional components to regulate expression of the transgene, for example, a regulatory switch, which are described herein in the section entitled “Regulatory Switches” for controlling and regulating the expression of a desired transgene or therapeutic protein, and can include if desired, a regulatory switch which is a kill switch to enable controlled cell death of a cell comprising a ceDNA vector.

[00160] The expression cassette can comprise more than 4000 nucleotides, 5000 nucleotides, 10,000 nucleotides or 20,000 nucleotides, or 30,000 nucleotides, or 40,000 nucleotides or 50,000 nucleotides, or any range between about 4000-10,000 nucleotides or 10,000-50,000 nucleotides, or more than 50,000 nucleotides. In some embodiments, the expression cassette can comprise a transgene in the range of 500 to 50,000 nucleotides in length. In some embodiments, the expression cassette can comprise a transgene in the range of 500 to 75,000 nucleotides in length. In some embodiments, the expression cassette can comprise a transgene which is in the range of 500 to 10,000 nucleotides in length. In some embodiments, the expression cassette can comprise a transgene which is in the range of 1000 to 10,000 nucleotides in length. In some embodiments, the expression cassette can comprise a transgene which is in the range of 500 to 5,000 nucleotides in length. The ceDNA

vectors do not have the size limitations of encapsidated AAV vectors, thus enable delivery of a large-size expression cassette to provide efficient transgene expression. In some embodiments, the ceDNA vector is devoid of prokaryote-specific methylation.

[00161] ceDNA expression cassette can include, for example, an expressible exogenous sequence (*e.g.*, open reading frame) or transgene that encodes a protein that is either absent, inactive, or insufficient activity in the recipient subject or a gene that encodes a protein having a desired biological or a therapeutic effect. The transgene can encode a gene product that can function to correct the expression of a defective gene or transcript. In principle, the expression cassette can include any gene that encodes a protein, polypeptide or RNA that is either reduced or absent due to a mutation or which conveys a therapeutic benefit when overexpressed is considered to be within the scope of the disclosure.

[00162] The expression cassette can comprise any transgene (*e.g.*, encoding therapeutic protein), for example, a desired transgene or therapeutic protein useful for treating a disease in a subject. A ceDNA vector can be used to deliver and express any a desired transgene or therapeutic protein of interest in the subject, alone or in combination with nucleic acids encoding polypeptides, or non-coding nucleic acids (*e.g.*, RNAi, miRs etc.), as well as exogenous genes and nucleotide sequences, including virus sequences in a subjects' genome, *e.g.*, HIV virus sequences and the like. Preferably a ceDNA vector disclosed herein is used for therapeutic purposes (*e.g.*, for medical, diagnostic, or veterinary uses) or immunogenic polypeptides. In certain embodiments, a ceDNA vector is useful to express any gene of interest in the subject, which includes one or more polypeptides, peptides, ribozymes, peptide nucleic acids, siRNAs, RNAis, antisense oligonucleotides, antisense polynucleotides, or RNAs (coding or non-coding; *e.g.*, siRNAs, shRNAs, micro-RNAs, and their antisense counterparts (*e.g.*, antagoMiR)), antibodies, fusion proteins, or any combination thereof.

[00163] The expression cassette can also encode polypeptides, sense or antisense oligonucleotides, or RNAs (coding or non-coding; *e.g.*, siRNAs, shRNAs, micro-RNAs, and their antisense counterparts (*e.g.*, antagoMiR)). Expression cassettes can include an exogenous sequence that encodes a reporter protein to be used for experimental or diagnostic purposes, such as β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, and others well known in the art.

[00164] Sequences provided in the expression cassette, expression construct of a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein can be codon optimized for the target host cell. As used herein, the term "codon optimized" or "codon optimization" refers to the process of modifying a nucleic acid sequence for enhanced expression in the cells of the vertebrate of interest, *e.g.*, mouse or human, by replacing at least one, more than one, or a significant number of codons of the native sequence (*e.g.*, a prokaryotic sequence) with codons that are more frequently or most frequently used in the genes of that vertebrate. Various species exhibit particular bias for certain codons of a particular amino acid. Typically, codon optimization does not alter the amino acid

sequence of the original translated protein. Optimized codons can be determined using *e.g.*, Aptagen's Gene Forge® codon optimization and custom gene synthesis platform (Aptagen, Inc., 2190 Fox Mill Rd. Suite 300, Herndon, Va. 20171) or another publicly available database. In some embodiments, the nucleic acid encoding a desired transgene or therapeutic protein is optimized for human expression, and/or is a human therapeutic protein, or functional fragment thereof, as known in the art.

[00165] A transgene expressed by the ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein encodes a therapeutic protein. There are many structural features of ceDNA vectors for expression of a desired transgene or therapeutic protein that differ from plasmid-based expression vectors. ceDNA vectors may possess one or more of the following features: the lack of original (*i.e.* not inserted) bacterial DNA, the lack of a prokaryotic origin of replication, being self-containing, *i.e.*, they do not require any sequences other than the two ITRs, including the Rep binding and terminal resolution sites (RBS and TRS), and an exogenous sequence between the ITRs, the presence of ITR sequences that form hairpins, and the absence of bacterial-type DNA methylation or indeed any other methylation considered abnormal by a mammalian host. In general, it is preferred for the present vectors not to contain any prokaryotic DNA but it is contemplated that some prokaryotic DNA may be inserted as an exogenous sequence, as a non-limiting example in a promoter or enhancer region. Another important feature distinguishing ceDNA vectors from plasmid expression vectors is that ceDNA vectors are single-strand linear DNA having closed ends, while plasmids are always double-strand DNA.

[00166] ceDNA vectors for expression of a desired transgene or therapeutic protein produced by the methods provided herein preferably have a linear and continuous structure rather than a non-continuous structure, as determined by restriction enzyme digestion assay (**FIG. 4D**). The linear and continuous structure is believed to be more stable from attack by cellular endonucleases, as well as less likely to be recombined and cause mutagenesis. Thus, a ceDNA vector in the linear and continuous structure is a preferred embodiment. The continuous, linear, single strand intramolecular duplex ceDNA vector can have covalently bound terminal ends, without sequences encoding AAV capsid proteins. These ceDNA vectors are structurally distinct from plasmids (including ceDNA plasmids described herein), which are circular duplex nucleic acid molecules of bacterial origin. The complimentary strands of plasmids may be separated following denaturation to produce two nucleic acid molecules, whereas in contrast, ceDNA vectors, while having complimentary strands, are a single DNA molecule and therefore even if denatured, remain a single molecule. In some embodiments, ceDNA vectors as described herein can be produced without DNA base methylation of prokaryotic type, unlike plasmids. Therefore, the ceDNA vectors and ceDNA-plasmids are different both in term of structure (in particular, linear versus circular) and also in view of the methods used for producing and purifying these different objects (see below), and also in view of their DNA methylation which is of prokaryotic type for ceDNA-plasmids and of eukaryotic type for the ceDNA vector.

[00167] There are several advantages of using a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein over plasmid-based expression vectors, such advantages include, but are not limited to: 1) plasmids contain bacterial DNA sequences and are subjected to prokaryotic-specific methylation, *e.g.*, 6-methyl adenosine and 5-methyl cytosine methylation, whereas capsid-free AAV vector sequences are of eukaryotic origin and do not undergo prokaryotic-specific methylation; as a result, capsid-free AAV vectors are less likely to induce inflammatory and immune responses compared to plasmids; 2) while plasmids require the presence of a resistance gene during the production process, ceDNA vectors do not; 3) while a circular plasmid is not delivered to the nucleus upon introduction into a cell and requires overloading to bypass degradation by cellular nucleases, ceDNA vectors contain viral cis-elements, *i.e.*, ITRs, that confer resistance to nucleases and can be designed to be targeted and delivered to the nucleus. It is hypothesized that the minimal defining elements indispensable for ITR function are a Rep-binding site (RBS; 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60) for AAV2) and a terminal resolution site (TRS; 5'-AGTTGG-3' (SEQ ID NO: 64) for AAV2) plus a variable palindromic sequence allowing for hairpin formation; and 4) ceDNA vectors do not have the over-representation of CpG dinucleotides often found in prokaryote-derived plasmids that reportedly binds a member of the Toll-like family of receptors, eliciting a T cell-mediated immune response. In contrast, transductions with capsid-free AAV vectors disclosed herein can efficiently target cell and tissue-types that are difficult to transduce with conventional AAV virions using various delivery reagent.

II. ITRs

[00168] As disclosed herein, ceDNA vectors for expression of a desired transgene or therapeutic protein contain a transgene or heterologous nucleic acid sequence positioned between two inverted terminal repeat (ITR) sequences, where the ITR sequences can be an asymmetrical ITR pair or a symmetrical- or substantially symmetrical ITR pair, as these terms are defined herein. A ceDNA vector as disclosed herein can comprise ITR sequences that are selected from any of: (i) at least one WT ITR and at least one modified AAV inverted terminal repeat (mod-ITR) (*e.g.*, asymmetric modified ITRs); (ii) two modified ITRs where the mod-ITR pair have a different three-dimensional spatial organization with respect to each other (*e.g.*, asymmetric modified ITRs), or (iii) symmetrical or substantially symmetrical WT-WT ITR pair, where each WT-ITR has the same three-dimensional spatial organization, or (iv) symmetrical or substantially symmetrical modified ITR pair, where each mod-ITR has the same three-dimensional spatial organization, where the methods of the present disclosure may further include a delivery system, such as but not limited to a liposome nanoparticle delivery system.

[00169] In some embodiments, the ITR sequence can be from viruses of the Parvoviridae family, which includes two subfamilies: Parvovirinae, which infect vertebrates, and Densovirinae, which infect insects. The subfamily Parvovirinae (referred to as the parvoviruses) includes the genus Dependovirus, the members of which, under most conditions, require coinfection with a helper virus

such as adenovirus or herpes virus for productive infection. The genus Dependovirus includes adeno-associated virus (AAV), which normally infects humans (*e.g.*, serotypes 2, 3A, 3B, 5, and 6) or primates (*e.g.*, serotypes 1 and 4), and related viruses that infect other warm-blooded animals (*e.g.*, bovine, canine, equine, and ovine adeno-associated viruses). The parvoviruses and other members of the Parvoviridae family are generally described in Kenneth I. Berns, “Parvoviridae: The Viruses and Their Replication,” Chapter 69 in *FIELDS VIROLOGY* (3d Ed. 1996).

[00170] While ITRs exemplified in the specification and Examples herein are AAV2 WT-ITRs, one of ordinary skill in the art is aware that one can as stated above use ITRs from any known parvovirus, for example a dependovirus such as AAV (*e.g.*, AAV1, AAV2, AAV3, AAV4, AAV5, AAV 5, AAV7, AAV8, AAV9, AAV10, AAV 11, AAV12, AAVrh8, AAVrh10, AAV-DJ, and AAV-DJ8 genome. *E.g.*, NCBI: NC 002077; NC 001401; NC001729; NC001829; NC006152; NC 006260; NC 006261), chimeric ITRs, or ITRs from any synthetic AAV. In some embodiments, the AAV can infect warm-blooded animals, *e.g.*, avian (AAAV), bovine (BAAV), canine, equine, and ovine adeno-associated viruses. In some embodiments the ITR is from B19 parvovirus (GenBank Accession No: NC 000883), Minute Virus from Mouse (MVM) (GenBank Accession No. NC 001510); goose parvovirus (GenBank Accession No. NC 001701); snake parvovirus 1 (GenBank Accession No. NC 006148). In some embodiments, the 5' WT-ITR can be from one serotype and the 3' WT-ITR from a different serotype, as discussed herein.

[00171] An ordinarily skilled artisan is aware that ITR sequences have a common structure of a double-stranded Holliday junction, which typically is a T-shaped or Y-shaped hairpin structure (see *e.g.*, **FIG. 2A** and **FIG. 3A**), where each WT-ITR is formed by two palindromic arms or loops (B-B' and C-C') embedded in a larger palindromic arm (A-A'), and a single stranded D sequence, (where the order of these palindromic sequences defines the flip or flop orientation of the ITR). See, for example, structural analysis and sequence comparison of ITRs from different AAV serotypes (AAV1-AAV6) and described in Grimm *et al.*, J. Virology, 2006; 80(1); 426-439; Yan *et al.*, J. Virology, 2005; 364-379; Duan *et al.*, Virology 1999; 261; 8-14. One of ordinary skill in the art can readily determine WT-ITR sequences from any AAV serotype for use in a ceDNA vector or ceDNA-plasmid based on the exemplary AAV2 ITR sequences provided herein. See, for example, the sequence comparison of ITRs from different AAV serotypes (AAV1-AAV6, and avian AAV (AAAV) and bovine AAV (BAAV)) described in Grimm *et al.*, J. Virology, 2006; 80(1); 426-439; that show the % identity of the left ITR of AAV2 to the left ITR from other serotypes: AAV-1 (84%), AAV-3 (86%), AAV-4 (79%), AAV-5 (58%), AAV-6 (left ITR) (100%) and AAV-6 (right ITR) (82%).

A. Symmetrical ITR pairs

[00172] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein comprises, in the 5' to 3' direction: a first adeno-associated virus (AAV) inverted terminal repeat (ITR), a nucleotide sequence of interest (for example an expression cassette as described herein) and a second AAV ITR, where the first ITR (5' ITR) and the

second ITR (3' ITR) are symmetric, or substantially symmetrical with respect to each other – that is, a ceDNA vector can comprise ITR sequences that have a symmetrical three-dimensional spatial organization such that their structure is the same shape in geometrical space, or have the same A, C-C' and B-B' loops in 3D space. In such an embodiment, a symmetrical ITR pair, or substantially symmetrical ITR pair can be modified ITRs (*e.g.*, mod-ITRs) that are not wild-type ITRs. A mod-ITR pair can have the same sequence which has one or more modifications from wild-type ITR and are reverse complements (inverted) of each other. In alternative embodiments, a modified ITR pair are substantially symmetrical as defined herein, that is, the modified ITR pair can have a different sequence but have corresponding or the same symmetrical three-dimensional shape.

[00173] (i) *Wildtype ITRs*

[00174] In some embodiments, the symmetrical ITRs, or substantially symmetrical ITRs are wild type (WT-ITRs) as described herein. That is, both ITRs have a wild type sequence, but do not necessarily have to be WT-ITRs from the same AAV serotype. That is, in some embodiments, one WT-ITR can be from one AAV serotype, and the other WT-ITR can be from a different AAV serotype. In such an embodiment, a WT-ITR pair are substantially symmetrical as defined herein, that is, they can have one or more conservative nucleotide modification while still retaining the symmetrical three-dimensional spatial organization.

[00175] Accordingly, as disclosed herein, ceDNA vectors contain a transgene or heterologous nucleic acid sequence positioned between two flanking wild-type inverted terminal repeat (WT-ITR) sequences, that are either the reverse complement (inverted) of each other, or alternatively, are substantially symmetrical relative to each other – that is a WT-ITR pair have symmetrical three-dimensional spatial organization. In some embodiments, a wild-type ITR sequence (*e.g.* AAV WT-ITR) comprises a functional Rep binding site (RBS; *e.g.* 5'-GCGCGCTCGCTCGCTC-3' for AAV2, SEQ ID NO: 60) and a functional terminal resolution site (TRS; *e.g.* 5'-AGTT-3', SEQ ID NO: 62).

[00176] In one aspect, ceDNA vectors for expression of a desired transgene or therapeutic protein are obtainable from a vector polynucleotide that encodes a heterologous nucleic acid operatively positioned between two WT inverted terminal repeat sequences (WT-ITRs) (*e.g.* AAV WT-ITRs). That is, both ITRs have a wild type sequence, but do not necessarily have to be WT-ITRs from the same AAV serotype. That is, in some embodiments, one WT-ITR can be from one AAV serotype, and the other WT-ITR can be from a different AAV serotype. In such an embodiment, the WT-ITR pair are substantially symmetrical as defined herein, that is, they can have one or more conservative nucleotide modification while still retaining the symmetrical three-dimensional spatial organization. In some embodiments, the 5' WT-ITR is from one AAV serotype, and the 3' WT-ITR is from the same or a different AAV serotype. In some embodiments, the 5' WT-ITR and the 3' WT-ITR are mirror images of each other, that is they are symmetrical. In some embodiments, the 5' WT-ITR and the 3' WT-ITR are from the same AAV serotype.

[00177] WT ITRs are well known. In one embodiment the two ITRs are from the same AAV2 serotype. In certain embodiments one can use WT from other serotypes. There are a number of serotypes that are homologous, *e.g.* AAV2, AAV4, AAV6, AAV8. In one embodiment, closely homologous ITRs (*e.g.* ITRs with a similar loop structure) can be used. In another embodiment, one can use AAV WT ITRs that are more diverse, *e.g.*, AAV2 and AAV5, and still another embodiment, one can use an ITR that is substantially WT - that is, it has the basic loop structure of the WT but some conservative nucleotide changes that do not alter or affect the properties. When using WT-ITRs from the same viral serotype, one or more regulatory sequences may further be used. In certain embodiments, the regulatory sequence is a regulatory switch that permits modulation of the activity of the ceDNA, *e.g.*, the expression of the encoded desired transgene or therapeutic protein.

[00178] In some embodiments, one aspect of the technology described herein relates to a ceDNA vector for expression of a desired transgene or therapeutic protein, wherein the ceDNA vector comprises at least one heterologous nucleotide sequence encoding a desired transgene or therapeutic protein, operably positioned between two wild-type inverted terminal repeat sequences (WT-ITRs), wherein the WT-ITRs can be from the same serotype, different serotypes or substantially symmetrical with respect to each other (*i.e.*, have the symmetrical three-dimensional spatial organization such that their structure is the same shape in geometrical space, or have the same A, C-C' and B-B' loops in 3D space). In some embodiments, the symmetric WT-ITRs comprises a functional terminal resolution site and a Rep binding site. In some embodiments, the heterologous nucleic acid sequence encodes a transgene, and wherein the vector is not in a viral capsid.

[00179] In some embodiments, the WT-ITRs are the same but the reverse complement of each other. For example, the sequence AACG in the 5' ITR may be CGTT (*i.e.*, the reverse complement) in the 3' ITR at the corresponding site. In one example, the 5' WT-ITR sense strand comprises the sequence of ATCGATCG and the corresponding 3' WT-ITR sense strand comprises CGATCGAT (*i.e.*, the reverse complement of ATCGATCG). In some embodiments, the WT-ITRs ceDNA further comprises a terminal resolution site and a replication protein binding site (RPS) (sometimes referred to as a replicative protein binding site), *e.g.* a Rep binding site.

[00180] Exemplary WT-ITR sequences for use in the ceDNA vectors for expression of a desired transgene or therapeutic protein comprising WT-ITRs are shown in **Table 3** herein, which shows pairs of WT-ITRs (5' WT-ITR and the 3' WT-ITR).

[00181] As an exemplary example, the present disclosure provides a ceDNA vector for expression of a desired transgene or therapeutic protein comprising a promoter operably linked to a transgene (*e.g.*, heterologous nucleic acid sequence), with or without the regulatory switch, where the ceDNA is devoid of capsid proteins and is: (a) produced from a ceDNA-plasmid (*e.g.*, see **FIGS. 1F-1G**) that encodes WT-ITRs, where each WT-ITR has the same number of intramolecularly duplexed base pairs in its hairpin secondary configuration (preferably excluding deletion of any AAA or TTT terminal loop in this configuration compared to these reference sequences), and (b) is identified as ceDNA

using the assay for the identification of ceDNA by agarose gel electrophoresis under native gel and denaturing conditions in Example 1.

[00182] In some embodiments, the flanking WT-ITRs are substantially symmetrical to each other. In this embodiment the 5' WT-ITR can be from one serotype of AAV, and the 3' WT-ITR from a different serotype of AAV, such that the WT-ITRs are not identical reverse complements. For example, the 5' WT-ITR can be from AAV2, and the 3' WT-ITR from a different serotype (*e.g.* AAV1, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12. In some embodiments, WT-ITRs can be selected from two different parvoviruses selected from any to of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV13, snake parvovirus (*e.g.*, royal python parvovirus), bovine parvovirus, goat parvovirus, avian parvovirus, canine parvovirus, equine parvovirus, shrimp parvovirus, porcine parvovirus, or insect AAV. In some embodiments, such a combination of WT ITRs is the combination of WT-ITRs from AAV2 and AAV6. In one embodiment, the substantially symmetrical WT-ITRs are when one is inverted relative to the other ITR at least 90% identical, at least 95% identical, at least 96%...97%... 98%... 99%....99.5% and all points in between, and has the same symmetrical three-dimensional spatial organization. In some embodiments, a WT-ITR pair are substantially symmetrical as they have symmetrical three-dimensional spatial organization, *e.g.*, have the same 3D organization of the A, C-C'. B-B' and D arms. In one embodiment, a substantially symmetrical WT-ITR pair are inverted relative to the other, and are at least 95% identical, at least 96%...97%... 98%... 99%....99.5% and all points in between, to each other, and one WT-ITR retains the Rep-binding site (RBS) of 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60) and a terminal resolution site (trs). In some embodiments, a substantially symmetrical WT-ITR pair are inverted relative to each other, and are at least 95% identical, at least 96%...97%... 98%... 99%....99.5% and all points in between, to each other, and one WT-ITR retains the Rep-binding site (RBS) of 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60) and a terminal resolution site (trs) and in addition to a variable palindromic sequence allowing for hairpin secondary structure formation. Homology can be determined by standard means well known in the art such as BLAST (Basic Local Alignment Search Tool), BLASTN at default setting.

[00183] In some embodiments, the structural element of the ITR can be any structural element that is involved in the functional interaction of the ITR with a large Rep protein (*e.g.*, Rep 78 or Rep 68). In certain embodiments, the structural element provides selectivity to the interaction of an ITR with a large Rep protein, *i.e.*, determines at least in part which Rep protein functionally interacts with the ITR. In other embodiments, the structural element physically interacts with a large Rep protein when the Rep protein is bound to the ITR. Each structural element can be, *e.g.*, a secondary structure of the ITR, a nucleotide sequence of the ITR, a spacing between two or more elements, or a combination of any of the above. In one embodiment, the structural elements are selected from the group consisting of an A and an A' arm, a B and a B' arm, a C and a C' arm, a D arm, a Rep binding site (RBE) and an RBE' (*i.e.*, complementary RBE sequence), and a terminal resolution site (trs).

[00184] By way of example only, **Table 2** indicates exemplary combinations of WT-ITRs.

[00185] **Table 2:** Exemplary combinations of WT-ITRs from the same serotype or different serotypes, or different parvoviruses. The order shown is not indicative of the ITR position, for example, “AAV1, AAV2” demonstrates that the ceDNA can comprise a WT-AAV1 ITR in the 5’ position, and a WT-AAV2 ITR in the 3’ position, or vice versa, a WT-AAV2 ITR the 5’ position, and a WT-AAV1 ITR in the 3’ position. Abbreviations: AAV serotype 1 (AAV1), AAV serotype 2 (AAV2), AAV serotype 3 (AAV3), AAV serotype 4 (AAV4), AAV serotype 5 (AAV5), AAV serotype 6 (AAV6), AAV serotype 7 (AAV7), AAV serotype 8 (AAV8), AAV serotype 9 (AAV9), AAV serotype 10 (AAV10), AAV serotype 11 (AAV11), or AAV serotype 12 (AAV12); AAVrh8, AAVrh10, AAV-DJ, and AAV-DJ8 genome (*E.g.*, NCBI: NC 002077; NC 001401; NC001729; NC001829; NC006152; NC 006260; NC 006261), ITRs from warm-blooded animals (avian AAV (AAV), bovine AAV (BAAV), canine, equine, and ovine AAV), ITRs from B19 parvovirus (GenBank Accession No: NC 000883), Minute Virus from Mouse (MVM) (GenBank Accession No. NC 001510); Goose: goose parvovirus (GenBank Accession No. NC 001701); snake: snake parvovirus 1 (GenBank Accession No. NC 006148).

Table 2: Exemplary combinations of WT ITRs

AAV1,AAV1	AAV2,AAV2	AAV3,AAV3	AAV4,AAV4	AAV5,AAV5
AAV1,AAV2	AAV2,AAV3	AAV3,AAV4	AAV4,AAV5	AAV5,AAV6
AAV1,AAV3	AAV2,AAV4	AAV3,AAV5	AAV4,AAV6	AAV5,AAV7
AAV1,AAV4	AAV2,AAV5	AAV3,AAV6	AAV4,AAV7	AAV5,AAV8
AAV1,AAV5	AAV2,AAV6	AAV3,AAV7	AAV4,AAV8	AAV5,AAV9
AAV1,AAV6	AAV2,AAV7	AAV3,AAV8	AAV4,AAV9	AAV5,AAV10
AAV1,AAV7	AAV2,AAV8	AAV3,AAV9	AAV4,AAV10	AAV5,AAV11
AAV1,AAV8	AAV2,AAV9	AAV3,AAV10	AAV4,AAV11	AAV5,AAV12
AAV1,AAV9	AAV2,AAV10	AAV3,AAV11	AAV4,AAV12	AAV5,AAVRH8
AAV1,AAV10	AAV2,AAV11	AAV3,AAV12	AAV4,AAVRH8	AAV5,AAVRH10
AAV1,AAV11	AAV2,AAV12	AAV3,AAVRH8	AAV4,AAVRH10	AAV5,AAV13
AAV1,AAV12	AAV2,AAVRH8	AAV3,AAVRH10	AAV4,AAV13	AAV5,AAVDJ
AAV1,AAVRH8	AAV2,AAVRH10	AAV3,AAV13	AAV4,AAVDJ	AAV5,AAVDJ8
AAV1,AAVRH10	AAV2,AAV13	AAV3,AAVDJ	AAV4,AAVDJ8	AAV5,AVIAN
AAV1,AAV13	AAV2,AAVDJ	AAV3,AAVDJ8	AAV4,AVIAN	AAV5,BOVINE
AAV1,AAVDJ	AAV2,AAVDJ8	AAV3,AVIAN	AAV4,BOVINE	AAV5,CANINE
AAV1,AAVDJ8	AAV2,AVIAN	AAV3,BOVINE	AAV4,CANINE	AAV5,EQUINE
AAV1,AVIAN	AAV2,BOVINE	AAV3,CANINE	AAV4,EQUINE	AAV5,GOAT
AAV1,BOVINE	AAV2,CANINE	AAV3,EQUINE	AAV4,GOAT	AAV5,SHRIMP
AAV1,CANINE	AAV2,EQUINE	AAV3,GOAT	AAV4,SHRIMP	AAV5,PORCINE
AAV1,EQUINE	AAV2,GOAT	AAV3,SHRIMP	AAV4,PORCINE	AAV5,INSECT
AAV1,GOAT	AAV2,SHRIMP	AAV3,PORCINE	AAV4,INSECT	AAV5,OVINE

AAV1,SHRIMP	AAV2,PORCINE	AAV3,INSECT	AAV4,OVINE	AAV5,B19
AAV1,PORCINE	AAV2,INSECT	AAV3,OVINE	AAV4,B19	AAV5,MVM
AAV1,INSECT	AAV2,OVINE	AAV3,B19	AAV4,MVM	AAV5,GOOSE
AAV1,OVINE	AAV2,B19	AAV3,MVM	AAV4,GOOSE	AAV5,SNAKE
AAV1,B19	AAV2,MVM	AAV3,GOOSE	AAV4,SNAKE	
AAV1,MVM	AAV2,GOOSE	AAV3,SNAKE		
AAV1,GOOSE	AAV2,SNAKE			
AAV1,SNAKE				
AAV6,AAV6	AAV7,AAV7	AAV8,AAV8	AAV9,AAV9	AAV10,AAV10
AAV6,AAV7	AAV7,AAV8	AAV8,AAV9	AAV9,AAV10	AAV10,AAV11
AAV6,AAV8	AAV7,AAV9	AAV8,AAV10	AAV9,AAV11	AAV10,AAV12
AAV6,AAV9	AAV7,AAV10	AAV8,AAV11	AAV9,AAV12	AAV10,AAVRH8
AAV6,AAV10	AAV7,AAV11	AAV8,AAV12	AAV9,AAVRH8	AAV10,AAVRH10
AAV6,AAV11	AAV7,AAV12	AAV8,AAVRH8	AAV9,AAVRH10	AAV10,AAV13
AAV6,AAV12	AAV7,AAVRH8	AAV8,AAVRH10	AAV9,AAV13	AAV10,AAVDJ
AAV6,AAVRH8	AAV7,AAVRH10	AAV8,AAV13	AAV9,AAVDJ	AAV10,AAVDJ8
AAV6,AAVRH10	AAV7,AAV13	AAV8,AAVDJ	AAV9,AAVDJ8	AAV10,AVIAN
AAV6,AAV13	AAV7,AAVDJ	AAV8,AAVDJ8	AAV9,AVIAN	AAV10,BOVINE
AAV6,AAVDJ	AAV7,AAVDJ8	AAV8,AVIAN	AAV9,BOVINE	AAV10,CANINE
AAV6,AAVDJ8	AAV7,AVIAN	AAV8,BOVINE	AAV9,CANINE	AAV10,EQUINE
AAV6,AVIAN	AAV7,BOVINE	AAV8,CANINE	AAV9,EQUINE	AAV10,GOAT
AAV6,BOVINE	AAV7,CANINE	AAV8,EQUINE	AAV9,GOAT	AAV10,SHRIMP
AAV6,CANINE	AAV7,EQUINE	AAV8,GOAT	AAV9,SHRIMP	AAV10,PORCINE
AAV6,EQUINE	AAV7,GOAT	AAV8,SHRIMP	AAV9,PORCINE	AAV10,INSECT
AAV6,GOAT	AAV7,SHRIMP	AAV8,PORCINE	AAV9,INSECT	AAV10,OVINE
AAV6,SHRIMP	AAV7,PORCINE	AAV8,INSECT	AAV9,OVINE	AAV10,B19
AAV6,PORCINE	AAV7,INSECT	AAV8,OVINE	AAV9,B19	AAV10,MVM
AAV6,INSECT	AAV7,OVINE	AAV8,B19	AAV9,MVM	AAV10,GOOSE
AAV6,OVINE	AAV7,B19	AAV8,MVM	AAV9,GOOSE	AAV10,SNAKE
AAV6,B19	AAV7,MVM	AAV8,GOOSE	AAV9,SNAKE	
AAV6,MVM	AAV7,GOOSE	AAV8,SNAKE		
AAV6,GOOSE	AAV7,SNAKE			
AAV6,SNAKE				
AAV11,AAV11	AAV12,AAV12	AAVRH8,AAVRH8	AAVRH10,AAVRH10	AAV13,AAV13
AAV11,AAV12	AAV12,AAVRH8	AAVRH8,AAVRH10	AAVRH10,AAV13	AAV13,AAVDJ
AAV11,AAVRH8	AAV12,AAVRH10	AAVRH8,AAV13	AAVRH10,AAVDJ	AAV13,AAVDJ8
AAV11,AAVRH10	AAV12,AAV13	AAVRH8,AAVDJ	AAVRH10,AAVDJ8	AAV13,AVIAN

AAV11,AAV13	AAV12,AAVDJ	AAVRH8,AAVDJ8	AAVRH10,AVIAN	AAV13,BOVINE
AAV11,AAVDJ	AAV12,AAVDJ8	AAVRH8,AVIAN	AAVRH10,BOVINE	AAV13,CANINE
AAV11,AAVDJ8	AAV12,AVIAN	AAVRH8,BOVINE	AAVRH10,CANINE	AAV13,EQUINE
AAV11,AVIAN	AAV12,BOVINE	AAVRH8,CANINE	AAVRH10,EQUINE	AAV13,GOAT
AAV11,BOVINE	AAV12,CANINE	AAVRH8,EQUINE	AAVRH10,GOAT	AAV13,SHRIMP
AAV11,CANINE	AAV12,EQUINE	AAVRH8,GOAT	AAVRH10,SHRIMP	AAV13,PORCINE
AAV11,EQUINE	AAV12,GOAT	AAVRH8,SHRIMP	AAVRH10,PORCINE	AAV13,INSECT
AAV11,GOAT	AAV12,SHRIMP	AAVRH8,PORCINE	AAVRH10,INSECT	AAV13,OVINE
AAV11,SHRIMP	AAV12,PORCINE	AAVRH8,INSECT	AAVRH10,OVINE	AAV13,B19
AAV11,PORCINE	AAV12,INSECT	AAVRH8,OVINE	AAVRH10,B19	AAV13,MVM
AAV11,INSECT	AAV12,OVINE	AAVRH8,B19	AAVRH10,MVM	AAV13,GOOSE
AAV11,OVINE	AAV12,B19	AAVRH8,MVM	AAVRH10,GOOSE	AAV13,SNAKE
AAV11,B19	AAV12,MVM	AAVRH8,GOOSE	AAVRH10,SNAKE	
AAV11,MVM	AAV12,GOOSE	AAVRH8,SNAKE		
AAV11,GOOSE	AAV12,SNAKE			
AAV11,SNAKE				
AAVDJ,AAVDJ	AAVDJ8,AAVDJ8	AVIAN, AVIAN	BOVINE, BOVINE	CANINE, CANINE
AAVDJ,AAVDJ8	AAVDJ8,AVIAN	AVIAN,BOVINE	BOVINE,CANINE	CANINE,EQUINE
AAVDJ,AVIAN	AAVDJ8,BOVINE	AVIAN,CANINE	BOVINE,EQUINE	CANINE,GOAT
AAVDJ,BOVINE	AAVDJ8,CANINE	AVIAN,EQUINE	BOVINE,GOAT	CANINE,SHRIMP
AAVDJ,CANINE	AAVDJ8,EQUINE	AVIAN,GOAT	BOVINE,SHRIMP	CANINE,PORCIN E
AAVDJ,EQUINE	AAVDJ8,GOAT	AVIAN,SHRIMP	BOVINE,PORCINE	CANINE,INSECT
AAVDJ,GOAT	AAVDJ8,SHRIMP	AVIAN,PORCINE	BOVINE,INSECT	CANINE,OVINE
AAVDJ,SHRIMP	AAVDJ8,PORCINE	AVIAN,INSECT	BOVINE,OVINE	CANINE,B19
AAVDJ,PORCINE	AAVDJ8,INSECT	AVIAN,OVINE	BOVINE,B19	CANINE,MVM
AAVDJ,INSECT	AAVDJ8,OVINE	AVIAN,B19	BOVINE,MVM	CANINE,GOOSE
AAVDJ,OVINE	AAVDJ8,B19	AVIAN,MVM	BOVINE,GOOSE	CANINE,SNAKE
AAVDJ,B19	AAVDJ8,MVM	AVIAN,GOOSE	BOVINE,SNAKE	
AAVDJ,MVM	AAVDJ8,GOOSE	AVIAN,SNAKE		
AAVDJ,GOOSE	AAVDJ8,SNAKE			
AAVDJ,SNAKE				
EQUINE, EQUINE	GOAT, GOAT	SHRIMP, SHRIMP	PORCINE, PORCINE	INSECT, INSECT
EQUINE,GOAT	GOAT,SHRIMP	SHRIMP,PORCINE	PORCINE,INSECT	INSECT,OVINE
EQUINE,SHRIMP	GOAT,PORCINE	SHRIMP,INSECT	PORCINE,OVINE	INSECT,B19
EQUINE,PORCINE	GOAT,INSECT	SHRIMP,OVINE	PORCINE,B19	INSECT,MVM
EQUINE,INSECT	GOAT,OVINE	SHRIMP,B19	PORCINE,MVM	INSECT,GOOSE

EQUINE,OVINE	GOAT,B19	SHRIMP,MVM	PORCINE,GOOSE	INSECT,SNAKE
EQUINE,B19	GOAT,MVM	SHRIMP,GOOSE	PORCINE,SNAKE	
EQUINE,MVM	GOAT,GOOSE	SHRIMP,SNAKE		
EQUINE,GOOSE	GOAT,SNAKE			
EQUINE,SNAKE				
OVINE, OVINE	B19, B19	MVM, MVM	GOOSE, GOOSE	SNAKE, SNAKE
OVINE,B19	B19,MVM	MVM,GOOSE	GOOSE,SNAKE	
OVINE,MVM	B19,GOOSE	MVM,SNAKE		
OVINE,GOOSE	B19,SNAKE			
OVINE,SNAKE				

[00186] By way of example only, **Table 3** shows the sequences of exemplary WT-ITRs from some different AAV serotypes.

Table 3: WT-ITR sequences

AAV serotype	5' WT-ITR (LEFT)	3' WT-ITR (RIGHT)
AAV1	5'- TTGCCCCACTCCCTCTCTGCGCGCTCGC TCGCTCGGTGGGGCCTGCGGACCAAA GGTCCGCAGACGGCAGAGGTCTCCTC TGCCGGCCCCACCGAGCGAGCGACGC GCGCAGAGAGGGAGTGGGCAACTCCA TCACTAGGGTAA-3' (SEQ ID NO: 5)	5'- TTACCCTAGTGATGGAGTTGCCCCACTC CCTCTCTGCGCGCTCGCTCGCTCGGT GGGGCCGGCAGAGGAGACCTCTGCCG TCTGCGGACCTTTGGTCCGCAGGCCCC ACCGAGCGAGCGAGCGCGCAGAGAGG GAGTGGGCAA-3' (SEQ ID NO: 10)
AAV2	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCCGCGCGGCAAGGCC CGGGCGTCGGGCGACCTTTGGTTCGCC CGGCCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTGGCCAACCTCCATCAC TAGGGGTTTCCT (SEQ ID NO: 2)	AGGAACCCCTAGTGATGGAGTTGGCCA CTCCCTCTCTGCGCGCTCGCTCGCTCAC TGAGGCCGGGCGACCAAGGTTCGCC GACGCCCGGGCTTTGCGGGCGGCCT CAGTGAGCGAGCGAGCGCGCAGCTGC CTGCAGG (SEQ ID NO: 1)
AAV3	5'- TTGGCCACTCCCTCTATGCGCACTCGC TCGCTCGGTGGGGCCTGGCGACCAAA GGTCGCCAGACGGACGTGGGTTTCCA CGTCCGGCCCCACCGAGCGAGCGAGT GCGCATAGAGGGAGTGGCCAACCTCCA TCACTAGAGGTAT-3' (SEQ ID NO: 6)	5'- ATACCTCTAGTGATGGAGTTGGCCACT CCCTCTATGCGCACTCGCTCGCTCGGT GGGGCCGGACGTGGAAACCCACGTCC GTCTGGCGACCTTTGGTTCGCCAGGCCC CACCGAGCGAGCGAGTGC GCATAGAG GGAGTGGCCAA-3' (SEQ ID NO: 11)
AAV4	5'- TTGGCCACTCCCTCTATGCGCGCTCGC TCACTCACTCGGCCCTGGAGACCAAA GGTCTCCAGACTGCCGGCCTCTGGCC GGCAGGGCCGAGTGAGTGAGCGAGC GCGCATAGAGGGAGTGGCCAACCT-3'	5'- AGTTGGCCACATTAGCTATGCGCGCTC GCTCACTCACTCGGCCCTGGAGACCAA AGGTCTCCAGACTGCCGGCCTCTGGCC GGCAGGGCCGAGTGAGTGAGCGAGCG GCGCATAGAGGGAGTGGCCAA-3' (SEQ ID

	(SEQ ID NO: 7)	NO: 12)
AAV5	5'- TCCCCCTGTCGCGTTCGCTCGCTCGC TGGCTCGTTTGGGGGGGCGACGGCCA GAGGGCCGTCGTCTGGCAGCTCTTTG AGCTGCCACCCCCCAAACGAGCCAG CGAGCGAGCGAACGCGACAGGGGGG AGAGTGCCACACTCTCAAGCAAGGGG GTTTTGTAAG -3' (SEQ ID NO: 8)	5'- CTTACAAAACCCCCTTGCTTGAGAGTG TGGCACTCTCCCCCTGTCGCGTTCGCT CGCTCGCTGGCTCGTTTGGGGGGGTGG CAGCTCAAAGAGCTGCCAGACGACGG CCCTCTGGCCGTCGCCCCCAAACGA GCCAGCGAGCGAGCGAACGCGACAGG GGGGA-3' (SEQ ID NO: 13)
AAV6	5'- TTGCCCACTCCCTCTAATGCGCGCTCG CTCGCTCGGTGGGGCCTGCGGACCAA AGGTCCGCAGACGGCAGAGGTCTCCT CTGCCGGCCCCACCGAGCGAGCGAGC GCGCATAGAGGGAGTGGGCAACTCCA TCACTAGGGGTAT-3' (SEQ ID NO: 9)	5'- ATACCCCTAGTGATGGAGTTGCCCACT CCCTCTATGCGCGCTCGCTCGCTCGGT GGGGCCGGCAGAGGAGACCTCTGCCG TCTGCGGACCTTTGGTCCGCAGGCCCC ACCGAGCGAGCGAGCGCGCATTAGAG GGAGTGGGCAA (SEQ ID NO: 14)

[00187] In some embodiments, the nucleotide sequence of the WT-ITR sequence can be modified (*e.g.*, by modifying 1, 2, 3, 4 or 5, or more nucleotides or any range therein), whereby the modification is a substitution for a complementary nucleotide, *e.g.*, G for a C, and vice versa, and T for an A, and vice versa.

[00188] In certain embodiments of the present invention, the ceDNA vector for expression of a desired transgene or therapeutic protein does not have a WT-ITR consisting of the nucleotide sequence selected from any of: SEQ ID NOs: 1, 2, 5-14. In alternative embodiments of the present invention, if a ceDNA vector has a WT-ITR comprising the nucleotide sequence selected from any of: SEQ ID NOs: 1, 2, 5-14, then the flanking ITR is also WT and the ceDNA vector comprises a regulatory switch, *e.g.*, as disclosed herein and in International application PCT/US18/49996 (*e.g.*, see Table 11 of PCT/US18/49996). In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein comprises a regulatory switch as disclosed herein and a WT-ITR selected having the nucleotide sequence selected from any of the group consisting of: SEQ ID NO: 1, 2, 5-14.

[00189] The ceDNA vector for expression of a desired transgene or therapeutic protein as described herein can include WT-ITR structures that retains an operable RBE, trs and RBE' portion. **FIG. 2A** and **FIG. 2B**, using wild-type ITRs for exemplary purposes, show one possible mechanism for the operation of a trs site within a wild type ITR structure portion of a ceDNA vector. In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein contains one or more functional WT-ITR polynucleotide sequences that comprise a Rep-binding site (RBS; 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60) for AAV2) and a terminal resolution site (TRS; 5'-AGTT (SEQ ID NO: 62)). In some embodiments, at least one WT-ITR is functional. In alternative embodiments, where a ceDNA vector for expression of a desired transgene or therapeutic protein

comprises two WT-ITRs that are substantially symmetrical to each other, at least one WT-ITR is functional and at least one WT-ITR is non-functional.

B. Modified ITRs (mod-ITRs) in general for ceDNA vectors comprising asymmetric ITR pairs or symmetric ITR pairs

[00190] As discussed herein, a ceDNA vector for expression of a desired transgene or therapeutic protein can comprise a symmetrical ITR pair or an asymmetrical ITR pair. In both instances, one or both of the ITRs can be modified ITRs – the difference being that in the first instance (*i.e.*, symmetric mod-ITRs), the mod-ITRs have the same three-dimensional spatial organization (*i.e.*, have the same A-A', C-C' and B-B' arm configurations), whereas in the second instance (*i.e.*, asymmetric mod-ITRs), the mod-ITRs have a different three-dimensional spatial organization (*i.e.*, have a different configuration of A-A', C-C' and B-B' arms).

[00191] In some embodiments, a modified ITR is an ITRs that is modified by deletion, insertion, and/or substitution as compared to a wild-type ITR sequence (*e.g.* AAV ITR). In some embodiments, at least one of the ITRs in the ceDNA vector comprises a functional Rep binding site (RBS; *e.g.* 5'-GCGCGCTCGCTCGCTC-3' for AAV2, SEQ ID NO: 60) and a functional terminal resolution site (TRS; *e.g.* 5'-AGTT-3', SEQ ID NO: 62.) In one embodiment, at least one of the ITRs is a non-functional ITR. In one embodiment, the different or modified ITRs are not each wild type ITRs from different serotypes.

[00192] Specific alterations and mutations in the ITRs are described in detail herein, but in the context of ITRs, “altered” or “mutated” or “modified”, it indicates that nucleotides have been inserted, deleted, and/or substituted relative to the wild-type, reference, or original ITR sequence. The altered or mutated ITR can be an engineered ITR. As used herein, “engineered” refers to the aspect of having been manipulated by the hand of man. For example, a polypeptide is considered to be “engineered” when at least one aspect of the polypeptide, *e.g.*, its sequence, has been manipulated by the hand of man to differ from the aspect as it exists in nature.

[00193] In some embodiments, a mod-ITR may be synthetic. In one embodiment, a synthetic ITR is based on ITR sequences from more than one AAV serotype. In another embodiment, a synthetic ITR includes no AAV-based sequence. In yet another embodiment, a synthetic ITR preserves the ITR structure described above although having only some or no AAV-sourced sequence. In some aspects, a synthetic ITR may interact preferentially with a wild type Rep or a Rep of a specific serotype, or in some instances will not be recognized by a wild-type Rep and be recognized only by a mutated Rep.

[00194] The skilled artisan can determine the corresponding sequence in other serotypes by known means. For example, determining if the change is in the A, A', B, B', C, C' or D region and determine the corresponding region in another serotype. One can use BLAST® (Basic Local Alignment Search Tool) or other homology alignment programs at default status to determine the corresponding sequence. The invention further provides populations and pluralities of ceDNA vectors comprising mod-ITRs from a combination of different AAV serotypes – that is, one mod-ITR can be from one

AAV serotype and the other mod-ITR can be from a different serotype. Without wishing to be bound by theory, in one embodiment one ITR can be from or based on an AAV2 ITR sequence and the other ITR of the ceDNA vector can be from or be based on any one or more ITR sequence of AAV serotype 1 (AAV1), AAV serotype 4 (AAV4), AAV serotype 5 (AAV5), AAV serotype 6 (AAV6), AAV serotype 7 (AAV7), AAV serotype 8 (AAV8), AAV serotype 9 (AAV9), AAV serotype 10 (AAV10), AAV serotype 11 (AAV11), or AAV serotype 12 (AAV12).

[00195] Any parvovirus ITR can be used as an ITR or as a base ITR for modification. Preferably, the parvovirus is a dependovirus. More preferably AAV. The serotype chosen can be based upon the tissue tropism of the serotype. AAV2 has a broad tissue tropism, AAV1 preferentially targets to neuronal and skeletal muscle, and AAV5 preferentially targets neuronal, retinal pigmented epithelia, and photoreceptors. AAV6 preferentially targets skeletal muscle and lung. AAV8 preferentially targets liver, skeletal muscle, heart, and pancreatic tissues. AAV9 preferentially targets liver, skeletal and lung tissue. In one embodiment, the modified ITR is based on an AAV2 ITR.

[00196] More specifically, the ability of a structural element to functionally interact with a particular large Rep protein can be altered by modifying the structural element. For example, the nucleotide sequence of the structural element can be modified as compared to the wild-type sequence of the ITR. In one embodiment, the structural element (*e.g.*, A arm, A' arm, B arm, B' arm, C arm, C' arm, D arm, RBE, RBE', and trs) of an ITR can be removed and replaced with a wild-type structural element from a different parvovirus. For example, the replacement structure can be from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV13, snake parvovirus (*e.g.*, royal python parvovirus), bovine parvovirus, goat parvovirus, avian parvovirus, canine parvovirus, equine parvovirus, shrimp parvovirus, porcine parvovirus, or insect AAV. For example, the ITR can be an AAV2 ITR and the A or A' arm or RBE can be replaced with a structural element from AAV5. In another example, the ITR can be an AAV5 ITR and the C or C' arms, the RBE, and the trs can be replaced with a structural element from AAV2. In another example, the AAV ITR can be an AAV5 ITR with the B and B' arms replaced with the AAV2 ITR B and B' arms.

[00197] By way of example only, **Table 4** indicates exemplary modifications of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in regions of a modified ITR, where X is indicative of a modification of at least one nucleic acid (*e.g.*, a deletion, insertion and/ or substitution) in that section relative to the corresponding wild-type ITR. In some embodiments, any modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in any of the regions of C and/or C' and/or B and/or B' retains three sequential T nucleotides (*i.e.*, TTT) in at least one terminal loop. For example, if the modification results in any of: a single arm ITR (*e.g.*, single C-C' arm, or a single B-B' arm), or a modified C-B' arm or C'-B arm, or a two arm ITR with at least one truncated arm (*e.g.*, a truncated C-C' arm and/or truncated B-B' arm), at least the single arm, or at least one of the arms of a two arm ITR (where one arm can be truncated) retains three sequential T nucleotides

(*i.e.*, TTT) in at least one terminal loop. In some embodiments, a truncated C-C' arm and/or a truncated B-B' arm has three sequential T nucleotides (*i.e.*, TTT) in the terminal loop.

[00198] **Table 4:** Exemplary combinations of modifications of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) to different B-B' and C-C' regions or arms of ITRs (X indicates a nucleotide modification, *e.g.*, addition, deletion or substitution of at least one nucleotide in the region).

Table 4

B region	B' region	C region	C' region
X			
	X		
X	X		
		X	
			X
		X	X
X		X	
X			X
	X	X	
	X		X
X	X	X	
X	X		X
X		X	X
	X	X	X
X	X	X	X

[00199] In some embodiments, mod-ITR for use in a ceDNA vector for expression of a desired transgene or therapeutic protein comprises an asymmetric ITR pair, or a symmetric mod-ITR pair as disclosed herein, can comprise any one of the combinations of modifications shown in **Table 4**, and also a modification of at least one nucleotide in any one or more of the regions selected from: between A' and C, between C and C', between C' and B, between B and B' and between B' and A. In some embodiments, any modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in the C or C' or B or B' regions, still preserves the terminal loop of the stem-loop. In some embodiments, any modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) between C and C' and/or B and B' retains three sequential T nucleotides (*i.e.*, TTT) in at least one terminal loop. In alternative embodiments, any modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) between C and C' and/or B and B' retains three sequential A nucleotides (*i.e.*, AAA) in at least one terminal loop. In some embodiments, a modified ITR for use herein can comprise any one of the combinations of modifications shown in **Table 4**, and also a modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in any one or more of the regions selected from: A', A and/or D. For example, in some embodiments, a modified ITR for use herein can comprise any one of the combinations of modifications shown in **Table 4**, and also a modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in the A region. In some embodiments, a modified ITR for use herein can comprise any one of the

combinations of modifications shown in **Table 4**, and also a modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in the A' region. In some embodiments, a modified ITR for use herein can comprise any one of the combinations of modifications shown in **Table 4**, and also a modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in the A and/or A' region. In some embodiments, a modified ITR for use herein can comprise any one of the combinations of modifications shown in **Table 4**, and also a modification of at least one nucleotide (*e.g.*, a deletion, insertion and/ or substitution) in the D region.

[00200] In one embodiment, the nucleotide sequence of the structural element can be modified (*e.g.*, by modifying 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or more nucleotides or any range therein) to produce a modified structural element. In one embodiment, the specific modifications to the ITRs are exemplified herein (*e.g.*, SEQ ID NOS: 3, 4, 15-47, 101-116 or 165-187, or shown in FIG. 7A-7B of PCT/US2018/064242, filed on December 6, 2018 (*e.g.*, SEQ ID Nos 97-98, 101-103, 105-108, 111-112, 117-134, 545-54 in PCT/US2018/064242, the contents of which are incorporated by reference in their entireties herein). In some embodiments, an ITR can be modified (*e.g.*, by modifying 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or more nucleotides or any range therein). In other embodiments, the ITR can have at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more sequence identity with one of the modified ITRs of SEQ ID NOS: 3, 4, 15-47, 101-116 or 165-187, or the RBE-containing section of the A-A' arm and C-C' and B-B' arms of SEQ ID NO: 3, 4, 15-47, 101-116 or 165-187, or shown in Tables 2-9 (*i.e.*, SEQ ID NO: 110-112, 115-190, 200-468) of International application PCT/US18/49996, which is incorporated herein in its entirety by reference.

[00201] In some embodiments, a modified ITR can for example, comprise removal or deletion of all of a particular arm, *e.g.*, all or part of the A-A' arm, or all or part of the B-B' arm or all or part of the C-C' arm, or alternatively, the removal of 1, 2, 3, 4, 5, 6, 7, 8, 9 or more base pairs forming the stem of the loop so long as the final loop capping the stem (*e.g.*, single arm) is still present (*e.g.*, see ITR-21 in **FIG. 7A** of PCT/US2018/064242, filed December 6, 2018, incorporated by reference in its entirety herein). In some embodiments, a modified ITR can comprise the removal of 1, 2, 3, 4, 5, 6, 7, 8, 9 or more base pairs from the B-B' arm. In some embodiments, a modified ITR can comprise the removal of 1, 2, 3, 4, 5, 6, 7, 8, 9 or more base pairs from the C-C' arm (see, *e.g.*, ITR-1 in **FIG. 3B**, or ITR-45 in **FIG. 7A** of PCT/US2018/064242, filed December 6, 2018). In some embodiments, a modified ITR can comprise the removal of 1, 2, 3, 4, 5, 6, 7, 8, 9 or more base pairs from the C-C' arm and the removal of 1, 2, 3, 4, 5, 6, 7, 8, 9 or more base pairs from the B-B' arm. Any combination of removal of base pairs is envisioned, for example, 6 base pairs can be removed in the C-C' arm and 2 base pairs in the B-B' arm. As an illustrative example, **FIG. 3B** shows an exemplary modified ITR with at least 7 base pairs deleted from each of the C portion and the C' portion, a substitution of a nucleotide in the loop between C and C' region, and at least one base pair deletion from each of the B

region and B' regions such that the modified ITR comprises two arms where at least one arm (*e.g.*, C-C') is truncated. In some embodiments, the modified ITR also comprises at least one base pair deletion from each of the B region and B' regions, such that the B-B' arm is also truncated relative to WT ITR.

[00202] In some embodiments, a modified ITR can have between 1 and 50 (*e.g.* 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50) nucleotide deletions relative to a full-length wild-type ITR sequence. In some embodiments, a modified ITR can have between 1 and 30 nucleotide deletions relative to a full-length WT ITR sequence. In some embodiments, a modified ITR has between 2 and 20 nucleotide deletions relative to a full-length wild-type ITR sequence.

[00203] In some embodiments, a modified ITR does not contain any nucleotide deletions in the RBE-containing portion of the A or A' regions, so as not to interfere with DNA replication (*e.g.* binding to an RBE by Rep protein, or nicking at a terminal resolution site). In some embodiments, a modified ITR encompassed for use herein has one or more deletions in the B, B', C, and/or C region as described herein.

[00204] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein comprising a symmetric ITR pair or asymmetric ITR pair comprises a regulatory switch as disclosed herein and at least one modified ITR selected having the nucleotide sequence selected from any of the group consisting of: SEQ ID NO: 3, 4, 15-47, 101-116 or 165-187.

[00205] In another embodiment, the structure of the structural element can be modified. For example, the structural element a change in the height of the stem and/or the number of nucleotides in the loop. For example, the height of the stem can be about 2, 3, 4, 5, 6, 7, 8, or 9 nucleotides or more or any range therein. In one embodiment, the stem height can be about 5 nucleotides to about 9 nucleotides and functionally interacts with Rep. In another embodiment, the stem height can be about 7 nucleotides and functionally interacts with Rep. In another example, the loop can have 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides or more or any range therein.

[00206] In another embodiment, the number of GAGY binding sites or GAGY-related binding sites within the RBE or extended RBE can be increased or decreased. In one example, the RBE or extended RBE, can comprise 1, 2, 3, 4, 5, or 6 or more GAGY binding sites or any range therein. Each GAGY binding site can independently be an exact GAGY sequence or a sequence similar to GAGY as long as the sequence is sufficient to bind a Rep protein.

[00207] In another embodiment, the spacing between two elements (such as but not limited to the RBE and a hairpin) can be altered (*e.g.*, increased or decreased) to alter functional interaction with a large Rep protein. For example, the spacing can be about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 21 nucleotides or more or any range therein.

[00208] The ceDNA vector for expression of a desired transgene or therapeutic protein as described herein can include an ITR structure that is modified with respect to the wild type AAV2

ITR structure disclosed herein, but still retains an operable RBE, trs and RBE' portion. **FIG. 2A** and **FIG. 2B** show one possible mechanism for the operation of a trs site within a wild type ITR structure portion of a ceDNA vector for expression of a desired transgene or therapeutic protein. In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein contains one or more functional ITR polynucleotide sequences that comprise a Rep-binding site (RBS; 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60) for AAV2) and a terminal resolution site (TRS; 5'-AGTT (SEQ ID NO: 62)). In some embodiments, at least one ITR (wt or modified ITR) is functional. In alternative embodiments, where a ceDNA vector for expression of a desired transgene or therapeutic protein comprises two modified ITRs that are different or asymmetrical to each other, at least one modified ITR is functional and at least one modified ITR is non-functional.

[00209] In some embodiments, the modified ITR (*e.g.*, the left or right ITR) of a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein has modifications within the loop arm, the truncated arm, or the spacer. Exemplary sequences of ITRs having modifications within the loop arm, the truncated arm, or the spacer are listed in Table 2 (*i.e.*, SEQ ID NOS: 135-190, 200-233); Table 3 (*e.g.*, SEQ ID Nos: 234-263); Table 4 (*e.g.*, SEQ ID NOS: 264-293); Table 5 (*e.g.*, SEQ ID Nos: 294-318 herein); Table 6 (*e.g.*, SEQ ID NO: 319-468; and Tables 7-9 (*e.g.*, SEQ ID Nos: 101-110, 111-112, 115-134) or Table 10A or 10B (*e.g.*, SEQ ID Nos: 9, 100, 469-483, 484-499) of International application PCT/US18/49996, which is incorporated herein in its entirety by reference.

[00210] In some embodiments, the modified ITR for use in a ceDNA vector for expression of a desired transgene or therapeutic protein comprising an asymmetric ITR pair, or symmetric mod-ITR pair is selected from any or a combination of those shown in Tables 2, 3, 4, 5, 6, 7, 8, 9 and 10A-10B of International application PCT/US18/49996 which is incorporated herein in its entirety by reference.

[00211] Additional exemplary modified ITRs for use in a ceDNA vector for expression of a desired transgene or therapeutic protein comprising an asymmetric ITR pair, or symmetric mod-ITR pair in each of the above classes are provided in **Tables 5A** and **5B**. The predicted secondary structure of the Right modified ITRs in **Table 5A** are shown in FIG. 7A of International Application PCT/US2018/064242, filed December 6, 2018, and the predicted secondary structure of the Left modified ITRs in **Table 5B** are shown in FIG. 7B of International Application PCT/US2018/064242, filed December 6, 2018, which is incorporated herein in its entirety by reference.

[00212] **Table 5A** and **Table 5B** show exemplary right and left modified ITRs.

[00213] **Table 5A:** Exemplary modified right ITRs. These exemplary modified right ITRs can comprise the RBE of GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60), spacer of ACTGAGGC (SEQ ID NO: 69), the spacer complement GCCTCAGT (SEQ ID NO: 70) and RBE' (*i.e.*, complement to RBE) of GAGCGAGCGAGCGCGC (SEQ ID NO: 71).

Table 5A: Exemplary Right modified ITRs		
ITR Construct	Sequence	SEQ ID NO:
ITR-18 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCGCACGCCCGGGTTTCCCGGGCGGCCTCAGTG AGCGAGCGAGCGCGCAGCTGCCTGCAGG	15
ITR-19 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGACGCCCGGGCTTGCCCGGGCGGCCTCA GTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	16
ITR-20 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCCCGGG CGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	17
ITR-21 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCTTTGCCTCAGTGAGCGAGCGAGCGCGCAGC TGCCTGCAGG	18
ITR-22 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACAAAGTCGCCCAGCGCCCGGGCT TTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC AGG	19
ITR-23 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGAAAATCGCCCAGCGCCCGGGCTTT GCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAG G	20
ITR-24 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGAAACGCCCAGCGCCCGGGCTTTGC CCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	21
ITR-25 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCAAAGCCCGACGCCCGGGCTTTGCC GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	22
ITR-26 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCCCGGG TTTCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC AGG	23
ITR-27 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCCCGGT TTCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAG G	24
ITR-28 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCCCGTT TCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	25
ITR-29 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCCCTTT GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	26
ITR-30 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCCTTTG GCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	27
ITR-31 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCAGCGCTTTGC GGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	28

ITR-32 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGTTTCGG CCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	29
ITR-49 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGGCCTCA GTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	30
ITR-50 right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCCGG CGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	31

[00214] TABLE 5B: Exemplary modified left ITRs. These exemplary modified left ITRs can comprise the RBE of GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60), spacer of ACTGAGGC (SEQ ID NO: 69), the spacer complement GCCTCAGT (SEQ ID NO: 70) and RBE complement (RBE') of GAGCGAGCGAGCGCGC (SEQ ID NO: 71).

Table 5B: Exemplary modified left ITRs		
ITR-33 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG AAACCCGGGCGTGCGCCTCAGTGAGCGAGCGAGCGCGCAGAGAG GGAGTGGCCAACTCCATCACTAGGGGTTTCCT	32
ITR-34 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGTCGGGC GACCTTTGGTTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACTCCATCACTAGGGGTTTCCT	33
ITR-35 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG CAAAGCCCGGGCGTCGGCCTCAGTGAGCGAGCGAGCGCGCAGAG AGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT	34
ITR-36 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG GTCGGGCGACCTTTGGTTCGCCCCGGCCTCAGTGAGCGAGCGAGCGC GCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT	35
ITR-37 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCAAAGCCTC AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA CTAGGGGTTTCCT	36
ITR-38 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG CAAAGCCCGGGCGTCGGGCGACTTTGTCGCCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGT TCCT	37
ITR-39 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG CAAAGCCCGGGCGTCGGGCGATTTTCGCCCCGGCCTCAGTGAGCGA GCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTC CT	38
ITR-40 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG CAAAGCCCGGGCGTCGGGCGTTTCGCCCCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT	39
ITR-41 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG CAAAGCCCGGGCGTCGGGCTTTGCCCGGCCTCAGTGAGCGAGCGA GCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT	40
ITR-42 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGG AAACCCGGGCGTCGGGCGACCTTTGGTTCGCCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGT TCCT	41

ITR-43 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGA AACCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGA GCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTC CT	42
ITR-44 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGAA ACGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	43
ITR-45 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCAAA GGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGA GCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	44
ITR-46 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCAAAG GCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGC GCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	45
ITR-47 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCAAAGC GTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGC GCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	46
ITR-48 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGAAACGT CGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	47

[00215] In one embodiment, a ceDNA vector for expression of a desired transgene or therapeutic protein comprises, in the 5' to 3' direction: a first adeno-associated virus (AAV) inverted terminal repeat (ITR), a nucleotide sequence of interest (for example an expression cassette as described herein) and a second AAV ITR, where the first ITR (5' ITR) and the second ITR (3' ITR) are asymmetric with respect to each other – that is, they have a different 3D-spatial configuration from one another. As an exemplary embodiment, the first ITR can be a wild-type ITR and the second ITR can be a mutated or modified ITR, or vice versa, where the first ITR can be a mutated or modified ITR and the second ITR a wild-type ITR. In some embodiment, the first ITR and the second ITR are both mod-ITRs, but have different sequences, or have different modifications, and thus are not the same modified ITRs, and have different 3D spatial configurations. Stated differently, a ceDNA vector with asymmetric ITRs comprises ITRs where any changes in one ITR relative to the WT-ITR are not reflected in the other ITR; or alternatively, where the asymmetric ITRs have a modified asymmetric ITR pair can have a different sequence and different three-dimensional shape with respect to each other. Exemplary asymmetric ITRs in the ceDNA vector for expression of a desired transgene or therapeutic protein and for use to generate a ceDNA-plasmid are shown in **Table 5A and 5B**.

[00216] In an alternative embodiment, a ceDNA vector for expression of a desired transgene or therapeutic protein comprises two symmetrical mod-ITRs - that is, both ITRs have the same sequence, but are reverse complements (inverted) of each other. In some embodiments, a symmetrical mod-ITR pair comprises at least one or any combination of a deletion, insertion, or substitution relative to wild type ITR sequence from the same AAV serotype. The additions, deletions, or substitutions in the symmetrical ITR are the same but the reverse complement of each other. For example, an insertion of 3 nucleotides in the C region of the 5' ITR would be reflected in the insertion of 3 reverse

complement nucleotides in the corresponding section in the C' region of the 3' ITR. Solely for illustration purposes only, if the addition is AACG in the 5' ITR, the addition is CGTT in the 3' ITR at the corresponding site. For example, if the 5' ITR sense strand is ATCGATCG with an addition of AACG between the G and A to result in the sequence ATCGAACGATCG (SEQ ID NO: 51). The corresponding 3' ITR sense strand is CGATCGAT (the reverse complement of ATCGATCG) with an addition of CGTT (*i.e.* the reverse complement of AACG) between the T and C to result in the sequence CGATCGTTCGAT (SEQ ID NO: 49) (the reverse complement of ATCGAACGATCG) (SEQ ID NO: 51).

[00217] In alternative embodiments, the modified ITR pair are substantially symmetrical as defined herein - that is, the modified ITR pair can have a different sequence but have corresponding or the same symmetrical three-dimensional shape. For example, one modified ITR can be from one serotype and the other modified ITR be from a different serotype, but they have the same mutation (*e.g.*, nucleotide insertion, deletion or substitution) in the same region. Stated differently, for illustrative purposes only, a 5' mod-ITR can be from AAV2 and have a deletion in the C region, and the 3' mod-ITR can be from AAV5 and have the corresponding deletion in the C' region, and provided the 5' mod-ITR and the 3' mod-ITR have the same or symmetrical three-dimensional spatial organization, they are encompassed for use herein as a modified ITR pair.

[00218] In some embodiments, a substantially symmetrical mod-ITR pair has the same A, C-C' and B-B' loops in 3D space, *e.g.*, if a modified ITR in a substantially symmetrical mod-ITR pair has a deletion of a C-C' arm, then the cognate mod-ITR has the corresponding deletion of the C-C' loop and also has a similar 3D structure of the remaining A and B-B' loops in the same shape in geometric space of its cognate mod-ITR. By way of example only, substantially symmetrical ITRs can have a symmetrical spatial organization such that their structure is the same shape in geometrical space. This can occur, *e.g.*, when a G-C pair is modified, for example, to a C-G pair or vice versa, or A-T pair is modified to a T-A pair, or vice versa. Therefore, using the exemplary example above of modified 5' ITR as a ATCGAACGATCG (SEQ ID NO: 51), and modified 3' ITR as CGATCGTTCGAT (SEQ ID NO: 49) (*i.e.*, the reverse complement of ATCGAACGATCG (SEQ ID NO: 51)), these modified ITRs would still be symmetrical if, for example, the 5' ITR had the sequence of ATCGAACCATCG (SEQ ID NO: 50), where G in the addition is modified to C, and the substantially symmetrical 3' ITR has the sequence of CGATCGTTCGAT (SEQ ID NO: 49), without the corresponding modification of the T in the addition to a. In some embodiments, such a modified ITR pair are substantially symmetrical as the modified ITR pair has symmetrical stereochemistry.

[00219] **Table 6** shows exemplary symmetric modified ITR pairs (*i.e.* a left modified ITRs and the symmetric right modified ITR) for use in a ceDNA vector for expression of a desired transgene or therapeutic protein. The bold (red) portion of the sequences identify partial ITR sequences (*i.e.*, sequences of A-A', C-C' and B-B' loops), also shown in FIGS 31A-46B. These exemplary modified ITRs can comprise the RBE of GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60), spacer of

ACTGAGGC (SEQ ID NO: 69), the spacer complement GCCTCAGT (SEQ ID NO: 70) and RBE' (i.e., complement to RBE) of GAGCGAGCGAGCGCGC (SEQ ID NO: 71).

LEFT modified ITR (modified 5' ITR)		Symmetric RIGHT modified ITR (modified 3' ITR)	
SEQ ID NO: 32 (ITR-33 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGAAACCCGGGCGTGCGC CTCAGTGAGCGAGCGAGCGC GCAGAGAGGGAGTGGCCAACT CCATCACTAGGGGTTTCCT	SEQ ID NO: 15 (ITR-18, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCGCACGC CCGGGTTTCCCGGGCG GCCTCAGTGAGCGAGC GAGCGCGCAGCTGCCT GCAGG
SEQ ID NO: 33 (ITR-34 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGTC GGGCGACCTTTGGTCGCCCCG GCCTCAGTGAGCGAGCGAGC GCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTTCCT	SEQ ID NO: 48 (ITR-51, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGGCCTCAGTGAGCGA GCGAGCGCGCAGCTGC CTGCAGG
SEQ ID NO: 34 (ITR-35 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCG GCCTCAGTGAGCGAGCGAGC GCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTTCCT	SEQ ID NO: 16 (ITR-19, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGACGC CCGGGCTTTGCCCGGG CGGCCTCAGTGAGCGA GCGAGCGCGCAGCTGC CTGCAGG
SEQ ID NO: 35 (ITR-36 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC GGGCGTCGGGCGACCTTTGG TCGCCCCGGCCTCAGTGAGCG AGCGAGCGCGCAGAGAGGGA GTGGCCAACTCCATCACTAGG GGTTCCT	SEQ ID NO: 17 (ITR-20, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCCCGGGCGCCTCAG TGAGCGAGCGAGCGCG CAGCTGCCTGCAGG
SEQ ID NO: 36 (ITR-37 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCAAAG CCTCAGTGAGCGAGCGAGCG CGCAGAGAGGGAGTGGCCAACT TCCATCACTAGGGGTTTCCT	SEQ ID NO: 18 (ITR-21, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCTTTGCC TCAGTGAGCGAGCGAG CGCGCAGCTGCCTGCAG G
SEQ ID NO: 37 (ITR-38 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCG GGCGACTTTGTGCCCCGGCC TCAGTGAGCGAGCGAGCGCG CAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTTCCT	SEQ ID NO: 19 (ITR-22 right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACAAAGTCGCCCCGACG CCCGGGCTTTGCCCGG GCGGCCTCAGTGAGCG AGCGAGCGCGCAGCTG

			CCTGCAGG
SEQ ID NO: 38 (ITR-39 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCTG GGCGATTTTCGCCCCGGCCTC AGTGAGCGAGCGAGCGCGCA GAGAGGGAGTGGCCAACTCCA TCACTAGGGGTTTCCT	SEQ ID NO: 20 (ITR-23, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG AAAATCGCCCGACGCC CGGGCTTTGCCCCGGGC GGCCTCAGTGAGCGAG CGAGCGCGCAGCTGCC TGCAGG
SEQ ID NO: 39 (ITR-40 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCTG GGCGTTTCGCCCCGGCCTCAG TGAGCGAGCGAGCGCGCAGAG GAGGGAGTGGCCAACTCCATC ACTAGGGGTTTCCT	SEQ ID NO: 21 (ITR-24, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG AAACGCCCCGACGCCCG GGCTTTGCCCCGGGCGG CCTCAGTGAGCGAGCG AGCGCGCAGCTGCCTGC AGG
SEQ ID NO: 40 (ITR-41 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCTG GGCTTTGCCCCGGCCTCAGTG AGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACTCCATCAC TAGGGGTTTCCT	SEQ ID NO: 22 (ITR-25 right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCA AAGCCCCGACGCCCGGG CTTTGCCCCGGGCGGCC TCAGTGAGCGAGCGAG CGCGCAGCTGCCTGCAG G
SEQ ID NO: 41 (ITR-42 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGGAAACCCGGGCGTCGGG CGACCTTTGGTCGCCCCGGCC TCAGTGAGCGAGCGAGCGCG CAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTTCCT	SEQ ID NO: 23 (ITR-26 right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCCCCGGTTTCCCGG GCGGCCTCAGTGAGCG AGCGAGCGCGCAGCTG CCTGCAGG
SEQ ID NO: 42(ITR-43 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGGAAACCGGGCGTCGGGCG ACCTTTGGTCGCCCCGGCCTC AGTGAGCGAGCGAGCGCGCA GAGAGGGAGTGGCCAACTCCA TCACTAGGGGTTTCCT	SEQ ID NO: 24 (ITR-27 right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCCCCGGTTTCCGGGC GGCCTCAGTGAGCGAG CGAGCGCGCAGCTGCC TGCAGG
SEQ ID NO: 43 (ITR-44 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CGAAACGGGCGTCGGGCGAC CTTTGGTCGCCCCGGCCTCAG TGAGCGAGCGAGCGCGCAGAG GAGGGAGTGGCCAACTCCATC ACTAGGGGTTTCCT	SEQ ID NO: 25 (ITR-28 right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCCCCGTTTCGGGCGG CCTCAGTGAGCGAGCG AGCGCGCAGCTGCCTGC

			AGG
SEQ ID NO:44 (ITR-45 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC CAAAGGGCGTCGGGCGACCT TTGGTCGCCCCGGCCTCAGTG AGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACTCCATCAC TAGGGGTTTCCT	SEQ ID NO:26 (ITR-29, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCCCTTTGGGCGGGCC TCAGTGAGCGAGCGAG CGCGCAGCTGCCTGCAG G
SEQ ID NO:45 (ITR-46 left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCC AAAGGGCGTCGGGCGACCTTT GGTCGCCCCGGCCTCAGTGAG CGAGCGAGCGCGCAGAGAGG GAGTGGCCAACTCCATCACTA GGGGTTCCT	SEQ ID NO: 27(ITR-30, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCCTTTGGCGGGCCTC AGTGAGCGAGCGAGCG CGCAGCTGCCTGCAGG
SEQ ID NO: 46 (ITR-47, left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGCA AAGCGTCGGGCGACCTTTGG TCGCCCCGGCCTCAGTGAGCG AGCGAGCGCGCAGAGAGGGA GTGGCCAACTCCATCACTAGG GGTTCCT	SEQ ID NO: 28 (ITR-31, right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGCTTTGCGGCCTCAG TGAGCGAGCGAGCGCG CAGCTGCCTGCAGG
SEQ ID NO: 47 (ITR-48, left)	CCTGCAGGCAGCTGCGCGCTC GCTCGCTCACTGAGGCCGAA ACGTCGGGCGACCTTTGGTC GCCCCGGCCTCAGTGAGCGAG CGAGCGCGCAGAGAGGGAGT GGCCAACTCCATCACTAGGGG TTCCT	SEQ ID NO: 29 (ITR-32 right)	AGGAACCCCTAGTGATG GAGTTGGCCACTCCCTCT CTGCGCGCTCGCTCGC TCACTGAGGCCGGGGCG ACCAAAGGTCGCCCCGA CGTTTCGGCCTCAGTG AGCGAGCGAGCGCGCA GCTGCCTGCAGG

[00220] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein comprising an asymmetric ITR pair can comprise an ITR with a modification corresponding to any of the modifications in ITR sequences or ITR partial sequences shown in any one or more of **Tables 9A-9B** herein, or the sequences shown in FIG. 7A-7B of International Application PCT/US2018/064242, filed December 6, 2018, which is incorporated herein in its entirety, or disclosed in Tables 2, 3, 4, 5, 6, 7, 8, 9 or 10A-10B of International application PCT/US18/49996 filed September 7, 2018 which is incorporated herein in its entirety by reference.

III. Exemplary ceDNA vectors

[00221] As described above, the present disclosure relates to recombinant ceDNA expression vectors and ceDNA vectors that encode a desired transgene or therapeutic protein, comprising any one of: an asymmetrical ITR pair, a symmetrical ITR pair, or substantially symmetrical ITR pair as described above. In certain embodiments, the disclosure relates to recombinant ceDNA vectors for expression of a desired transgene or therapeutic protein having flanking ITR sequences and a

transgene, where the ITR sequences are asymmetrical, symmetrical or substantially symmetrical relative to each other as defined herein, and the ceDNA further comprises a nucleotide sequence of interest (for example an expression cassette comprising the nucleic acid of a transgene) located between the flanking ITRs, wherein said nucleic acid molecule is devoid of viral capsid protein coding sequences.

[00222] The ceDNA expression vector for expression of a desired transgene or therapeutic protein may be any ceDNA vector that can be conveniently subjected to recombinant DNA procedures including nucleotide sequence(s) as described herein, provided at least one ITR is altered. The ceDNA vectors for expression of a desired transgene or therapeutic protein of the present disclosure are compatible with the host cell into which the ceDNA vector is to be introduced. In certain embodiments, the ceDNA vectors may be linear. In certain embodiments, the ceDNA vectors may exist as an extrachromosomal entity. In certain embodiments, the ceDNA vectors of the present disclosure may contain an element(s) that permits integration of a donor sequence into the host cell's genome. As used herein "transgene" and "heterologous nucleotide sequence" are synonymous, and encode a desired transgene or therapeutic protein, as described herein.

[00223] Referring now to **FIGS 1A-1G**, schematics of the functional components of two non-limiting plasmids useful in making a ceDNA vector for expression of a desired transgene or therapeutic protein are shown. **FIG. 1A, 1B, 1D, 1F** show the construct of ceDNA vectors or the corresponding sequences of ceDNA plasmids for expression of a desired transgene or therapeutic protein. ceDNA vectors are capsid-free and can be obtained from a plasmid encoding in this order: a first ITR, an expressible transgene cassette and a second ITR, where the first and second ITR sequences are asymmetrical, symmetrical or substantially symmetrical relative to each other as defined herein. ceDNA vectors for expression of a desired transgene or therapeutic protein are capsid-free and can be obtained from a plasmid encoding in this order: a first ITR, an expressible transgene (protein or nucleic acid) and a second ITR, where the first and second ITR sequences are asymmetrical, symmetrical or substantially symmetrical relative to each other as defined herein. In some embodiments, the expressible transgene cassette includes, as needed: an enhancer/promoter, one or more homology arms, a donor sequence, a post-transcription regulatory element (*e.g.*, WPRE, *e.g.*, SEQ ID NO: 67)), and a polyadenylation and termination signal (*e.g.*, BGH polyA, *e.g.*, SEQ ID NO: 68).

[00224] **FIG. 5** is a gel confirming the production of ceDNA from multiple plasmid constructs using the method described in the Examples. The ceDNA is confirmed by a characteristic band pattern in the gel, as discussed with respect to **FIG. 4A** above and in the Examples.

A. Regulatory elements.

[00225] The ceDNA vectors for expression of a desired transgene or therapeutic protein as described herein comprising an asymmetric ITR pair or symmetric ITR pair as defined herein, can further comprise a specific combination of cis-regulatory elements. The cis-regulatory elements

include, but are not limited to, a promoter, a riboswitch, an insulator, a mir-regulatable element, a post-transcriptional regulatory element, a tissue- and cell type-specific promoter and an enhancer. Exemplary Promoters are listed in **Table 7**. Exemplary enhancers are listed in **Table 8**. In some embodiments, the ITR can act as the promoter for the transgene, *e.g.*, a desired transgene or therapeutic protein. In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein as described herein comprises additional components to regulate expression of the transgene, for example, regulatory switches as described herein, to regulate the expression of the transgene, or a kill switch, which can kill a cell comprising the ceDNA vector encoding a desired transgene or therapeutic protein thereof. Regulatory elements, including Regulatory Switches that can be used in the present invention are more fully discussed in International application PCT/US18/49996, which is incorporated herein in its entirety by reference.

[00226] In embodiments, the second nucleotide sequence includes a regulatory sequence, and a nucleotide sequence encoding a nuclease. In certain embodiments the gene regulatory sequence is operably linked to the nucleotide sequence encoding the nuclease. In certain embodiments, the regulatory sequence is suitable for controlling the expression of the nuclease in a host cell. In certain embodiments, the regulatory sequence includes a suitable promoter sequence, being able to direct transcription of a gene operably linked to the promoter sequence, such as a nucleotide sequence encoding the nuclease(s) of the present disclosure. In certain embodiments, the second nucleotide sequence includes an intron sequence linked to the 5' terminus of the nucleotide sequence encoding the nuclease. In certain embodiments, an enhancer sequence is provided upstream of the promoter to increase the efficacy of the promoter. In certain embodiments, the regulatory sequence includes an enhancer and a promoter, wherein the second nucleotide sequence includes an intron sequence upstream of the nucleotide sequence encoding a nuclease, wherein the intron includes one or more nuclease cleavage site(s), and wherein the promoter is operably linked to the nucleotide sequence encoding the nuclease.

[00227] The ceDNA vectors for expression of a desired transgene or therapeutic protein produced synthetically, or using a cell-based production method as described herein in the Examples, can further comprise a specific combination of cis-regulatory elements such as WHP posttranscriptional regulatory element (WPRE) (*e.g.*, SEQ ID NO: 67) and BGH polyA (SEQ ID NO: 68). Suitable expression cassettes for use in expression constructs are not limited by the packaging constraint imposed by the viral capsid.

(i). Promoters:

[00228] It will be appreciated by one of ordinary skill in the art that promoters used in the ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein should be tailored as appropriate for the specific sequences they are promoting. Exemplary promoters operatively linked to a transgene useful in a ceDNA vector are disclosed in **Table 7**, herein.

Table 7

Table 7: promoters						
Genetic Element_Type	Description	Length	Tissue Specificity	CG Content	SEQ ID NO	Sequence
promoter	chicken B-actin core promoter ; part of constitutive CAG promoter set	278	Constitutive	33	200	TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCATCTCCCCCCTCCCCACCCCAATTTTGTATTTATTTTAAATATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAGCGAAGCGCGGCGGGGCG
promoter	hAAT promoter ; part of hAAT promoter Set	348	Liver	12	201	GATCTTGCTACCACTGGAACAGCCACTAAGGATTCTGCAGTGAGAGCAGAGGGCCAGCTAAGTGGTACTCTCCAGAGACTGTCTGACTCACGCCACCCCTCCACCTTGACACAGGACGCTGTGGTTCTGAGCCAGGTACAATGACTCTTTTGGTAAGTGCAGTGGAAAGCTGTACTGCCCAGGCAAAGCGTCCGGGCGAGCGTAGGCGGGCGACTCAGATCCCAGCCAGTGGAAGTGGCCCTGTTTGTCTCTCCGATAACTGGGGTGACCTTGTTAATATTCACCAGCAGCTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGG
promoter	CpG-free human EF1a core promoter (3' sequence AAGCTT may be a spacer/restriction enzyme cut site and was absorbed); part of CET promoter set	226	Constitutive	0	202	GTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCTCAGTGGGAGAGAGCAGCATGGCCACAGTCCCTGAGAAGTTGGGGGAGGGGTGGGCAATTGAAGTGGTGCCTAGAGAAGTGGGGCTTGGGTAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCAAGCTT
promoter	murine TTR liver specific promoter (3' CTCCTG may be a spacer/restriction enzyme cut site and was absorbed); part of CRM8 VandenDriessche promoter set	225	Liver	5	203	CCGTCTGTCTGCACATTTCTAGAGCGAGTGTTCCGATACTCTAACTCTCCCTAGGCAAGGTTTATTTGTGTAGGTTACTTATCTCCTTTGTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCTGGGTTGGAAGGAGGGGGTATAAAGCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCCTG
promoter	HLP promoter derived from BMN270	143	Liver	5	204	GGCGACTCAGATCCAGCCAGTGGACTTAGCCCTGTTTGCTCTCCGATAACTGGGGTGACCTTGGTTAATATTCACCAGCAGCCTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTC
promoter	Mutant TTR promoter derived from SPK-8011	222	Liver	4	205	GTCTGTCTGCACATTTCTAGAGCGAGTGTTCCGATACTCTAATCTCCCTAGGCAAGGTTTATTTGACTTAGGTTACTTATCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTGGCAGGGATCAGCAGCTGGGTTGGAAGGAGGGGGTATAAAGCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCCT

Table 7: promoters						
promoter	TTR promoter derived from Sangamo CRMSBS2-Intron3	223	Liver	4	206	GTCTGTCTGCACATTTCTGAGAGCGAGTGTCCGATACTCTAATC TCCTAGGCAAGGTTTCATATTTGTGTAGGTTACTTATTCTCCTTTT GTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCT TGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAA AGCCCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCC TG
promoter	Endogenous hFVIII promoter (-3000 to -1 of 5' flanking genomic sequence)	3000	Endogenous	21	207	GTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACT ACAGGCACGTGCCACCATGCCCGCTAATTTTTGTATTTTTAGT AGAGGAGGAGTTTCATCTTGTAGCTAGGATGGTCTAGATCTCC TGACCTCGTGATCTGCCCGCTCAGCCTCCCAAAGTGCTGGGAT TACAGGTGTGAGCCACCGTGCCCGGCCATATTTGATTTAAAAA TTAGCAATAATAGATAAAATTTTCAATCAACTAAGCCCTTGGGCC AGGGAATGCTATTCTTAAAAAGTGCTTCTATCAATATAGCCTCT GACTCATTACTTTGTTAATTTTTAAATTGATTTTCATTCTGATTA ACATTCACCCAGATTATTAATTATACAATCTGTTAACTGTAGA ACCTCAAACATGTTGGATTGTACTGTATTTGTCTGGAAGACACAT TTTTAAACATTGTAATCGCTATAAGAGAAGCACTGGGAAAGAA AGGAGCTTCTATGCCTGCAGTGCTGAGGAGCCCTTAACAGTG TGCCCGCCCCCTAAGCTACTCATGCAGTCATCCCCATCCAGTTA GTCAACTTTATTCCAAAAAAGCTGGTGTTCCAAATTTTCTTCTC AAAGCCACAGATCCAAATTCATCAGCAGTTCCACAAACGTT ACCCTCACATGAATCCAGCCATTTTACCCTCTCCAGTGGTAC CATCATAGCCCAAGCCGCCACCATTTCTACCCCCGGTTAACAG GCCACCTCCTTCTACCCTTATCCTGTAGAGTTTGTATATCTAC AGTGATCAGAAAGATCAGCCTAAAAGATAATTCTGATCACCACC CTCCTCTACTCACAACCCGGCCGTGTCTCCCATGCCCCTCAGTG TAGAAGTCAATGTCCCTTTGCTGAAATGCAACCTTAGTGAAACTT TCCATGACTAACCTCCTTTAAATTTGCAACCTGGTCCACCCTTAC TCCCCCTTACCCACTTCTCTTTTTGCACAGCACTTATTTACCTT CTAACATACTGTATAATGTACTCATGTATTGTAATTATTGCTTATC ATCCCTCTTTCAGTTGCTTATATTTTTCATCAATGTGTACCCAGTG CCTAGGACAATATCTGTCTAGGACAAATGGGTAGTTATGTGGCT GTAGGCAAGCCATTTAACCTCTCTGTACCTCAGTTACTTTATCTG TATCCACTTTGCGGTGTTGTCTGAGGATTAATCAGATAGCCT ATGTGTAGCACCTGGCAGTGAATTTATCACCTGTACTGTAAC GTCTACTTTCTGTCTCCTCCATTGGACTGTCTTCCAGGGGGT TGGGAAGTGGGATTTCTTCTTCTGAGGCATAGAAGTATAGCA TAGTGGTTAGGAGCATGACTTCTGGAGCCAGAGTACATGGGTTT GAATGCTACCACTCACAAGCTGTGTGCCATGGAGAAGTTGCCCT AACCTCTCCGTGCTTCAGTTTCATCACCATAAAATGAAGGTAAG AATAGTACCTGTATTTAAAGCACCTAGAACAGTTCTGGCATA TAGTGTCAGCTGTCTCTGTCATCCTGTACCTGTACAGAGAGG AGTGTTTATCAAAGGGGCTTCTGTGCTGTTTCCAAACAGTCT GACAATATACCAATTGCTCCCTAACACATTCTGTTGTGCAGAA CTGAGCTCAATGATAACATTTTATAGCAACCCTGATCAAGTTTC TTCTATAATCTTACACTTTGAGGCCCTGCAGGGGCCCTCAC TCTCCCTAATAAACATTAACCTGAGTAGGGTGTGAGCTACCA TGGCTACATTCTGATGTAAAGAGATATCCTATACCTGGGCCA AATGTAAACAGCCTGAAAAAGTGTAGGTTAAAAACAAACAA AATAAATAAATGAATAAATGCCAGGTGGTTATGAGTGCTATTGA GAAAAATGAAGCCAAGAGGGATATCAGTGATGCAGGTGGGGG TAAAGAGCTTACAACATAAATGTGGTGTCCATATTTAAACCTCA TTCAACAGGGAAGATTGGAGCTGAAATGTGAAGGAGTTGTGGG AGTGGAACCTACGTGAAATCTGGGGGAAAGGTGTTTGGGTAA AAGAAATAGCAAGTGTGAGGTCCAGGGGCATGAGTGTGCTTG ATATTTTAGGGAAGAGTAAGGAGACCAAGTATAACCAGAGTGAG ATGAGACTACAGAGGTGAGGAGAAAGGGCATGCAGACCATGTG GGATGCTCTAGGACCTAGGCCATGGTAAAGATGTAGGGTTTAC CCTGATGGAGGTGAGAAGCCATTGGAGGATTCTGAGAAGAGGA GTGACAGGACTCGCTTTATAGTTTAAATTATAACTATAAATTAT AGTTTTTAAACAATAGTTCCTAACCTCATGTTATATGTAAAC

Table 7: promoters						
						TACAGTTTTAAAACTATAAATTCCTCATACTGGCAGCAGTGTGA GGGGCAAGGGGAAAAGCAGAGAGACTAACAGGTTGCTGGTTA CTCTTGCTAGTGCAAGTGAATTCAGAACTTCGACAACATCCAG AATTCTCTTGCTGCTGCCACTCAGGAAGAGGGTTGGAGTAGGC TAGGAATAGGAGCACAAATTAAGCTCCTGTTCACTTTGACTTCT CCATCCCTCTCCTCCTTTCCTTAAAGGTTCTGATTAAAGCAGACT ATGCCCCTACTGCTCTCAGAAGTGAATGGGTTAAGTTTAGCAGC CTCCCTTTTGCTACTTCAGTTCCTGTGGCTGCTCCCACTGAT AAAAAGGAAGCAATCCTATCGGTTACTGCTTAGTGCTGAGCACA TCCAGTGGGTAAAGTTCCTTAAATGCTCTGCAAAGAAATTTGGG ACTTTTCATTAAATCAGAAATTTACTTTTTCCCTCCTGGGAGC TAAAGATATTTAGAGAAGAATTAACCTTTTGCTTCCAGTTGA ACATTTGTAGCAATAAGTC
promoter	hAAT promoter derived from Nathwani_hFlX	205	Liver	10	208	AATGACTCCTTCGGTAAGTGACGTGGAAGCTGTACACTGCCCA GGCAAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCCA GCCAGTGGACTTAGCCCTGTTTGCTCCTCCGATACTGGGGTG ACCTTGGTTAATATTCACCAGCAGCCTCCCCGTTGCCCTCTGG ATCCACTGCTTAAATACGGACGAGGACAGG
promoter	hAAT promoter derived from SPK9001	397	Liver	12	209	GATCTTGCTACCAAGTGAACAGCCACTAAGGATTCTGCAGTGAG AGCAGAGGGCCAGCTAAGTGGTACTCTCCAGAGACTGTCTGA CTCAGGCCACCCCTCCACCTTGACACAGGACGCTGTGGTTCT GAGCCAGGTACAATGACTCCTTCGGTAAGTGACGTGGAAGCT GTACACTGCCAGGCAAAGCGTCCGGGCAGCGTAGGCGGGCG ACTCAGATCCCAGCCAGTGGACTTAGCCCTGTTTGCTCCTCGA TAACTGGGGTGACCTTGTTAATATTCACCAGCAGCCTCCCCG TTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGC CCTGTCTCCTCAGCTTCAAGCACCACTGACCTGGGACAGTG AAT
promoter	Endogenous hG6Pase promoter (-2864 to -1 of 5' flanking)	2864	Endogenous (Liver)	28	210	CCTTTGAGAATCCACGGTGTCTCGATGCAGTCAGCTTTCTAACA GCTGGGGCCTCACCTGTTTTCCACGGATAAAAACGTGCTGGAG GAAGCAGAAAGGGGCTGGCAGGTGGAAGATGAGGACCAGCT CATCGTCTCATGACTATGAGGTTGCTCTGATCCAGAGGGTCCCC CTGCCTGGTGGCCACCAGGCAAGACTCCCACTGTCCCTGGA TGCCAGAGTGGGATGTCAACTCCATCACTTATCAACTCCTTATC CATAGGGGTATTCTTCTGAGGCGTCTCAGAAAACAGGGCCCTC CCCATATGCTGACCACATAATAGAACCCCTCCCACTCAGAGAC CCTGGCTGCTAGCTGCCCTGGCATGACCCAGACAGTGGCCTTTG TATATGTTTTAGACTCACCTTGACTCACCTGACCATAGAAAC TCTCATCCCAGAGGTCACTGCAATAGTTACTCCACAACAGAGGC TTATCTGGGTAGAGGGAGGCTCCCTACCTATGGCCAGCAGCCC TGACAGTGCAGATCACATATACCCACGCCCCAGCACTGCCTGC CACGCATGGGCTTACTTTACACCCACCCACAGTCACCAACACATT ACCTGCTCTCAAGGTTAGGCGTGGCAGGAGAAGTTTGCTTGG ACCAGCAGAAACCATGCAGTCAAGGACAACCTGGAGTCAGCATG GGCTGGGTGCGAGCCCTTGGTGGGTGGGGAGGAGACTCCAG GTCATACCTCCTGGAGGATGTTTTAATCATTTCCAGCATGGAATG CTGTCAACTTTTGCCACAGATTCATTAGCTCTGAGTTCTTTTTTC TGTCCTCCAGTACCCCTTACATGTCAATATGGACTTAATGATGGG AAATTCAGGCAAGTTTTTAAACATTTTATCCCCCTGGCTCTTATC CTCAAAAATGCATGAATTTGGAGGAGTGCGTCAATGCTGTAA TCCCAATGCTTTGCTAGGTTGAGGCGGGAGGATCACTGAAGCC AGGAATTTGAGACCAGCCTGGGCCGCATAGTGAGACCCCGTTTC TACAAAAATAAATAAATAAATAAATAAATAGTGATATGAAGC ATGATTAAATAGCCCTATTTTTTAAATGCATGAGTTCGTTACCT GATTCATTCCCTGGTTCCTTTCACAGTCCTCCGTGACCCAGTGT TAGGGTTTTGGTCTCTCTACTATTGTAGGCTGATATATAGTATA CACACACACACACACATATACACACACAGTGTATCTTG AGCTTTCTTTGTATATCTACACACATATGTATAAGAAAGCTCAA GATATAGAAGCCCTTTTCAAAAATAACTGAAAGTTTCAAACTCT

Table 7: promoters

						TTAAGTCTCCAGTTACCATTTTCTGGTATTCTTATTTGGAACCAT ACATTCATCATATTGTTGCACAGTAAGACTATACATTCAATTTTT GCTTAAACGTATGAGTTAAACACTTGGCCAGGCATGGTGGTTC ACACCTGTAATCCCAGAGCTTTGGGAAGCCAAGACTGGCAGATC TCTTGAGCTCAGGAATTCAAGACCAGCCTGGGCAACATGAAAA AACCCCATCTCTACAAAAGATAGAAAAATTAGCCAGGCATGGTG GCGTGTGCCTGTGGTCCCAGCTACTCAGGAGGCTGAGGTGGGA GGATCACATTAGCCCAGGAGGTTGAGGCTGCAGTGAGCCGTGA TTATGCCACTGCACTCCAGCCTGGGAGACAGAGTGAGACCCTGT TTCAAAAAAGAGAGAGAAAAATTTAAAAAGAAAAACAACACC AAGGGCTGTAACTTAAGGTCAATTAATGAATTAATCACTGCAT TCAAAAAACGATTACTTTCTGGCCCTAAGAGACATGAGGCCAATA CCAGGAAGGGGTTGATCTCCCAAACAGAGGCAGACCCTAGA CTCTAATACAGTTAAGGAAAGACCAGCAAGATGATAGTCCCA TACAATAGAAGTTACTATATTTATTTGTTGTTTTCTTTGTTTT GTTTTGTTTTGTTTTGTTTTGTTTTAGAGACTGGGGTCTTGCTCG ATTGCCAGGCTGTAGTGCAGCGGTGGGACAATAGCTCACTGC AGACTCCAACCTCTGGGCTCAAGCAATCCTCCTGCCTCAGCCTCC TGAATAGCTGGGACTACAAGGTACACCATCACACACCAAAAA CAATTTTTTAAATTTTTGTGTAGAAACGAGGTCTTGCTTTGTG CCCAGGCTGGTCTCCAACCTCTGGCTTCAAGGGATCCTCCACCT CAGCCTCCCAAATTGCTGGGATTACAGGTGTGAGCCACCACAAC CAGCCAGAACTTTACTAATTTTAAATTAAGAACTTAAACTTGA ATAGCTAGAGCACAAGATTTTTCTTTGTCCTCCAAATAAGTGCA GTTGCAGGCATAGAAAATCTGACATCTTGAAGAATCATCGTG GATGTAGACTCTGTCTGTGTCTCTGGCCTGGTTTCGGGGACCA GGAGGGCAGACCCTTGCACTGCCAAGAAGCATGCCAAAGTTAA TCATTGGCCCTGCTGAGTACATGGCCGATCAGGCTGTTTTGTG TGCTGTTTTCTATTTTACGTAAATCACCTGAACATGTTTGCAT CAACCTACTGGTGATGCACCTTTGATCAATACATTTTAGACAAAC GTGGTTTTTGAGTCCAAAGATCAGGGCTGGGTTGACCTGAATAC TGGATACAGGGCATATAAAACAGGGGCAAGGCACAGACTCATA GCAGAGCAATCACCACCAAGCCTGGAATAACTGCAAGGGCTCT GCTGACATCTTCTGAGGTGCCAAGGAATGAGG
promoter	Human Rhodopsin kinase (GRK1) promoter (1793-2087 of genbank entry AY327580)	295	Photoreceptors	11	211	GGGCCCCAGAAGCCTGGTGGTTGTTGTCCTTCTCAGGGGAAAA GTGAGGCGGGCCCTTGAGGAAGGGGCGGGCAGAATGATCT AATCGGATTCCAAGCAGCTCAGGGGATTGTCTTTTCTAGCACCT TCTTGCCACTCTAAGCGTCTCCGTGACCCCGCTGGGATTTAG CCTGGTGCTGTGTGAGCCCCGGCTCCAGGGGCTTCCAGTG TCCCCAGGGAACCTCGACAGGGCCAGGGCGTCTCTCTGTC GCAAGGGCAGGGACGGGCCACAGGCAAGGGC
promoter	Truncated hAAT Core promoter; Part of LP1 promoter set	206	Liver	10	212	GAATGACTCCTTCGGTAAGTGCAGTGGAAGCTGTACTGCCC AGGCAAAGCGTCCGGGCAGCGTAGGCGGGCAGACTCAGATCCCA GCCAGTGGACTTAGCCCTGTTGCTCCTCCGATAACTGGGGTG ACCTTGTTAATATTACCAGCAGCCTCCCCGTTGCCCTCTGG ATCCACTGCTTAAATACGGACGAGGACAGG
promoter	Human EF-1a promoter (contains EF-1a intron A)	1179	Constitutive	94	213	GGCTCCGGTGCCCGTCACTGGGCGAGCGCATCGCCACAG TCCCCGAGAAGTTGGGGGAGGGGTGCGCAATTGAACCGGTG CCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCG TGTAAGTGCAGTAGTCGCGTGAACGTTCTTTTCGCAACGGGT TGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCGCGGG CCTGGCCTCTTACGGGTTATGGCCCTGCGTGCTTGAATTA TCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT GGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCC CTTCGCTCGTGCTTGAGTTGAGGCTGGCTGGGCGCTGGGG CCGCCGCTGCGAATCTGGTGGCACCTTCGCGCTGTCTCGCTG CTTTCGATAAGTCTTAGCCATTTAAATTTTATGATGACCTGCTG CGACGCTTTTTTCTGGCAAGATAGTCTTGTAATGCGGGCCAA GATCTGCACACTGGTATTTCTGGTTTTGGGGCCGGGGCGGCGA

Table 7: promoters

						CGGGGCCCGTGCCTCCAGCGCACATGTTCTGGCGAGGCGGGGCTGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCTGGTCTCGCGCCGCCGTGTATCGCCCCCCTGGGCGGCAAGGCTGGCCCGGTGGGCACCAAGTTCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGA GTCACCCACACAAAGGAAAAGGGCCTTCCGTCTCAGCCGTGCTTCATGTGACTCCACGGAGTACCGGGCGCGCTCCAGGCACCTCGATTAGTTCTCGAGCTTTGGAGTACGTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCACTGAGTGGGTGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCCCTTTTGGAGTTGGATCTTGGTTCAATCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTCTCCATTTAGGTGTCGTGA
promoter	hRK promoter- Nearly identical to human rhodopsin kinase (GRK1) promoter (1793-2087 of genbank entry AY327580), but with a few indels of unknown origin.	292	Photoreceptors	11	214	GGGCCCCAGAAGCCTGGTGGTTGTTTGTCTTCTCAGGGGAAAA GTGAGGCGGGCCCTTGAGGAAGGGGCGGGCAGAATGATCTAATCGGATCCAAGCAGCTCAGGGGATTGTCTTTTCTAGCACCTTCTTGCCACTCCTAAGCGTCTCCGTGACCCCGGTGGGATTAGCCTGGTGCTGTGTAGCCCCGGTCTCCAGGGGCTTCCAGTGGTCCCCAGGAACCCTCGACAGGGCCCGGTCTCTCTGTCCAGCAAGGCAGGACGCGGCCACAGGCCAAGGGC
promoter	Interphotoreceptor retinoid-binding protein (IRBP) promoter sequence	1325	Photoreceptors	14	215	GCTGCCTACTGAGGCACACAGGGGCGCCTGCCTGCTGCCCGCTCAGCCAAGGCGGTGTTGCTGGAGCCAGCTTGGGACAGCTCTCCCAACGCTCTGCCCTGGCCTTGGACCACTCTCTGGGCGTAGTTGTCTGTCTGTAAAGTGAGGAAAGTGCCCATCTCCAGAGGCATTAGCGGCAAGCAGGGCTTCCAGGTTCCGACCCCATAGCAGGAC TTCTTGGATTCTACAGCCAGTCAGTTGCAAGCAGCACCCAAATTATTTCTATAAGAAGTGGCAGGAGCTGGATCTGAAGAGTCAGCAGTCTACCTTCCCTGTTTCTGTGCTTATGCAGTCAGGAGGAATGATCTGGATTCCATGTGAAGCCTGGGACCAGGAGACCCAAGACTTCTGCTTGATTCTCCCTGCGAACTGCAGGCTGTGGGCTGAGCCTTCAAGAAGCAGGAGTCCCTCTAGCCATTAACCTCTCAGAGCTAACCTCATTTGAATGGGAACACTAGTCTGTGATGTCTGGAAGGTGGGGGCTCTACACTCCACACCTACATGGTGGTCCAGACACATCATTCCAGCATTAGAAAGCTAGGGGGACCCGTTCTGTTCCCTGAGGCATTAAAGGGACATAGAAATAAATCTCAAGCTCTGAGGCTGATGCCAGCCTCAGACTCAGCCTCTGCACTGTATGGGCCAATGTAGCCCCAAGGACTTCTTCTGTGTCACCCCTATCTGCCACACCTAAAACGATGGGCTTCTATTAGTTACAGAACTCTCTGGCCTGTTTTGTTTGCTTTGCTTTGTTTTGTTTTGTTTTTGTTTTTTGT TTTTAGCTATGAAACAGAGGTAATATCTAATACAGATAACTTACAGTAATGAGTGCTTCTACTTACTGGTACTGGGAAGAAGTGC TTTACAGATATTTTCTCATTTAATCTACACAATAAGTAATTAAGCATTTCCCTGAGGCCACGGGAGAGACAGTGGCAGAACAGTTCTCCAAGGAGGACTTGCAAGTTAATAACTGGACTTTGCAAGGCTCTGTGGAACCTGTGAGCTTGTAAAGGATGGAGCACAGTGTCTGGCATGTAGCAGGAACCTAAAATAATGGCAGTGATTAATGTTATGATATGCAGACACACAGCAAGATAAGATGCAATGTACCTCTGGGTCAAACACCTGGCCACTCTCCCGATACCCAGGGTTGATGTCTTGAATTAGACAGGATTAAGGCTTACTGGAGCTGGAAGCCTGGCCCAACTCAGGAGTTTAGCCCCAGACCTTCTGTCCACCAG

Table 7: promoters						
promoterSet	promoter set containing CpGmin CME Enhancer, SV40_Enhancer_Invivogen, and CpG-free hEF1a core promoter	883	Constitutive	0	216	GAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCAATAGGGACTTTCCATTGGGTTTTGCCAGTACATAAGGTCAATA GGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATT GAGTCAATAGGGACTTTCCAATGGGTTTTGCCAGTACATAAGG TCAATGGGAGGTAAGCCAATGGGTTTTCCATTACTGACATGT ATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCAGTACAT AAGGTCAATAGGGGTGAATCAACAGGAAAGTCCCATTGGAGCC AAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCAG TACAAAAGGTCAATAGGGGGTGAGTCAATGGGTTTTCCATTAT TGGCACATACATAAGGTCAATAGGGGTGGGGCTGAAATAAC CTCTGAAAGAGGAACCTGGTTAGGTACCTCTGAGGCTGAAAG AACAGCTGTGGAATGTGTGTCAGTTAGGGGTGGGAAAGTCCC CAGGCTCCCCAGCAGGCAGAAAGTATGAAAGCATGCATCTCAAT TAGTCAGCAACCAGGTGTGGAAGTCCCCAGGCTCCCCAGCAG GCAGAAAGTATGAAAGCATGCATCTCAATTAGTCAGCAACCATA GTCCCACTAGTGAGAAGAGCATGCTTGAGGGCTGAGTGCCCC TCAGTGGGCAGAGAGCACATGGCCACAGTCCCTGAGAAAGTTG GGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGG GGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCAC CTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAAGTGT CTCTGTGAACATTCTCTGTGAACATTCTCTGTGAACATTCAAGCTTCTGCTTCTCC CTCCTGTGAGTTGGTAAGTCACTGACTGTCTATGCCTGGGAAA GGGTGGGCAGGAGATGGGGCAGTGCAGGAAAAGTGGCACTAT GAACCCTGCAGCCCTAGACAATTGTACTAACCTTCTTCTTTCTC CTCCTGACAGGTTGGTGTACAGTAGCTTCC
promoterSet	promoter set containing SV40_Enhancer_Invivogen, CpG-free hEF1a core promoter, and CET Intron	639	Constitutive	0	217	GGGCTGAAATAACCTCTGAAAGAGGAACCTGGTTAGGTACCTT CTGAGGCTGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGG TGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAAGTATGCAAA GCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAGTCCCCA GGCTCCCCAGCAGGCAGAAAGTATGAAAGCATGCATCTCAATTA GTCAGCAACCATAAGTCCCACTAGTGGAAGAGCATGCTTGAG GGCTGAGTGCCCTCAGTGGGCAGAGAGCACATGGCCACAGT CCCTGAGAAGTTGGGGGAGGGGTGGGCAATTGAACTGGTGC CTAGAGAAGGTGGGGCTTGGGTAACTGGGAAAGTGATGTGG TGTACTGGCTCACCTTTTTCCCCAGGGTGGGGGAGAACCATAT ATAAGTGCAAGTAGTCTCTGTGAACATTCAAGCTTCTGCTTCTCC CTCCTGTGAGTTGGTAAGTCACTGACTGTCTATGCCTGGGAAA GGGTGGGCAGGAGATGGGGCAGTGCAGGAAAAGTGGCACTAT GAACCCTGCAGCCCTAGACAATTGTACTAACCTTCTTCTTTCTC CTCCTGACAGGTTGGTGTACAGTAGCTTCC
promoterSet	CpGmin hAAT promoter Set; contains CpGmin APOe-CR hAAT enhancer, hAAT core promoter, and CpGmin hAAT-Intron	1272	Liver	24	218	AGGCTCAGAGGCACACAGGAGTTCTGGGCTCACCTGCCCTT TCCAACCCCTCAGTCCCATCCTCCAGCAGCTGTTGTGTGCTGC CTCTGAAGTCCACACTGAACAACTTCAGCCTACTCATGTCCCTA AAATGGGCAACATTGCAAGCAGCAAAACAGCAAAACACACAGCC CTCCTGCTGCTGACCTTGGAGCTGGGGCAGAGGTGAGAGAC CTCTCTGGGGCCATGCCACCTCCAACCTCACTCGACCCCTTGG AATTTCGGTGGAGAGGAGCAGAGGTTGTCTGGCTGTTTAGG TAGTGTGAGAGGGTCCGGGTTCAAACCACTTGTCTGGGTGGGG AGTCGTAGTAAGTGGCTATGCCCCGACCCGAAGCCTGTTTCC CCATCTGTACAATGGAATGATAAAGACGCCCATCTGATAGGGT TTTTGTGGCAATAAACATTGGTTTTTTGTTTTGTTTTGTTTTG TTTTTGAGATGGAGTTTGTCTGTGCGCCAGGCTGGAGTGCA GTGACACAATCTCATCTACCACAACCTTCCCTGCCTCAGCCTC CCAAGTAGCTGGGATTACAAGCATGTGCCACCACACCTGGCTAA TTTTCTATTTTTAGTAGAGACGGGTTTCTCCATGTTGGTCAAGCCT CAGCCTCCCAAGTAAGTGGGATTACAGGCCTGTGCCACCACACC CGGCTAATTTTTCTATTTTTGACAGGGACGGGTTTACCATGT TGGTCAGGCTGGTCTAGAGGTAAGTCTGCTACCAAGTGGAA CAGCCACTAAGGATTCTGCAGTGAGAGCAGAGGGCCAGCTAAG TGGTACTCTCCAGAGACTGTCTGACTCAGCCACCCCTCCACC TTGGACACAGGACGCTGTGGTTTCTGAGCCAGGTACAATGACTC CTTTCGGTAAGTGCAAGTGGAGCTGTACACTGCCAGGCAAAAG CGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCAGCCAGTGG ACTTAGCCCTGTTTGTCTCTCGATAACTGGGGTGACCTTGGTT AATATTCACAGCAGCTCCCCGTTGCCCTCTGGATCCACTGC

Table 7: promoters						
						TTAAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTTCAGGC ACCACCCTGACCTGGGACAGTGAATAATTACTCTAAGGTAAT ATAAAATTTTAAAGTGATAATGTGTTAACTACTGATTCTAATT GTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoterSet	LP1 promoter Set; contains hAAT-HCR_LP1_Enhancer, hAAT_LP1_promoter, and hAAT-Intron	547	Liver	14	219	CCCTAAATGGGCAAACATTGCAAGCAGCAACAGCAAACACA CAGCCCTCCCTGCCTGCTGACCTTGGAGCTGGGGCAGAGGTCA GAGACCTCTCTGGGCCCATGCCACCTCCAACATCCAACGACCCC TTGGAATTTTTCGGTGGAGAGGAGCAGAGGTTGCTCTGGCGTG GTTTAGGTAGTGTGAGAGGGGAATGACTCCTTTCGGTAAGTGC AGTGGAAGCTGTACTGCCCAGGCAAAGCGTCCGGGCAGCGT AGGCGGGGCGACTCAGATCCCAGCCAGTGGACTTAGCCCTGTTT GCTCCTCCGATAACTGGGTGACCTTGGTTAATATTCACCAGCA GCCTCCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGAGA GGACAGGGCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCT GGGACAGTGAATCCGGACTCTAAGGTAATATAAAATTTTAAAG TGATAATGTGTTAACTACTGATTCTAATTGTTTCTCTCTTTAG ATTCCAACCTTTGGAAGTGA
promoterSet	Synthetic CRM8 TBG promoter set with 5 CpGs; contains 2 copies of HS-CRM8_SERP_Enhancer, TBG promoter, and MVM intron	709	Liver	5	220	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTACCCCAAGTT ATCGGAGGAGCAAACAGGGGGCTAAGTCCACATACGGGGGAGG CTGCTGGTGAATATTAACCAAGGTACCCCAAGTTATCGGAGGAG CAAACAGGGGGCTAAGTCCACATAGGGCTGGAAGCTACCTTTGA CATCATTTCTCTGCGAATGCATGTATAATTTCTACAGAACCTAT TAGAAAGGATCACCCAGCCTCTGCTTTGTACAACTTTCCCTTAA AAACTGCCAATTCCTAGCTGTTTGGCCCAATAGTGAGAACTTT TTCCTGCTGCCTCTTGGTGCTTTGCCTATGGCCCTATTCTGCCT GCTGAAGACACTCTTGCAGCATGGACTTAAACCCCTCCAGCTC TGACAATCCTCTTTCTTTTGTGTTTACATGAAGGGTCTGGCAGC CAAAGCAATCACTCAAAGTTCAAACCTTATCATTTTGTCTTTGT CCTCTTGGCCTTGGTTTGTACATCAGCTTTGAAAATACCATCCC AGGGTTAATGCTGGGGTTAATTTATACTAAGAGTGCTCTAGTT TTGCAATACAGGACATGCTATAAAATGGAAGATCTCCTGAAG AGGTAAGGGTTTAAAGGATGGTGGTGGTGGGGTATTAAATGT TTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	TBG core promoter (Thyroxine Binding Globulin; Liver Specific)	460	Liver	1	221	GGGCTGGAAGCTACCTTTGACATCATTTCTCTGCAATGCATG TATAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCTCTGC TTTTGTACAACCTTTCCCTTAAAAAACTGCCAATTCCACTGCTTT GGCCCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTGCTTTTG CCTATGGCCCTATTCTGCCTGCTGAAGACACTCTGCCAGCATG GACTTAAACCCCTCCAGCTCTGACAATCCTCTTTCTTTTGTGTT ACATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAAC CTTATCATTTTGTCTTTGTTCTCTTGGCCTTGGTTTGTACATC AGCTTTGAAAATACCATCCAGGGTTAATGCTGGGGTTAATTTA TAACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAAA AATGGAAAGAT
promoterSet	Synthetic CRM8 LP1 promoter set with 18 CpGs; contains 2 copies of HS-CRM8_SERP_Enhancer, hAPO-HCR_LP1_Enhancer, hAAT_LP1_promoter, and hAAT-Intron	699	Liver	18	222	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTACCCCAAGTT ATCGGAGGAGCAAACAGGGGGCTAAGTCCACATACGGGGGAGG CTGCTGGTGAATATTAACCAAGGTACCCCAAGTTATCGGAGGAG CAAACAGGGGGCTAAGTCCACATACCCTAAAATGGGCAAACATTG CAAGCAGCAAACAGCAAACACAGCCCTCCCTGCCTGCTGACC TTGGAGCTGGGGCAGAGGTGAGACCTCTCTGGGCCCATGGCC ACCTCCAACATCCAACCTGACCCCTTGAATTTTTCGGTGGAGAG GAGCAGAGGTTGCTCTGGCGTGGTTAGGTAGTGTGAGAGGGG AATGACTCCTTTCGGTAAGTGCAGTGAAGCTGTACTGCCCCA GGCAAAGCGTCCGGGCAGCGTAGGGCGGCGACTCAGATCCCA GCCAGTGGACTTAGCCCTGTTTGTCTCCTCCGATAACTGGGGTG ACCTTGGTTAATATTCACCAGCAGCCTCCCCGTTGCCCTCTGG ATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCTCA GCTTCAGGCACCACCACTGACCTGGGACAGTGAATCCGGACTCT AAGGTAATATAAAATTTTAAAGTGATAATGTGTTAACTACTG ATTCTAATTGTTTCTCTTTTAGATTCCAACCTTTGGAAGTGA

Table 7: promoters						
promoterSet	Synthetic mic/bik TBG promoter set; contains 2 copies of mic/bik enhancer, TBG core promoter; does not contain an intron	681	Liver	1	223	AGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGGCCCTTGG CAGCATTTACTCTCTGTTTGTCTGGTTAATAATCTCAGGAGC ACAAACATTCCAGATCCAGGTTAATTTTTAAAAAGCAGTCAAAA GTCCAAGTGGCCCTTGGCAGCATTTACTCTCTGTTTGTCTGG TTAATAATCTCAGGAGCACAAACATTCAGATCCTGCTCCAG GGCTGGAAGCTACCTTTGACATCATTTCTCTGCGAATGCATGTA TAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCCTCTGCTT TTGTACAACCTTCCCTTAAAAAACTGCCAATCCACTGCTGTTTG GCCCAATAGTGAGAACTTTTTCTGCTGCCTCTTGGTGCTTTTGC CTATGGCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGG ACTTAAACCCCTCCAGCTCTGACAATCCTCTTTCTCTTTGTTTTA CATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACC TTATCATTTTTTGTCTTTGTTCTCTGGCCTTGGTTTTGTACATCA GCTTTGAAAATACCATCCAGGGTTAATGCTGGGGTTAATTTAT AACTAAGAGTGCTCTAGTTTGCAATACAGGACATGCTATAAAA ATGGAAAGAT
promoterSet	Synthetic human CEF1 promoter set; contains human_CMV _Enhancer and hEF1a core promoter	532	Constitutive	0	224	GTTACATAACTTATGGTAAATGGCCTGCCTGGCTGACTGCCCAA TGACCCCTGCCCAATGATGTCAATAATGATGTATGTTCCCATGTA ATGCCAATAGGGACTTTCCATTGATGTCAATGGGTGGAGTATTT ATGGTAAGTGGCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTATGCCCCCTATTGATGTCAATGATGGTAAATGGCCTGCCT GGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGC AGTACATCTATGTATTAGTCAATTGCTATTACCATTGGGAATCACT AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCTCAGTGGG CAGAGAGACATGGCCACAGTCCCTGAGAAGTTGGGGGGAG GGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGG GTAACTGGGAAAGTATGTGGTGTACTGGCTCCACCTTTTCC CCAGGGTGGGGGAGAACCATATATAAGTGACAGTAGTCTCTGTG AACATTC
promoterSet	Synthetic human CEF1 promoter set; contains murine_CMV _Enhancer, human_CMV _Enhancer, and hEF1a core promoter (In that order)	955	Constitutive	0	225	GAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCA ATAGGGACTTTCCATTGGGTTTTGCCAGTACATAAGGTCAATA GGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATT GAGTCAATAGGGACTTTCCAATGGGTTTTGCCAGTACATAAGG TCAATGGGAGGTAAGCCAATGGGTTTTCCATTACTGACATGT ATCTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCAGTACAT AAGGTCAATAGGGGTGAATCAACAGGAAAGTCCCATTGGAGCC AAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCAG TACAAAAGGTCAATAGGGGGTGAGTCAATGGGTTTTCCATTA TTGGCACATACATAAGGTCAATAGGGGTGGTTACATAACTTATG GTAAATGGCCTGCCTGGCTGACTGCCCAATGACCCCTGCCCAAT GATGTCAATAATGATGTATGTTCCCATGTAATGCCAATAGGGAC TTTCCATTGATGTCAATGGGTGGAGTATTTATGGTAACTGCCAC TTGGCAGTACATCAAGTGTATCATATGCCAAGTATGCCCCCTATT GATGTCAATGATGGTAAATGGCCTGCCTGGCATTATGCCAGTA CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTATGTATT AGTCATTGCTATTACCATGGGAATCACTAGTGGAGAAGAGCAT GCTTGAGGGCTGAGTGCCCTCAGTGGGCAGAGAGCACATGGC CCACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAA CTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAGTGGGAAAGT GATGTGGTGTACTGGCTCCACCTTTTCCAGGGGTGGGGGAGA ACCATATATAAGTGACAGTAGTCTCTGTGAACATTC
promoterSet	Synthetic human CEF1 promoter set; contains human_CMV _Enhancer, murine_CMV _Enhancer, and hEF1a core	955	Constitutive	0	226	GTTACATAACTTATGGTAAATGGCCTGCCTGGCTGACTGCCCAA TGACCCCTGCCCAATGATGTCAATAATGATGTATGTTCCCATGTA ATGCCAATAGGGACTTTCCATTGATGTCAATGGGTGGAGTATTT ATGGTAAGTGGCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTATGCCCCCTATTGATGTCAATGATGGTAAATGGCCTGCCT GGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGC AGTACATCTATGTATTAGTCAATTGCTATTACCATTGGGAGTCAATG GGAAAAACCCATTGGAGCCAAGTACACTGACTCAATAGGGGACTT TCCATTGGGTTTTGCCAGTACATAAGGTCAATAGGGGGTGAGT CAACAGGAAAGTCCCATTGGAGCCAAGTACATTGAGTCAATAG

Table 7: promoters						
	promoter (In that order)					GGACTTTCCAATGGGTTTTGCCAGTACATAAGGTCAATGGGAG GTAAAGCCAATGGGTTTTTCCATTACTGACATGTATACTGAGTCA TTAGGGACTTTCCAATGGGTTTTGCCAGTACATAAGGTCAATA GGGGTGAATCAACAGGAAAGTCCCATTTGGAGCCAAGTACACTG AGTCAATAGGGACTTTCCATTGGGTTTTGCCAGTACAAAAGGT CAATAGGGGGTGAGTCAATGGGTTTTTCCATTATTGGCACATA CATAAGGTCAATAGGGGTGGAATCACTAGTGGAGAAGAGCAT GCTTGAGGGCTGAGTGCCCTCAGTGGGCAGAGAGCACATGGC CCACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAA CTGGTGCTAGAGAAGGTGGGGCTTGGGTAACTGGGAAAGT GATGTGGTGTACTGGCTCCACCTTTTTCCCGAGGGTGGGGGAGA ACCATATATAAGTGCAGTAGTCTCTGTGAACATTC
promoterSet	Constitutive promoter Set containing CMV enhancer, gB-actin_promoter, and CAG-intron	1923	Constitutive	192	227	TCATATTGGCCATTAGCCATATTATTGTTTATATAGCATA AATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCA TAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATG TTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG GTCAATTAGTTCATAGCCATATATGGAGTTCGCGTTACATAACT TACGGTAAATGGCCGCTGGCTGACCGCCCAACGACCCCGCC CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAAC TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC CCCCTATTGACGTCAATGACGGTAAATGGCCGCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGGTGCAGGTGAGCCCCAC GTTCTGTTCACTCTCCCATCTCCCCCCTCCCCACCCCAATT TTGTATTTATTTATTTTTAATTATTTTGTGCAGCGATGGGGGCG GGGGGGGGGGGGGGCGCGGCCAGCGGGGGGGGGGGCGGG GCGAGGGGGGGGGGGGGGGGGCGAGGGCGAGAGGTGCGGCGCA GCCAATCAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAG GCGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GGCGGGAGTCGCTGCGACGCTGCCTTCGCCCCGTGCCCGCTCC GCCGCGGCTCGCGCGCGCGCGCGCGCTGACTGACCGCGTT ACTCCCACAGGTGAGCGGGCGGGGACGGCCCTTCTCCTCGGGG TGTAATTAGCGCTTGGTTAATGACGGCTGTTTCTTTCTGTGG CTCGTGAAAGCCTTGAGGGGCTCCGGGAGGGCCCTTTGTGCG GGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGCGTG GGGAGCGCGCGTGCAGCCGCGCTGCCCGCGGCTGTGAGC GCTGCGGGGCGGGCGCGGGGCTTTGTGCGTCCGAGTGTGCG CGAGGGGAGCGCGCGCGGGGGCGGTGCCCGCGGTGCGGGG GGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCG TGGGGGGGTGAGCAGGGGGGTGTGGGCGCGGCGGTGCGGCTG TAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCC GGCTTCGGGTGCGGGGCTCCGTACGGGGCTGCGCGGGGCT CGCCGTGCCGGGCGGGGGGTGCGGCGAGGTGGGGGTGCCGG GCGGGGCGGGGCCCTCGGGCGGGGAGGGCTCGGGGGAG GGGCGCGCGCGCCCCGAGCGCCGCGGCTGTGAGGCGCG GCGAGCCGAGCCATTGCCTTTATGGTAATCGTGCGAGAGGG CGCAGGGACTTCTTTGTCCCAATCTGTGCGAGCCGAAATCT GGGAGGCGCGCGCCGACCCCTCTAGCGGGCGCGGGGCGAAG CGGTGCGGCGCGCGCAGGAAGGAAATGGGCGGGGAGGGCCTT CGTGCGTGCAGCGCGCGCGCTCCCTTCTCCCTCTCCAGCTCGG GGCTGTCCGCGGGGGGACGGCTGCCTCGGGGGGACGGGGC AGGGCGGGGTTGCGCTTCTGGCGTGTGACCGCGGCTCTAGAG CCTCTGTAACCATGTTTTCAGCTTCTTCTTTCTACAGCTCCT GGGCAACGTGCTGTTATTGTGCTGTCTCATTTGTGACAG AATTCCTCGAAGATCGAAGGGGTTCAAGCTTGGCATTCCGGTA CTGTTGGTAAAGCCA

Table 7: promoters						
promoterSet	hAAT promoter Set; contains APOe-CR hAAT enhancer, hAAT core promoter, and hAAT-intron (Composed of hAAT 5' UTR and modSV40 intron)	1272	Liver	26	228	AGGCTCAGAGGCACACAGGAGTTTCTGGGCTCACCTGCCCTTCCAACCCCTCAGTTCCTCCAGCAGCTGTTGTGTCTGCTCTCTGAAGTCCACACTGAACAACTTCAGCCTACTCATGTCCCTAAATGGGCAACATTGCAAGCAGCAACAGCAACACACAGCCCTCCCTGCGCTGCTGACCTTGGAGCTGGGGCAGAGGTGAGAGACCTCTCTGGGCCCAGTCCACCTCCAACATCCACTCGACCCCTTGGAATTCGGTGGAGAGGAGCAGAGGTTGCTCTGGCGTGGTTAGGTAGTGTGAGAGGGTCCGGGTTCAAAACCACTTGCTGGGTGGGAGTCGTGAGTAAAGTGGCTATGCCCCGACCCGAAGCCTGTTCCCATCTGTACAATGAAATGATAAAGACGCCCATCTGATAGGGT TTTTGTGGCAATAAACATTTGGTTTTTTGTTTTGTTTTGTTTTGTTTTGAGATGGAGGTTTGCTGTGCGCCAGGCTGGAGTGACGTACACAACTCTCATCTCACCACAACCTTCCCTGCTCAGCTCCAAGTAGCTGGGATTACAAGCATGTGCCACCACCTGGCTAA TTTTCTATTTTTAGTAGAGACGGGTTTCTCATGTTGGTCAGCCTCAGCCTCCCAAGTAAGTGGGATTACAGGCCTGTGCCACCACACCCGGCTAATTTTTCTATTTTTGACAGGGACGGGGTTTACCATGT TGGTCAGGCTGGTCTAGAGGTACCGGATCTTGCTACCAAGTGAACAGCCACTAAGGATTCTGCAGTGAGAGCAGAGGGCCAGCTAAGTGCTACTCTCCAGAGACTGTGACTCAGCCACCCCTCCACCTTGACACAGGACGCTGTGGTTTCTGAGCCAGGTACAATGACTCCTTCGGTAAGTGCAAGTGGAGCTGTACACTGCCAGGCAAA GCGTCCGGGCAGCGTAGGGCGGGCGACTCAGATCCAGCCAGTG GACTAGCCCTGTTTGTCTCCGATAACTGGGGTGACCTGGT TAATATTCACCAGCAGCCTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTTCAGG CACCACCACTGACCTGGGACAGTGAATCCGGACTTAAGGTAAA TATAAAATTTTTAAGTGATAATGTGTTAACTACTGATTCTAAT TGTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoterSet	CpG-free CET promoter Set; containing murine_CMV_Enhancer, hEF1a core promoter, and CET synthetic intron	826	Constitutive	0	229	GAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCA ATAGGGACTTTCCATTGGGTTTGGCCAGTACATAAGGTCAATA GGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATT GAGTCAATAGGGACTTTCCAATGGGTTTTGCCAGTACATAAGG TCAATGGGAGGTAAGCCAATGGGTTTTCCATTACTGACATGT ATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCAGTACAT AAGGTCAATAGGGGTGAATCAACAGGAAAGTCCCATTGGAGCC AAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCAG TACAAAAGGTCAATAGGGGGTGAGTCAATGGGTTTTCCATTAT TGGCACATACATAAGGTCAATAGGGGTGACTAGTGGAGAAGA GCATGCTTGAGGGCTGAGTGCCCTCAGTGGGCAGAGAGCACA TGGCCACAGTCCCTGAGAAGTTGGGGGAGGGGTGGGCAATT GAAGTGGTGCCTAGAGAAGGTGGGGCTGGGTAACTGGGAA AGTGATGTGGTGTACTGGCTCCACCTTTTCCAGGGTGGGGG AGAACCATATATAAGTGCAAGTAGTCTGTGAACATTCAAGCTT CTGCCTTCTCCTCTGTGAGTTTGGTAAGTCACTGACTGTCTAT GCCTGGGAAAGGTGGGCAGGAGATGGGGCAGTGACAGGAAA AGTGGCACTATGAACCTGCAGCCCTAGACAATTGTACTAACCT TCTTCTTCTCTCTCTGACAGGTTGGTGTACAGTAGCTTCC
promoterSet	Canonical VandenDriessche promoter set; contains 1 copy of HS-SERP_Enhancer, TTR liver specific promoter, and MVM intron	399	Liver	9	230	CGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAACAGGGGCTAAGTCCACACGCGTGGTACCG TCTGTCTGCACATTCGTAGAGCGAGTGTTCCGATACTCTAATCT CCTAGGCAAGGTTTCAATTTGTGTAGGTTACTTATTCTCCTTT GTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCT TGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAA AGCCCTTCACCAGGAGAAGCCGTACACAGATCCACAAGCTCC TGAAGAGGTAAGGGTTAAGGGATGGTTGGTTGGTGGGGTATT AATGTTTTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTCA GGTG

Table 7: promoters						
promoterSet	Constitutive promoter Set containing CMV enhancer and CMV promoter (no Intron)	654	Constitutive	33	231	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGT CATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTA CGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCCGCCA TTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG GACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC CTTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGC CCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCGA AGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCAAAATGTCGTAACAACCTCCGCCCAT TGACGCAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATAT AAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGC TTATCGAAATTAATACGACTCACTATAGGGAGACCC
promoter	Murine Phosphoglycerate Kinase (PGK) promoter	500	Constitutive	39	232	GGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCTGGAGCATGCGC TTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCCTCT GGCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGCT CCGTTCTTTGGTGGCCCCCTTCGCGCACCTTCTACTCTCCCTA GTCAGGAAGTTCCCCCGCCCCGAGCTCGCGTGTGCAGGAC GTGACAAATGGAAGTAGCACGTCTCACTAGTCTCGTGCAATG GACAGCACCGCTGAGCAATGGAAGCGGGTAGGCCTTTGGGGCA GCGGCCAATAGCAGCTTTGCTCTTCGCTTCTGGGCTCAGAGG CTGGGAAGGGTGGGTCCGGGGCGGGCTCAGGGGCGGGGCTC AGGGGCGGGGCGGGCGCCGAAGGTCCTCCGGAGGCGCGGCA TTCTGCAGCTTCAAAGCGCACGTCTGCCGCGTGTCTCTCTCT CTCTATCTCCGGGCTTTTCG
promoterSet	SV40 + Human albumin Invivogen promoter set; containing SV40 enhancer (Invivogen) and huAlb promoter (Invivogen)	450	Liver	3	233	GGGCTGAAATAACCTCTGAAAGAGGAACCTGGTTAGGTACCTT CTGAGGCTGAAAGAACCAGCTGTGGAATGTGTGCAGTTAGGG TGTGGAAAGTCCCAGGCTCCCAGCAGGCAGAAAGTATGCAAA GCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAGTCCCCA GGCTCCCAGCAGGCAGAAAGTATGCAAGCATGCATCTCAATTA GTCAGCAACCATAGTCCCACTAGTTCAGATGGTAAATATACAC AAGGGATTAGTCAAACAATTTTTGGCAAGAATATTATGAATTT TGTAAATCGGTTGGCAGCCAATGAAATACAAAGATGAGTCTAGTT AATAATCTACAATTATTGGTTAAAGAAGTATATTAGTGCTAATTT CCCTCCGTTTGTCTAGCTTTTCTCTTCTGTCAACCCACACGCT TTGGCACC
promoterSet	CMV enhancer + Human albumin Invivogen promoter set; contains CMV enhancer and huAlb promoter (Invivogen)	594	Liver	22	234	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGT CATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTA CGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCCGCCA TTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG GACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC CTTATTGACGTCAATGACGGTAAATGGCCGCTGGCATTATGC CCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGACTAGTTCAGATGGTAAAT ATACACAAGGGATTAGTCAAACAATTTTTGGCAAGAATATTA TGAATTTTGAATCGTTGGCAGCCAATGAAATACAAAGATGAG TCTAGTTAATAATCTACAATTATTGGTTAAAGAAGTATATTAGTG CTAATTTCCCTCCGTTTGTCTAGCTTTTCTCTTCTGTCAACCCA CACGCTTTGGCACC

Table 7: promoters						
promoter	Human UBC promoter	1210	Constitutive	95	235	GGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCCCTC CTCACGGCGAGCGTGCCACGTGACACGAAGGGCGCAGGAGC GTCTGATCCTTCCGCCCCGACGCTCAGGACAGCGGCCCCGTGC TCATAAGACTCGGCCTTAGAACCCAGTATCAGCAGAAGGACAT TTTAGGACGGGACTTGGGTGACTCTAGGGCACTGGTTTTCTTTC CAGAGAGCGGAACAGGCGAGGAAAAGTAGTCCCTTCTCGGCGA TTCTGCGGAGGGATCTCCGTGGGGCGGTGAACGCCGATGATTA TATAAGGACGCGCCGGGTGTGGCACAGCTAGTCCGTGCGAGC CGGGATTTGGGTGCGGTTCTTGTGTTGGATCGCTGTGATCGT CACTTGGTGAGTAGCGGGCTGCTGGGCTGGCCGGGGCTTTCGT GGCCGCCGGGCCGCTCGGTGGGACGGAAGCGTGTGGAGAGAC CGCAAGGGCTGTAGTCTGGGTCCGCGAGCAAGGTTGCCCTGA ACTGGGGGTTGGGGGAGCGCAGCAAAATGGCGGCTGTTCCC GAGTCTTGAATGGAAGACGCTGTGAGGCGGGCTGTGAGGTCG TTGAAACAAGGTGGGGGCGATGGTGGGCGGCAAGAACCCAAG GTCTTGAGGCCTTCGCTAATGCGGGAAAGCTCTTATTCGGGTGA GATGGGCTGGGGCACCCTCTGGGGACCCTGACGTGAAGTTTGT CACTGACTGGAGAACTCGGTTTGTCTGTGTCGGGGGGCGGC AGTTATGCGGTGCCGTTGGGCACTGCACCCGTACCTTTGGGAGC GCGCGCCCTCGTGTGTCGTGACGTCACCCGTTCTGTTGCTTAT AATGCAGGGTGGGGCCACCTGCCGGTAGGTGTGCGGTAGGCTT TTCTCCGTGCGAGGACGCGAGGGTTCGGGCCTAGGGTAGGCTCTC CTGAATCGACAGGCGCCGACCTCTGGTGAGGGGAGGGATAA GTGAGGCGTCAGTTTCTTGGTGGTTTTATGTACCTATCTTCTT AAGTAGCTGAAGCTCCGGTTTTGAACTATGCGCTCGGGGTTGGC GAGTGTGTTTTGTGAAGTTTTTAGGCACCTTTGAAATGTAATC ATTTGGGTCAATATGTAATTTTCAGTGTTAGACTAGTAAATTGTC CGCTAAATTCTGGCCGTTTTTGGCTTTTTTGTAGAC
promoter	Endogenous hGFAP promoter (5' 3kb region)	3000	Muller Cell	44	236	TTAAGGGTTGAGTGTGAGGAAAGGTCTGAGGGTTGAGAAGGG GTGGAGGATGCACCTGGGCCTATGACAGGGGTCCACGGAGGT GGCTGATGGCAAAAGCTGGGGGACTCCAAGTCTGATGCTGAA ACAAGCTTGTGTCTCACATACACAGGGACAGTTCAGTGAGCTT AATGACAGGCACCTCCTGCTCATCATCTTTCTCTCTAGGACA GCTTGGCCCTTATTTAACTAGACTTCCCTTGAACCAAAAGGGAA GGGTACATGCTGTGACTTGTGGGAGCCTGGAAGGCGGGCC ACTCCTAGCCACAGAGATGAGACAGAGTTCAGACAAGAGCTTAT CCCCAGTCTTCTTTCTATTTTGTATTTTATTTTATTTTATTT TATTGAGACAGAGTCTGTGACCCAGGCTGGGGTGCAGTGAT GCGACATTGGCTTACTGCACTCCACCTCCTGGGCTCAGGTGA TCCTCCACCTCAGCCTCCCAATAGCTGGGATCACAGTAGTGC ACCACCATACCTGGCTAATTTTTTTGTATTTTTGTACAGACAAA ATTTACCAACATTGCCAGGCTGGTCTCGAACTCCTGGACTCAA GCGATCCGCCACCTCAGCCTCCCAAGTGTGCGATTACAGGC ATGAGCCACTATGCCAGCCTTGTCTTCTTTAAAGCCTCCTGT CCTTCCCAGGTCCCAGTTCATAGCAGGATCAAAGGTCACTGG GCGCTACCCCCGTCTCAAGATGCTTTTCTATGTCACTGCTTA CGCCAGGTGAGTGTGACTAGAGCCTAAGGAGCTCCCACCTCC CTCTCTGTGCTGGGACTCACAGAGGGAGACCTCAGGAGGCGAT CTGTCCATCACATGTCCAAATGCAGAGCATACCCTGGGCTGGGC GCAGTGGCGCACAAGTGAATCCAGCACTTTGGGAGGCTGAT GTGGAAGGATCACTTGAGCCAGAAGTTCTAGACCAGCCTGGG CAACATGGCAAGACCCTATCTACAAAAAAGTTAAAAATCA GCCACGTGTGGTGACACACACCTGTAGTCCCAGCTATTCAAGGAG GCTGAGGTGAGGGGATCACTTAAGGCTGGGAGGTTGAGGCTG CAGTGAGTCGTGGTTGCGCCTGCACTCCAGCCTGGGCAACA GTGAGACCCTGTCTCAAAGACAAAAAAGGTTAAAAAATCA AAGAACATATCCTGTTGGTGTGAGTAGGGGACGCTGCTCTGACAG AGGCTCGGGGGCCTGAGCTGGCTCTGTGAGCTGGGGAGGAGG CAGACAGCCAGGCCTTGTCTGCAAGCAGACCTGGCAGCATTGG GCTGGCCGCCCCCAGGGCCTCTTCTCATGCCAGTGAATGAC TCACCTTGGCACAGACACAATGTTCCGGGTGGGCACAGTGCCT

Table 7: promoters

						GCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAG CCCATTGAGCAGGGGGCTTGCATTGCACCCAGCCTGACAGCCT GGCATCTTGGGATAAAAGCAGCACAGCCCCATAGGGGCTGCC TTGCTGTGTGGCGCCACCGGCGGTGGAGAACAAGGCTCTATT AGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCAGGCATG GACAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCC CAGAAGTCCAAGGACACAAATGGGTGAGGGGACTGGGCAGGG TTCTGACCTGTGGGACCAGAGTGGAGGGCGTAGATGGACCTG AAGTCTCCAGGGACAACAGGGGCCAGGTCTCAGGCTCCTAGTT GGGCCAGTGGCTCCAGCGTTTCCAAACCCATCCATCCCAGAG GTTCTTCCCATCTCTCAGGCTGATGTGTGGGAACTCGAGGAAA TAAATCTCAGTGGGAGACGGAGGGGTGGCCAGGGAACGGG GCGCTGCAGGAATAAAGACGAGCCAGCACAGCCAGCTCATGTG TAACGGCTTTGTGGAGCTGTCAAGGCCTGGTCTCTGGGAGAGA GGCACAGGGAGGCCAGACAAGGAAGGGGTGACCTGGAGGGA CAGATCCAGGGGCTAAAGTCTGATAAGGCAAGAGAGTGCCGG CCCCTCTTGCCCTATCAGGACCTCCACTGCCACATAGAGGCCAT GATTGACCTTAGACAAAGGGCTGGTGCCAATCCCAGCCCCCA GCCCCAGAACTCCAGGGAATGAATGGGCAGAGAGCAGGAATGT GGGACATCTGTGTTCAAGGGAAGGACTCAGGAGTCTGCTGGG AATGAGGCCTAGTAGGAAATGAGGTGGCCCTTGAGGGTACAGA ACAGGTTCACTTCTGCCAAATCCCAGCACCTTGACGGCACTTA CAGCTGAGTGAGATAATGCCTGGGTTATGAAATCAAAAAGTTG GAAAGCAGGTGAGAGTCTCTGGTACAGCCCTCCTTCCCTTTT TTTTTTTTTTTTGTGAGACAAGGTCTCTCTGTTGCCAGGCT GGAGTGGCGCAACACAGCTCACTGCAGCCTCAACCTACTGGG CTAAGCAATCTCCAGCCTCAGCCTCCAAAGTGCTGGGATTA CAAGCATGAGCCACCCCACTCAGCCCTTCTCTCTTTTAATTG ATGCATAATAATTGTAAGTATTCATCATGGTCCAACCAACCTTT CTTGACCCACCTTCTAGAGAGAGGGTCTCTTGCTTCAGCGGT CAGGGCCCCAGACCATGGTCTGGCTCCAGGTACCACTGCCTC ATGCAGGAGTTGGCGTGCCAGGAAGCTCTGCCTCTGGGCACA GTGACCTCAGTGGGGTGGGGGAGCTCTCCCATAGCTGGGCT GCGGCCCAACCCACCCCTCAGGCTATGCCAGGGGGTGTGCC AGGGGCACCCGGGCATCGCCAGTCTAGCCCACTCCTCATAAAG CCTCGCATCCAGGAGCGAGCAGAGCCAGAGCAGG
promoter	Endogenous hRLBP1 promoter (5' 3kb region)	3000	Muller Cell	32	237	ACGATTTCCCTTCACTCTTATTACCTGGTGGTGGTGGGG GGGGGGGGGTGTCTCTCAGCAACCCACCCGGGATCTTGAG GAGAAAGAGGGCAGAGAAAAGAGGGAATGGGACTGGCCAG ATCCAGCCCCACAGCCGGGCTTCCACATGGCCGAGCAGGA CCAGAGCAGGAGCACAAAGGAGGGCTTTGATGCGCCTCCAG CCAGGCCAGGCCTCTCCCTCTCCCTTTCTCTGGGTCTCT TTGCCCCACTGAGGGCCTCTGTGAGCCGATTTAACGAAACT GTGGGCGGTGAGAAGTTCCTATGACACACTAATCCCAACCTGC TGACCGGACCACGCTCCAGCGGAGGGAACCTCTAGAGCTCCA GGACATTCAGGTACCAAGGTAGCCCAAGGAGGAGCTGCCGACC TGGCAGGTAAGTCAATACCTGGGGCTTGCTGGGCCAGGGAGC CCAGGACTGGGGTGAGGACTCAGGGGAGCAGGGAGACCACTG CCAAGATGCCTGTAAACTGAAACCACTGGCCATTCTCCAGG TTGAGCCAGACCAATTTGATGGCAGATTTAGCAAATAAAAATAC AGGACACCCAGTTAAATGTGAATTTAGATGAACAGCAAATACT TTTTAGTATTAAGAAAGTTACATTTAGGCTCACGCTGTAATC CCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACCTGAGGTCA GGAGTTCGAGACCAGCCTGGCCAACATGGTGAACCCCATCTCC ACTAAAAATACCAAAAATTAGCCAGGCGTGCTGGTGGGCACCT GTAGTTCAGCTACTCAGGAGGCTAAGGCAGGAGAATTGCTTG AACCTGGGAGGCAGAGGTTGCAGTGAGCTGAGATCGACCACTT GCACTCTAGCCTGGGCGACAAGAACAACAACTCCATCTCAAAAA AAAAAAAAAAAAAAGTTACATTTAACTGGGCATTCTGTATTT AATTGGTAATCTGAGATGGCAGGGAACAGCATCAGCATGGTGT GAGGGATAGGCATTTTTTCATTGTGTACAGCTGTAAATCAGTA

Table 7: promoters

						TTTTTAAACTCAAAGTTAATGGCTTGGGCATATTTAGAAAAGA GTTGCCGCACGGACTTGAACCTGTATTCTAAATCTAGGATCT TGTTCGTATGGTCTGCACAACTGGCTGGGGGTGTCCAGCCTCTG TCCCTCTTGCCTGGGCTCCCCAGGGCAGTTCTGTACGCTCTCCA TTTCATTCTGTTCAGCAAAACCAACTGATAGCACAGCAGCA TTTCAGCCTGTCTACCTCTGTGCCACATACCTGGATGTCTACCA GCCAGAAAGGTGGCTTAGATTTGGTTCCTGTGGGTGGATTATG GCCCCAGAACTTCCCTGTGCTTGTGGGGGTGTGGAGTGGAA AGAGCAGGAAATGGGGGACCCTCCGATACTCTATGGGGGTCT CCAAGTCTCTTTGTGCAAGTTAGGGTAATAATCAATATGGAGCT AAGAAAGAGAAGGGGAACTATGCTTTAGAACAGGACACTGTGC CAGGAGCATTGCAGAAATTATATGGTTTTACGACAGTTCTTTTT GGTAGGTACTGTTATTATCCTCAGTTTGAGATGAGGAAACTGA GACCCAGAAAGGTTAAATAACTTGCTAGGGTCACACAAGTCATA ACTGACAAAGCCTGATTCAAACCCAGGTCTCCCTAACCTTTAAG GTTTCTATGACGCCAGCTCTCCTAGGGAGTTGTCTTCAGATGTC TTGGCTCTAGGTGTCAAAAAAGACTTGGTGTGAGGCAGGCAT AGGTTCAAGTCCCACTCTGTCACTTACCAACTGTGACTAGGTG ATTGAACTGACCATGGAACCTGGTCACATGCAGGAGCAGGATG GTGAAGGGTTCTTGAAGGCACTTAGGCAGGACATTTAGGCAGG AGAGAAAACCTGGAACAGAGAGCTGTCTCAAAAAATACCCA CTGGGGAAGCAGGTTGTCTGTGGGCCATGAATGGGACCTGTT CTGGTAACCAAGCATTGCTTATGTGTCCATTACATTTATAACAC TTCCATCCTACTTTACAGGGAACAACCAAGACTGGGGTTAAATC TCACAGCCTGCAAGTGAAGAGAAGAACTTGAACCCAGGTCCA ACTTTGCGCCACAGCAGGCTGCCTCTTGGTCTGACAGGAAGT CACAACTGGGTCTGAGTACTGATCCCTGGCTATTTTTGGCTGT GTTACCTTGGACAAGTCACTTATTCTCTCCGTTTCTCTCTATG TAAATGGAAATAATAATGTTGACCTGGGTCTGAGAGAGTGG ATTTGAAAGTACTTAGTGCATCACAAGCACAGAACACACTTCC AGTCTCGTGATTATGTACTTATGTAAGTGGTCATCACCCATCTTG AGAATGAATGCATTGGGGAAGGGCCATCCACTAGGCTGCGAA GTTTCTGAGGGACTCCTTCGGGCTGGAGAAGGATGGCCACAGG AGGGAGGAGAGATTGCCTTATCCTGCAGTGATCATGTCAATGAG AACAGAGCCAGATTCTTTTTTCTGGCAGGGCCAACCTGTTTAA ACATCTAAGGACTGAGCTATTTGTGTCTGTGCCCTTTGTCCAAGC AGTGTTCCTCAAAGTGTAGCCCAAGAACCATCTCCCTCAGAGCC ACCAGGAAGTGCTTAAATTGCAAGTTCCTAGGCCACAGCCTGC ACCTGCAGAGTCAGAATCATGGAGGTTGGGACCCAGGCACCTG CGTTTCTAACAAATGCCTCGGGTGATTCTGATGCAATTGAAAGT TTGAGATCCACAGTTCTGAGACAATAACAGAATGGTTTTCTAA CCCCTGCAGCCCTGACTTCTATCCTAGGGAAGGGGCCGGCTGG AGAGGCCAGGACAGAGAAAGCAGATCCCTCTTTTTCCAAGGA CTCTGTGCTTCCATAGGCAAC
promoter	Murine RPE65 promoter	718	RPE Cells	2	238	GAACAAAAGCAATGGTGAAGACAGTGATGGACAACAGGCAAG CAGTGGTGATAAGCAAAAACATGTAGTGTTTCTCTTTAATAAG TTCTCAGCTAAAGTTCTCAGCCTTGTTGAAAGGACCTGGATACT GAACTGTGCCGAAGAAGGATAGCAGGGTTAAACATGCAAGA CAGCACCTCATATACCTCTAATGTTGTTAACAATAGCTAACTTTT ATCAACAGTGTCCTGTCAACCATGACAGTTACAACATAATGATA ATGACTGTACTTTCTAACCAGGTCTAGATCACTTATAATAAAT ATATCTTTTAGTAATTGAGTAAATGAATTACAGTGAGGATAACA GCAAAGAAATGGTGACAGATGTTTACACCAAGAAAGTATGAT GACTGAGGTGAGTCTCAGGACTGCATGGCAGGCCCACATGGCTC TTTTTATCCAACCTCACTACTCCCTCTCCCTTGAAGGATCCAAGT CTGGAAAATAGCCAAAACACTGTTATGTAAACACCAAGTCCAAA TAATGTGCAAGCATCTAAAGTATTGAAAGCCACTTTTGTACCTT CCATCAGCTGAGGGGTGGAGAGGGTCCCAGAGCCGAGGCTC CTCCAATAAGGATTAGATTGCATACAAAAAGCCCTGGCTAAGA ACTTGCTTCTCATCTACAGCTGGTACCAGAAGTCTCTAATCT TCACTGGAAGAAA

Table 7: promoters						
promoter	Rat EF-1a promoter	1313	Constitutive	102	239	GGAGCCGAGAGTAATTCATACAAAAGGAGGGATCGCCTTCGCA AGGGGAGAGCCCAGGGACCGTCCCTAAATTCACAGACCCAA ATCCCTGTAGCCGCCACGACAGCGGAGGAGCATGCGCCCA GGGCTGAGCGCGGGTAGATCAGAGCACAAAGCTCACAGTCCC CGGCGGTGGGGGAGGGGCGCGCTGAGCGGGGGCCAGGGAG CTGGCGCGGGGCAAACTGGGAAAGTGGTGTCTGTCTGGCTC CGCCCTCTCCCGAGGGTGGGGGAGAACGGTATATAAGTGCGG TAGTCGCCTTGGACGTTCTTTTCGCAACGGGTTGCCGTAGAA CGCAGGTGAGTGGCGGGTGTGGCTTCCGCGGGCCCCGGAGCTG GAGCCCTGCTCTGAGCGGGCCGGGCTGATATGCGAGTGTCTG CGCAGGGTTTAGCTGTGAGCATTCCCACTTCGAGTGGCGGGCG GTGCGGGGGTGAGAGTGGCAGGCCTAGCGGCAACCCCGTAGC CTGCCTCGTGTCCGGCTTGAGGCCTAGCGTGGTGTCCCGGCC GCGTGCCACTCCGGCCGCACTATGCGTTTTTGTCTTGTGCCCC TCGATTGCCTCCAGCAGCATGGGCTAACAAAGGGAGGGTGTG GGGCTCACTCTTAAGGAGCCCATGAAGCTTACGTTGGATAGGA ATGGAAGGGCAGGAGGGGCGACTGGGGCCCGCCCGCTTCGG AGCACATGTCCGACGCCACCTGGATGGGCGGAGGCTGTGGCT TTCCGAAGCAATCGGGCGTGAGTTTAGCCTACCTGGGCCATGTG GCCCTAGCACTGGGCACGGTCTGGCCTGGCGGTGCCGCTTCCC TTGCCTCCCAACAAGGGTGAGGCCGTCCCGCCCGGCACCAAGTTG CTTGCGCGGAAAGATGGCCGCTCCCGGGGCCCTGTTGCAAGGA GCTCAAAATGGAGGACGCGGCAGCCCGTGAGCGGGCGGGT GAGTCACCCACACAAAGGAAGAGGGCCTTGCCCTCGCGGCC GCTGCTTCTGTGACCCGTGGTCTATCGGCCGATAGTCACCTC GGGCTTCTTGTGAGACCGCTCGTCGCGCGGGGGGAGGGGAT CTAATGGCGTTGGAGTTTGTTCACATTTGTTGGGTGGAGACTAG TCAGGCCAGCCTGGCGCTGGAAGTCATTCTTGAATTTGCCCT TTGAGTTTGGAGCGAGGCTAATTCTCAAGCCTTAGCGGTTCA AAGGTATTTTCTAAACCCGTTTCCAGGTGTGTGAAAGCCACCG CTAATTCAAAGCAA
promoterSet	Human EF-1a promoter Set composed of SV40_Enhancer_Oz and human_FullLength_EF1a promoter	1420	Constitutive	95	240	GGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTC TGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGT GTGGAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAAG CATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAGTCCCCAG GCTCCCCAGCAGGCAGAAGTATGCAAGCATGCATCTCAATTAG TCAGCAACCATAGTCCCACTAGTGGCTCCGGTGCCCGTCAGTGG GCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAG GGGTGCGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGG TAACTGGGAAAGTGATGTCTGTACTGGCTCCGCTTTTTCCC GAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTG AACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAG TGCCGTGTGTGTTCCCGCGGGCCTGGCCTCTTACGGGTTATG GCCCTTGCCTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGA TTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTC GAGGCCTTGCCTTAAGGAGCCCTTCGCCTCGTGCTTGAGTTG AGGCCTGGCCTGGGCGCTGGGGCCGCGCGTGCGAATCTGGTG GCACCTTCGCGCTGTCTCGCTGCTTCGATAAGTCTCTAGCCAT TTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTCTGGCAAGA TAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGG TTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCCTCCAGCG CACATGTTGCGCGAGGCGGGGCTGCGAGCGCGGCCACCGAG AATCGGACGGGGTAGTCTCAAGCTGGCCGCGCTGCTCTGGTG CCTGGTCTCGCGCCCGGTGTATCGCCCCGCCCTGGGCGGCAAG GCTGGCCCCGGTCCGCCACCAAGTTGCGTGAGCGGAAAGATGGCCG CTTCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGC GCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAA GGGCCTTTCGCTCCTCAGCCGTGCTTCATGTGACTCCACGGAG TACCGGGCGCGCTCCAGGCACCTCGATTAGTTCTCGAGCTTTTG GAGTACGTGCTTTAGGTTGGGGGGAGGGGTTTTATGCGATG GAGTTTCCCACTGAGTGGGTGGAGACTGAAGTTAGGCCAG

Table 7: promoters

						CTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCCCTTTTGGATT TGGATCTTGGTTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTT TTTTCTTCCATTTCAAGTGTCGTGA
promoterSet	Rat EF-1a promoter Set composed of CMV_Enhancer and rat_FullLength_EF1a promoter	1831	Constitutive	124	241	TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATA AATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCA TAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATG TTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG GTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACT TACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCC CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAAC TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC CCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTCTCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGGGGAGCCGAGAGTAATT CATACAAAAGGAGGGATCGCCTTCGCAAGGGGAGAGCCAGG GACCGTCCCTAAATTCTCACAGACCCAAATCCCTGTAGCCGCCCC ACGACAGCGCGAGGAGCATGCGCCAGGGCTGAGCGCGGGTA GATCAGAGCACACAAGCTCACAGTCCCGCGGTGGGGGGAG GGGCGCGCTGAGCGGGGCCAGGGAGCTGGCGCGGGGCAAA CTGGGAAAGTGTTGTCGTGTGCTGGCTCCGCCCTTTCCGAGG GTGGGGGAGAACGGTATATAAGTGCGGTAGTCGCCTTGGACGT TCTTTTTCGCAACGGGTTTCCGCTCAGAACGCAGGTGAGTGGCG GGTGTGGCTTCCGCGGGCCCGGAGCTGGAGCCCTGCTCTGAG CGGGCCGGGTGATATGCGAGTGTGCTCCGACGGGTTTAGCTG TGAGCATTCCCACTTCGAGTGGCGGGCGGTGCGGGGGTGGAGAG TGCGAGGCCTAGCGGCAACCCGTAGCCTCGCTCGTGTCCGGC TTGAGGCCTAGCGTGGTGTCCGCGCGCGTGCCTCCGCGCCG CACTATGCGTTTTTGTCTTGTGCTCGCTCGATTGCCTCCAGCA GCATGGGCTAACAAAGGGAGGGTGTGGGGCTCACTCTTAAGGA GCCCATGAAGCTTACGTTGGATAGGAATGGAAGGGCAGGAGG GGCGACTGGGGCCCGCCGCTTCGGAGCACATGTCGACGCC ACCTGGATGGGGCGAGGCCTGTGGCTTTCCGAAGCAATCGGGC GTAGTTTAGCCTACCTGGGCCATGTGGCCCTAGCACTGGGCAC GGTCTGGCCTGGCGGTGCCGCTTCCCTTGCCTCCCAACAAGGG TGAGGCCGTCCCGCCGGCACCAGTTGCTTGCAGCGAAAGATG GCCGCTCCCGGGGCCCTGTTGCAAGGAGCTCAAATGGAGGAC GCGGCAGCCCGGTGGAGCGGGCGGGTGAAGTCAACACACAAA GGAAGAGGGCCTTGGCCCTCGCCGGCCGCTGCTTCTGTGACCC CGTGGTCTATCGGCCGCATAGTCACCTCGGGCTTCTCTTGAGCA CCGCTCGTCGCGCGGGGGAGGGGATCTAATGGCGTTGGAG TTTGTTCACATTTGGTGGGTGGAGACTAGTCAGGCCAGCCTGGC GCTGGAAGTCATTCTTGAATTTGCCCTTTGAGTTTGGAGCGA GGCTAATTCTCAAGCCTCTAGCGGTTCAAAGGTATTTCTAAAC CCGTTTCCAGGTGTTGTGAAAGCCACCGCTAATTCAAAGCAA
promoter	Endogenous hABC11 promoter (5' 3kb region)	3000	Endogenous (Liver)	21	242	CCAGGCATGGTGGCTCATACCCGTAATCCAGCTACTCAGGAGG CTGAGGCAGGAGAATCACATGAACCAAGAGGTGGAGGTTGCA GTGAGCCAAGATTGAGCCACTGTACTCCAGCCTGGGCAACAGA GCAAGACTTGGTCTCAGAAAAAAAAAAAAAGTATGTCTTGAC TTTAAAAAATTCAATAAACTGACCTGTCTTTTTTAAAAAACAGC CTTTTGGGGTATAATTTACATATCACAGAGTTCACCTATGTAAA GTATTCAATGGTTTTCAATATATTAACAGAGTTGTGCGACCATCA CCATAATCTAACTTTAGAACATTTTCTCATCCCCAAAAGAAACC TTATATCTGTTACCAAGTCACTCCTCATTCCCTCCCAACCCCTACCC CTACCCAGCATTAGGCAACCACTTATTTATTTCTGTCCCTATAG ATTTGCCATCTTGGACATTTTCATGAAATGGAATCATACAGTAT GTGGTCTTTTGGAGCCGTCTTCTTTCAGTAGCATGATTTTGGAG TTCATCTGTGTAGCATGTATCAGTACTTCAATCTACATACATTTAC CGTAATTAAGTGAACCGTTTGGACTATTTTCAATAATATTCATTAT GTTTTCTGTTGTTATGCTTTTTTGTAGTTCTTATGTTTTTTTAAAC TTTTGTTGGATTGATGACATTTTCTACATACTAGTTTTTAAATCCT

Table 7: promoters

					TTGCTTATTTAGAACTATAGATTTTACTGGTACTTTTTTCATTGCT TTTTCTTAAAAATTTTCAGATATTGGTTGAACCTTGTTCAGATATTA GTTGAACCTTGTAAATTAATAAATGGTTAAATATTGGCAATTTCTT TTGGTTAATCAAACATATATTTAATTATAGTTGTATAAATATGT ATTTAATTATAATTATAAAACAATGCTCCTCAGATTGTCATAACAA TGAACCTAACATACTTTATCTGCATATCGAACACCTTATCTTGTG TCAAGTTACACTCATATCTACATACTGTGTAGAGTTTAAATATG TTCTTTTGAATATAAAAGGTTTACTTGGTATCAATATTTGATT GGCCGCTCCTGACATATTTTGTAACTCTTGTGCTCACCTTGTTC TCTCTTCATGGCTCCCTTCTGGATACTCTTCTGGCTAAGGCAC ATCCTCTAGTTGTTGTTTATGCAAGGTCTGTAAGTGTAACCTC TGACTTTGAATGTCTGTAAAGATGCTGAATAATTTTGGCTCAG TGAAAAATTCTAAGTTAAAGATTACTTTTTTCTCATCACTTTGA AGACATTACGCCACTGTTTCTAGCCTCTATTGCTGATGAGAAAA CTTCTGTCAGTCTGTTCTTATATTTGAATATGCATTTCCCTTTC ACAGTGTTTAGGATGGATTTTGTATTCTTGATGCTTACTACA GTTTGATTCTTGAACAACACAGGTTGCAACTGTGGAGGTCCACT TGTATGGGGATTGTTTTCAACCAATCTCAGATGAAAAATATAGT ATTCTCAGGATGCAAAACCAAGTGGATATGTAGAGCCAATTTTC CTATGCACAAGTTCTGCAAGCCAAGTGTAGGACTTGTGTATACC TGGATTTTGGTATATGCAAATTTTGGTATACATGGGAGTGCTAG AACCAATCTCCTGCATATACTGAGGGACATTTCTATATAATGTAT CTAAGTTTGTACTGATATCTATTCCAATCAATTCTTGGTGTCTACT GTTAATTTGAAGAATCAGGTAATTGCTTCTGGAAAATCTTAGC AATTATCTCTTAATTATTACACTTCTGTCACTTCCACTCTCTGCT TCTGGGATTCCAATTAGGTGAATTTAGAAGATTTTCATAACTCCC CCTTCTCTCTTTTATTGTACATGTGTATATATGTATGTAATA CATATCACGGTCTCCTCCTGTGACCTCCATGGGTCTGCATTTCT CATAAGGAATAGATGCTTCAATGGTGGCCAGCAGTTTCTCAGG GTCTTCTCAGCAGTGCATGGGGCCACATTAGCTCCTCTGGCTCC AAGCGAAGAGATGGTCTTAGCCCCCTGTTTGATTGGGGCACT TACAGTCTCTCGCCAGCTAACTCTCACACTCGTCAGCATCCAG ACGCTGAGGGGAAAAATACCAGTCTTCTGTGCTCTGCTTACTC TTCGGTACTTCTCTGCCATTTCTGGTTCCTGAAGATGTTTATTTT ATTTATTTGAGTCTGACTGTATCTCTTTTAAAAACATGTTATCCA CCATTGCTATATATTTGAAGCAGAGAAAGTTAGTGAAGCATAAAA CTTGCTGCTGAATCGAGTGTCTATATCTGGAATTCTCAGCCTGT ACCCTCTATAAACTAATTTTCCACTGTGAATAAGACTAATCATG ACTCTGTCGACATTTACATTTTATTTAGAAAATGTCTTCTTCTGT TCCTTTGATCCAAGCTTGACTCACCTTACCTTGAGGTTGCATTTA CAAAGGAACACTGAAGGTTACCAACAGTATGTGGGTGTCGTTT ATCAACTACAGTGAAGTCAAGAATATCACCAGTTGGTTTGCCTTTC TCATGGTTTAAATGTTTCTCATTAATAAATAAAGCACAGAT AAGCAGAAAGAATAACCATCCATCCAACACTAGAGGAAAAATTT ATCAATGGTTTGTCTTATCTTCTATAAATTAAGCTATAAAAAAC AACCATCCATGTAACAACTAGAGAAAACCTTATCAATGACTGT GGCTTATCTTCTGATAATTAGGCTCTTTCAGGGAGTTATTAAC CGATTTTAAAACTTTGTCTGAGATTGATTAGTAAAGATTATTT TTGAACCAAAATTTGTTCTTTCGTTTGGTACTTTGATTAAAGAAGA AAGAAGAGATAATAATTGCAATGATTCTTTATTTTATTTATAG GGTCGTTGGCTGTGGGTGCAATTACC
--	--	--	--	--	---

Table 7: promoters

promoter	Endogenous hPAH promoter (5' 3kb region)	309 5	Endogenous (Liver)	37	243	TATGGCACAAGCAATCTCTATTTTATCTTAGTGCATAAATAAA TTTTCTCTTTTGGCAGAATAATTTTTTAAAGAACGATTAGTT TTTCTCTCTCAGATAGCAATGATGTGCTTCTCTCAACCTAGAT TTAGGGCATTATGTGAGATAGGATTAAAAATCCATTTTGT ACAACCACTATGGAGAACAGTTTGGCAGTTCCCAAAAACTAA AAATAGAGCTACTATATGATCTAGTGATCCCACTGCTGGGTATA TACCTATAAGAAAGGAAATCAGTATATCAAAGAGATGTCTGTTC TTTTATGTTTGTGCAGCACTGTTCAATAGCCAAGATCTGGAA GCAACCCCAAGTCTCCATCAACATGGGTTTTAAAAAATGTGGTA CTTTAATACCAATGGAGTACTATTAGCAATAAAAAAGAATGA GATCCTGTTATTTGCAATAACATGGACAGAACTGGAGGTCATTA TGTCAAATGAAATAAGCCAGGCACAGAAAGCCAAACATCACAT ATTCTCACTCATATGTGGGGTCTAAAAATCAAAACAATCTGATTC ATGGAGCTAGAGAGTAGAGAGCTAATTACCAGAGGTGGGGAA GGGTAGTAGGGGCTGGAGGGGAGGTGGAGATGGTTAATAGG TACAAAAAATATAGAAAGATGAATAAGACCTAGTATTTGATA GTACAACAGGGCGAATATAGTCAAAATAATTTAATTATACATTT AAAAATACCTGAAAGAGTATAATTGGCTTGTTCGAACACAAAA GATAAATGCTTGAGGGGATGGATGCCCATTTTCAATGATGTGA TTATTACACATTGCATGCCTGTATCAAAATATTGCACATACTCCA TGAATGCATACATCTACTATGTTCCCAAAAAATTAACATTAGA AAAAAGAGTTGCATTTTCAGCTGTTATGGGGAGAAGAAAGAAA AGCTATCATTTTGTTCCTAAAAATTATGTTGCCTCATTTCAAA CAGGAAAGCAAAAGTATTTGAGAGCCAGTGCACTGCCTTGGTG TTGGGTGAAACATAGATTGAATTTGGGCCATTTGTTAACTTCC TAGGCCTCAGTTTCTTGCCTATTAAGGGAGTGCATAGTTTAT GGGATTGTTAAGAGGAAGAAGTGAACCATGCACGTGGAGAG CGTGGCAGAGTGTCTAAGACAGAGTGTGCATGCAATAAGTAG ATAATATTCTTGTCTTTCTTATTGCATGCCTGTAATATTTTGG AGTTGTCACATTCATTGCCCTCAAGTAGCATCAAGGGATGAAAT TATGTTTGTAAGAAAATCCTGAGGCTGAGGAATACAACATGTTT TATGTCTACTACACTGAAAAATGCCGGAGTCAGATAAAGAATAC AGATTCTCCTGAGGATGGAAATCAAGATCTTCGCCTTCAATATT AACCAACATTGAGCTTCAACTTACTATGGGAAATATTCATCAGG CCCTAAAGGTTCTTTTGGACAGAAATTGCACTTGTTATATCTG TATTCTAGCAGACAGTAGACAGCCTGGCAGATCATAAAGGCTT AAGGAATCCTAAATATCCCTTAAATTTCTATTTAAAGACAAAA ACAAAACAAAAAACAACAAAAAATGAGGCATGGGCT TGACCAATCAGTGGTAGAACCAAGAGTTAAACCACTTGTGTTG AATCCTAAACCTGAGTTTTATTTACTTATTTATTTATTTGTT TATTTATTTTTCAGATGCTTGGTCAAAGAACAGTGGGAGGAGAG GGATGGGCTTCCAGCAACCTTTATTATTGGCTTATTTCTTACAG CCCATTAATTTCTTGGGAAAATATTAAGCAGGCACTCAAGGCT TGAGGCCCTGAGTTTTACATCTTTCTGAACCTCTGAACCTGC TTTCAGCATCTTTTATACTTTGTTTTACCTCTGGTCAGTAATG CCTCACCTCAGTCTTCTCTAAAAGTGTGGTTAATGGCATCTCC TGACTATTTGAAGACCACTGGCCAAATCCCACCACTCACTCATA GACCATCCCCCTACTTTACTTTCTTCAAAAGACTTAGCCCTACCTA AAGTTATTTATATGTTTATTTTCTGCCACCAGAATGGCAGCATA GCTGGGGAGGCAGAGTCTGTTTGTTCATTGCTGTATTCCCAA GACTAGAACACCACCAAGCACACGGTACAGGTCTCAGTAATTAT TGTCAAATTTATGTGGATTTGCTTTTAAACAATATCTTCCATTTAC TGAGTGTTTATGTGGAAGAACTGTAATAATTTTAAATGCAATTTCT TTATTCCTATTCTTAAACCTTCCAGCAAGGTGGCTCTACCACTC TCTTTCCGAGCTTCAGGAGCAGTTGTGCGAATAGCTGGAGAAC ACCAGGCTGGATTAAACCCAGATCGCTCTTACATTTGCTCTTTA CCTGCTGTGCTCAGCGTTACGTGCCCTCTAGCTGTAGTTTCTG AAGTCAGCGCACAGCAAGGCAGTGTGCTTAGAGGTTAACAGAA GGGAAAACAACAACAACAAAAATCTAAATGAGAATCCTGACTG TTTCAGCTGGGGTAAGGGGGGCGGATTATTCATATAATTGTTA TACCAGACGGTCGAGGCTTAGTCCAATTGCAGAGAACTCGCTT
----------	--	-------	--------------------	----	-----	--

Table 7: promoters

						CCCAGGCTTCTGAGAGTCCCGGAAGTGCCTAAACCTGTCTAATC GACGGGGCTTGGGTGGCCCGTCGCTCCCTGGCTTCTTCCCTTTA CCCAGGGCGGGGAGCGAAGTGGTGCCTCCTGCGTCCCCACAC CCTCCCTCAGCCCCCTCCCTCCGGCCCGTCTGGGCAGGTGACCT GGAGCATCCGGCAGGCTGCCCTGGCCTCCTGCGTCAGGACAAC GCCCACGAGGGGCGTTACTGTGCGGAGATGCACCACGCAAGAG ACACCTTTGTAACCTCTTCTCCTCCCTAGTGCGAGGTTAAAC CTTAGCCCCACGTGCTGTTTGAAACCTGCCTGTACCTGAGGCC CTAAGAACCCAGAGACCTCACTCCCGGGGAGCCAGC
promoter	Murine CD4 Promoter sequence	1807	Muller Cell	34	244	AGCTTGTAGATACTCGGAACAAATGCAATTCTTACGAATACTTT AGTCTATACACAGAAAAAGCTGGCTGAAAAATAAATGATTATT TTTAATATTTTAAACAGTTATTAATTGTGTATGTGGCAGGCTG TGACAGGTAGAGGACAACCTTGCCTAAGGCACCATGTGGGTTC GAAGGATCTAAGTGTCCCATGCTTGGCAGCAAGCACTTATCAC TGGCCATCTTCCAGTCTAGCTGTAGTTTGCAGTATATTTTATA CTGCAGCAGCCACTGGCTTGTGTGGGAGCTAGTGCCTAGACCA AACCAGGATTGCTTCTCTTGAACCCCTCTGGCACTCATTACGTGC TTGATGAATAAATGGATGGACAGGTGGCTGTGTACATTTCTCTC ACTTCTCAGTTTCTTTCAGTAAATCCCAAAATATCATTTTCTTCA GAAATTCTGGCATGATTCAATCCGGGTCTGCCCCTGGCCATGCCT TCTGTGTTTCTCATTAGTAAGAAGTCCACTCAGATTTAGTTCAC ATTAAGAAATAAACAGAGCTTTGATATCCAAATGTCAACTTGCA GGGTATTAGAGAAGATAGGGAATTGCAATTTTACATACGATTTT CCCCATTTCAGCCTTGAGATTCGTCTTGAAGCATATGGCA AATGTGCATCCCTCTTGAATGTACTAAGATGTAAAGGGGAAT TTGAATGTATTAAAGTTTGCAGCAAGAGAATATAAATGTAAAC AAGAAAGAACAGTTAAATGTGTGAGTGGATATGGGGATGGGTA GAATGAGAGACGGGAACCATGTATGTGCGTCGGGATGGATAG GAAATATGATGAACAGATATAGCTGAGGAGGGGTGTGAAAAG GATTGAAAAGTTGTGAGGTGGGCAATACAAGAATTGGTGGG CAGGTGTAGTATGGCTAGATTAGTGCATTGTCAGAAGGAAGAT GGGTGGACAGAGGAATGGATGGGTGGATTGTGAGTCGAGAAG GATTTAAGAAATTGGTAGATATTTTGAAGCATGAATGAAATGT GTTGAGCACCCCTTGGGTTTTCCCGGATCAAAGATCAGATGAGC GGTTTGGACTTCTCTCAGAGGGAAAGAGGAAAGAACTCCCA CAAGTCCCCACTTTTCACTCCCACTTGGCCAGGAAAGCACTC TCCACTAGGATGGATCTCTAGTCTCTCTCTCCCTTCAGCCTC TTTCTTCTCAGTTCCTCCCTAAGATAAGTCCAGCTTCTCAGCT TCCTGGGAAAACAGTCTTTCCTAGCCAGGTTCCCAAGTTTAGT GGGAAAGGAGAACTGGAAGATTTAACTGAGAGGGGCGAGGT CTTAGAACTCAGTCATTCTCCTTGTCCAGGCAGCGCTTCTCATA GGCTGGTAGGCTGGGCCAGGGTAGGAAGCCTGTGGAGTGGCC CTGGAGAACGTGGGGCGGCACGGGGGCTGGGGGGGAGGGG GGCGGCCATTCTTCTGTCCAAGAGAGCAGGGCAGGAGTGCA GGGGAGTAGCGAAAGCAGGCTGGTGTGTCTTAAACTTCCGTT GGCTGCTTAGTCACAGCCCCCTCGCTTGGGTGTGCTTTCGCG CGCTCCCTCCCTCTAGGTCACTCACTCTTCAAAGCCTGGAATA AAAACCAAGCCAACTTCCGAAGCGGTCTCATTGCCAGCAGCC CCCAGCCAGTGACAGGTTCCATTCACTTCCCTTCTCCCTTCTCCC ACGACCTTTTCCAGAGGCGACTAGATCCCTCCGTTTCATCCAGC ACGC

Table 7: promoters

promoter	Endogenous hABC4 promoter (5' 3kb region)	3000	Endogenous (Liver)	91	245	GAAAATTTGTCACAACTAAAGAAAACAAGAAAGAGACAGTAG ATGAAAGAGTGCTCATTAGGTGAAAGGAAAATGATCCAAGAGG GTAGCTTTGAGATGTAGGAAGAAACAAAAGCAAGAAAATGAT AAATGTTTTGATAAAGCTAAATAAGTATCAACTCATAAAGAAAT AATATTTCCAGAAGAGTCATGAATATACAGAGAAAATTAAGTA CATGACAATGGCAATGTAAAAGTTAGGGGTGAATAAAAAGAG ACTTAAGAGTTCTAAAATCATTGCATTGTCCTGGAAGAGGAAAA AGTACAATGATTAGTCAAAGATACATGTCATAATCCCTAGAAAG GAGATCATTATTAAATAGAAAATAAAAGAATACATCTTATAGAA AGGAAATCTAAATGATAATATTAACAGATCTAAAATAAGGCAA AAGTGAGGATAAAAAAGAAAGATGGAACCAATGGGGCAAATA GAAAAAGTAAGATAGCGTGGTAGGGCATTAAATCCAGCCTTAC ATCAATGCATAAGTATCTCAATATTCTACTGTAAGGGGAAAGTA AAGATTTCTTACAGCCTGAGTGTAATGGAGAAATCTAGTTTATC ATAGTGCTTTAAATATTGTAAGTCTTCAACTCTAGTTGATGAAT AAATGATGGAATTCTCAGTGACTGCACTGTTATCAAATAAAT ATAAAAGGAGCTCCTGGAATTGGATGTAATACAGGTAAAGAAG TAAACACAGCCATATAGGCATGGCTTCTTGAGGGACAACCTTG TGAATCGGCTCAGACAGACAGACAGGCAAATACACCTCATTGCC TCATACATGTTATTGCTTTAGTTTTGTTCTGAACCTTCCTACTC CTTCAAGTATCTGCATTTACTTTATCAAATCTCTTTTATTAGAGA CTGAAGAACTGTCATCTCTTATGTGCTAATGAGTTTAATAATG TCCTCCAGTCACCACAAGCCTTCTTCAAACACACAATCCAAC TGCTCCGTCTCAGAGTATCTTGAAATAATGATCTGACCGCCTGT TAGACCAGTGAAGGGAAGGAATTTGGGTTGATTAAGAAGAGA ATCCTCATGGTCATGGTAGACTGATATGGAGAGAAAACATTTTG AGGAAAAATACTCAACTAAATTCATTTCTACTCCAGCATGCAGTT TCAAGTCAAGTTCCACCTTAGCTCCAGGTGGCAGGCAGAGCAG GATGCAGAGGCACAGCACAGTAAGGGGTGAGTGCCGAAGCT GCTGGCTCCTGTTCCAGTCTTTCTTCTTGGCTCGCTGAACCTT TACTATAATAATAGTCACCAATTTATTAGGTGTCTCCTACGTGCAG GACACTTTACACACAGTATCCCTAATCCTAATAACACCCCTATTTT ATAGATCCAATGACTGAGTCAAGAAATTACATAACCTGGCCAGAC AGCTGGTACATGGGAAAGGTGAGATTCACACCAGGGTCCACCC AGCATCTCTACTTATACCATGCTCTGCTTAAAGGTTCTCTGAGAA CTCAGACAAGCCTTGGGCTAACAATTGTGTTAACAGGACATAGC AGGTGCAAGGACCCACTGGTCATCCTGCTACCTGATCAGAAGG AAGGAAAGTTGTATTTGTTGCTCACCTACTATGTTTTAGGCATAG TACTAGGTGCTTTACCTAGTACTTAATCCCTTATCCTCAACTCA TTTATTCCTCGCAATAACCTGATAAGGGAGATGTTTTATCCTCA TTTTACATATAAGGAAACAGGCCTAGAGAAATGAGCACAGTGTG CAAAGTCATAGTTAATAAGATGTGAAGCTCTGAGTTTGAAG TCTCCGTTTCAAAGCCATGAACTTATGGCTCCCCGTTTTAGAC ACTTCCTTTTGGGAAGAGTGTGGAGGAATTAATCAGAAAGAAG AAAGTCATACTCAAATAGGTGGTAGGAGCAGAGACAATCAAT ACAGACAGAAGTCTTAGATGAGAGCAGTGAAGCCAGGGCACTGG ACTGGGACTCAGGAGGCTTCCCCTAGACTCTGGTTCCACCGATG CAGCCTCAGGCAGGACTCACCTCTCTGGGCATCCGTTTCTTCAT ATGTTAAACATACGGGGTTTTAATTAGATGATCGCTGAAGACCC CTCTAGCCCTAAAACCTGTGTCTCTTAAGTGCTAAGAGGGCAC CAACAGCGTTCCTCCTCCCCAAGGAGCATAATGTGATGGTTCCT GCCGGCCCTGGCTGACTCTCGCGTCTTGGAGATAATTGGGTT CAGTGCCACCTGGACCAGAACTGGGGATCGGGAAGCAAGAGG CGAGTCTATTGCTCTCTCTCGGTCTGGGCCGCTGTGATTGTT GGGCGTCCGGAAACTGTCTCCCTATGGGTTAAAAACAAAAC GAGCGCCATGGGGTGTGACAGTCATCTGAGGGGCTTGGGTG GCCCATCAGGCGAGGCTTCTCGGCACCCGAGGCTCCAGCCTGA TCTCGGTCTTATCCTGCGACCGGGCTGGTCTGGCGGGTCCGCA GGGTGGGCGGCGGCCAGCCGGGCGCCCCGGCGGCAAGAGC GGCAGGCTGCGCCCTGGCCGCGCCTAGCCTGGGGAGAGAGC TGGGCGGGCGGCGGGAGCTGCTCTCGCGGGCGCGGCCCTCGC
----------	---	------	--------------------	----	-----	--

Table 7: promoters						
						CCTGGCTGCAACGGTAGGCGTTTCCCGGGCCGGACGCGCGTGG GGGGCGGGGGCGGGGCGGGGGCGAGGCCGCGCGAGCAAA GTCCAGGCCCTCTGCTGCAGCGCCCGCGCTCCAGAGGCCCTG CCAGACACGCGCGAGGTTTCGAGGTGAGAGAGGTCCGGGCGCG TCTGGCCTCGAAGGGAGACCCGGGACGTGGGGCGCGGGGCGG GAGTGGCCGACCTCCACCCAGTGCCCCGGGGCCCGCGACTC GTGCGCCGGGCGCGGAGAGGGTGTACTTGGTTCTGAGGCTG TGGTTTCTCTCAGGCTGAG
promoter	Human RPE65 Promoter (-742 : +15) of NG_008472.1	757	RPE Cells	1	246	TGAATTGATGCTGTATACTCTCAGAGTGCCAAACATATACCAAT GGACAAGAAGGTGAGGCAGAGAGCAGACAGGCATTAGTGACA AGCAAAGATATGCAGAATTTCTTCTCAGCAAATCAAAAGTCCT CAACCTGGTTGGAAGAATATTGGCACTGAATGGTATCAATAAG GTTGCTAGAGAGGGTTAGAGGTGCACAATGTGCTTCCATAACAT TTTATACTTCTCCAATCTTAGCACTAATCAAACATGGTTGAATAC TTTGTTTACTATAACTCTTACAGAGTTATAAGATCTGTGAAGACA GGGACAGGGACAATACCCATCTGTCTGGTTCATAGGTGGTAT GTAATAGATATTTTTAAAAATAAGTGAGTTAATGAATGAGGGTG AGAATGAAGGCACAGAGGTATTAGGGGGAGGTGGGCCCCAGA GAATGGTGCCAAGGTCCAGTGGGGTGAAGTGGGATCAGCTCAGG CCTGACGCTGGCCACTCCACCTAGCTCCTTTCTTTCTAATCTGTT CTCATTCTCCTTGGGAAGGATTGAGGTCTCTGAAAAACAGCCAA ACAACTGTTATGGGAACAGCAAGCCCAATAAAGCCAAGCATC AGGGGGATCTGAGAGCTGAAAGCAACTTCTGTTCCCCCTCCCTC AGCTGAAGGGGTGGGAAGGGCTCCCAAGCCATAACTCGTTT TAAGGGATTAGAAGGCATAAAAAGGCCCTGGCTGAGAACTT CCTTCTTCATTCTG
promoter	tMCK Promoter. Triplet repeat of 2R5S enhancer sequence followed by [-80 : +7] of murine MCK promoter	720	Muscle	16	247	CCACTACGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTGG GGACACCCGAGATGCTGTTATAATTAACCCCAACACCTGCTG CCCCCCCCCCAACACCTGCTGCCTGAGCCTGAGCGGTTACCC CACCCCGGTGCTGGGTCTTAGGCTCTGTACACCATGGAGGAGA AGCTCGCTCTAAAAATAACCTGTCCCTGGTGGGCCCACTACGG GTCTAGGCTGCCATGTAAGGAGGCAAGGCCTGGGGACACCCG AGATGCCTGGTTATAATTAACCCCAACACCTGCTGCCCCCCCCC CCAACACCTGCTGCCTGAGCCTGAGCGGTTACCCACCCCGGT GCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAAGCTCGCTC TAAAAATAACCTGTCCCTGGTGGGCCACTACGGGTCTAGGCTG CCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGG TTATAATTAACCCCAACACCTGCTGCCCCCCCCCCCCAACACCT GCTGCCTGAGCCTGAGCGGTTACCCACCCCGGTGCCTGGGTCT TAGGCTCTGTACACCATGGAGGAGAAGCTCGCTCTAAAAATAAC CCTGTCCCTGGTGGGCCCTCCCTGGGGACAGCCCTCCTGGCT AGTCACACCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGG GCTGCCCCGGGTAC

Table 7: promoters						
promoter	MHCK7 Promoter	772	Muscle	16	248	ACCCTTCAGATTAATAAATACTGAGGTAAGGGCTGGGTAGGG GAGGTGGTGTGAGACGCTCCTGTCTCTCTATCTGCCATCG GCCCTTTGGGGAGGAGGAATGTGCCAAGGACTAAAAAAGGC CATGGAGCCAGAGGGGCGAGGGCAACAGACCTTTTATGGGCA AACCTTGGGGCCCTGCTGTCTAGCATGCCCACTACGGGTCTAG GCTGCCCATGTAAGGAGGCAAGGCTGGGGACACCGAGATGC CTGGTTATAATTAACCCAGACATGTGGCTGCCCCCCCCCCCCAA CACCTGCTGCCTCTAAAAATAACCTGTCCCTGGTGGATCCCCTG CATGCGAAGATCTTCGAACAAGGCTGTGGGGGACTGAGGGCAG GCTGTAACAGGCTTGGGGGCCAGGGCTTATACGTGCCTGGGAC TCCCAAAGTATTACTGTTCCATGTTCCCGGCGAAGGGCCAGCTG TCCCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCACTGAG CAAGTCAGCCCTTGGGGCAGCCATACAAGGCATGGGGCTGG GCAAGCTGCACGCTGGGTCCGGGGTGGGCACGGTGCCCGGG CAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCTCCTGGG GACAGCCCCCTCTGGCTAGTCACACCTGTAGGCTCCTATATA ACCCAGGGGCACAGGGGCTGCCCTCATTCTACCACCACTCCAC AGCAGACAGACACTCAGGAGCCAGCCAGC
promoter	MCK Promoter derived from rAAVrh74.MCK GALGT2 (Serepta's dystroglycan modifying therapy to promote Utrophin usage). Derived from mouse MCK core enhancer (206bp) fused to the MCK core promoter (351bp)	558	Muscle	12	249	CAGCCACTATGGGTCTAGGCTGCCATGTAAGGAGGCAAGGCC TGGGGACACCCGAGATGCCTGGTTATAATTAACCCAGACATGTG GCTGCTCCCCCCCCCAACACCTGCTGCTGAGCCTACCCCCA CCCCCGTGCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAA GCTCGCTCTAAAAATAACCTGTCCCTGGTGGGCTGTGGGGGAC TGAGGGCAGGCTGTAACAGGCTTGGGGGCCAGGGCTTATACGT GCCTGGGACTCCCAAAGTATTACTGTTCCATGTTCCCGGCGAAG GGCCAGCTGTCCCCGCCAGCTAGACTCAGCACTTAGTTAGGA ACCAGTGAGCAAGTCAGCCCTTGGGGCAGCCATACAAGGCCA TGGGGCTGGGCAAGCTGCACGCTGGGTCCGGGGTGGGCACG GTGCCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCC CCTCCTGGGGACAGCCCTCCTGGCTAGTCACACCTGTAGGC TCCTCTATATAACCCAGGGGCACAGGGGCTGCCCCC
promoterSet	MCK Promoter/5p UTR derived from rAAVrh74.MCK GALGT2 (Serepta's dystroglycan modifying therapy to promote Utrophin usage)	766	Muscle	21	250	CAGCCACTATGGGTCTAGGCTGCCATGTAAGGAGGCAAGGCC TGGGGACACCCGAGATGCCTGGTTATAATTAACCCAGACATGTG GCTGCTCCCCCCCCCAACACCTGCTGCTGAGCCTACCCCCA CCCCCGTGCCTGGGTCTTAGGCTCTGTACACCATGGAGGAGAA GCTCGCTCTAAAAATAACCTGTCCCTGGTGGGCTGTGGGGGAC TGAGGGCAGGCTGTAACAGGCTTGGGGGCCAGGGCTTATACGT GCCTGGGACTCCCAAAGTATTACTGTTCCATGTTCCCGGCGAAG GGCCAGCTGTCCCCGCCAGCTAGACTCAGCACTTAGTTAGGA ACCAGTGAGCAAGTCAGCCCTTGGGGCAGCCATACAAGGCCA TGGGGCTGGGCAAGCTGCACGCTGGGTCCGGGGTGGGCACG GTGCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCC CCTCCTGGGGACAGCCCTCCTGGCTAGTCACACCTGTAGGC TCCTCTATATAACCCAGGGGCACAGGGGCTGCCCCGGGTACCC ACCACCTCCACAGCACAGACAGACACTCAGGAGCCAGCCAGCC AGGTAAGTTTAGTCTTTTTGTCTTTATTTAGGTCCCGGATCCG GTGGTGGTGCAATCAAAGAACTGCTCCTCAGTGGATGTTGCCT TTACTTCTAGGCCTGTACGGAAGTGTACTTCTGCTCTAAAAGCT GCGGAATTGTACCCGCGGCCGCG

Table 7: promoters						
promoterSet	Contains MHCK7 Promoter linked to SV40intron	961	Muscle	25	251	GTTTAAACAAGCTTGCATGTCTAAGCTAGACCTTCAGATTA AATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTGGTGTGAGA CGCTCCTGTCTCCTCTATCTGCCCATCGGCCCTTTGGGGAGGA GGAATGTGCCCAAGGACTAAAAAAGGCCATGGAGCCAGAGG GGCGAGGGCAACAGACCTTTCATGGGCAACCTTGGGGCCCTG CTGTCTAGCATGCCCACTACGGGTCTAGGCTGCCCATGTAAGG AGGCAAGGCCTGGGGACACCCGAGATGCCTGTTATAATTAAC CCAGACATGTGGCTGCCCCCCCCCCCCAACACCTGCTGCCTCTA AAAATAACCTGTCCCTGGTGGATCCCTGCATGCGAAGATCTT CGAACAAGGCTGTGGGGGACTGAGGGCAGGCTGTAACAGGCT TGGGGGCCAGGGCTTATACGTGCCTGGGACTCCCAAAGTATTAC TGTTCCATGTTCCCGGCGAAGGGCCAGCTGTCCCCGCCAGCTA GACTCAGCACTTAGTTTAGGAACAGTGAGCAAGTCAGCCCTTG GGGCAGCCCATACAAGGCCATGGGGCTGGGCAAGCTGCACGCC TGGGTCCGGGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAG CTCATCTGCTCTAGGGGCCCTCCTGGGACAGCCCTCCTG GCTAGTCACACCTGTAGGCTCCTCTATAAACCAGGGGCACA GGGGCTGCCCTATTCTACCACCACCTCCACAGCACAGACAGAC ACTCAGGAGCCAGCCAGCGCGGCCAGGTAAGTTTAGTCTTT TTGTCTTTTATTTTTCAGGTCCCGGATCCGGTGGTGGTGCAAA AGAAGTGTCTCCTCAGTGGATGTTGCCTTTACTTCTAGGCCTGTAC GGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGCG
promoter	Muscle Specific Promoter derived from the human Desmin gene. Contains a ~1.7kb human DES promoter/enhancer region extending from 1.7kb upstream of the transcription start site to 35bp downstream within exon I of DES.	1736	Muscle	39	252	AAAAGAGTGCAGTAACAAAGCCCCCTTTACAATTTACCCGGCAC ATTACACCCATCCTGAGGCCAAAGCCACAGGCTGTGAGGTCTC ACTGTCTCAGCTTCTGAGCTATAAAATGGGAATGATGCTAGTG TCTACCTCCTAGGGTTGGAGAATTGGGGGTGATGGGTGTGAAG TGCTCAGCAGCTTGGCCCACTAGGTGGTCAGTACATGTAAGG TATTATTGTTGCTACATACATTAGTAGGGCTGGGCCTCTTTAAA CCTTTATAGGGTAGCATGGCAAGGCTAACCATCCTCACTTTATAT CTGACAAGCTGGGGCTCAGAGAGGACGTGCCTGAGCTGGGGCT CAGACAAGGACACACCTACTAGTAACCCTCCAGCTGGTGATGG CAGGTCTAGGGTAGGACAGTGAAGTGGCTCCTAATCGAGCACTC TATTTTCAGGGTTTGCAATCCAAAGGGTCAGGTCCAAGAGGGA CCTGGAGTGCCAAGTGGAGGTGTAGAGGCACGGCCAGTACCCA TGGAGAATGGTGGATGTCCTTAGGGGTTAGCAAGTGCCCTGTG CTAAGGAGGGGGCTTTGGAGGTTGGGCAGGCCTCTGTGGGGC TCCATTTTTGTGGGGGTGGGGGCTGGAGCATTATAGGGGGTGG GAAGTGATTGGGGGTGTACCCCTAGCCTTCTTATCTGACGCCC ACCATGCCTCCTCAGGTACCCCTGCCCCCCACAGCTCCTCTCC TGTGCCTTGTTCAGCCATGCGTTCTCCTCTATAAATACCCGCT CTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAGATGGTT GGGTTGACATGCGGCTCCTGACAAACACAAACCCCTGGTGTGT GTGGGCGTGGGTGGTGTGAGTAGGGGATGAATCAGGGAGGG GGCGGGGGACCCAGGGGGCAGGAGCCACACAAAGTCTGTGCG GGGGTGGGAGCGCACATAGCAATTGGAAGTGAAGCTTATCA GACCCCTTCTGGAATCAGCCCACTGTTTATAAAGTCTGAGGCCCC ACCTCGACAGTACCGGGGAGGAAGAGGGCCTGCACTAGTCCA GAGGGAAGTGAAGGCTCAGGGCTAGCTCGCCATAGACATACA TGGCAGGCAGGCTTTGGCCAGGATCCCTCCGCTGCCAGGCGTC TCCCTGCCCTCCCTTCTGCCTAGAGACCCCAACCTCAAGCCTG GCTGGTCTTTCCTGAGACCAACCTCTTCGACTTCAAGAGAA TATTTAGGAACAAGGTGGTTTAGGGCCTTCTGGAACAGGCC TTGACCTTTAAGAAATGACCCAAAGTCTCTCCTGACCAAAAA GGGGACCTCAAATAAGGGGAAGCCTCTCTCTGCTGTCTCCC CTGACCCCACTCCCCCAACCCAGGACGAGGAGATAACCAAGGG CTGAAAGAGGCCCCGCTGGGGGCTGCAGACATGCTTGCTGCCT GCCCTGGCGAAGGATTGGCAGGCTTGCCGTCACAGGACCCCC GCTGGCTGACTCAGGGGCGCAGGCCTCTTGCGGGGGAGCTGGC CTCCCCCCCCCACGGCCACGGGCGGCCCTTCTGGCAGGACA GCGGGATCTTGACGTGTAGGGGAGGGGAGGCGGGGGCTGA TGTCAGGAGGGATACAAATAGTCCGACGGCTGGGGGCCCTGT

Table 7: promoters						
						CTCCCCCTCGCCGCATCCACTCTCCGGCCGGCCG
promoterSet	CMV enhancer + CMV Promoter + 5pUTR + Kozak Used in Stargen pONY8.95CM VABCR construct	807	Constitutive	48	253	GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGT CATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTA CGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCA TTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG GACTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTG CCCACTTGGCAGTACATCAAGTGATCATATGCCAAGTACGCCC CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGC CCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCGA AGTCTCCACCCCATGACGTCAATGGGAGTTTGTGGCACCA AATCAACGGGACTTTCAAAATGTCGTAACAACTCCGCCCAT TGACGCAATGGGCGGTAGGCATGTACGGTGGGAGGTCTATAT AAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCTGGAGACGC CATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATC CAGCCTCCGCGGCCCAAGCTTCAGCTGCTCGAGGGCGCGCCTC TAGAGCTAGCGTTGCGGCCGCTGGCTCTAACGGCGTTTATGT CTTTGCTGTCTGAGGGGCCTCAGCTCTGACCAATCTGGTCTTCG TGTGGTCATTAGC
promoter	Endogenous hPAH ORF (-973 to -3)	973	Endogenous (Photoreceptors)	17	254	AAGTCAGCATCCATTCTCTGTGGTTCTCCCTCCGCCCATCC AGGTCTCAAGGGTCTAGAGTCTTTCAAAGAGAACACATTCTGAG ATTTGAGGAGGCAGAGACAAAAGTTCCACTGCGAAGTGCCAG GGAGGCTTCTGTTGGGGTGCCCTTGGGATCACAGATCCCCCA CCTGGTGATGAGTCAACCCAGCACCCCATTCAGGGGCTGGA ATGACAGTAATGGGCCACCTGCTGCCTCTCTCATACCCGCACC CCAGTCAGACATTGCAAGTCAGTCACGGCTCTGTCTGTGGGC CTGGAGTGTTCCAGTGCCCTTTCCATCACAGCACAAGCAGCCA CTACTAGTCGATCAATTTAGCACAAGAGATAAACATCATTACCC TCTGCTAAGCTCAGAGATAACCCAACTAGCTGACCATAATGACT TCAGTCATTACGGAGCAAGATAAAAGACTAAAAGAGGGAGGGA TCACCTCAGATCTGCCGAGTGAGTCGATTGGACTTAAAGGGCCA GTCAAACCTGACTGCCGGCTCATGGCAGGCTCTTGCCGAGGAC AAATGCCAGCCTATATTTATGCAAAGAGATTTGTTCAAACTT AAGGTCAAAGATACCTAAAGACATCCCCTCAGGAACCCCTCTC ATGGAGGAGAGTGCTGAGGGTCTTGTTTCCATTGCATCCCC CACCTCAATTTCCCTGGTGCCAGCCACTTGTGTCTTAGGGTTC TCTTCTCTCCATAAAAGGGAGCCAACACAGTGTCGGCCTCCTCT CCCCACTAAGGGCTTATGTGTAATTAAGGGGATTATGCTTTG AAGGGGAAAAGTAGCCTTAATCACCAGGAGAAGGACACAGCG TCCGGAGCCAGAGGCGCTTAACGGCGTTTATGTCCTTTGCTG TCTGAGGGCCTCAGCTCTGACCAATCTGGTCTTCGTGTGGTCA TT
promoter	Muscle Specific CK8 Promoter	450	Muscle	9	255	CTAGACTAGCATGCTGCCCATGTAAGGAGGCAAGGCCTGGGGA CACCCGAGATGCCTGGTTATAATTAACCCAGACATGTGGCTGCC CCCCCCCCCAACACCTGCTGCCTCTAAAAATAACCTGCATGC CATGTTCCCGGCGAAGGGCCAGCTGTCCCCGCCAGCTAGACTC AGCACTAGTTTAGGAACAGTGAGCAAGTCAGCCCTTGGGGC AGCCCATACAAGCCATGGGGCTGGGCAAGCTGCACGCTGGG TCCGGGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAGCTCAT CTGCTCTCAGGGGCCCTCCCTGGGGACAGCCCTCTGGCTAG TCACACCCTGTAGGCTCTCTATATAACCCAGGGGCACAGGGGC TGCCCTCATTCTACCACCACCTCCACAGCACAGACAGACTCAG GAGCCAGCCAGC

Table 7: promoters						
promoter	Muscle Specific human cTnT_Promoter	455	Muscle	4	256	CTGCTCCAGCTGGCCCTCCAGGCCTGGGTTGCTGGCCTCTGC TTTATCAGGATTCTCAAGAGGGACAGCTGGTTTATGTTGCATGA CTGTTCCCTGCATATCTGCTCTGGTTTTAAATAGCTTATCTGCTAG CCTGCTCCAGCTGGCCCTCCAGGCCTGGGTTGCTGGCCTCTG CTTTATCAGGATTCTCAAGAGGGACAGCTGGTTTATGTTGCATG ACTGTTCCCTGCATATCTGCTCTGGTTTTAAATAGCTTATCTGAG CAGCTGGAGGACCACATGGGCTTATATGGGGCACCTGCCAAAA TAGCAGCCAACACCCCCCTGTGCGACATTCTCCCTGGCTCAC CAGGCCCCAGCCACATGCCTGCTTAAAGCCCTCTCCATCCTCTG CCTCACCCAGTCCCCGCTGAGACTGAGCAGACGCCTCCAGGATC TGTCGGCAGCT
promoter	Endogenous hABC4 promoter (5' 3050bp region)	3050	Endogenous (Liver)	91	257	ATTTTTCAAGATAAAAGTGAAATAAATTTTCAGGAAAAAAGC TGAGAAAAATTTGTACAAACTAAAGAAAAACAAGAAAGACAG TAGATGAAAGAGTGCTCATTAGGTGAAAGGAAAAATGATCCAAG AGGGTAGCTTTGAGATGTAGGAAGAAACAAAAAGCAAGAAAT GATAAATGTTTTGATAAAGCTAAATAAGTATCAACTCATAAGA AATAATATTCAGAGAGTCAATGATATACAGAGAAAAATTA GTACATGACAATGGCAATGTAAAGTTAGGGGTGAATAAAAA GAGACTTAAGAGTTCTAAATCATTGCATTGTCTGGAAGAGGA AAAAGTACAATGATTAGTCAAAGATACATGTATAATCCCTAGA AAGGAGATCATTATTAATAGAAAAATAAAGAATACATCTTATA GAAAGGAAATCTAAATGATAATATTAACAGATCTAAATAAGG CAAAAGTGAGGATAAAAAAGAAAGATGGAACCAATGGGGCAA ATAGAAAAAGTAAGATAGCGTGGTAGGGCATTAAATCCAGCCTT ACATCAATGCATAAGTATCTCAATATTCTACTGTAAAGGGAAAG TAAAGATTTCTTACAGCCTGAGTGTAATGGAGAAATCTAGTTTA TCATAGTGCTTTAAATATTGTAAGTCTTCACTTCTAGTTGATGA ATAAATGATGGAATTCAGTGATACTGCACTGTTATCAATAA ATATAAAGGAGCTCCTGGAATTGGATGTAATACAGGTAAAGA AGTAAACACAGCCATATAGGCATGGCTTCTTGCAGGGACAATT TGTGAATCGGCTCAGACAGACAGACAGGCAGGCAATACACCT CATTGCCTCATACATGTTATTTGCTTTAGTTTTGTTCTGAACCTT CCTACTCTCAAGTATCTGCATTTACTTTATCAAATCTCTTTAT TAGAGACTGAAGAACTGTCATCTCCTTATGTGCTAATGAGTTT AATAATGTCCTCCAGTACCACAAGCCTTCTTCAAACATACACAA TTCCAAGTCTTCCGTCTCAGAGTATCTTGAATAATGATCTGAC CGCTGTTAGACCAGTGAAGGGAAGGAATTTGGGTTGATTTAA GAAGAGAATCCTCATGGTCATGGTAGACTGATATGGAGAGAAA ACATTTTGAGGAAAAATACTCAACTAAATTCATTTCTACTCCAGC ATGCAGTTTCAAGTCAAGTTCCACCTTAGCTCCAGGTGGCAGGC AGAGCAGGATGCAGAGGCACAGCACAAAGTAAGGGGTGAGTGC CGAAGCTGCTGGCTCCTGTTCCAGTCTTCTTCTTGGCCTCGCC TGAACCTTTACTATAATAAGTACCATTATTAGGTGTCTCCTA CGTGCAGGACATTTACACACAGTATCCCTAATCCTAATACACCC TTATTTTATAGATCCAATGACTGAGTCAAGAATTACATAACCTGG CCAGACAGCTGGTACATGGGAAAGGTGAGATTACACCAAGGGT CCACCCAGCATCTCTACTTATACCATGCTCTGCTTAAGGTTCTCT GAGAACTCAGACAAGCCTTGGGCTAACAATTGTGTTAACAGGA CATAGCAGGTGCAAGGACCACTGGTCATCCTGCTACCTGATCA GAAGGAAGGAAAGTTGATTGTTGCTCACCTACTATGTTTTAG GCATAGTACTAGGTGCTTTTACCTAGTACTTAATCCCTTATCCTC AACTCATTTATCCTCGCAATAACCTGATAAGGGAGATGTTTTTA TCCTCATTTTACATATAAGGAAACAGGCCTAGAGAAATGAGCAC AGTGTCAAAGTCACATAGTTAATAAGATGTGAAGCTCTGAGTT TGAAAGTCTCCGGTTTCAAAGCCATGAACTTATGGCTCCCCGTT TTAGACACTTCTTTTGGGAAGAGTGTGGAGGAATTAATCAGAA AGAAGAAAGTCATACTCAAATAGGTGGTAGGAGCAGAGACAAT TCAATACAGACAGAAGTCTTAGATGAGAGCAGTGAGCCAGGGC ACTGGACTGGGACTCAGGAGGCTCCCCTAGACTCTGGTCCAC CGATGCAGCCTCAGGCAGGACTTACCTCTCTGGGCATCCGTTT CTTCATATGTTAAACATACGGGGTTTTAATTAGATGATCGCTGAA

Table 7: promoters

						GACCCCTCTAGCCCTAAACTCTGTGTCTCTTAAGTGCTAAGAG GGCACCAACAGCGTTCCTCCTCCCAAGGAGCATAATGTGATGG TTCCTGCCGGCCCTGGCTGACTCTCGCCGTCTTGAGATAATTG GGTTCAGTGCCACCTGGACCAGAACTGGGGATGCGGAAGCAAG AGGCGAGTCTATTGCTCTCTCGGTCTTGGGCCCTGTGAT TGTTGGGCGTCCGAAACTGTCTCCCTATGGGTTTAAAAACAA AACTGAGCGCCCATGGGGTGTGACAGTCATCTGCAGGGGCTTG GGTGGCCCATCAGGCGAGGCTTCTCGGCACCCGAGGCTCCAG CCTGATCTCGGTCTTATCCTGCGACCGGGCTGTTCTGGCGGGT CGCCAGGGTGGGCGGCGGCCAGCCGGGCGCCCGCGCGCA AGAGCGGAGGCTGCGCCCTGGCCCGCGCTAGCTGGGGAG AGAGCTGGGCGGGCGGGGAGTGCTCTCGCGGCCGCGGC CCTCGCCCTGGCTGCAACGTAAGCGTTTCCGGGGCGGACGC GCGTGGGGGGCGGGGCGGGGCGGGGCGAGGCCGCGGCG AGCAAAGTCCAGGCCCTCTGCTGCAGCGCCGCGCTCCAGA GGCCTGCCAGACACGCGGAGGTTGAGGTGAGAGAGGTCC GGGCGCTCTGGCTCGAAGGGAGACCCGGGACGTGGGGCGC GGGGCGGAGTGGCCGACCTCCACCAAGTCCCCCGGGCCCC GCGACTGTGCGCGGGCGCGGAGAGGTGTAAGTGTCT GAGGTGTGGTTTCTCCTCAGGCTGAG
promoter	Endogenous hUSH1b promoter (5' 3kb region)	3000	Endogenous (Photoreceptors)	49	258	GGGTGGCTCCAGTCAGCTGGTTTGGCAAAGTTTCTGGATGATT ACGGAATAACATGTGTCCCAACCCGAGAGCAGGTTGTGGGG GCAATGTTGCATTGACCAGCGTCAGAGAACACATCAGAGGC AAGGGTGGGTGTGACGAGGGAGAAGGCGCAGAAGGCAGGG CTTAGCTCAGCACTCTCCTCCTGCCATGCTGCTGACCGTTC CCTCTGAGTCCCAACAGCCAGGTAGAGGAGGAAGAAATGG GGCTGAGACCCAGCACATCAGTGATTAAGTCAGGATCAGGTG CGGTTTCTGCTCAGGTGCTGAGACAGCAGGCGGTGTCTGCAA ACAACAGGAGGCACCTGAAGCTAGCCTGGGGGGCCACGCCCA GGTGCGGTGCAATCAGCAGCACAGCCAGAGACAGACCCCAATG ACCCGCGCTCCCTGTGCGCAGCCAGTGTCTGCACAGAGCCCTG AGCAGCCTCTGGACATTAGTCCAGCCCCAGCACGGCCCGTCCC CCACGCTGATGTACCGCACCCAGACCTTGAGGGCCCCCTCCGG CTCCGCTCCTGGGAGAAGGCTCTGGAGTGAGGAGGGGAGGG CAGCAGTGCTGGCTGGACAGCTGCTCTGGGCAGGAGAGAGAG GGAGAGACAAGAGACACACAGAGAGACGCGCAGGAAGGG AAAGACCCAGAGGGACGCTAGAACGAGACTTGAGCCAGACA GAGGAAGAGGGGACGTGTGTTTGCAGACTGGCTGGGCCCGTG ACCCAGCTTCTGAGTCTCCGTGCAGGTGGCAGCTGTACAGG CTGGCAGGTCACTGAGAGTGGGCAGCTGGGCCCCAGGTAAGG ATGGGCTGCCACTGTCTGGGCATTGGGAGGGGTTTGGATGT GGAGGAGTCATGGACTTGAGCTACCTCTAGAGCCTCTGCCAC AGCCACTTGCTCCTGGGACTGGGCTTCTGCCACCTTGAGGGC TCAGCCACCACAGCCACTGAATGAACTGTCCGAGCCTGGGAA GATGGATGTGTGCTCCCTGGAGGAGGAAGGCCAAGGAGCA TGTTGTCCATCGAATCTTCTGAGCTGGGGCTGGGGTTAGTGG CATCCTGGGGCCAGGGGAATAGACATGCTGTGGTGGCAGAGA GAAGAGTCCGTTCTCTGTCTCTTTGCTTCTCTGACACTCT TTATCTCCGTTTTTGGATAAGTCACTTCCTCCTATGCCCAAA TATCCCATCTGTGAAATGGGAGTATGAAGCCCAACAGCCAGG GTTGTAGTGGGGAAGAGGTAAATCAGGTATAGACATAGAAAT ACAAATACAGTCTATGCCCCCTGTTGTGAGTTGAAAAAGAAAT AACTGAAGGTGGTCTAGTTCTCATTTTGAAGAAATGAAATGTCTG TCTGGTCATTTTAAATGTGGCCCTTAAATTTACGCCCTCACCA CTCTCCCCATCCCTTGGAGCCCCATGTCTAGTGAAAGCACTG GCTCTGCCCCAGCCCTCATGGCTCATGCTGGCATAGGGCGCT GCTCCACAGCCTGGGCACCATCTTCAGACAAGTGCCCGTGCA ACTGCTGCTGGCCCTGTTGAATCCACATCTCCACCAGGCATCCA GACTAGTTCAGGTCTCTGGAAGGACTGTGGGTTTGTGTGTCCC AGAGCTCCAGGGCAGGGGTGAGGGCTCGGATGTGGGCGAGTG TCATGGGCAGAGGATCGAATGCCCGGCGGCTCTGAATGGGCC

Table 7: promoters

						CTTGTGAAAAATTGATGCGCATTCTAGGAGACAGGTTGGGAGC CAGAGGGGCTCATACCAAGGCTCTGAGGCTGGGGCTGCCTTT AAGCTCCTTCTGAGGCCGTCTCTGGGCTGGCCCTGTGCTGGA CAAGGCTGGAGACAAGGCAATGTCTCAGACCCTCTCCATTGGC CACATCTGCCCTGGATCACTCGCCAATTTGGGGCAGAGGT GGGACTGACCCTTACCCTGACAACATAATGCATATAGTCAAAAT GGGATAAAGGGGAATATAGAGGCTCTGGCAGCTGGGAGTG GTCAGGGAAGGCTTCTGGAGGAGGTATCATCTGAACTGAGCC ATGAACCATAAGTGGAAATTCAGTCAAAATTTAGGTAGAA GGGCCAGTGTGTGAAGGCCAGGAGATGGCAAGAGCTGGCGTA TTTCAGGAACAGTGAGTCACTGAGGATGTCCAAGTATAAGGGT AGGAAAGGGAGTGAGCAGTGAGAGAAAAGACCGAGGCATCAG CAGGGGCCAGATTGTGCTGGGCTAGCGGGCGGGCCGGGC CCGGGCCAGGCCAGGTGCGGTGCATTAGCAGCACAGCCAG AGACAGACCCCAATGACCCTGCCTCCCGTCAGCAGCCAGTGCT CTGCACAGAGCCATCCTGAGGGCAGTGGGTGCTCTTGAGAGGT TTCAGGCAGGGTGTGCTGTGAGCAGGTATGCCAGCCCTTGAC CTTCTGCTCAGTCAGGCTTGTCTTGTACCCACATTCCTGGGGC AGTCCCTAAGCTGAGTGCCGGAGATTAAGTCTAGTCCTAAAT TGCTCTGGCTAGCTGTGTGACCCTGGCAAGTCTTGGTCCCTCT CTGGGCCCTTTGCCGTAGGTCCTGGTGGGGCCAGACTTGCTA CTTTCTAGGAGCCCTTTGGGAATCTCTGAATGACAGTGGCTGAG AGAAGAATTCAGTGCTCTGGGCAGTGGTGTGAGCAGTGG CTGAGGCTCAGGTACACAGGCTGGGCAGTGGTCAGAGGGAG AGAAGCCAAGGAGGGTCCCTGAGGGAGGAGGAGCTGGGGC TTTGGGAGGAGCCAGGTGACCCAGCCAGGCTCAAGGCTTCC AGGGCTGGCCTGCCAGAAGCATGACATGGTCTCTCCTCGCA GAACTGTGCTGGCCAGTGGGCAGCAGGAGCTCTGACTTGG GACC
promoter	Endogenous hUSH2a promoter (5' 3kb region)	3000	Endogenous (Photoreceptors)	21	259	TAATAGGCAGAGTTTCTTAATGTGGACTAGAGTTGCTAATCTTA GATTATCCATTTGAGTCATGATTTCTACTATACAAAGCAGGAGT TGTTATGGGGTAGAAGAATTTTATCCAGGAATGACAAAGATA AGTTGAAGCACTACAGTAAAAAATTAGAGTTAGACATGGACAC GTAGAAGGGAACAACAGACTCTACAGACTCTAGGACCTACTTGA GGCTGAAGGGTGGGAGGAGGTGGAAGATTGAAAACTACCTA TCAGGTAAGTGTCTTATTACCTGGATGATGACATAATCTGTACAT CTAACCCCATGACACACAATTTACCTATATAACAAACCTCCAAA TGTACCCCTGAACCTAAAATAAAAGTTAGAAAAATGAGAATT AGTTCTTGGAATCACAAGATATAAAGAGAAGCCAGCCATTGAAT ACCTGTTTGAAAGTAGGTTGACTTCATGTTTTGTAGCAGGCTCG AATAATCCATTTGTCTAATCACTGTGCTCTATAATACCTATTTTC AAAGATAGTTTCCCAAGTTCTGAGAAGTCTTACATATTAGCTG ACTTTATACTAAAATTTGGGTTTAAAAAATTTTTTTTAGAGAC ATGGTCTCACTCTGTCATCCAGGTTAAAGTGCAGTGGTGGTGTG ATAATAGTTTACTGCAGCCTCGAAATCCTGGGCTCAACAACCTC CCACCTCAGCATCTAAGTAGCTGGGACTACGAGTGTGTGCCAC CATGCCTGGCTAAATTTTTTATTTTATTTTATTTTATTTT TTTGGAGACGTGGGATTTCACTATGTTGCACAGCATGGTCTTGA ACTCCTGGCTTCAAGCAATCCTCCACCTTGGCTCCCAATCCC TAGGAGGCACAAGCATGAGCAATTTGTGCTTGGCCTAAAAATTTG TTTTAAATTAAGTTTTCTGGTAAGAATGTAATAGCGTATTTTG ACAAAGGGTGAGAAAGGCTTCTTGGGAAGCACTAATGCTAA TTGATAAAATTGATATATAAATGGGTTGTGTTTCCAGCTCTCTT CTGGGAGAGAAATAAAGGGAATCTAATAAAGAACAATGTTGG TTTTCTCTGGCTGCTTTACTAACAAGAAACACCATGAAACATTT CTCTATTTCTAACATTTCTATAAAAAAGATAACTTATAGAGAA CAAAATCACAATCGACCAGTTATTTCCCAACAAATTTTCCATTT TTACAATACAAAGGGAAGCTACAAGTATTAGCTGATTTAGAAT ATTTCTCATCTAGGATGAGATGTCCAGATGGCAGAGTAGAGA GAGTTTTGGATATAATTGAACTCTATAGAATTGGTGGCAATG TGCACATATACACACACACAGTTCCTATCCAATTAAGCAGCC

Table 7: promoters

						AAAAAGTCAGCAATCCCATTGCTTCTTTAGTTTAATTAAAGTCAC TGATTTTCCAAACCAACATTTAGAGATCACATCAGATGCTACTC ATAATGTAAGGAAGCATGTATTATGGAGAGGTTATCCTGGGTG AAAGGTACAGCAACAACTGAATAGTCAACCGAACTTCTATCAA TGGGCCAAGCTTTGGGAGCATCAATATATAAAAGTTTAGAATTC CATTTTGTATCCTCTTCTCCCCAAAAAGAAAGAGCACTGGAAAT TATTCCTTGTGTGGTGTTAATAGTGGTAGATCATTTTGATTAAG GAATTAATGGATTGAGGTGCATGAGAGCAAGAAAGAGGAGG GGCAAGAGGGGGGATTATAGGATAAGGTGTACTGCTACTTTAA AATTATGTATGCATGATCCCATCCAGGTCCTCCCACTGCTTGAG GTACCAGCGGAAAGCTTGGGCAGCTCAGTTCCAAGAGGGCCAC CAAGCAGACCACGCTCTGAGCTTCAGGTAACCAAGTGTTCCTC TGCAGAATACTTTACCTGGGCACCAAGTCTTCTTCAGCATTC CTGCTGTACAGCCTATTTGCTGAGTAACCAAGGGTTACAGCAG CGTTGCCAGGCAACGAGGGACAGCGGTCTGTTGAAGAGCCAT TTGTCACACTGAGGGGACTGGTTGAAATGCAATAAAGAAATGG TAACTCAGCTTATTTATCAATACAATTACTTGCACAGTATTAGGG ATCCATGTGTAACTACAAATTCATAGTCATATGAGGAAACACA GAAACATTTTGCTAAATATTAAAGCATAGGACAGACAGATGGTG TTGGGTTTCTAATCAGCTTTACTCTGAGCTTAAAGTTGCTGCACA TGCTGGGATAAGGGGAAAGGCCAAAGTCTTTGCCAGCTTTAT TTTGGGCATCTGTAAGTTAGCTCTGGGTACAATGTACAGTGCA TGTGTAAGAAAATCTACAAGATCTTTCCCTGTTAAGTAGAGC TGGTAATGCCATTGCTAATCCCTGGGGTGAAGTAACAACACAA AATTATTGTATGTGAATATATTATAAATTATATATATATAAA ACACACACATATATTATATAAATATTATGTATACTGGTTATAA ATATTACTGGTTGCTCTGTGGACTTATAAAGTGCTTGATTGCC AATGCAATCAAGAGATTTACCAAAAGGATGAGTATTTTACTCTG AGCACTGTGCTTCAAATGTTTTTGAGAAGTTCAGTAGTGTTC TTCTAGGAGCTCAAAGTCTCAGGCCTGGGATGAGCTTCAGTTT TAAAGGTGCAGCAGCTTTCCCTTGACGCCCTACGTTTTTGATTCC CAGATACCAGCAGCTACTCATGTCTTCGCCATTGCTAAGAACGT CGTTGGTATTACCTTACTCTGAGAACGTGTCTGCAGTTCCAGAA AATGGAGTATCGCAACATCACTTAAAGTACCCTGCTTCAAAGTA TTGCTGGCAAGTGCGGTGGGCTGATTATTTAGAAATGCT TTATCAGGAGGAGAATGCTTTTTTGTAAC
promoterSet	CASI promoter set containing a CMV enhancer, ubiquitin C enhancer elements, and Chicken B-actin core promoter	1053	Constitutive	99	260	CGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCCA ACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCAT GTAAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTA TTTACGGTAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTT GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAG GTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCTCCC CACCCCCAATTTGTATTTATTTATTTTAAATATTTTGTGACG GATGGGGGCGGGGGGGGGGGGGGGCGCGCGCCAGGCGGG GCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAG GTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCC TTTTATGGCGAGGCGGGCGGGCGGGCGGCTATAAAAAGCGA AGCGCGGGCGGGCGGGAGTCGCTGCGCGCTGCCTTCGCCCG TGCCCCGCTCCGCCCGCCTCGCGCGCCCGCCCGGCTCTGA CTGACCGGTTACTAAAACAGGTAAGTCCGGCTCCGCGCCGG GTTTTGGCGCTCCCGCGGGCGCCCCCTCTCACGGCGAGCGC TGCCACGTGAGCAGGAAGGGCGCAGCGAGCGTCTGATCCTTCC GCCCCGACGCTCAGGACAGCGGCCGCTGCTCATAAGACTCGG CCTTAGAACCCAGTATCAGCAGAAGGACATTTTAGGACGGGA CTTGGGTGACTTAGGGCACTGTTTTCTTCCAGAGAGCGGAA CAGGCGAGGAAAAGTAGTCCCTTCTCGGCGATTCTGCGGAGGG ATCTCCGTGGGGCGGTGAACGCCGATGATGCCTCTACTAACCAT GTTTCATGTTTTCTTTTTTTCTACAGGTCTGGGTGACGAACAG GCTAGC

Table 7: promoters

promoter	Endogenous hABC4 promoter (5' 3kb region)	3000	Endogenous (Liver)	38	261	GCTTGCTACTGAAAAGCTAAGGCCAGAGGTAAAGACTATGGAT TTGGGGAATGAATATTCTGTGAAGCCATAAGATAATGGCCTGA GGTGCTGAGGACCAAGTAGTGCTAGGAACCTTGCATCCATGACTA TAGGGCTCTTTAGAAGTGTGCCACAGTACAGCATCATGCAGTAG AATCTAAGTTGTTCTTTGTAATAATGAATGCCAGCAATATTTTAA AATAATAATAATACCATTAAAAAGTGGGCAAAGGACATGAATA GACATTTTTCAAAGGAAACATACAAATCGCCAAGAAGTATATG AAAAATTAACAGTTAATGTTCAATGAATACTTATTGCAGGCTAG GTACTGAGTTGAGCATTTTGCATGCATCATCTCACTTAAAAATAT GTATGTCCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTA AAAATACAAAATTAGCCAGACATGGTGGTACATGCCTGTAATC CCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAATCTG GGAGGCAGAGGTTGCAGTGAGCCGAGATTGCACCACTGCAATC TAGCCTGGGTGACAGAGCGATACTCTGTCTCAAAAGATAATAAT AATAAAATAATGTATGTCAATTGTTGAAATTTGGAAAATGAAC AAGTGTGTGTGTGAATAACTGGGTGTATTCTATACATATGGCTT TATAACTACCTATTAACCTAAGGTCATTAAATGCAATGTATCAAA ATACTCTTTGGATCATCTAGATTGTTGCACATTATCCTATAATAT GAGATGCCACAATTTATTTACACAGTCGACAATTGTAACCCAGC TTGCTTTTGGCTTTTACTGTTTTACATAACTTGGTAAAAATCCT CATATAAATATTTGAAAATTTCTAAGTGTCCATTTGTGAATGTA AAAATATTTTAGAGATCTAAGATTGGTGCAAAACTTGCAATC AGCTACATAGTTCTACTTGAGGCAATTTTCACTCAAAATATATCA TAAACCATAGTACAAAATAGAGCATAGACCTCTCCTTGTGAAG CAGTTGTTTTGCCTTACATTTTTTTTTTTTTTTTTTTTTGAGAT GGAGTCTCGCTCTGTGCGCCGGGCTGGAGTGCAGTGGCGCAAT CTCAGCTCACTGCAAGCTCCGCTCCCGGGTTCACGCCATTCTCC TGCTCAGCCTCCCGAGCAGCTGGGACTACAGGTGCCCGCTACC ACGTCTGGCTAATTTTTATTTTTAGTAGAAACGGGGTTTAC TGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGC CCACCTCGGCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCAC CGTGCGTGGCTGCCTAAATTTTAATAATCATTGTGCAAATTAT TTAGCACTCCAGTGTTTTGATTTTCTCCTCTGCTGGGTAGGAAT AACATAAATACTGTTATTCACCATGGTGGTGGGAAGTTTCAA AGAGCACATGTCTATAAAGTGCTTAGTGCAAGGCTTGGCATGCA GTTAACACAAAATAAATGCGAGCTGCTGTCTTAACAATACTGA CTACACGGCACTGTGATGCTTATGTAAATGCCAGGCTGTGTGTC TGTAACCTGAGGTATTTGTGTAATATTTTCTAAAATAAATCTA ACTAAGGTTGTTCTTCTCACTTGTATGGGGTCATCTTATGCGGTA GATGCTCAAAACACAAATTCAGATACAGAGTGGGCAGTGGTAG TTAGGAAGATAGAAAGGCTAGGGAGTGTTCTGGGAAGTCAGT AAACTTGGAAGATCTAAGGTTATATTAATAATGTTGTATCAGAA CAAAGGCTCAGGACGTTAGTGTTAGCAGAAACCAGATATCTTAG AGCAGTGGTTTGTCAACTTTGCCAGCAATCCACAGTAAGAAATT CACTCCGGCCGGGCGCGGGCCTGTAATCCAGCACTTTGGGA AGCCGAGGCGGGTGGATGACTTGAGGTGAGGAGTTCGAGACC ATCTGGCTAACACAGTGAAACCCGCTCTACTAAAAATACAA AAATTAGCCGGGCGTGGTGGTGTGTCCTGTAATCCAGCTACT TGGGAGGTTGAGGCAGGAGAATCACTTGAACACAGGAGGCGG AGGTGACAGTGAGCCGAGATCGTGCCATTGCACTCCAGCCTGG GTGACAGAGGGAGACTCTATCTCAAAAAAGAAAAAAGAAA TTCAACTCCACTAACACCCACAATGCAATAAATGTGTGAATGT GTACAACTATTTATCAAGCAGTACTTATTATATGTCTGTAATC TGATATTTTATAGCCTGTTTCATTTTATTTAATGTTGATTGTTAC CCACTAAATTTATTTTATTGAGACCCCTAATTTGAAATATTGCCT TGAATATATATACATATATATACACATATATACATATATATA ACACATATATACACATATATACACATATATACACATATATATA CATATATACACATATATACATATATACACATATATACATATATA ATATATATACACATATATACATATATACACATATATACATATATA CACATATATACATATATACACATATATACATATATACATATATAT ACATATATACACATATATACATATATACATATATATACATATATA
----------	---	------	--------------------	----	-----	---

Table 7: promoters

						CATATATATATACACATACATATATATATATACCTTGTAAAAA TAAAGGTTTGCAGCTCCATATTTTTAAAAAATCTTACCCAAG CATTTAATCAGTACTGAATGGTTTTGTCTTGTCTTCATGTCAAGT TGAATTTGGGGGTACTATTCCAGAATATTTACATGTTAGACAAT GTTCTGTAAAAGGGGCATTGTAGCAGCATGCAGGCAGTATTCA ACCAAAAAGTGGGCAAGAGTCATAATTCATCTGGTTTCTCTTC CTTTAAGCAGGTAGTTCCAATTTGCCAGCAGA
promoter	Endogenous hABC4 promoter (5' 3.1kb region)	310 2	Liver	33	262	CGGGAGTCTGAGGGTAGCAGAAGGGTGCGGATTTAAAGTTAC TGTTAGAGTGGCTGGAAAATGGGAGACCGGTTTACAGAGACATTT TATCTACTTAAAAAGTGTGCTTTTGTATCAGCTCAAAGTGAATG CAAAACAAGAACAAAAGGGTTAAAGGCTCAGGTTTAAATCCC AGGTATATGTACATTTCAATTGAGGTATTTTTTTTTCTTTCTAA ATGATCAGTACATTTCTTTCTAAAGAAAATCTTTCTTAACT ACTCTATTTTTAACTTCTCCACAAAGATGAGAAAACATTTA AAAATCATTGGGGCTATTTTTCTGTTACCGAGTAAAGAGAATCT CTAAACCATATTTATAACTCTTACTCTAAATATTTGCATTTACCT CATGCCAGAGCCCGTTGATGACTGACTAAACAGAGTTTCAAAGT TTGAAGAACAGGAAATTTAGAAATGACTAACAATTATGTAGGTT TATTTCTCTCAGTATAGAATGTTTATAGAATTAATGCCAGAGG TTTTCAGAGAAAAATGCAGAAATTTTACTTTGCAAATCCAGAA GATGCAATTGTTCAAGTATTTGTTAAGAAACATTAATTTAAGTA TGCAGATATCATTGAGAATTAATATTTAATTTCTAACTATTA ATCTTTAGTAGGATGCACATATGCAAAATGCCTCATTAGTACTG TAAGAAAAGATTCTTGGCCGGGCGCGGTGGCTCATGACTGTAA TCCCAGCACTTAGGGAGGCCGAGGTGGGCGGATGACGAGGTC AGGAGATCGAGACCACCCTGGCACACGGTCAAACCCGCTCTCTA CTAAAGATACAAAAAATTAGCCGGGCGTGATGGCGGGCGCTG TAGTCCCAGCTACTCGGGAGGCTGAGGCAGAGAATGGCGTGA ACTCGGGAGGCGGAGCTTGCAAGTGAAGCGAGATAGTGCCACT GCACTCCAGTCTGGGCGAAAGAGCGAGACTCCATCTCAAAAAA AAAAAAAAAAAAAGAAAAGATTCTTTAGGTTTCATCAATTTGT TTTAAAGCTAGGGCTCTTCATTAGATATAGGAAAATCAATTCAA AGTTTCTATTCAGTCATGATGAATTTGAGATTTTTTAGGTTTCTT TGATTTAACAATATATTACATTATAATGTTGTGGTGAAAATAA ATGGAATAATATTATTCTTTTCTTTGTTAAATGAAAAGTATGC ACAAAGTATATGTGAGAGTGACAAAGGCTGAATTTGTCAATTA GTAACAATTGTATTCAACAGTAAGGATTTATGTTGGGTAGGC CTTCCCAGGGACTTCTACAAGGAAAAGCTAGAGTTGGTTACT GACTTCTAATAAATAATGCCTACAATTTCTAGGAAGTTAAAAGTT GACATAATTTATCCAAGAAAGAATTATTTCTTAACTTAGAATAG TTCTTTTTCTTTTCAGATGTAGGTTTTCTGGCTTTAGAAAAA TGCTTGTTTTCTTCAATGGAAAATAGGCACACTGTTTTATGCT GTTTCTGTAGTCAGAAAGACAAGTCTGGTATTTCTTTTCAGG ACTCCCTGAGTCATTAAAAAAATCTTCTATCTATCTATGTATC TATCATCCATCTAGCTTTGATTTTTCTCTCTGTGCTTTATTAGT TAATTAGTACCATTTCTGAAGAAGAAATAACATAAGATTATAG AAAAAATTTCTTTCATTGTAAGACTGAATAGAAAAAATTTCTT TCATTATAAGACTGAGTAGAAAAAATAACTTTGTTAGTCTCTG TGCCTCTATGTCCATGAGGAAATTTGACTACTGGTTTTGACTGA CTGAGTTATATAATTAAGTAAAATAACTGGCTTAGTACTAATTAT TGTTCTGTAGTATCAGAGAAAGTTGTTCTTCTACTGGTTGAGCT CAGTAGTTCTTCATATTCTGAGCAAAAGGGCAGAGGTAGGATA GCTTTTCTGAGGTAGAGATAAAGAACCTTGGGTAGGGAAGGAAG ATTTATGAAATATTTAAAAAATTTCTTCTTCTGCTTTGTTTTA GACATAATGTTAAATTTATTTGAAATTTAAAGCAACATAAAAGA ACATGTGATTTTCTACTTATTGAAAGAGAGAAAGGAAAAAAT ATGAAACAGGGATGGAAAGAATCCTATGCCTGGTGAAGGTCAA GGGTTCTCATAACCTACAGAGAATTTGGGGTCAGCCTGTCTAT TGTATATTATGGCAAAGATAATCATCATCTCATTGGGTCCATTT TCCTCTCCATCTCTGCTTAACTGAAGATCCCATGAGATATACTCA CACTGAATCTAAATAGCCTATCTCAGGGCTTGAATCACATGTGG

Table 7: promoters

						GCCACAGCAGGAATGGGAACATGGAATTTCTAAGTCCTATCTTA CTTGTTATTGTTGCTATGCTTTTTCTAGTTTGCATCTGAGGCCAA CATCAGCTTTTTCAGACAGAATGGCTTTGGAATAGTAAAAAGA CACAGAAGCCCTAAAATATGTATGTATGTATGTGTGTGTGCG TGGTGAGTACTTGTGTGTAATTTTTCATTATCTATAGGTAAAA GCACACTTGGAATTAGCAATAGATGCAATTTGGGACTTAACTCT TTCAGTATGCTTATTTCTAAGCAAAGTATTAGTTTGGTTAGTA ATTACTAAACACTGAGAACTAAATTGCAAAACCAAGAAGCTAAA ATGTTCAAGTGGGAAATTACAGTTAAATACCATGGTAATGAATA AAAGGTACAAATCGTTTTAACTCTTATGTAAAATTTGATAAGATG TTTTACACAACCTTAATACATTGACAAGGTCTTGTGGGAGAAAAC AGTTCAGATGGTAAATATACACAAGGGATTAGTCAAACAATT TTTTGGCAAGAATATTATGAATTTTGAATCGGTGGCAGCAAAT GAAATACAAAGATGAGTCTAGTTAATAATCTACAATTATTGGTT AAAGAAGTATATTAGTGCTAATTCCTCCGTTTGTCTAGCTTT TCTCTCTGTCAACCCACACGCCTTTGGCACA
promoter	Murine Albumin Promoter (muAlb Enhancer region + core muAlb Promoter)	2337	Liver	15	263	TCTAGCTTCCTTAGCATGACGTTCCACTTTTTCTAAGGTGGAGC TTACTCTTTGATTTGATCTTTTGTGAACTTTTGGAAATTACCCA TCTTCCTAAGCTTCTGCTTCTCAGTTTTCTGCTTGCTCATTCCA CTTTTCCAGCTGACCCTGCCCTACCAACATTGCTCCACAGCA CAAATTCATCCAGAGAAAATAAATTCTAAGTTTTATAGTTGTTTG GATCGCATAGGTAGCTAAAGAGGTGGCAACCCACACATCCTTA GGCATGAGCTTGATTTTTTTGATTTAGAACCTTCCCTCTCTGTT CCTAGACTACACTACACATTCTGCAAGCATAGCACAGAGCAATG TTCTACTTTAATTACTTTTCTTTCTGTATCCTCACAGCCTAGAA AATAACCTGCGTTACAGCATCCACTCAGTATCCCTTGAGCATGA GGTGACACTACTTAACATAGGGACGAGATGGTACTTTGTGTCTC CTGCTCTGTGACGAGGGCACTGTACTTGCTGATACCAGGGAATG TTTGTTCTTAAATACCATCATTCCGGACGTGTTGCCTTGCCAG TTTCCATGTACATGCAGAAAGAAGTTGGACTGATCAATACAG TCCTCTGCCTTTAAGCAATAGGAAAAGGCCAAGTTGTCTACGTT TAGTATGTGGCTGTAGAAAAGGTATAGATATAAAAAATTTAACT AATGAAATGGCAGTCTTACACATTTTTGGCAGCTTATTTAAAGTC TTGGTGTTAAGTACGCTGGAGCTGTACAGCTACCAATCAGGCA TGTCTGGGAATGAGTACCGGGACCATAAGTTACTGACATTCG TTTCCATTCCATTTGAATACACACTTTTGTGATGGTATTGCTTGC TGAAATGTTTTGCAAAAAAACCCTTCAAATTCATATATATTA TTTTAATAAATGAATTTTAAATTTATCTCAATGTTATAAAAAAGTC AATTTTAATAATTAGGTACTTATATACCAATAATATCTAACAAT CATTTTTAAACATTTGTTTATTGAGCTTATTATGGATGAATCTATC TCTATATACTCTATATACTCTAAAAAGAAGAAAGACCATAGAC AATCATCTATTTGATATGTGTAAGTTTACATGTGAGTAGACATC AGATGCTCCATTTCTCACTGTAATACCATTTATAGTTACTTGCAA AACTAACTGGAATTCTAGGACTTAAATATTTAAGTTTATAGCTGG GTGACTGGTTGGAAAATTTAGGTAAGTACTGAAACCAAGAGA TTATAAAACAATAAATTCTAAAGTTTTAGAAAGTATCATAATCAA ATATTACCCTCTAATGAAAATATTCCAAAGTTGAGCTACAGAAAT TTCAACATAAGATAATTTTAGCTGTAACAATGTAATTTGTTGTCT ATTTTCTTTGAGATACAGTTTTTTCTGTCTAGCTTTGGCTGTCCT GGACCTTGCTCTGTAGACCAGTTGGTCTTGAACCTCAGAGATCT GCTTGCCTCTGCCTTGCAAGTGCTAGGATTAAGCATGTGCCA CCACTGCCTGGCTACAATCTATGTTTATAAGAGATTATAAGCT CTGGCTTTGTGACATTAATCTTTCAGATAATAAGTCTTTTGGATT GTGTCTGGAGAACATACAGACTGTGAGCAGATGTTTCAAGGTA TATTTGCTTAGGGGTGAATTCATCTGCAGCAATAATTATGAGC AGAATTACTGACACTTCCATTTTATACATTCTACTTGCTGATCTAT GAAACATAGATAAGCATGCAGGCATTCTCATAGTTTTCTTTATC TGGAAAAACATTAAATATGAAAGAAGCACTTTATTAATACAGTT TAGATGTGTTTTGCCATCTTTAATTTCTAAGAAATACTAAGCT GATGCAGAGTGAAGAGTGTGTGAAAAGCAGTGGTGCAGCTTG GCTTGAACCTCGTTCTCCAGCTTGGGATCGACCTGCAGGCATGCT

Table 7: promoters						
						TCCATGCCAAGGCCCACTGAAATGCTCAAATGGGAGACAAA GAGATTAAGCTCTTATGTAAAATTTGCTGTTTTACATAACTTTAA TGAATGGACAAAGTCTTGTCATGGGGGTGGGGGTGGGGTTA GAGGGGAACAGCTCCAGATGGCAAACATACGCAAGGGATTAG TCAAACAACCTTTTTGGCAAAGATGGTATGATTTTGTAAATGGGGT AGGAACCAATGAAATGCGAGGTAAGTATGGTTAATGATCTACA GTTATTGGTTAAAGAAGTATATTAGAGCGAGTCTTCTGCACAC AGATCACCTTTCCTATCAACCCC
promoter	Chimeric Promoter hAPOe Enhancer + TBG core promoter + modSV40intron	1330	Liver	14	264	AGGCTCAGAGGCACACAGGAGTTTCTGGGCTCACCTGCCCCCT TCCAACCCTCAGTTCCCATCCTCCAGCAGCTGTTGTGTGCTGC CTCTGAAGTCCACACTGAACAACTTCAGCCTACTCATGTCCTA AAATGGGCAACATTGCAAGCAGCAAACAGCAAACACACAGCC CTCCCTGCCTGCTGACCTTGGAGCTGGGGCAGAGGTAGAGAC CTCTCTGGGCCATGCCACCTCCAACATCCACTCGACCCCTTGGA ATTCGTTGGAGAGGAGCAGAGGTTGCTGGCGTGGTTTAGG TAGTGTGAGAGGGTCCGGGTTCAAACCACTTGCTGGGTGGGG AGTCGTAGTAAGTGGCTATGCCCCGACCCGAAGCCTGTTTCC CCATCTGTACAATGGAATGATAAAGACGCCCATCTGATAGGGT TTTTGTGGCAAATAAACATTTGGTTTTTTGTTTTGTTTTGTTTTG TTTTTGAGATGGAGTTTGTCTGTGCGCCAGGCTGGAGTGCA GTGACACAATCTCATCTACCACAACCTTCCCTGCCTCAGCCTC CCAAGTAGCTGGGATTACAAGCATGTGCCACCACCTGGCTAA TTTTCTATTTTAGTAGAGACGGGTTTCTCCATGTTGGTCAGCCT CAGCCTCCCAAGTAAGTGGGATTACAGGCCTGTGCCACCACACC CGGCTAATTTTTCTATTTTACAGGGACGGGGTTTACCATGT TGGTCAGGCTGGTCTAGAGGTACCGGGGCTGGAAGTACCTTT GACATCATTTCTCTGCGAATGCATGTATAATTTCTACAGAACC ATTAGAAAAGGATACCCAGCCTCTGCTTTGTACAACCTTCCCTT AAAAAAGTCCAATTCAGTCTGTTGGCCCAATAGTGAGAAC TTTTCTGCTGCCTCTTGGTCTTTTGCCTATGGCCCTATTCTG CCTGCTGAAGACACTCTGCCAGCATGGACTTAAACCCCTCCAG CTCTGACAATCCTTTTCTCTTTGTTTTACATGAAGGGTCTGGCA GCCAAAGCAATCACTCAAAGTTCAAACCTTATCATTTTTGCTTT GTTCTCTTGGCCTTGGTTTTGTACATCAGCTTTGAAAATACC CCAGGGTTAATGCTGGGGTTAATTTATACTAAGAGTGCTCTA GTTTTGCAATACAGGACATGCTATAAAATGGAAAGATCTCTAA GGTAAATATAAAATTTTAAAGTGTATAATGTGTTAACTACTGAT TCTAATTGTTTCTCTTTTAGATTCCAACCTTTGGAAGTGA
Pomoter	mCMV enhancer + EF-1a core promoter + SI 126 Intron	937	Constitutive	21	265	AGATTGTACCTGCCCGTACATAAGGTCAATAGGGGGTGAATCA ACAGGAAAGTCCCATTGGAGCCAAGTACACTGCGTCAATAGGG ACTTCCATTGGGTTTTGCCCGGTACATAAGGTCAATAGGGGAT GAGTCAATGGGAAAAACCAATTGGAGCCAAGTACACTGACTCA ATAGGGACTTTCCATTGGGTTTTGCCAGTACATAAGGTCAATA GGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATT GAGTCAATAGGGACTTTCCATGGGTTTTGCCAGTACATAAGG TCAATGGGAGGTAAGCCAATGGGTTTTTCCATTACTGGCACGT ATACTGAGTCATTAGGGACTTTCCATGGGTTTTGCCAGTACAT AAGGTCAATAGGGGTGAATCAACAGGAAAGTCCCATTGGAGCC AAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCAG TACAAAAGGTCAATAGGGGGTGAGTCAATGGGTTTTTCCATT TTGGCACGTACATAAGGTCAATAGGGGTGACTAGTCAGTGGGC AGAGCGCACATCGCCACAGTCCCGAGAAAGTTGGGGGGAGG GGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGT AACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCCG AGGGTGGGGGAGAACCGTATATAAGTGACAGTAGTTGCCGTGAA CGTTCTTTTTCGCAACGGGTTTCCCGCAGAACACAGCTGAAGC TTCTGCCTTCTCCCTCCTGTGAGTTTGGTAAGTCACTGACTGCT ATGCCTGGGAAAGGTGGGCAGGAGATGGGGCAGTGCAGGAA AAGTGGCACTATGAACCTGCAGCCCTAGACAATTGACTAACCC TCTTCTCTTCTCTCTCTGACAG

Table 7: promoters						
promoter	LSP Promoter #2 - Synthetic mTTRenh-promoter Shire	367	Liver	11	266	GAGCTTGGGCTGCAGGTCGAGGGCACTGGGAGGATGTTGAGT AAGATGGAAACTACTGATGACCTTGACAGAGACAGAGTATTA GGACATGTTTGAACAGGGGCGGGCGATCAGCAGGTAGCTCTA GAGGATCCCCGTCTGTCTGCACATTTCTGAGAGCGAGTGTCCG ATACTCTAATCTCCCTAGGCAAGGTTTATTTGTAGGTTACT TATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTT TGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGA GGGGGTATAAAAGCCCTTCACCAGGAGAAGCCGTACACAGA CTAGGCGCGCCACCGCCACC
promoter	LSP Promoter #4 - HS-CRM8 2x SerpEnh TTRmin MVMIntron	468	Liver	9	267	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATACGGGGGAGG CTGCTGGTGAATATTAACCAAGGTCACCCAGTTATCGGAGGAG CAAACAGGGGCTAAGTCCACATACCGTCTGTCTGCACATTTCT AGAGCGAGTGTTCCGATACTCTAATCTCCCTAGGCAAGGTTTCT ATTTGTGTAGGTTACTTATTCTCCTTTTGTGACTAAGTCAATAAT CAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGC CTGGGTTGGAAGGAGGGGGTATAAAAGCCCTTCACCAGGAGA AGCCGTCACACAGATCCACAAGCTCCTGAAGAGGTAAGGGTTT AAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGA GCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #5 - HS-CRM1 AlbEnh TTRmin MVM	426	Liver	7	268	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATT ATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTC TAGCTTTTCTCATGCGTGTACCGTCTGTCTGCACATTTCTGTA GCGAGTGTTCCGATACTCTAATCTCCCTAGGCAAGGTTTCTATTT GTGTAGGTTACTTATTCTCCTTTTGTGACTAAGTCAATAATCAG AATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTG GGTGGAAGGAGGGGGTATAAAAGCCCTTCACCAGGAGAAG CCGTCACACAGATCCACAAGCTCCTGAAGAGGTAAGGGTTTAA GGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGAGC ACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #6 - HS-CRM2 Apo4Enh TTRmin MVM	396	Liver	7	269	GAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGGCAAACATT GCAAGCAGCAAACAGCAAACATAGATGCGTGTTACCGTCTGT CTGCACATTTCTGAGAGCGAGTGTTCCGATACTCTAATCTCCCTA GGCAAGGTTTCTATTTGTGTAGGTTACTTATTCTCCTTTTGTGA CTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCA GGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAAGCCC CTTACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGAAG AGGTAAGGGTTTAAAGGATGGTTGGTTGGTGGGGTATTAATGT TTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #7 - HS-CRM10 Enh TTRmin MVM	495	Liver	6	270	GATGCTCTAATCTCTAGACAAGGTTTCTATTTGTATGGGTTAC TTATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGT TTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTAGATATAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTG TTACCGTCTGTCTGCACATTTCTGAGAGCGAGTGTTCCGATACTC TAATCTCCCTAGGCAAGGTTTCTATTTGTGTAGGTTACTTATTCT CCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAG TCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGG TATAAAAGCCCTTCACCAGGAGAAGCCGTCACACAGATCCACA AGCTCCTGAAGAGGTAAGGGTTTAAAGGATGGTTGGTTGGTGGG GGTATTAATGTTTAATTACCTGGAGCACCTGCCTGAAATCACTTT TTTTCAGGTTG
promoter	LSP Promoter #8 - HS-CRM8 SerpEnh huTBGpro MVM	640	Liver	4	271	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATGCGTGTTAGGG CTGGAAGCTACCTTGACATCATTTCTCTGCGAATGCATGTATA ATTTCTACAGAACCTATTAGAAAGGATCACCAGCCTCTGCTTTT GTACAACCTTCCCTAAAAAACTGCCAATTCCTGCTGTTTGGC CCAATAGTGAGAACTTTTCTGCTGCCTCTGGTGCTTTTGCCT ATGGCCCTATTCTGCTGCTGAAGACACTCTTGCCAGCATGGA CTAAACCCCTCCAGCTCTGACAATCCTCTTCTTTTGTTTTAC ATGAAGGGTCTGGCAGCAAAGCAATCACTCAAAGTTCAAACCT TATCATTTTTGCTTTGTTCTCTTGGCCTTGGTTTGTACATCAG

Table 7: promoters						
						CTTTGAAAATACCATCCCAGGGTTAATGCTGGGGTTAATTTATA ACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAAAAA TGGAAGATCTCCTGAAGAGGTAAGGGTTAAGGGATGGTTGG TTGGTGGGGTATTAATGTTTAATTACCTGGAGCACCTGCCTGAA ATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #9 - HS-CRM1 AlbEnh huTBGpro MVM	667	Liver	3	272	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATT ATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCC TAGCTTTTCTCATGCGTGTTAGGGCTGGAAGCTACCTTTGACATC ATTTCTCTGCGAATGCATGTATAATTTCTACAGAACCTATTAGA AAGGATCACCCAGCCTCTGCTTTGTACAACCTTCCCTTAAAAAA CTGCCAATCCACTGCTGTTTGGCCCAATAGTGAGAACTTTTTCC TGCTGCCTCTTGGTGCTTTTGCCTATGGCCCTATTCTGCCTGCT GAAGACACTCTTGCCAGCATGGACTTAAACCCCTCCAGCTCTGA CAATCCTCTTCTCTTTGTTTTACATGAAGGGTCTGGCAGCCAA AGCAATCACTCAAAGTTCAAACCTATCATTTTTTGTCTTGTCTCT CTTGGCCTTGGTTTTGTACATCAGCTTTGAAAATACCATCCCAGG GTTAATGCTGGGGTTAATTTATAACTAAGAGTGCTCTAGTTTTGC AATACAGGACATGCTATAAAAAATGGAAGATCTCCTGAAGAGG TAAGGGTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAA TTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #10 - HS- CRM2 Apo4Enh huTBGpro MVM	637	Liver	3	273	GAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGGCAAACTT GCAAGCAGCAAACAGCAAACACATAGATGCGTGTTAGGGCTGG AAGTACCTTTGACATCATTTCTCTGCGAATGCATGTATAATTT CTACAGAACCTATTAGAAAGGATCACCCAGCCTCTGCTTTGTAC AACTTTCCCTTAAAAAACTGCCAATCCACTGCTGTTTGGCCCAA TAGTGAGAACTTTTTCTGCTGCCTCTTGGTGCTTTTGCCTATGG CCCCTATTCTGCTGCTGAAGACACTCTTGCCAGCATGGACTTAA ACCCCTCCAGCTCTGACAATCCTCTTTCTCTTTGTTTTACATGAA GGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACCTATCA TTTTTTGCTTTGTTCTCTTGGCCTTGGTTTTGTACATCAGCTTTG AAAATACCATCCCAGGGTTAATGCTGGGGTTAATTTATAACTAA GAGTGCTCTAGTTTTGCAATACAGGACATGCTATAAAATGGAA AGATCTCCTGAAGAGGTAAGGGTTAAGGGATGGTTGGTTGGT GGGGTATTAATGTTTAAATTACCTGGAGCACCTGCCTGAAATCAC TTTTTTTCAGGTTG
promoter	LSP Promoter #11 - HS- CRM10 Enh huTBGpro MVM	736	Liver	2	274	GATGCTCTAATCTCTAGACAAGGTTTCATATTTGATGGGTTAC TTATTCTCTCTTGTGACTAAGTCAATAATCAGAATCAGCAGGT TTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTGAGTATAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTG TTAGGGCTGGAAGCTACCTTTGACATCATTTCTCTGCGAATGCA TGATAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCCTCT GCTTTGTACAACCTTTCCCTTAAAAAACTGCCAATCCACTGCT TTTGGCCCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTCTT TTGCCTATGGCCCTATTCTGCCTGCTGAAGACACTCTGCCAGC ATGGACTTAAACCCCTCCAGCTCTGACAATCCTCTTCTCTTTGT TTACATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCA AACCTTATCATTTTTTGTCTTGTCTCTTGGCCTTGGTTTTGTAC ATCAGCTTTGAAAATACCATCCCAGGGTTAATGCTGGGGTTAAT TTATACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATA AAAATGGAAAGATCTCCTGAAGAGGTAAGGGTTAAGGGATGG TTGGTTGGTGGGGTATTAATGTTTAAATTACCTGGAGCACCTGCC TGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #12 - HS- CRM8 SerpEnh muAlbpro MVM	515	Liver	6	275	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATGCGTGTTAGGC ATGCTTCCATGCCAAGGCCACACTGAAATGCTCAATGGGAGA CAAAGAGATTAAGCTCTTATGTAAATTTGCTGTTTTACATAACT TTAATGAATGGACAAAGTCTTGTCATGGGGGTGGGGTGGGG TTAGAGGGGAACAGCTCCAGATGGCAAACATACGCAAGGGATT TAGTCAAACAACCTTTTGGCAAAGATGGTATGATTTTGAATGG GGTAGGAACCAATGAAATGCGAGGTAAGTATGGTTAATGATCT ACAGTTATTGGTTAAAGAAGTATATTAGAGCGAGTCTTTCTGCA

Table 7: promoters						
						CACAGATCACCTTTCTATCAACCCCTCCTGAAGAGGTAAGGG TTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCT GGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #13 - HS- CRM1 AlbEnh muAlbpro MVM	542	Liver	5	276	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATT ATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCC TAGCTTTTCTCATGCGTGTTAGGCATGCTTCCATGCCAAGGCCCA CACTGAAATGCTCAAATGGGAGACAAAGAGATTAAGCTCTTATG TAAAATTTGCTGTTTACATAACTTTAATGAATGGACAAAGCTT GTGCATGGGGGTGGGGGTGGGGTTAGAGGGGAACAGCTCCAG ATGGCAAACATACGCAAGGGATTAGTCAAACAACTTTTGGCA AAGATGGTATGATTTTGAATGGGGTAGGAACCAATGAAATGC GAGGTAAGTATGGTTAATGATCTACAGTTATTGGTTAAAGAAGT ATATTAGAGCGAGTCTTTCTGCACACAGATCACCTTTCTATCAA CCCCCTCCTGAAGAGGTAAGGGTTAAGGGATGGTTGGTTGGT GGGGTATTAATGTTAATTACCTGGAGCACCTGCCTGAAATCAC TTTTTTTCAGGTTG
promoter	LSP Promoter #14 - HS- CRM2 Apo4Enh muAlbpro MVM	512	Liver	5	277	GAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGGCAAACATT GCAAGCAGCAAACAGCAAACACATAGATGCGTGTTAGGCATGC TTCCATGCCAAGGCCACACTGAAATGCTCAAATGGGAGACAAA GAGATTAAGCTCTTATGTAAAATTTGCTGTTTACATAACTTTAA TGAATGGACAAAGTCTTGTGCATGGGGGTGGGGGTGGGGTTA GAGGGGAACAGCTCCAGATGGCAAACATACGCAAGGGATTTAG TCAAACAACCTTTTTGGCAAAGATGGTATGATTTTGAATGGGGT AGGAACCAATGAAATGCGAGGTAAGTATGGTTAATGATCTACA GTTATTGGTTAAAGAAGTATATTAGAGCGAGTCTTTCTGCACAC AGATCACCTTTCTATCAACCCCTCCTGAAGAGGTAAGGGTTA AGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGAG CACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #15 - HS- CRM10 Enh muAlbpro MVM	611	Liver	4	278	GATGCTCTAATCTCTAGACAAGGTTCAATTTGTATGGGTTAC TTATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGT TTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTGAGTATAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTG TTAGGCATGCTTCCATGCCAAGGCCACACTGAAATGCTCAAAT GGGAGACAAAGAGATTAAGCTCTTATGTAAAATTTGCTGTTTAA CATAACTTTAATGAATGGACAAAGTCTTGTGCATGGGGGTGGG GGTGGGGTTAGAGGGGAACAGCTCCAGATGGCAAACATACGC AAGGGATTTAGTCAAACAACCTTTTTGGCAAAGATGGTATGATT TGTAATGGGGTAGGAACCAATGAAATGCGAGGTAAGTATGGTT AATGATCTACAGTTATTGGTTAAAGAAGTATATTAGAGCGAGTC TTTCTGCACACAGATCACCTTTCTATCAACCCCTCCTGAAGAG GTAAGGGTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTT AATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #16 - CRM8 SerpEnh huAlbpro MVM	355	Liver	5	279	CGGGGGAGGCTGCTGGTGAATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATGCGTGTTAAAC AGTTCAGATGGTAAATATACACAAGGGATTAGTCAAACAATT TTTTGGCAAGAATATTATGAATTTTGAATCGGTTGGCAGCCAAT GAAATACAAAGATGAGTCTAGTTAATAATCTACAATTATTGGTT AAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGCCTCTCCTG AAGAGGTAAGGGTTAAGGGATGGTTGGTTGGTGGGGTATTAA TGTTTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGG TTG
promoter	LSP Promoter #17 - HS- CRM1 AlbEnh huAlbpro MVM	382	Liver	4	280	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATT ATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCC TAGCTTTTCTCATGCGTGTTAAACAGTTCAGATGGTAAATATAC ACAAGGGATTTAGTCAAACAATTTTTGGCAAGAATATTATGAA TTTTGTAAATCGGTTGGCAGCCAATGAAATACAAAGATGAGCTA GTTAATAATCTACAATTATTGGTTAAAGAAGTATATTAGTGCTAA

Table 7: promoters						
						TTTCCTCCGTTTGTCTCTCTGAAGAGGTAAGGGTTTAAGGG ATGGTTGGTTGGTGGGGTATTAAATGTTAATTACCTGGAGCACC TGCTGAAATCACTTTTTTTCAGTTG
promoter	LSP Promoter #18 - HS- CRM2 Apo4Enh huAlbpro MVM	352	Liver	4	281	GAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGGCAAACATT GCAAGCAGCAAACAGCAAACACATAGATGCGTGTTAAACAGTT CCAGATGGTAAATATACACAAGGGATTTAGTCAAACAATTTTT GGCAAGAATATTATGAATTTTGAATCGGTTGGCAGCCAATGAA ATACAAAGATGAGTCTAGTTAATAATCTACAATTATTGGTTAAA GAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCTCTCTGAAG AGGTAAGGGTTTAAGGGATGGTTGGTTGGTGGGGTATTAAATGT TTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGTTG
promoter	LSP Promoter #19 - HS- CRM10 Enh huAlbpro MVM	451	Liver	3	282	GATGCTCTAATCTCTAGACAAGGTTCAATTTGTATGGGTTAC TTATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGT TTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTGAGTATAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTG TTAAACAGTTCAGATGGTAAATATACACAAGGGATTTAGTCAA ACAATTTTTTGGCAAGAATATTATGAATTTTGAATCGGTTGGCA GCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATTA TTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCT CTCCTGAAGAGGTAAGGGTTTAAGGGATGGTTGGTTGGTGGGG TATTAATGTTTAATTACCTGGAGCACCTGCCTGAAATCACTTTTT TCAGTTG
promoter	LSP Promoter #20 - HS- CRM8 SerpEnh huAATpro MVM	430	Liver	13	283	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATGCGTGTTAAAT GACTCCTTTCGGTAAGTGCACTGGAAGCTGTACACTGCCCAGGC AAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCCAGCCA GTGGACTTAGCCCCTGTTGCTCCTCCGATAACTGGGGTGACCTT GGTTAATATTACCAGCAGCCTCCCCGTTGCCCTCTGGATCCA CTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTTC AGGCACCACCACTGACCTGGGACAGTCTCCTGAAGAGGTAAGG GTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACC TGGAGCACCTGCCTGAAATCACTTTTTTTCAGTTG
promoter	LSP Promoter #21 - HS- CRM1 AlbEnh huAATpro MVM	457	Liver	12	284	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATT ATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCC TAGCTTTTCTCATGCGTGTTAAATGACTCCTTTCGGTAAGTGACAG TGGAAGCTGTACACTGCCCAGGCAAAGCGTCCGGGCAGCGTAG GCGGGCGACTCAGATCCCAGCCAGTGACTTAGCCCCTGTTTGC TCCTCCGATAACTGGGGTGACCTTGGTTAATATTACCAGCAGC CTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAG GACAGGGCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTG GGACAGTCTCCTGAAGAGGTAAGGGTTTAAGGGATGGTTGGTT GGTGGGGTATTAATGTTTAATTACCTGGAGCACCTGCCTGAAAT CACTTTTTTTCAGTTG
promoter	LSP Promoter #22 - HS- CRM2 Apo4Enh huAATpro MVM	427	Liver	12	285	GAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGGCAAACATT GCAAGCAGCAAACAGCAAACACATAGATGCGTGTTAAATGACT CCTTTCGGTAAGTGCACTGGAAGCTGTACACTGCCCAGGCAAA GCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCCAGCCAGTG GACTTAGCCCCTGTTGCTCCTCCGATAACTGGGGTGACCTTGGT TAATATTACCAGCAGCCTCCCCGTTGCCCTCTGGATCCACTG CTTAAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTTCAGG CACCACCACTGACCTGGGACAGTCTCCTGAAGAGGTAAGGGTTT AAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGA GCACCTGCCTGAAATCACTTTTTTTCAGTTG
promoter	LSP Promoter #23 - HS- CRM10 Enh huAATpro MVM	526	Liver	11	286	GATGCTCTAATCTCTAGACAAGGTTCAATTTGTATGGGTTAC TTATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGT TTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTGAGTATAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTG TTAATGACTCCTTTCGGTAAGTGCACTGGAAGCTGTACACTGC CCAGGCAAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCC CAGCCAGTGGACTTAGCCCCTGTTGCTCCTCCGATAACTGGGG TGACCTTGGTTAATATTACCAGCAGCCTCCCCGTTGCCCTCT

Table 7: promoters						
						GGATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCCT CAGCTTCAGGCACCACCACTGACCTGGGACAGTCTCTGAAGAG GTAAGGGTTAAGGGATGGTTGGTTGGTGGGGTATTAAATGTTT AATTACCTGGAGCACCTGCCTGAAATCACTTTTTTTCAGGTTG
promoter	LSP Promoter #24 - HS- CRM8 SerpEnh huAATpro SV40in	435	Liver	14	287	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATGCGTGTTAAAT GACTCCTTTTCGGTAAGTGACGTGGAAGCTGTACACTGCCAGGC AAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCAGCCA GTGGACTTAGCCCTGTTTGTCTCTCCGATAACTGGGGTGACCTT GGTTAATATTACCAGCAGCCTCCCCGTTGCCCCCTCTGGATCCA CTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCTCAGCTTC AGGCACCACCACTGACCTGGGACAGTGAATCCGGACTCTAAGG TAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATT TAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #25 - HS- CRM1 AlbEnh huAATpro SV40in	462	Liver	13	288	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATT ATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTC TAGCTTTTCTCATGCGTGTTAAATGACTCCTTCGGTAAGTGACAG TGGAAGCTGTACACTGCCAGGCAAAGCGTCCGGGCAGCGTAG GCGGGCGACTCAGATCCCAGCCAGTGGACTTAGCCCTGTTTGC TCCTCCGATAACTGGGGTGACCTTGTTAATATTACCAGCAGC CTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAG GACAGGGCCCTGTCTCTCAGCTTCAGGCACCACCACTGACCTG GGACAGTGAATCCGGACTCTAAGGTAATATAAAATTTTTAAGT GTATAATGTGTTAAACTACTGATTCTAATTGTTTCTCTTTTGA TTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #26 - HS- CRM2 Apo4Enh huAATpro SV40in	448	Liver	16	289	GCGGCCGCAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGG CAAACATTGCAAGCAGCAAACAGCAAACATAGATGCGTGTTA AATGACTCCTTTTCGGTAAGTGCAAGTGAAGCTGTACACTGCCA GGCAAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCCA GCCAGTGGACTTAGCCCTGTTTGTCTCTCCGATAACTGGGGTG ACCTTGTTAATATTACCAGCAGCCTCCCCGTTGCCCTCTGG ATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCTCA GCTTCAGGCACCACCACTGACCTGGGACAGTGAATCCGGACTCT AAGGTAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTG ATTCTAATTGTTTCTCTTTTAGATTCCAACCTTTGGAAGTGA TTAAAC
promoter	LSP Promoter #27 - HS- CRM10 Enh huAATpro SV40in	531	Liver	12	290	GATGCTCTAATCTCTCTAGACAAGGTTTCATATTGTATGGGTTAC TTATTCTCTTTTGTGACTAAGTCAATAATCAGAATCAGCAGGT TTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTGAGTATAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTG TTAAATGACTCCTTTTCGGTAAGTGCAAGTGAAGCTGTACACTGC CCAGGCAAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCC CAGCCAGTGGACTTAGCCCTGTTTGTCTCTCCGATAACTGGGG TGACCTTGTTAATATTACCAGCAGCCTCCCCGTTGCCCTCT GGATCCACTGCTTAAATACGGACGAGGACAGGGCCCTGTCTCT CAGCTTCAGGCACCACCACTGACCTGGGACAGTGAATCCGGACT CTAAGGTAATATAAAATTTTTAAGTGTATAATGTGTTAAACTAC TGATTCTAATTGTTTCTCTCTTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #28 - HS- CRM8 SerpEnh huTBGpro SV40in	636	Liver	4	291	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTCACCCAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACATGCGTGTTAGGG CTGGAAGCTACCTTTGACATCATTTCTCTGCCAATGCATGTATA ATTTCTACAGAACCTATTAGAAAGGATACCCAGCCTCTGCTTTT GTACAACTTTCCCTTAAAAAACTGCCAATTCAGTCTGTTTGGC CCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTGCTTTTGCCT ATGGCCCTATTCTGCTGTGAAGACACTCTTGCCAGCATGGA CTTAAACCCCTCCAGCTCTGACAATCCTCTTTCTTTTGTTTTAC ATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACCT TATCATTTTTTGTCTTTGTTCTCTTGGCCTTGGTTTTGTACATCAG CTTTGAAAAATACCATCCCAGGGTTAATGCTGGGGTTAATTTATA ACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAAAAA TGGAAGATCTCTAAGGTAATATAAAATTTTTAAGTGTATAAT

Table 7: promoters						
						GTGTTAACTACTGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #29 - HS-CRM1 AlbEnh huTBGpro SV40in	663	Liver	3	292	AGCCAATGAAATACAAAGATGAGTCTAGTTAATAATCTACAATTATTGGTTAAAGAAGTATATTAGTGCTAATTTCCCTCCGTTTGTCC TAGCTTTTCTCATGCGTGTTAGGGCTGGAAGCTACCTTTGACATCATTTCCCTCTGCGAATGCATGTATAATTTCTACAGAACCATTAGAAAGGATCAGCCAGCCTCTGCTTTGTACAACCTTTCCCTTAAAAAACTGCCAATCCACTGCTGTTGGCCCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTGCTTTTGCCTATGGCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGGACTTAAACCCCTCCAGCTCTGA CAATCCTCTTTCTCTTTGTTTTACATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACCTTATCATTTTTGCTTTGTTCTCTTGGCCTGGTTTTGTACATCAGCTTTGAAAATACCATCCAGGGTTAATGCTGGGGTTAATTTATACTAAGAGTGCTCTAGTTTTGC AATACAGGACATGCTATAAAAAATGGAAGATCTCTAAGGTAAATATAAAATTTTAAAGTGATAATGTGTTAACTACTGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #30 - HS-CRM2 Apo4Enh huTBGpro SV40in	633	Liver	3	293	GAATGACCTTCAGCCTGTTCCCGTCCCTGATATGGGCAACATTGCAAGCAGCAACAGCAACACATAGATGCGTGTTAGGGCTGG AAGCTACCTTTGACATCATTTCTCTGCGAATGCATGTATAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCCTCTGCTTTGTACAACTTTCCCTTAAAAAACTGCCAATCCACTGCTGTTGGCCCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTGCTTTTGCCTATGGCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGGACTTAAACCCCTCCAGCTCTGACAATCCTCTTTCTCTTTGTTTTACATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACCTTATCATTTTTGCTTTGTTCTCTTGGCCTTGGTTTTGTACATCAGCTTTGAAAATACCATCCAGGGTTAATGCTGGGGTTAATTTATACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAAAAAATGGAAAGATCTCTAAGGTAAATATAAAATTTTAAAGTGATAATGTGTTAACTACTGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #31 - HS-CRM10 Enh huTBGpro SV40in	732	Liver	2	294	GATGCTCTAATCTCTCTAGACAAGGTTTCATATTGTATGGGTTACTTATTCTCTCTTTGTTGACTAAGTCAATAATCAGAATCAGCAGGTTGCAGTCAGATTGGCAGGGATAAGCAGCCTAGCTCAGGAGAA GTGAGTATAAAAAGCCCCAGGCTGGGAGCAGCCATCAATGCGTGTTAGGGCTGGAAGCTACCTTTGACATCATTTCTCTGCGAATGCA TGATAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCCTCTGCTTTGTACAACTTTCCCTTAAAAAACTGCCAATCCACTGCTGTTTGGCCCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTGCTTTGCCTATGGCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGGACTTAAACCCCTCCAGCTCTGACAATCCTCTTTCTCTTTGT TTTACATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACCTTATCATTTTTGCTTTGTTCTCTTGGCCTTGGTTTTGTACATCAGCTTTGAAAATACCATCCAGGGTTAATGCTGGGGTTAATTTATACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAAAAATGGAAGATCTCTAAGGTAAATATAAAATTTTAAAGTGTA TAATGTGTTAACTACTGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #32 - AMPBenh2x-huTBGpro SV40in	762	Liver	4	295	AGGTTAATTTTTTAAAAAGCAGTCAAAGTCCAAGTGCCCTTGGCAGCATTTACTCTCTGTTTGTCTGGTTAATAATCTCAGGAGCACAAACATTCCAGATCCAGGTTAATTTTTAAAAAGCAGTCAAAGTCCAAGTGCCCTTGGCAGCATTTACTCTCTGTTTGTCTGGTTAATAATCTCAGGAGCACAAACATTCCAGATCCGCGCGCCAGGGCTGGAAGCTACCTTTGACATCATTTCTCTGCGAATGCATGTA TAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCCTCTGCTTTGTACAACTTTCCCTTAAAAAACTGCCAATCCACTGCTGTTTGGCCCAATAGTGAGAACTTTTCTGCTGCCTCTTGGTGCTTTGCTATGGCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGGACTTAAACCCCTCCAGCTCTGACAATCCTCTTTCTCTTTGTTTTACATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACCTTTGGAAGTGA

Table 7: promoters						
						TTATCATTTTTTGCTTTGTTCTCTTGGCCTTGGTTTTGTACATCA GCTTTGAAAATACCATCCCAGGGTTAATGCTGGGGTTAATTAT AACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAACT CTAAGGTAAATATAAAATTTTAAAGTGATAATGTGTTAAACTAC TGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGA
promoter	LSP Promoter #33 - AMPBenh2x-huTBGpro MVM	766	Liver	4	296	AGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGGCCCTTGG CAGCATTTACTCTCTGTTTGTCTGCTGGTTAATAATCTCAGGAGC ACAAACATTCAGATCCAGGTTAATTTTTAAAAAGCAGTCAAAA GTCCAAGTGGCCCTTGGCAGCATTTACTCTCTGTTTGTCTGG TTAATAATCTCAGGAGCACAAACATTCAGATCCGGCGCGCCAG GGCTGGAAGCTACCTTTGACATCATTTCTCTGCGAATGCATGTA TAATTTCTACAGAACCTATTAGAAAGGATCACCCAGCCTCTGCTT TTGTACAACCTTCCCTTAAAAAACTGCCAATCCACTGCTGTTTG GCCCAATAGTGAGAACTTTTCTCTGCTGCCTCTTGGTGCTTTTGC CTATGGCCCTATTCTGCCTGCTGAAGACACTCTTGCCAGCATGG ACTTAACCCCTCCAGCTCTGACAATCTCTTTCTCTTTGTTTTA CATGAAGGGTCTGGCAGCCAAAGCAATCACTCAAAGTTCAAACC TTATCATTTTTTGCTTTGTTCTCTTGGCCTTGGTTTTGTACATCA GCTTTGAAAATACCATCCCAGGGTTAATGCTGGGGTTAATTAT AACTAAGAGTGCTCTAGTTTTGCAATACAGGACATGCTATAACT CCTGAAGAGGTAAGGGTTAAGGGATGGTTGGTTGGTGGGGTA TTAATGTTTAATTACCTGGAGCACTGCCTGAAATCACTTTTTTTC AGGTTG

[00229] Expression cassettes of the ceDNA vector for expression of a desired transgene or therapeutic protein can include a promoter, *e.g.*, any of the promoter selected from **Table 7**, which can influence overall expression levels as well as cell-specificity. For transgene expression, *e.g.*, expression of a desired transgene or therapeutic protein n, they can include a highly active virus-derived immediate early promoter. Expression cassettes can contain tissue-specific eukaryotic promoters to limit transgene expression to specific cell types and reduce toxic effects and immune responses resulting from unregulated, ectopic expression. In some embodiments, an expression cassette can contain a promoter or synthetic regulatory element, such as a CAG promoter (SEQ ID NO: 72). The CAG promoter comprises (i) the cytomegalovirus (CMV) early enhancer element, (ii) the promoter, the first exon and the first intron of chicken beta-actin gene, and (iii) the splice acceptor of the rabbit beta-globin gene. Alternatively, an expression cassette can contain an Alpha-1-antitrypsin (AAT) promoter (SEQ ID NO: 73 or SEQ ID NO: 74), a liver specific (LP1) promoter (SEQ ID NO: 75 or SEQ ID NO: 76), or a Human elongation factor-1 alpha (EF1- α) promoter (*e.g.*, SEQ ID NO: 77 or SEQ ID NO: 78). In some embodiments, the expression cassette includes one or more constitutive promoters, for example, a retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), or a cytomegalovirus (CMV) immediate early promoter (optionally with the CMV enhancer, *e.g.*, SEQ ID NO: 79). Alternatively, an inducible promoter, a native promoter for a transgene, a tissue-specific promoter, or various promoters known in the art can be used.

[00230] Suitable promoters, including those described in **Table 7** and above, can be derived from viruses and can therefore be referred to as viral promoters, or they can be derived from any organism, including prokaryotic or eukaryotic organisms. Suitable promoters can be used to drive expression by any RNA polymerase (*e.g.*, pol I, pol II, pol III). Exemplary promoters include, but are not limited to the SV40 early promoter, mouse mammary tumor virus long terminal repeat (LTR) promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), a rous sarcoma virus (RSV) promoter, a human U6 small nuclear promoter (U6, *e.g.*, SEQ ID NO: 80) (Miyagishi *et al.*, *Nature Biotechnology* 20, 497-500 (2002)), an enhanced U6 promoter (*e.g.*, Xia *et al.*, *Nucleic Acids Res.* 2003 Sep. 1; 31(17)), a human H1 promoter (H1) (*e.g.*, SEQ ID NO: 81 or SEQ ID NO: 155), a CAG promoter, a human alpha 1-antitrypsin (hAAT) promoter (*e.g.*, SEQ ID NO: 82), and the like. In certain embodiments, these promoters are altered at their downstream intron containing end to include one or more nuclease cleavage sites. In certain embodiments, the DNA containing the nuclease cleavage site(s) is foreign to the promoter DNA.

[00231] In one embodiment, the promoter used is the native promoter of the gene encoding the therapeutic protein. The promoters and other regulatory sequences for the respective genes encoding the therapeutic proteins are known and have been characterized. The promoter region used may further include one or more additional regulatory sequences (*e.g.*, native), *e.g.*, enhancers, (*e.g.* SEQ ID NO: 79 and SEQ ID NO: 83), including a SV40 enhancer (SEQ ID NO: 126).

[00232] In some embodiments, a promoter may also be a promoter from a human gene such as human ubiquitin C (hUbC), human actin, human myosin, human hemoglobin, human muscle creatine, or human metallothionein. The promoter may also be a tissue specific promoter, such as a liver specific promoter, such as human alpha 1-antitrypsin (hAAT), natural or synthetic. In one embodiment, delivery to the liver can be achieved using endogenous ApoE specific targeting of the composition comprising a ceDNA vector to hepatocytes via the low density lipoprotein (LDL) receptor present on the surface of the hepatocyte.

[00233] Non-limiting examples of suitable promoters for use in accordance with the present invention include any of the promoters listed in **Table 7**, or any of the following: the CAG promoter of, for example (SEQ ID NO: 72), the hAAT promoter (SEQ ID NO: 82), the human EF1- α promoter (SEQ ID NO: 77) or a fragment of the EF1- α promoter (SEQ ID NO: 78), IE2 promoter (*e.g.*, SEQ ID NO: 84) and the rat EF1- α promoter (SEQ ID NO: 85), mEF1 promoter (SEQ ID NO: 59), or 1E1 promoter fragment (SEQ ID NO: 125).

(ii) *Enhancers*

[00234] In some embodiments, a ceDNA expressing a desired transgene or therapeutic protein comprises one or more enhancers. In some embodiments, an enhancer sequence is located 5' of the promoter sequence. In some embodiments, the enhancer sequence is located 3' of the promoter sequence. Exemplary enhancers are listed in **Table 8** herein.

Table 8: Exemplary Enhancer sequences

Table 8 (Enhancers)					
Description	Length	Tissue Specificity	CG Content	SEQ ID NO:	Sequence
cytomegalovirus enhancer	518	Constitutive	22	300	TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATA AATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCA TAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATG TTGGCATTGATTATTGACTAGTTAATAAGTAATCAATTACGGG GTCATTAGTTTCATAGCCATATATGGAGTTCGCGTTACATAACT TACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCC CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAAC TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC CCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTCTCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGG
Human apolipoprotein E/C-I liver specific enhancer	777	Liver	13	301	AGGCTCAGAGGACACAGGAGTTTCTGGGCTCACCTGCCCCCT TCCAACCCCTCAGTTCCCATCCTCCAGCAGCTGTTGTGTGCTGC CTCTGAAGTCCACACTGAACAACTTCAGCCTACTCATGTCCCTA AAATGGGCAAACATTGCAAGCAGCAAACAGCAAACACACAGCC CTCCCTGCCTGCTGACCTTGGAGCTGGGGCAGAGGTGAGAGAC CTCTCTGGGCCCATGCCACCTCCAACATCCACTCGACCCCTTGG ATTTCCGGTGGAGAGGAGCAGAGGTTGCTCTGGCGTGGTTTAGG TAGTGTGAGAGGGTCCGGGTTCAAACCACTTGTCTGGGTGGGG AGTCGTCAAGTGGCTATGCCCCGACCCGAAGCCTGTTTCC CCATCTGTACAATGGAAATGATAAAGACGCCCATCTGATAGGGT TTTTGTGGCAAATAAACATTTGTTTTTTTGTGTTTTGTTTTG TTTTTTGAGATGGAGGTTTGTCTGTGCGCCAGGCTGGAGTGCA GTGACACAATCTCATCTACCACAACCTTCCCCTGCCTCAGCCTC CCAAGTAGCTGGGATTACAAGCATGTGCCACCACACTGGCTAA TTTTCTATTTTGTAGTAGAGACGGGTTTCTCCATGTTGGTCAGCCT CAGCCTCCCAAGTAAGTGGGATTACAGGCCTGTGCCACCACACC CGGCTAATTTTTTCTATTTTTGACAGGGACGGGTTTACCATGT TGGTCAGGCTGGTCTAGAGGTACCG
CpG-free Murine CMV enhancer	427	Constitutive	0	302	GAGTCAATGGGAAAAACCCATTGGAGCCAAGTACACTGACTCA ATAGGGACTTTCCATTGGGTTTTGCCAGTACATAAGGTCAATA GGGGGTGAGTCAACAGGAAAGTCCCATTGGAGCCAAGTACATT GAGTCAATAGGGACTTTCCAATGGGTTTTGCCAGTACATAAGG TCAATGGGAGGTAAGCCAATGGGTTTTCCCATTAAGTACATGT ATACTGAGTCATTAGGGACTTTCCAATGGGTTTTGCCAGTACAT AAGGTCAATAGGGGTGAATCAACAGGAAAGTCCATTGGAGCC AAGTACACTGAGTCAATAGGGACTTTCCATTGGGTTTTGCCAG TACAAAAGGTCAATAGGGGGTGAATCAATGGGTTTTTCCATTA TTGGCACATACATAAGGTCAATAGGGGTGACTA
HS-CRM8 SERP enhancer	83	Liver	4	303	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTACCCCAAGTT ATCGGAGGAGCAAACAGGGGCTAAGTCCACACGCGTGGTA
Human apolipoprotein E/C-I liver specific enhancer	777	Liver	12	304	AGGCTCAGAGGACACAGGAGTTTCTGGGCTCACCTGCCCCCT TCCAACCCCTCAGTTCCCATCCTCCAGCAGCTGTTGTGTGCTGC CTCTGAAGTCCACACTGAACAACTTCAGCCTACTCATGTCCCTA AAATGGGCAAACATTGCAAGCAGCAAACAGCAAACACACAGCC CTCCCTGCCTGCTGACCTTGGAGCTGGGGCAGAGGTGAGAGAC CTCTCTGGGCCCATGCCACCTCCAACATCCACTCGACCCCTTGG ATTTCCGGTGGAGAGGAGCAGAGGTTGCTCTGGCGTGGTTTAGG TAGTGTGAGAGGGTCCGGGTTCAAACCACTTGTCTGGGTGGGG AGTCGTCAAGTGGCTATGCCCCGACCCGAAGCCTGTTTCC CCATCTGTACAATGGAAATGATAAAGACGCCCATCTGATAGGGT TTTTGTGGCAAATAAACATTTGTTTTTTTGTGTTTTGTTTTG TTTTTTGAGATGGAGGTTTGTCTGTGCGCCAGGCTGGAGTGCA GTGACACAATCTCATCTACCACAACCTTCCCCTGCCTCAGCCTC

					CCAAGTAGCTGGGATTACAAGCATGTGCCACCACCTGGCTAA TTTTCTATTTTAGTAGAGACGGGTTTCTCCATGTTGGTCAGCCT CAGCCTCCCAAGTAACTGGGATTACAGGCCTGTGCCACCACACC CGGCTAATTTTTCTATTTTGACAGGGACGGGTTTCACCATGT TGGTCAGGCTGGTCTAGAGGTAAGT
34bp APOe/c- 1 Enhancer and 32bp AAT X-region	66	Liver	1	305	GTTTGCTGCTTGAATGTTGCCATTTTAGGGTGGACACAGGA CGCTGTGGTTTCTGAGCCAGGG
Insulting sequence and hAPO-HCR Enhancer	212	Liver	4	306	GGAGGGGTGGAGTCGTGACCCCTAAAATGGGCAAACATTGCAA GCAGCAAACAGCAAACACACAGCCCTCCCTGCTGTGACCTTG GAGCTGGGGCAGAGGTGAGAGACCTCTCTGGGCCATGCCACC TCCAACATCCACTCGACCCCTTGAATTCGGTGGAGAGGAGCA GAGGTTGCTCTGGCGTGTTTAGGTAGTGTGAGAGGGG
hAPO-HCR Enhancer derived from SPK9001	330	Liver	4	307	AGGCTCAGAGGCACACAGGAGTTTCTGGGCTCACCCTGCCCCCT TCCAACCCCTCAGTCCCATCTCCAGCAGCTGTTGTGTGCTGC CTCTGAAGTCCACACTGAACAACTTCAGCCTACTCATGTCCCTA AAATGGGCAAACATTGCAAGCAGCAAACAGCAAACACACAGCC CTCCCTGCCTGCTGACCTTGGAGCTGGGGCAGAGGTGAGAGAC CTCTCTGGGGCCATGCCACCTCCAACATCCACTCGACCCCTTGA ATTTGGTGGAGAGGAGCAGAGGTTGTCTGGCGTGTTTAGG TAGTGTGAGAGGGGTACCCGGG
hAPO-HCR Enhancer	194	Liver	3	308	CCCTAAAATGGGCAAACATTGCAAGCAGCAAACAGCAAACACA CAGCCCTCCCTGCTGCTGACCTTGGAGCTGGGGCAGAGGTCA GAGACCTCTCTGGGGCCATGCCACCTCCAACATCCACTCGACCC TTGGAATTTTTCGGTGGAGAGGAGCAGAGGTTGTCTGGCGTG GTTTAGGTAGTGTGAGAGGG
SV40 Enhancer Invivogen	240	Constitutive	0	309	GGGCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTT CTGAGGCTGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGG TGTGAAAGTCCCAGGCTCCCAGCAGGCAGAGTATGCAAA GCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCCA GGCTCCCAGCAGGCAGAGTATGCAAGCATGCATCTCAATTA GTCAGCAACCATAGTCCCACTA
HS-CRM8 SERP enhancer with all spacers/cutsit es removed	73	Liver	2	310	CGGGGGAGGCTGCTGGTGAATATTAACCAAGGTACCCCACTT ATCGGAGGAGCAAACAGGGGCTAAGTCCAC
Alpha mic/bik Enhancer	100	Liver	0	311	AGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGGCCCTTGG CAGCATTACTCTCTGTTTGTCTGGTTAATAATCTCAGGAGC ACAAACATTCC
CpG-free Human CMV Enhancer v2	296	Constitutive	0	312	GTTACATAACTTATGGTAAATGGCCTGCCTGGCTGACTGCCCAA TGACCCTGCCCAATGATGTCAATAATGATGTATGTTCCCATGTA ATGCCAATAGGGACTTTCCATTGATGTCAATGGGTGGAGTATTT ATGGTAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTATGCCCCCTATTGATGTCAATGATGGTAAATGGCCTGCCT GGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGC AGTACATCTATGTATTAGTCATTGCTATTA
SV40 Enhancer	235	Constitutive	1	313	GGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTC TGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTCAGTTAGGGT GTGGAAGTCCCCAGGCTCCCCAGCAGGCAGAGTATGCAAAAG CATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCCAG GCTCCCCAGCAGGCAGAGTATGCAAGCATGCATCTCAATTAG TCAGCAACCATAGTCCC

(iii) 5' UTR sequences and intron sequences

[00235] In some embodiments, a ceDNA vector comprises a 5' UTR sequence and/or an intron sequence that located 3' of the 5' ITR sequence. In some embodiments, the 5' UTR is located 5' of

the transgene, *e.g.*, sequence encoding a desired transgene or therapeutic protein. Exemplary 5' UTR sequences listed in Table 9A.

Table 9A: Exemplary 5' UTR sequences and intron sequences

Table 9A: 5' UTR and intron sequences					
Description	Length	Reference	CG Content	SEQ ID NO:	Sequence
synthetic 5' UTR element composed of chicken B-actin 5'UTR/Intron and rabbit B-globin intron and 1st exon	1127		137	315	GGAGTCGCTGCGACGCTGCCTTCGCCCCGTGCCCCGCTCCGC CGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTT ACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGG GCTGTAATTAGCGCTTGGTTAATGACGGCTTGTTCCTTTCT GTGGCTGCGTGAAAGCCTTGAGGGGCTCCGGGAGGGCCCTT TGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCCTGTGTG TGTGCGTGGGGAGCGCCGCTGCGGCCGCGCTGCCCGGCG GCTGTGAGCGCTGCGGGCGCGGCGGGGCTTGTGCGCTC CGCAGTGTGCGCGAGGGGAGCGCGCCGGGGGCGGTGCC CGGGTGTGCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGT GCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGGTGTGGGC GCGGCGGTGCGGGCTGTAACCCCCCTGCACCCCCCTCCCG AGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTA CGGGGCGTGGCGGGGGCTCGCCGTGCCGGCGGGGGGTG GCGGCAGGTGGGGGTGCCGGCGGGGCGGGGCCGCTCGG GCCGGGAGGGCTCGGGGAGGGGCGCGCGGCCCGCGGA GCGCCGCGGCTGTGAGGCGCGCGAGCCGAGCCATTGC CTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTG TCCCAAATCTGTGCGGAGCCGAAATCTGGGAGGCGCCGCG CACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCG GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTGC GCGCCGCGTCCCCTTCTCCTCTCCAGCCTCGGGGCTGTCCG CGGGGGGACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGG GGTTCGGCTTCTGGCGTGTGACCGCGGCTCTAGAGCCTCTG CTAACCATGTTTAGCCTTCTCTTTTCTACAGCTCCTGGGC AACGTGCTGGTTATTGTGCTGTCTCATCTTTGTCGACAGAAT TCCTCGAAGATCCGAAGGGTTCAAGCTTGGCATTCCGGTAC TGTTGGTAAAGCCA
modified SV40 Intron	93		0	316	CTCTAAGGTAAATATAAAATTTTAAAGTGATAATGTGTTAA CTACTGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGG AACTGA
5' UTR of hAAT just upstream of ORF (3' CGGA may be spacer/restriction enzyme cut site, and was absorbed into the sequence)	54		1	317	GCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTGGGAC AGTGAATCCGGA
CET promotor set synthetic intron	173		0	318	CTGCCTTCTCCTCCTGTGAGTTTGGTAAGTCACTGACTGTCT ATGCCTGGGAAAGGGTGGGCAGGAGATGGGGCAGTGCAGG AAAAGTGGCACTATGAACCCTGCAGCCCTAGACAATTGTA AACCTTCTTCTTCTCTCTCTGACAGGTTGGTGACAGTAG CTTCC
Minute Virus Mice (MVM) Intron	91		0	319	AAGAGGTAAGGGTTTAAGGGATGGTTGGTTGGTGGGGTATT AATGTTTAATTACCTGGAGCACCTGCCTGAAATCACTTTTTTC AGGTTG
5' UTR of hAAT	54		0	320	GCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTGGGAC AGTGAATAATTA
5' UTR of hAAT combined with modSV40 intron	147		1	321	GCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTGGGAC AGTGAATCCGGACTTAAGGTAATATAAAATTTTAAAGTGT ATAATGTGTTAAACTACTGATTCTAATTGTTTCTCTCTTTAGA TTCCAACCTTTGGAAGTGA

5' UTR of hAAT (3' TAATTA may be spacer/restriction enzyme cut site, and was absorbed into the sequence) combined with modSV40 intron	147		0	322	GCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTGGGACAGTGAATAATTACTCTAAGGTAAATATAAAATTTTAAAGTGTAATGTGTAAACTACTGATTCTAATTGTTTCTCTCTTTTAGATTCCAACCTTTGGAAGTGA
42bp of 5' UTR of AAT derived from BMN270 - includes Kozak	48	https://www.ncbi.nlm.nih.gov/pubmed/29292164	1	323	TCCTCAGCTTCAGGCACCACCACTGACCTGGGACAGTGAATCGCCACC
Intron/Enhancer from EF1a1	128	US2017/0216408	6	324	GCTAGCAGGTAAGTGCCGTGTGTGGTTCCCGCGGCGCTGGCCTCTTTACGGGTATGGCCCTTGCCTGCGCTGAATTACTGACACTGACATCCACTTTTTCTTTTCTCCACAGGTTAAACGCCACC
Synthetic SBR intron derived from Sangamo CRMSBS2-Intron3 -- includes kozak	98	WO2017074526	2	325	AAGAGGTAAGGGTTAAGTTATCGTTAGTTCGTGCACCATTATGTTTAATTACCTGGAGCACCTGCCTGAAATCATTTTTTTTCAGGTGGCTAGT
Endogenous hFVIII 5' UTR	172	NG_011403.1	0	326	GCTTAGTGCTGAGCACATCCAGTGGGTAAAGTTCCTTAAATGCTCTGCAAAGAAATTGGGACTTTTCATTAATCAGAAATTTACTTTTTCCCCTCCTGGGAGCTAAAGATATTTAGAGAAGAA TTAACCTTTTGCTTCTCCAGTTGAACATTTGTAGCAATAAGTCA
hAAT 5' UTR + modSV40 + kozak	160	http://www.bloodjournal.org/content/early/2005/12/01/blood-2005-10-4035?ssoc-checked=true	1	327	GCCCTGTCTCCTCAGCTTCAGGCACCACCACTGACCTGGGACAGTGAATCCGGACTCTAAGGTAAATATAAAATTTTAAAGTGTATAATGTGTAAACTACTGATTCTAATTGTTTCTCTCTTTAGATTCCAACCTTTGGAAGTGAATTCTAGACCACC
hFIX 5' UTR and Kozak	29	US20160375110	0	328	ACCACCTTCACAATCTGCTAGCAAAGGTT
Chimeric Intron	133	U47119.2	2	329	GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAACTGGGCTTGTGAGACAGAGAAGACTCTTGCCTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTGCCTTTCTCTCCACAG
Large fragment of Human Alpha-1 Antitrypsin (AAT) 5' UTR	341		9	330	TGGGCAGGAAGTGGGCACTGTGCCAGGGCATGCACTGCCTCCACGCAGCAACCCTCAGAGTCTGAGCTGAACCAAGAAGGAGGAGGGGGTGGGCTCCGAGGAAGGCCATAGCCGTGCTGCTGCCAGGAATCCAGGTTGGAGGGGCGCAACCTCCTGCCAGCCTCAGGCCACTCTCCTGTGCTGCCAGAAGAGACAGAGCTTGAGGAGAGCTTGAGGAGAGCAGGAAAGCCTCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAGGGCCGTGCTCCTCAGCTTCAGGCACCACCACTGACCTGGGACAGTGAATCGACA
5pUTR	316	US9644216	6	331	TCTAGAGAAGCTTTATTGCGGTAGTTATCACAGTTAAATTGCTAACGCAGTCAGTGCTTCTGACACAACAGTCTCGAAGTTAAGCTGCAGTGACTCTTAAGGTAGCCTTGACAGAAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTAAGGAGACCAATAGAACTGGGCTTGTGAGACAGAGAAGACTCTTGCCTTTCTGATAGGCACCTATTGGTCTTACTGACATCCTTTGCCTTTCTCTCCACAGGTGTCCACTCCAGTTCAATTACAGTCTTAAGGCCCTGCAG
Human cDNA ABCB4 5pUTR	76	NM_000443	8	332	CAAAGTCCAGGCCCTCTGTGTCAGCGCCCGCGCTCCAGAGGCCCTGCCAGACACGCGGAGGTTGAGGCTGAG

(Variant A, predominant Isoform)					
Human cDNA ABCB11 5pUTR	127	NM_003742	2	333	AGAATGATGAAAACCGAGGTTGGAAAAGGTTGTGAAACCTT TTAACCTCCACAGTGGAGTCCATTATTTCTCTGGCTTCTCA AATTCATATTCACAGGGTCGTTGGCTGTGGGTTGCAATTACC
Human G6Pase 5pUTR	80	NM_000151.3	0	334	ATAGCAGAGCAATCACCACCAAGCCTGGAATAACTGCAAGG GCTCTGCTGACATCTTCTGAGGTGCCAAGGAAATGAGG
MCK 5pUTR derived from rAAVirh74.MCK GALGT2. Contains 53bp of endogenous mouse MCK Exon1 (untranslated), SV40 late 16S/19S splice signals, 5pUTR derived from plasmid pCMVB.	208	https://patentimage.s.storage.googleapis.com/4f/8a/d6/b915c650f5eeb5/WO2017049031A1.pdf	8	335	GGGTCACCACCCTCCACAGCACAGACAGACTCAGGAGC CAGCCAGCCAGGTAAGTTTAGTCTTTTTGTCTTTATTTTCAGG TCCCGGATCCGGTGGTGGTGCAAATCAAAGAAGTGTCTCTCA GTGGATGTTGCCTTACTTCTAGGCCTGTACGGAAGTGTTACT TCTGCTCTAAAAGCTGCGGAATTGTACCCGCGGCCGCG
CpG Free 5' UTR synthetic (SI 126) Intron	159		0	336	AAGCTTCTGCCTTCTCCCTCCTGTGAGTTGGTAAGTCACTGA CTGTCTATGCCTGGGAAAGGGTGGGCAGGAGATGGGGCAGT GCAGGAAAAGTGGCACTATGAACCTGCAGCCCTAGACAATT GTACTAACCTTCTTCTCTTCTCTCTCTGACAG
5' UTR of Human Cytochrome b-245 alpha chain (CYBA) gene	36	(NM_000101.4)	5	337	CGCGCCTAGCAGTGTCCAGCCGGGTTCTGTCTGCC
5' UTR of Human 2,4-dienoyl-CoA reductase 1 (DECR1) gene	141	(NM_001330575.1)	14	338	ACGCCGCTGGGTCCCAGTCCCGTCCCATCCCCGGCGGCC TAGGCAGCGTTTCCAGCCCCGAGAAGTTGTCTTTTGTCCC GCCCCTGCGCCCAACCGCCTGCGCCGCTTCCGGCCCGAGT TCTGGAGACTCAAC
5' UTR of Human glia maturation factor gamma (GMFG) gene	110	(NM_001301008.1)	4	339	GTTGGATGAAACCTTCTCTACTGCACAGCCCGCCCCCTAC AGCCCCGGTCCCCACGCCTAGAAGACAGCGGAAGTAAAGAAA AGAAGAGGCCTGTGGACAGAACAATC
5' UTR of Human late endosomal/lysosomal adaptor, MAPK and MTOR activator 2 (LAMTOR2)	164	(NM_001145264.1)	13	340	GGTGGGGCGGGGTTGAGTCGGAACCACAATAGCCAGGCGA AGAACTACAACCTCCAGGGCGTCCCGGAGCAGGCCAACGG GACTACGGGAAGCAGCGGGCAGCGGCCCGGGAGGCACC TCGGAGATCTGGGTGAAAAGCCAGGGTTAGGAACCGTAG GC
5' UTR of Human myosin light chain 6B (MYL6B)	127	(NM_002475.4)	8	341	GGCCACCGGAATTAACCTTCAGGGCTGGGGGCCGCGCTAT GCCCCGCCCCCTCCCAGCCCCAGACAGGACCCCGCAGGAG ATGGGTGCCCCCATCCGCACACTGTCTTTGCCACCGGACA TC
Large fragment of Human Alpha-1 Antitrypsin (AAT) 5' UTR	341		9	342	TGGGCAGGAACTGGGCACTGTGCCAGGGCATGCACTGCCT CCACGCAGCAACCCTCAGAGTCTGTAGCTGAACCAAGAAGG AGGAGGGGGTGGGGCTCCGAGGAAGGCTAGCCGTGCTG CTGCCAGGAATTCAGGTTGGAGGGGCGGCAACCTCCTGCC AGCCTTCAGGCCACTCTCTGTGCTGCCAGAAGAGACAGAG CTTGAGGAGAGCTTGAGGAGAGCAGGAAAGCCTCCCCGTT GCCCTCTGGATTCACTGCTTAAATACGGACGAGGACAGGGC CTTGTCTCTCAGCTTCAGGCACCACCTGACCTGGGACAGT GAATCGACA

(iv) 3' UTR sequences

[00236] In some embodiments, a ceDNA vector comprises a 3' UTR sequence that located 5' of the 3' ITR sequence. In some embodiments, the 3' UTR is located 3' of the transgene, *e.g.*, sequence encoding a desired transgene or therapeutic protein. Exemplary 3' UTR sequences listed in **Table 9B**.

Table 9B: Exemplary 3' UTR sequences and intron sequences

Table 9B (3' UTRs)					
Description	Length	Reference	CG Content	SEQ ID NO:	Sequence
WHP Posttranscriptional Response Element	581		20	345	GAGCATCTTACCGCCATTTATCCCATATTTGTTCTGTTTTCTTGATTTGG GTATACATTTAAATGTTAATAAAACAAAATGGTGGGGCAATCATTACAT TTTTAGGGATATGTAATTACTAGTTCAGGTGATTGCCACAAGACAAAACA TGTTAAGAACTTTCCCGTTATTTACGCTCTGTTCTGTTAATCAACCTCT GGATTACAAAATTTGTGAAAGATTGACTGATATCTTAACTATGTTGCTC CTTTACGCTGTGTGGATATGCTCTTTATAGCCTCTGTATCTAGCTATTG CTTCCCGTACGGCTTTGTTTTCTCTCTGTATAAATCCTGGTTGCTGT CTCTTTAGAGGAGTTGTGGCCGTTGTCCGTCAACGTGGCGTGGTGTG CTCTGTGTTGTGACGCAACCCCACTGGCTGGGGCATTGCCACCACCT GTCAACTCCTTTCTGGGACTTTCGCTTTCCCTCCCGATGCCACGGCA GAACTCATCGCCGCTGCTTGGCCGCTGCTGGACAGGGGCTAGTTGC TGGGCACTGATAATCCGTGGTGTGTC
Triplet repeat of mir-142 binding site	77		1	346	TCCATAAAGTAGGAAACACTACACGATTCCATAAAGTAGGAAACACTAC ATCACTCCATAAAGTAGGAAACACTACA
hFIX 3' UTR and polyA spacer derived from SPK9001	88	US2016/0375110	0	347	TGAAAGATGGATTTCCAAGGTTAATTCATTGGAATTGAAAATTAACAGA GATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCT
Human hemoglobin beta (HBB) 3pUTR	395		1	348	TAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGA ATCCTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCA ATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTAAGATAT AGTGATTTTCCCAAGGTTTGAAGTCTCTTCTTTATGTTTAA TGCACTGACCTCCACATCCCTTTTAGTAAAATATTCAGAAATAATTTA AATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAAATCCAGA TGCTCAAGGCCCTTCATAATATCCCCAGTTTGTAGTTGGACTTAGGGA ACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGC
Interferon Beta S/MAR (Scaffold/matrix-associated Region)	800		0	349	AGTCAATATGTTACCCCCAAAAAGCTGTTGTTAACTTGCCAACCTCATT CTAAAATGTATATAGAAGCCCAAGACAATAACAAAAATATTCTTGTA GAACAAAATGGGAAAGAATGTTCCACTAAATATCAAGATTTAGAGCAAA GCATGAGATGTGTGGGATAGACAGTGAGGCTGATAAATAGAGTAGA GCTCAGAAACAGACCCATTGATATATGTAAGTGACCTATGAAAAAATA TGGCATTTTACAATGGGAAAATGATGGTCTTTTCTTTTGTAGAAAAACA GGGAAATATATTATATGTAATAAATAAGGGAACCCATATGTCATAC CATACACACAAAAAATCCAGTGAATTATAAGTCTAAATGGAGAAGGC AAAACTTTAAATCTTTAGAAAATAATAGAAGCATGCCATCAAGACTT CAGTGATAGAGAAAAATTTCTTATGACTCAAAGTCCTAACCAAAAGAAA AGATTGTTAATTAGATTGCATGAATATTAAGACTTATTTTAAATTTAA AAACCATTAAGAAAAGTCAGGCCATAGAATGACAGAAAATTTGCAAC ACCCAGTAAAGAGAATTGTAATATGAGATTATAAAAGAGTCTTAC AAATCAGTAAAAATAAACTAGACAAAAATTTGAACAGATGAAAGAG AAACTCTAAATAATCATTACACATGAGAACTCAATCTCAGAAATCAGAG AACTATCATTGCATATACACTAAATTAGAGAAATTTAAAGGCTAAGTA ACATCTGTGGC

Beta-Globulin MAR (Matrix-associated region)	407		0	350	AATTATCTCTAAGGCATGTGAACCTGGCTGTCTTGGTTTTTCATCTGTACTTCATCTGCTACCTCTGTGACCTGAAACATATTTATAATTCCATTAAAGCTGTGCATATGATAGATTTATCATATGTATTTTCTTAAAGGATTTTGTAAAGAACTAATTGAATTGATACCTGTAAAGTCTTTATCACACTACCCAATAAATAATAAATCTCTTTGTTGAGCTCTGTCTTCTATAAATATGTACCAGTTTTATTGTTTGTAGTGGTAGTGATTTTATTCTCTTCTATATATATACACACACATGTGTGCATTATAAATATATACAAATTTTATGAATAAAAAAATTATTAGCAATCAATATTGAAAACCACTGATTTTGTATGTGAGCAACAGCAGATTAAGG
Human Albumin 3' UTR Sequence	186		1	351	CATCACATTTAAAGCATCTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAAAAGCTTATTCATCTGTTTTCTTTTCGTTGGTGTAAGCAACACCCTGTCTAAAAACATAAATTTCTTAATCATTTTGCCTCTTTCTCTGTGCTTCAATTAATAAAAAATGGAAAGAATCT
CpG minimized HBB 3pUTR	395		0	352	TAAAATACAGCATAGCAAACTTTAACCTCCAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCAATTAGCTGTTTGAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGATTTTCCCAAGGTTGAACTAGCTCTTCATTTCTTATGTTTTAAATGCACTGACCTCCCACTTCCCTTTTAGTAAAATATTCAGAAATAATTTAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCCAGC
WHP Posttranscriptional Response Element. Missing 3' Cytosine.	580		20	353	GAGCATCTTACCGCCATTTATTCCCATATTTGTTCTGTTTTCTTGATTTGGGTATACATTTAAATGTTAATAAAACAAATGGTGGGGCAATCATTTACATTTTATAGGGATATGTAATTACTAGTTCAGGTGTATTGCCACAAGACAAACATGTTAAGAACTTTCCCGTTATTTACGCTCTGTTCTGTAAATCAACCTCTGGATTACAAAATTTGTGAAGATTGACTGATTTCTTAACATATGTTGCTCTTTTACGCTGTGTGGATATGCTGCTTTATAGCCTCTGTATCTAGCTATTGCTTCCCGTACGGCTTTGTTTTCTCCTCTGTATAAATCCTGGTTGCTGTCTCTTTAGAGGAGTTGTGGCCGTTGTCCGTCAACGTGGCGTGGTGTGCTCTGTGTTTGTGACGCAACCCCACTGGCTGGGGCATTGCCACCACCTGTCAACTCCTTTCTGGGACTTTGCTTTCCCTCCCGATCGCCACGGCAAACTCATCGCCGCTGCCTTGCCTGCTGGACAGGGGCTAGGTTGCTGGGCACTGATAATTCGTGGTGTGT
3' UTR of Human Cytochrome b-245 alpha chain (CYBA) gene	64	(NM_000101.4)	5	354	CCTCGCCCGGACCTGCCCTCCCGCCAGGTGCACCCACCTGCAATAAATGCAGCGAAGCCGGGA
Shortened WPRE3 sequence with minimal gamma and alpha elements	247	WPRE 3 ref https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3975461/	10	355	GATAATCAACCTCTGGATTACAAAATTTGTGAAGATTGACTGGTATTCTAACTATGTTGCTCCTTTACGCTATGTGGATACGCTGCTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTCATTTTCTCCTCCTGTATAAATCCTGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCTGCCTTGCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGTGG
Human hemoglobin beta (HBB) 3pUTR	144		1	356	AAATACATCATTGCAATGAAAATAAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGC
First 62bp of WPRE 3pUTR element	62		1	357	GAGCATCTTACCGCCATTTATTCCCATATTTGTTCTGTTTTCTTGATTTGGGTATACATTT

(v). *Polyadenylation Sequences:*

[00237] A sequence encoding a polyadenylation sequence can be included in the ceDNA vector for expression of a desired transgene or therapeutic protein to stabilize an mRNA expressed from the ceDNA vector, and to aid in nuclear export and translation. In one embodiment, the ceDNA vector does not include a polyadenylation sequence. In other embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein includes at least 1, at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, least 45, at least 50 or more adenine dinucleotides. In some embodiments, the polyadenylation sequence comprises about 43 nucleotides, about 40-50 nucleotides, about 40-55 nucleotides, about 45-50 nucleotides, about 35-50 nucleotides, or any range there between.

[00238] The expression cassettes can include any poly-adenylation sequence known in the art or a variation thereof. In some embodiments, a poly-adenylation (polyA) sequence is selected from any of those listed in **Table 10**. Other polyA sequences commonly known in the art can also be used, *e.g.*, including but not limited to, naturally occurring sequence isolated from bovine BGHpA (*e.g.*, SEQ ID NO: 68) or a virus SV40pA (*e.g.*, SEQ ID NO: 86), or a synthetic sequence (*e.g.*, SEQ ID NO: 87). Some expression cassettes can also include SV40 late polyA signal upstream enhancer (USE) sequence. In some embodiments, a USE sequence can be used in combination with SV40pA or heterologous poly-A signal. PolyA sequences are located 3' of the transgene encoding a desired transgene or therapeutic protein.

[00239] The expression cassettes can also include a post-transcriptional element to increase the expression of a transgene. In some embodiments, Woodchuck Hepatitis Virus (WHP) posttranscriptional regulatory element (WPRE) (*e.g.*, SEQ ID NO: 67) is used to increase the expression of a transgene. Other posttranscriptional processing elements such as the post-transcriptional element from the thymidine kinase gene of herpes simplex virus, or hepatitis B virus (HBV) can be used. Secretory sequences can be linked to the transgenes, *e.g.*, VH-02 and VK-A26 sequences, *e.g.*, SEQ ID NO: 88 and SEQ ID NO: 89.

Table 10: Exemplary polyA sequences

Table 10: Exemplary polyA sequences					
Description	Length	Reference	CG Content	SEQ ID NO:	Sequence
bovine growth hormone Terminator and poly-adenylation sequence.	225		3	360	TGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTCCCCCTCCCCCG TGCCCTTCTTGACCTGGAAGGTGCCACTCCCACTGTCTTTT CTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTG TCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGG GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGC GGTGGGCTCTATGGC
Synthetic polyA derived from BMN270	49	https://www.ncbi.nlm.nih.gov/pubmed/29292164	0	361	AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTT GTGTG

Synthetic polyA derived from SPK8011	54	US2017/021 6408	2	362	GCGGCCGCCAATAAAAGATCAGAGCTCTAGAGATCTGTGTGTTGGTTTTTGTGT
Synthetic polyA and insulating sequence derived from Sangamo_CRM SBS2-Intron3	74	WO2017074 526	2	363	GGATCCAATAAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTGTGTGTTTTCTGTAAACGATCGGG
SV40 Late polyA and 3' Insulating sequence derived from Nathwani hFIX	143	http://www.bloodjournal.org/content/early/2005/12/01/blood-2005-10-4035?sso-checked=true	1	364	CTCGATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATGTCATTCATTTATGTTTTCAGGTTTCAGGGGGAGGTGTGGGAGTTTTTTAAACTAGT
bGH polyA derived from SPK9001	228	US2016/037 5110	0	365	CTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCTCTGCCTTCCTTGACCTGGAAGGTGCCACTCCCCTGCTTTCTCTAATAAAATGAGGAAATTGCATCACATTGCTGAGTAGGTGTCACTTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGG
CpGfree SV40 polyA	222		0	366	CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAACCATTATAAGCTGCAATAAACAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGATGTGGGAGGTTTTTAAAGCAAGTAAACCTCTACAAATGTGGTA
SV40 late polyA	226		0	367	CCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACA ACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAACCATTATAAGCTGCAATAAAC AAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCA GGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAACCTCT ACAATGTGGTATGG
C60pAC30HSL polyA containing A64 polyA sequence and C30 histone stem loop sequence	129		0	368	GTTAACAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATGCATCCCCCCCC CCCCCCCCCCCCCCCCCCCCCCAAGGCTCTTTTCAGAGCCA CCA
polyA used in J. Chou G6Pase constructs containing a SV40 polyA	232	US9644216	4	369	GCGGCCGCGGGGATCCAGACATGATAAGATACATTGATGAG TTTGGACAAACCACA ACTAGAATGCAGTGAAAAAATGCTTT ATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTAT AAGCTGCAATAAACCAAGTTAACAACAACAATTGCATTCATTTT ATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAGTCG ACCATGCTGGGGAGAGATCT
SV40 polyadenylation signal	135		0	370	GATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACC ACAACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTT GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAA ACAAGTT
herpesvirus thymidine kinase polyadenylation signal	49		4	371	CGGCAATAAAAAGACAGAATAAAACGCACGGGTGTTGGGTC GTTTGTTTC

SV40 late polyadenylation signal	226		0	372	CCATACCACATTGTAGAGGTTTACTTGCTTAAAAACCTC CCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG TTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATA AAGCAATAGCATCACAATTTACAAATAAAGCATTTTTTTCA CTGCATTCTAGTTGTGGTTGTCCAAACTCATCAATGTATCTTA TCATGTCTGG
Human Albumin 3' UTR and Terminator/pol yA Sequence	416		2	373	CATCACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAG AAAGAAAATGAAGATCAAAAGCTTATTCATCTGTTTTCTTTT TCGTTGGTGTAAGCCAACACCCTGTCTAAAAACATAAATTT CTTTAATCATTTTGCCTCTTTTCTGTGCTTCAATTAATAAAAA AATGGAAGAATCTAATAGAGTGGTACAGCACTGTTATTTT CAAAGATGTGTGCTATCCTGAAAATCTGTAGGTTCTGTGG AAGTTCAGTGTTCTCTTATTCCACTTCGGTAGAGGATTCT AGTTTCTGTGGGCTAATTAATAAATCATTAACTCTTCTA AGTTATGGATTATAAACATTCAAAATAATTTTTGACATTATG ATAATTCTGAATAAAAGAACAAAAACCATG
Human Albumin 3' UTR and Terminator/pol yA Sequence	415		2	374	ATCACATTTAAAAGCATCTCAGCCTACCATGAGAATAAGAGA AAGAAAATGAAGATCAAAAGCTTATTCATCTGTTTTCTTTT CGTTGGTGTAAGCCAACACCCTGTCTAAAAACATAAATTTT TTAATCATTTTGCCTCTTTTCTGTGCTTCAATTAATAAAAA ATGGAAGAATCTAATAGAGTGGTACAGCACTGTTATTTTCT AAAGATGTGTGCTATCCTGAAAATCTGTAGGTTCTGTGGA AGTTCAGTGTTCTCTTATTCCACTTCGGTAGAGGATTCTA GTTTCTGTGGGCTAATTAATAAATCATTAACTCTTCTAA GTTATGGATTATAAACATTCAAAATAATTTTTGACATTATGA TAATTCTGAATAAAAGAACAAAAACCATG
CpGfree, Short SV40 polyA	122		0	375	TAAGATACATTGATGAGTTTGGACAAACCACAAGTGAATGC AGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGC TTTATTTGTAACCATTATAAGCTGCAATAACAAGTT
CpGfree, Short SV40 polyA	133		0	376	TGCTTTATTTGTGAAATTTGTGATGCTATTGCTTATTTGTAAC CATTATAAGCTGCAATAACAAGTTAACAACAACAAATTGCATT CATTTTATGTTTCAAGTTCAGGGGGAGGTGTGGGAGGTTTTT TAAA

(vi). *Nuclear Localization Sequences*

[00240] In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein comprises one or more nuclear localization sequences (NLSs), for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs. In some embodiments, the one or more NLSs are located at or near the amino-terminus, at or near the carboxy-terminus, or a combination of these (*e.g.*, one or more NLS at the amino-terminus and/or one or more NLS at the carboxy terminus). When more than one NLS is present, each can be selected independently of the others, such that a single NLS is present in more than one copy and/or in combination with one or more other NLSs present in one or more copies. Non-limiting examples of NLSs are shown in **Table 11**.

Table 11: Nuclear Localization Signals

<i>SOURCE</i>	<i>SEQUENCE</i>	<i>SEQ ID NO.</i>
SV40 virus large T-antigen	PKKKRKV (encoded by CCCAAGAAGAAGAGGAAGGTG; SEQ ID NO: 91)	90
nucleoplasmin	KRPAATKKAGQAKKKK	92
c-myc	PAAKRVKLD	93
	RQRRNELKRSP	94
hRNPA1 M9	NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY	95
IBB domain from	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV	96

importin-alpha		
myoma T protein	VSRKRPRP	97
	PPKKARED	98
human p53	PQPKKKPL	99
mouse c-abl IV	SALIKKKKKMAP	100
influenza virus NS1	DRLRR	117
	PKQKKRK	118
Hepatitis virus delta antigen	RKLKKKIKKL	119
mouse Mx1 protein	REKKKFLKRR	120
human poly(ADP- ribose) polymerase	KRKGDEV DGVDEVAKKKSKK	121
steroid hormone receptors (human) glucocorticoid	RKCLQAGMNLEARKTKK	122

B. Additional Components of ceDNA vectors

[00241] The ceDNA vectors for expression of a desired transgene or therapeutic protein of the present disclosure may contain nucleotides that encode other components for gene expression. For example, to select for specific gene targeting events, a protective shRNA may be embedded in a microRNA and inserted into a recombinant ceDNA vector designed to integrate site-specifically into the highly active locus, such as an albumin locus. Such embodiments may provide a system for *in vivo* selection and expansion of gene-modified hepatocytes in any genetic background such as described in Nygaard *et al.*, *A universal system to select gene-modified hepatocytes in vivo*, *Gene Therapy*, June 8, 2016. The ceDNA vectors of the present disclosure may contain one or more selectable markers that permit selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, NeoR, and the like. In certain embodiments, positive selection markers are incorporated into the donor sequences such as NeoR. Negative selection markers may be incorporated downstream the donor sequences, for example a nucleic acid sequence HSV-tk encoding a negative selection marker may be incorporated into a nucleic acid construct downstream the donor sequence.

C. Regulatory Switches

[00242] A molecular regulatory switch is one which generates a measurable change in state in response to a signal. Such regulatory switches can be usefully combined with the ceDNA vectors for expression of a desired transgene or therapeutic protein as described herein to control the output of expression of a desired transgene or therapeutic protein from the ceDNA vector. In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein comprises a regulatory switch that serves to fine tune expression of the a desired transgene or therapeutic protein. For example, it can serve as a biocontainment function of the ceDNA vector. In

some embodiments, the switch is an “ON/OFF” switch that is designed to start or stop (*i.e.*, shut down) expression a desired transgene or therapeutic protein in the ceDNA vector in a controllable and regulatable fashion. In some embodiments, the switch can include a “kill switch” that can instruct the cell comprising the ceDNA vector to undergo cell programmed death once the switch is activated. Exemplary regulatory switches encompassed for use in a ceDNA vector for expression of a desired transgene or therapeutic protein can be used to regulate the expression of a transgene, and are more fully discussed in International application PCT/US18/49996, which is incorporated herein in its entirety by reference

(i) Binary Regulatory Switches

[00243] In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein comprises a regulatory switch that can serve to controllably modulate expression of a desired transgene or therapeutic protein. For example, the expression cassette located between the ITRs of the ceDNA vector may additionally comprise a regulatory region, *e.g.*, a promoter, cis-element, repressor, enhancer etc., that is operatively linked to the nucleic acid sequence encoding a desired transgene or therapeutic protein, where the regulatory region is regulated by one or more cofactors or exogenous agents. By way of example only, regulatory regions can be modulated by small molecule switches or inducible or repressible promoters. Non-limiting examples of inducible promoters are hormone-inducible or metal-inducible promoters. Other exemplary inducible promoters/enhancer elements include, but are not limited to, an RU486-inducible promoter, an ecdysone-inducible promoter, a rapamycin-inducible promoter, and a metallothionein promoter.

(ii) Small molecule Regulatory Switches

[00244] A variety of art-known small-molecule based regulatory switches are known in the art and can be combined with the ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein to form a regulatory-switch controlled ceDNA vector. In some embodiments, the regulatory switch can be selected from any one or a combination of: an orthogonal ligand/nuclear receptor pair, for example retinoid receptor variant/LG335 and GRQCIMFI, along with an artificial promoter controlling expression of the operatively linked transgene, such as that as disclosed in Taylor, *et al.* BMC Biotechnology 10 (2010): 15; engineered steroid receptors, *e.g.*, modified progesterone receptor with a C-terminal truncation that cannot bind progesterone but binds RU486 (mifepristone) (US Patent No. 5,364,791); an ecdysone receptor from *Drosophila* and their ecdysteroid ligands (Saez, *et al.*, PNAS, 97(26)(2000), 14512–14517; or a switch controlled by the antibiotic trimethoprim (TMP), as disclosed in Sando R 3rd; Nat Methods. 2013, 10(11):1085-8. In some embodiments, the regulatory switch to control the transgene or expressed by the ceDNA vector is a pro-drug activation switch, such as that disclosed in US patents 8,771,679, and 6,339,070.

(iii) “Passcode” Regulatory Switches

[00245] In some embodiments the regulatory switch can be a “passcode switch” or “passcode circuit”. Passcode switches allow fine tuning of the control of the expression of the transgene from the

ceDNA vector when specific conditions occur – that is, a combination of conditions need to be present for transgene expression and/or repression to occur. For example, for expression of a transgene to occur at least conditions A and B must occur. A passcode regulatory switch can be any number of conditions, *e.g.*, at least 2, or at least 3, or at least 4, or at least 5, or at least 6 or at least 7 or more conditions to be present for transgene expression to occur. In some embodiments, at least 2 conditions (*e.g.*, A, B conditions) need to occur, and in some embodiments, at least 3 conditions need to occur (*e.g.*, A, B and C, or A, B and D). By way of an example only, for gene expression from a ceDNA to occur that has a passcode “ABC” regulatory switch, conditions A, B and C must be present. Conditions A, B and C could be as follows; condition A is the presence of a condition or disease, condition B is a hormonal response, and condition C is a response to the transgene expression. For example, if the transgene edits a defective EPO gene, Condition A is the presence of Chronic Kidney Disease (CKD), Condition B occurs if the subject has hypoxic conditions in the kidney, Condition C is that Erythropoietin-producing cells (EPC) recruitment in the kidney is impaired; or alternatively, HIF-2 activation is impaired. Once the oxygen levels increase or the desired level of EPO is reached, the transgene turns off again until 3 conditions occur, turning it back on.

[00246] In some embodiments, a passcode regulatory switch or “Passcode circuit” encompassed for use in the ceDNA vector comprises hybrid transcription factors (TFs) to expand the range and complexity of environmental signals used to define biocontainment conditions. As opposed to a deadman switch which triggers cell death in the presence of a predetermined condition, the “passcode circuit” allows cell survival or transgene expression in the presence of a particular “passcode”, and can be easily reprogrammed to allow transgene expression and/or cell survival only when the predetermined environmental condition or passcode is present.

[00247] Any and all combinations of regulatory switches disclosed herein, *e.g.*, small molecule switches, nucleic acid-based switches, small molecule-nucleic acid hybrid switches, post-transcriptional transgene regulation switches, post-translational regulation, radiation-controlled switches, hypoxia-mediated switches and other regulatory switches known by persons of ordinary skill in the art as disclosed herein can be used in a passcode regulatory switch as disclosed herein. Regulatory switches encompassed for use are also discussed in the review article Kis *et al.*, J R Soc Interface. 12: 20141000 (2015), and summarized in Table 1 of Kis. In some embodiments, a regulatory switch for use in a passcode system can be selected from any or a combination of the switches disclosed in Table 11 of International Patent Application PCT/US18/49996, which is incorporated herein in its entirety by reference.

(iv). Nucleic acid-based regulatory switches to control transgene expression

[00248] In some embodiments, the regulatory switch to control the expression of a desired transgene or therapeutic protein by the ceDNA is based on a nucleic-acid based control mechanism. Exemplary nucleic acid control mechanisms are known in the art and are envisioned for use. For example, such

mechanisms include riboswitches, such as those disclosed in, *e.g.*, US2009/0305253, US2008/0269258, US2017/0204477, WO2018026762A1, US patent 9,222,093 and EP application EP288071, and also disclosed in the review by Villa JK *et al.*, Microbiol Spectr. 2018 May;6(3). Also included are metabolite-responsive transcription biosensors, such as those disclosed in WO2018/075486 and WO2017/147585. Other art-known mechanisms envisioned for use include silencing of the transgene with an siRNA or RNAi molecule (*e.g.*, miR, shRNA). For example, the ceDNA vector can comprise a regulatory switch that encodes a RNAi molecule that is complementary to the part of the transgene expressed by the ceDNA vector. When such RNAi is expressed even if the transgene (*e.g.*, a desired transgene or therapeutic protein) is expressed by the ceDNA vector, it will be silenced by the complementary RNAi molecule, and when the RNAi is not expressed when the transgene is expressed by the ceDNA vector the transgene (*e.g.*, a desired transgene or therapeutic protein) is not silenced by the RNAi.

[00249] In some embodiments, the regulatory switch is a tissue-specific self-inactivating regulatory switch, for example as disclosed in US2002/0022018, whereby the regulatory switch deliberately switches transgene (*e.g.*, a desired transgene or therapeutic protein) off at a site where transgene expression might otherwise be disadvantageous. In some embodiments, the regulatory switch is a recombinase reversible gene expression system, for example as disclosed in US2014/0127162 and US Patent 8,324,436.

(v). Post-transcriptional and post-translational regulatory switches.

[00250] In some embodiments, the regulatory switch to control the expression of a desired transgene or therapeutic protein by the ceDNA vector is a post-transcriptional modification system. For example, such a regulatory switch can be an aptazyme riboswitch that is sensitive to tetracycline or theophylline, as disclosed in US2018/0119156, GB201107768, WO2001/064956A3, EP Patent 2707487 and Beilstein *et al.*, ACS Synth. Biol., 2015, 4 (5), pp 526–534; Zhong *et al.*, Elife. 2016 Nov 2;5. pii: e18858. In some embodiments, it is envisioned that a person of ordinary skill in the art could encode both the transgene and an inhibitory siRNA which contains a ligand sensitive (OFF-switch) aptamer, the net result being a ligand sensitive ON-switch.

(vi). Other exemplary regulatory switches

[00251] Any known regulatory switch can be used in the ceDNA vector to control the expression of a desired transgene or therapeutic protein by the ceDNA vector, including those triggered by environmental changes. Additional examples include, but are not limited to; the BOC method of Suzuki *et al.*, Scientific Reports 8; 10051 (2018); genetic code expansion and a non-physiologic amino acid; radiation-controlled or ultra-sound controlled on/off switches (see, *e.g.*, Scott S *et al.*, Gene Ther. 2000 Jul;7(13):1121-5; US patents 5,612,318; 5,571,797; 5,770,581; 5,817,636; and WO1999/025385A1. In some embodiments, the regulatory switch is controlled by an implantable system, *e.g.*, as disclosed in US patent 7,840,263; US2007/0190028A1 where gene expression is

controlled by one or more forms of energy, including electromagnetic energy, that activates promoters operatively linked to the transgene in the ceDNA vector.

[00252] In some embodiments, a regulatory switch envisioned for use in the ceDNA vector is a hypoxia-mediated or stress-activated switch, *e.g.*, such as those disclosed in WO1999060142A2, US patent 5,834,306; 6,218,179; 6,709,858; US2015/0322410; Greco *et al.*, (2004) Targeted Cancer Therapies 9, S368, as well as FROG, TOAD and NRSE elements and conditionally inducible silence elements, including hypoxia response elements (HREs), inflammatory response elements (IREs) and shear-stress activated elements (SSAEs), *e.g.*, as disclosed in U.S. Patent 9,394,526. Such an embodiment is useful for turning on expression of the transgene from the ceDNA vector after ischemia or in ischemic tissues, and/or tumors.

(vii). Kill Switches

[00253] Other embodiments described herein relate to a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein comprising a kill switch. A kill switch as disclosed herein enables a cell comprising the ceDNA vector to be killed or undergo programmed cell death as a means to permanently remove an introduced ceDNA vector from the subject's system. It will be appreciated by one of ordinary skill in the art that use of kill switches in the ceDNA vectors for expression of a desired transgene or therapeutic protein would be typically coupled with targeting of the ceDNA vector to a limited number of cells that the subject can acceptably lose or to a cell type where apoptosis is desirable (*e.g.*, cancer cells). In all aspects, a "kill switch" as disclosed herein is designed to provide rapid and robust cell killing of the cell comprising the ceDNA vector in the absence of an input survival signal or other specified condition. Stated another way, a kill switch encoded by a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein can restrict cell survival of a cell comprising a ceDNA vector to an environment defined by specific input signals. Such kill switches serve as a biological biocontainment function should it be desirable to remove the ceDNA vector e expression of a desired transgene or therapeutic protein in a subject or to ensure that it will not express the encoded a transgene or therapeutic protein.

[00254] Other kill switches known to a person of ordinary skill in the art are encompassed for use in the ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein, *e.g.*, as disclosed in US2010/0175141; US2013/0009799; US2011/0172826; US2013/0109568, as well as kill switches disclosed in Jusiak *et al.*, Reviews in Cell Biology and molecular Medicine; 2014; 1-56; Kobayashi *et al.*, PNAS, 2004; 101; 8419-9; Marchisio *et al.*, Int. Journal of Biochem and Cell Biol., 2011; 43; 310-319; and in Reinshagen *et al.*, Science Translational Medicine, 2018, 11.

[00255] Accordingly, in some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein can comprise a kill switch nucleic acid construct, which comprises the nucleic acid encoding an effector toxin or reporter protein, where the expression of the effector toxin (*e.g.*, a death protein) or reporter protein is controlled by a predetermined condition. For example, a predetermined condition can be the presence of an environmental agent, such as, *e.g.*, an

exogenous agent, without which the cell will default to expression of the effector toxin (*e.g.*, a death protein) and be killed. In alternative embodiments, a predetermined condition is the presence of two or more environmental agents, *e.g.*, the cell will only survive when two or more necessary exogenous agents are supplied, and without either of which, the cell comprising the ceDNA vector is killed.

[00256] In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein is modified to incorporate a kill-switch to destroy the cells comprising the ceDNA vector to effectively terminate the *in vivo* expression of the transgene being expressed by the ceDNA vector (*e.g.*, expression of a desired transgene or therapeutic protein). Specifically, the ceDNA vector is further genetically engineered to express a switch-protein that is not functional in mammalian cells under normal physiological conditions. Only upon administration of a drug or environmental condition that specifically targets this switch-protein, the cells expressing the switch-protein will be destroyed thereby terminating the expression of the therapeutic protein or peptide. For instance, it was reported that cells expressing HSV-thymidine kinase can be killed upon administration of drugs, such as ganciclovir and cytosine deaminase. See, for example, Dey and Evans, Suicide Gene Therapy by Herpes Simplex Virus-1 Thymidine Kinase (HSV-TK), in *Targets in Gene Therapy*, edited by You (2011); and Beltinger *et al.*, *Proc. Natl. Acad. Sci. USA* 96(15):8699-8704 (1999). In some embodiments the ceDNA vector can comprise a siRNA kill switch referred to as DISE (Death Induced by Survival gene Elimination) (Murmah *et al.*, *Oncotarget*. 2017; 8:84643-84658. Induction of DISE in ovarian cancer cells *in vivo*).

IV. Production of a ceDNA Vector

A. Production in General

[00257] Certain methods for the production of a ceDNA vector for expression of a desired transgene or therapeutic protein comprising an asymmetrical ITR pair or symmetrical ITR pair as defined herein is described in section IV of International application PCT/US18/49996 filed September 7, 2018, which is incorporated herein in its entirety by reference. In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can be produced using insect cells, as described herein. In alternative embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can be produced synthetically and in some embodiments, in a cell-free method, as disclosed on International Application PCT/US19/14122, filed January 18, 2019, which is incorporated herein in its entirety by reference.

[00258] As described herein, in one embodiment, a ceDNA vector for expression of a desired transgene or therapeutic protein can be obtained, for example, by the process comprising the steps of: a) incubating a population of host cells (*e.g.* insect cells) harboring the polynucleotide expression construct template (*e.g.*, a ceDNA-plasmid, a ceDNA-Bacmid, and/or a ceDNA-baculovirus), which is devoid of viral capsid coding sequences, in the presence of a Rep protein under conditions effective and for a time sufficient to induce production of the ceDNA vector within the host cells, and wherein

the host cells do not comprise viral capsid coding sequences; and b) harvesting and isolating the ceDNA vector from the host cells. The presence of Rep protein induces replication of the vector polynucleotide with a modified ITR to produce the ceDNA vector in a host cell. However, no viral particles (*e.g.* AAV virions) are expressed. Thus, there is no size limitation such as that naturally imposed in AAV or other viral-based vectors.

[00259] The presence of the ceDNA vector isolated from the host cells can be confirmed by digesting DNA isolated from the host cell with a restriction enzyme having a single recognition site on the ceDNA vector and analyzing the digested DNA material on a non-denaturing gel to confirm the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA.

[00260] In yet another aspect, the invention provides for use of host cell lines that have stably integrated the DNA vector polynucleotide expression template (ceDNA template) into their own genome in production of the non-viral DNA vector, *e.g.* as described in Lee, L. *et al.* (2013) Plos One 8(8): e69879. Preferably, Rep is added to host cells at an MOI of about 3. When the host cell line is a mammalian cell line, *e.g.*, HEK293 cells, the cell lines can have polynucleotide vector template stably integrated, and a second vector such as herpes virus can be used to introduce Rep protein into cells, allowing for the excision and amplification of ceDNA in the presence of Rep and helper virus.

[00261] In one embodiment, the host cells used to make the ceDNA vectors for expression of a desired transgene or therapeutic protein as described herein are insect cells, and baculovirus is used to deliver both the polynucleotide that encodes Rep protein and the non-viral DNA vector polynucleotide expression construct template for ceDNA, *e.g.*, as described in **FIGS. 4A-4C** and Example 1. In some embodiments, the host cell is engineered to express Rep protein.

[00262] The ceDNA vector is then harvested and isolated from the host cells. The time for harvesting and collecting ceDNA vectors described herein from the cells can be selected and optimized to achieve a high-yield production of the ceDNA vectors. For example, the harvest time can be selected in view of cell viability, cell morphology, cell growth, etc. In one embodiment, cells are grown under sufficient conditions and harvested a sufficient time after baculoviral infection to produce ceDNA vectors but before a majority of cells start to die because of the baculoviral toxicity. The DNA vectors can be isolated using plasmid purification kits such as Qiagen Endo-Free Plasmid kits. Other methods developed for plasmid isolation can be also adapted for DNA vectors. Generally, any nucleic acid purification methods can be adopted.

[00263] The DNA vectors can be purified by any means known to those of skill in the art for purification of DNA. In one embodiment, ceDNA vectors are purified as DNA molecules. In another embodiment, the ceDNA vectors are purified as exosomes or microparticles.

[00264] Exosomes are small membrane vesicles of endocytic origin that are released into the extracellular environment following fusion of multivesicular bodies with the plasma membrane. Their surface consists of a lipid bilayer from the donor cell's cell membrane, they contain cytosol from

the cell that produced the exosome, and exhibit membrane proteins from the parental cell on the surface. Exosomes are produced by various cell types including epithelial cells, B and T lymphocytes, mast cells (MC) as well as dendritic cells (DC). Some embodiments, exosomes with a diameter between 10nm and 1 μ m, between 20nm and 500nm, between 30nm and 250nm, between 50nm and 100nm are envisioned for use. Exosomes can be isolated for a delivery to target cells using either their donor cells or by introducing specific nucleic acids into them. Various approaches known in the art can be used to produce exosomes containing capsid-free AAV vectors of the present invention.

[00265] Generally, lipid nanoparticles comprise an ionizable amino lipid (*e.g.*, heptatriacontan-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate, DLin-MC3-DMA, a phosphatidylcholine (1,2-distearoyl-sn-glycero-3-phosphocholine, DSPC), cholesterol and a coat lipid (polyethylene glycol-dimyristolglycerol, PEG-DMG), for example as disclosed by Tam *et al.* (2013). Advances in Lipid Nanoparticles for siRNA delivery. *Pharmaceuticals* 5(3): 498-507. In some embodiments, a lipid nanoparticle has a mean diameter between about 10 and about 1000 nm. In some embodiments, a lipid nanoparticle has a diameter that is less than 300 nm. In some embodiments, a lipid nanoparticle has a diameter between about 10 and about 300 nm. In some embodiments, a lipid nanoparticle has a diameter that is less than 200 nm. In some embodiments, a lipid nanoparticle has a diameter between about 25 and about 200 nm. In some other embodiments, the lipid particles comprising a therapeutic nucleic acid and/or an immunosuppressant typically have a mean diameter of from about 20 nm to about 100 nm, 30 nm to about 150 nm, from about 40 nm to about 150 nm, from about 50 nm to about 150 nm, from about 60 nm to about 130 nm, from about 70 nm to about 110 nm, from about 70 nm to about 100 nm, from about 80 nm to about 100 nm, from about 90 nm to about 100 nm, from about 70 to about 90 nm, from about 80 nm to about 90 nm, from about 70 nm to about 80 nm, or about 30 nm, 35 nm, 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm, 110 nm, 115 nm, 120 nm, 125 nm, 130 nm, 135 nm, 140 nm, 145 nm, or 150 nm to ensure effective delivery. Nucleic acid containing lipid particles and their method of preparation are disclosed in, *e.g.*, PCT/US18/50042, U.S. Patent Publication Nos. 20040142025 and 20070042031, the disclosures of which are herein incorporated by reference in their entirety for all purposes. In some embodiments, a lipid nanoparticle preparation (*e.g.*, composition comprising a plurality of lipid nanoparticles) has a size distribution in which the mean size (*e.g.*, diameter) is about 70 nm to about 200 nm, and more typically the mean size is about 100 nm or less.

[00266] According to some embodiments, a liquid pharmaceutical composition comprising a nucleic acid of the present invention may be formulated in lipid particles. In some embodiments, the lipid particle comprising a nucleic acid can be formed from a cationic lipid. In some other embodiments, the lipid particle comprising a nucleic acid can be formed from non-cationic lipid. In a preferred embodiment, the lipid particle of the invention is a nucleic acid containing lipid particle, which is formed from a cationic lipid comprising a nucleic acid selected from the group consisting of

mRNA, antisense RNA and oligonucleotide, ribozymes, aptamer, interfering RNAs (RNAi), Dicer-substrate dsRNA, small hairpin RNA (shRNA), asymmetrical interfering RNA (aiRNA), microRNA (miRNA), minicircle DNA, minigene, viral DNA (*e.g.*, Lentiviral or AAV genome) or non-viral synthetic DNA vectors, closed-ended linear duplex DNA (ceDNA / CELiD), plasmids, bacmids, doggybone (dbDNA™) DNA vectors, minimalistic immunological-defined gene expression (MIDGE)-vector, nonviral ministring DNA vector (linear-covalently closed DNA vector), or dumbbell-shaped DNA minimal vector (“dumbbell DNA”).

[00267] Various lipid nanoparticles known in the art can be used to deliver a closed-ended DNA vector, including a ceDNA vector as described herein. For example, various delivery methods using lipid nanoparticles are described in U.S. Patent Nos. 9,404,127, 9,006,417 and 9,518,272.

[00268] In some embodiments, a closed-ended DNA vector, including a ceDNA vector as described herein, is delivered by a gold nanoparticle. Generally, a nucleic acid can be covalently bound to a gold nanoparticle or non-covalently bound to a gold nanoparticle (*e.g.*, bound by a charge-charge interaction), for example as described by Ding *et al.* (2014). *Gold Nanoparticles for Nucleic Acid Delivery*. Mol. Ther. 22(6); 1075-1083. In some embodiments, gold nanoparticle-nucleic acid conjugates are produced using methods described, for example, in U.S. Patent No. 6,812,334.

[00269] The presence of the ceDNA vector for expression of a desired transgene or therapeutic protein can be confirmed by digesting the vector DNA isolated from the cells with a restriction enzyme having a single recognition site on the DNA vector and analyzing both digested and undigested DNA material using gel electrophoresis to confirm the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA. **FIG. 4C** and **FIG. 4D** illustrate one embodiment for identifying the presence of the closed ended ceDNA vectors produced by the processes herein.

B. ceDNA Plasmid

[00270] A ceDNA-plasmid is a plasmid used for later production of a ceDNA vector for expression of a desired transgene or therapeutic protein. In some embodiments, a ceDNA-plasmid can be constructed using known techniques to provide at least the following as operatively linked components in the direction of transcription: (1) a modified 5' ITR sequence; (2) an expression cassette containing a cis-regulatory element, for example, a promoter, inducible promoter, regulatory switch, enhancers and the like; and (3) a modified 3' ITR sequence, where the 3' ITR sequence is symmetric relative to the 5' ITR sequence. In some embodiments, the expression cassette flanked by the ITRs comprises a cloning site for introducing an exogenous sequence. The expression cassette replaces the rep and cap coding regions of the AAV genomes.

[00271] In one aspect, a ceDNA vector for expression of a desired transgene or therapeutic protein is obtained from a plasmid, referred to herein as a “ceDNA-plasmid” encoding in this order: a first adeno-associated virus (AAV) inverted terminal repeat (ITR), an expression cassette comprising a transgene, and a mutated or modified AAV ITR, wherein said ceDNA-plasmid is devoid of AAV

capsid protein coding sequences. In alternative embodiments, the ceDNA-plasmid encodes in this order: a first (or 5') modified or mutated AAV ITR, an expression cassette comprising a transgene, and a second (or 3') modified AAV ITR, wherein said ceDNA-plasmid is devoid of AAV capsid protein coding sequences, and wherein the 5' and 3' ITRs are symmetric relative to each other. In alternative embodiments, the ceDNA-plasmid encodes in this order: a first (or 5') modified or mutated AAV ITR, an expression cassette comprising a transgene, and a second (or 3') mutated or modified AAV ITR, wherein said ceDNA-plasmid is devoid of AAV capsid protein coding sequences, and wherein the 5' and 3' modified ITRs have the same modifications (*i.e.*, they are inverse complement or symmetric relative to each other).

[00272] In a further embodiment, the ceDNA-plasmid system is devoid of viral capsid protein coding sequences (*i.e.* it is devoid of AAV capsid genes but also of capsid genes of other viruses). In addition, in a particular embodiment, the ceDNA-plasmid is also devoid of AAV Rep protein coding sequences. Accordingly, in a preferred embodiment, ceDNA-plasmid is devoid of functional AAV cap and AAV rep genes GG-3' for AAV2) plus a variable palindromic sequence allowing for hairpin formation.

[00273] A ceDNA-plasmid of the present invention can be generated using natural nucleotide sequences of the genomes of any AAV serotypes well known in the art. In one embodiment, the ceDNA-plasmid backbone is derived from the AAV1, AAV2, AAV3, AAV4, AAV5, AAV 5, AAV7, AAV8, AAV9, AAV10, AAV 11, AAV12, AAVrh8, AAVrh10, AAV-DJ, and AAV-DJ8 genome. *E.g.*, NCBI: NC 002077; NC 001401; NC001729; NC001829; NC006152; NC 006260; NC 006261; Kotin and Smith, The Springer Index of Viruses, available at the URL maintained by Springer (at www.oesys.springer.de/viruses/database/mkchapter.asp?virID=42.04.) (note -references to a URL or database refer to the contents of the URL or database as of the effective filing date of this application) In a particular embodiment, the ceDNA-plasmid backbone is derived from the AAV2 genome. In another particular embodiment, the ceDNA-plasmid backbone is a synthetic backbone genetically engineered to include at its 5' and 3' ITRs derived from one of these AAV genomes.

[00274] A ceDNA-plasmid can optionally include a selectable or selection marker for use in the establishment of a ceDNA vector-producing cell line. In one embodiment, the selection marker can be inserted downstream (*i.e.*, 3') of the 3' ITR sequence. In another embodiment, the selection marker can be inserted upstream (*i.e.*, 5') of the 5' ITR sequence. Appropriate selection markers include, for example, those that confer drug resistance. Selection markers can be, for example, a blasticidin S-resistance gene, kanamycin, geneticin, and the like. In a preferred embodiment, the drug selection marker is a blasticidin S-resistance gene.

[00275] An exemplary ceDNA (*e.g.*, rAAV0) vector for expression of a desired transgene or therapeutic protein is produced from an rAAV plasmid. A method for the production of a rAAV vector, can comprise: (a) providing a host cell with a rAAV plasmid as described above, wherein both the host cell and the plasmid are devoid of capsid protein encoding genes, (b) culturing the host cell

under conditions allowing production of an ceDNA genome, and (c) harvesting the cells and isolating the AAV genome produced from said cells.

C. Exemplary method of making the ceDNA vectors from ceDNA plasmids

[00276] Methods for making capsid-less ceDNA vectors for expression of a desired transgene or therapeutic protein are also provided herein, notably a method with a sufficiently high yield to provide sufficient vector for *in vivo* experiments.

[00277] In some embodiments, a method for the production of a ceDNA vector for expression of a desired transgene or therapeutic protein comprises the steps of: (1) introducing the nucleic acid construct comprising an expression cassette and two symmetric ITR sequences into a host cell (*e.g.*, Sf9 cells), (2) optionally, establishing a clonal cell line, for example, by using a selection marker present on the plasmid, (3) introducing a Rep coding gene (either by transfection or infection with a baculovirus carrying said gene) into said insect cell, and (4) harvesting the cell and purifying the ceDNA vector. The nucleic acid construct comprising an expression cassette and two ITR sequences described above for the production of ceDNA vector can be in the form of a ceDNA plasmid, or Bacmid or Baculovirus generated with the ceDNA plasmid as described below. The nucleic acid construct can be introduced into a host cell by transfection, viral transduction, stable integration, or other methods known in the art.

D. Cell lines

[00278] Host cell lines used in the production of a ceDNA vector for expression of a desired transgene or therapeutic protein can include insect cell lines derived from *Spodoptera frugiperda*, such as Sf9 Sf21, or *Trichoplusia ni* cell, or other invertebrate, vertebrate, or other eukaryotic cell lines including mammalian cells. Other cell lines known to an ordinarily skilled artisan can also be used, such as HEK293, Huh-7, HeLa, HepG2, Hep1A, 911, CHO, COS, MeWo, NIH3T3, A549, HT1 180, monocytes, and mature and immature dendritic cells. Host cell lines can be transfected for stable expression of the ceDNA-plasmid for high yield ceDNA vector production.

[00279] CeDNA-plasmids can be introduced into Sf9 cells by transient transfection using reagents (*e.g.*, liposomal, calcium phosphate) or physical means (*e.g.*, electroporation) known in the art. Alternatively, stable Sf9 cell lines which have stably integrated the ceDNA-plasmid into their genomes can be established. Such stable cell lines can be established by incorporating a selection marker into the ceDNA -plasmid as described above. If the ceDNA -plasmid used to transfect the cell line includes a selection marker, such as an antibiotic, cells that have been transfected with the ceDNA-plasmid and integrated the ceDNA-plasmid DNA into their genome can be selected for by addition of the antibiotic to the cell growth media. Resistant clones of the cells can then be isolated by single-cell dilution or colony transfer techniques and propagated.

E. Isolating and Purifying ceDNA vectors

[00280] Examples of the process for obtaining and isolating ceDNA vectors are described in **FIGS. 4A-4E** and the specific examples below. ceDNA-vectors for expression of a desired transgene or

therapeutic protein disclosed herein can be obtained from a producer cell expressing AAV Rep protein(s), further transformed with a ceDNA-plasmid, ceDNA-bacmid, or ceDNA-baculovirus. Plasmids useful for the production of ceDNA vectors include plasmids that encode a desired transgene or therapeutic protein, or plasmids encoding one or more REP proteins.

[00281] In one aspect, a polynucleotide encodes the AAV Rep protein (Rep 78 or 68) delivered to a producer cell in a plasmid (Rep-plasmid), a bacmid (Rep-bacmid), or a baculovirus (Rep-baculovirus). The Rep-plasmid, Rep-bacmid, and Rep-baculovirus can be generated by methods described above.

[00282] Methods to produce a ceDNA vector for expression of a desired transgene or therapeutic protein are described herein. Expression constructs used for generating a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein can be a plasmid (*e.g.*, ceDNA-plasmids), a Bacmid (*e.g.*, ceDNA-bacmid), and/or a baculovirus (*e.g.*, ceDNA-baculovirus). By way of an example only, a ceDNA-vector can be generated from the cells co-infected with ceDNA-baculovirus and Rep-baculovirus. Rep proteins produced from the Rep-baculovirus can replicate the ceDNA-baculovirus to generate ceDNA-vectors. Alternatively, ceDNA vectors for expression of a desired transgene or therapeutic protein can be generated from the cells stably transfected with a construct comprising a sequence encoding the AAV Rep protein (Rep78/52) delivered in Rep-plasmids, Rep-bacmids, or Rep-baculovirus. CeDNA-Baculovirus can be transiently transfected to the cells, be replicated by Rep protein and produce ceDNA vectors.

[00283] The bacmid (*e.g.*, ceDNA-bacmid) can be transfected into permissive insect cells such as Sf9, Sf21, Tni (*Trichoplusia ni*) cell, High Five cell, and generate ceDNA-baculovirus, which is a recombinant baculovirus including the sequences comprising the symmetric ITRs and the expression cassette. ceDNA-baculovirus can be again infected into the insect cells to obtain a next generation of the recombinant baculovirus. Optionally, the step can be repeated once or multiple times to produce the recombinant baculovirus in a larger quantity.

[00284] The time for harvesting and collecting ceDNA vectors for expression of a desired transgene or therapeutic protein as described herein from the cells can be selected and optimized to achieve a high-yield production of the ceDNA vectors. For example, the harvest time can be selected in view of cell viability, cell morphology, cell growth, etc. Usually, cells can be harvested after sufficient time after baculoviral infection to produce ceDNA vectors (*e.g.*, ceDNA vectors) but before majority of cells start to die because of the viral toxicity. The ceDNA-vectors can be isolated from the Sf9 cells using plasmid purification kits such as Qiagen ENDO-FREE PLASMID® kits. Other methods developed for plasmid isolation can be also adapted for ceDNA vectors. Generally, any art-known nucleic acid purification methods can be adopted, as well as commercially available DNA extraction kits.

[00285] Alternatively, purification can be implemented by subjecting a cell pellet to an alkaline lysis process, centrifuging the resulting lysate and performing chromatographic separation. As one

non-limiting example, the process can be performed by loading the supernatant on an ion exchange column (*e.g.* SARTOBIND Q®) which retains nucleic acids, and then eluting (*e.g.* with a 1.2 M NaCl solution) and performing a further chromatographic purification on a gel filtration column (*e.g.* 6 fast flow GE). The capsid-free AAV vector is then recovered by, *e.g.*, precipitation.

[00286] In some embodiments, ceDNA vectors for expression of a desired transgene or therapeutic protein can also be purified in the form of exosomes, or microparticles. It is known in the art that many cell types release not only soluble proteins, but also complex protein/nucleic acid cargoes via membrane microvesicle shedding (Cocucci *et al.*, 2009; EP 10306226.1) Such vesicles include microvesicles (also referred to as microparticles) and exosomes (also referred to as nanovesicles), both of which comprise proteins and RNA as cargo. Microvesicles are generated from the direct budding of the plasma membrane, and exosomes are released into the extracellular environment upon fusion of multivesicular endosomes with the plasma membrane. Thus, ceDNA vector-containing microvesicles and/or exosomes can be isolated from cells that have been transduced with the ceDNA-plasmid or a bacmid or baculovirus generated with the ceDNA-plasmid.

[00287] Microvesicles can be isolated by subjecting culture medium to filtration or ultracentrifugation at 20,000 x g, and exosomes at 100,000 x g. The optimal duration of ultracentrifugation can be experimentally-determined and will depend on the particular cell type from which the vesicles are isolated. Preferably, the culture medium is first cleared by low-speed centrifugation (*e.g.*, at 2000 x g for 5-20 minutes) and subjected to spin concentration using, *e.g.*, an AMICON® spin column (Millipore, Watford, UK). Microvesicles and exosomes can be further purified via FACS or MACS by using specific antibodies that recognize specific surface antigens present on the microvesicles and exosomes. Other microvesicle and exosome purification methods include, but are not limited to, immunoprecipitation, affinity chromatography, filtration, and magnetic beads coated with specific antibodies or aptamers. Upon purification, vesicles are washed with, *e.g.*, phosphate-buffered saline. One advantage of using microvesicles or exosome to deliver ceDNA-containing vesicles is that these vesicles can be targeted to various cell types by including on their membranes proteins recognized by specific receptors on the respective cell types. (See also EP 10306226)

[00288] Another aspect of the invention herein relates to methods of purifying ceDNA vectors from host cell lines that have stably integrated a ceDNA construct into their own genome. In one embodiment, ceDNA vectors are purified as DNA molecules. In another embodiment, the ceDNA vectors are purified as exosomes or microparticles.

[00289] **FIG. 5** of International application PCT/US18/49996 shows a gel confirming the production of ceDNA from multiple ceDNA-plasmid constructs using the method described in the Examples. The ceDNA is confirmed by a characteristic band pattern in the gel (**FIG. 4D**).

[00290] In some embodiments, the non-viral, capsid-free DNA vector has covalently-closed ends. Such a non-viral, capsid-free DNA vector is also referred to as ceDNA or ceDNA vectors.

Since the ceDNA vector has covalently closed ends, it is preferably resistant to exonuclease digestion (*e.g.* exonuclease I or exonuclease III), *e.g.* for over an hour at 37°C.

[00291] These non-viral capsid free ceDNA vectors can be produced in permissive host cells from an expression construct (*e.g.*, a plasmid, a Bacmid, a baculovirus, or an integrated cell-line) *e.g.*, see the Examples disclosed in International Patent Application PCT/US18/49996 filed on September 7, 2018, or using synthetic production, *e.g.*, see the Examples disclosed in International Patent Application PCT/US19/14122, filed December 6, 2018, each of which are incorporated herein in their entirety by reference. In some embodiments, the ceDNA vectors useful in the methods and compositions as disclosed herein comprise a heterologous gene positioned between two inverted terminal repeat (ITR) sequences. In some embodiments, at least one of the ITRs is modified by deletion, insertion, and/or substitution as compared to a wild-type ITR sequence (*e.g.* AAV ITR); and at least one of the ITRs comprises a functional terminal resolution site (*trs*) and a Rep binding site. The ceDNA vector is preferably duplex, *e.g.*, self-complementary, over at least a portion of the molecule, such as the expression cassette (*e.g.* ceDNA is not a double stranded circular molecular).

[00292] In some embodiments, at least one of the ITRs is an AAV ITR, *e.g.*, a wild type ITR. For example, the polynucleotide vector template described herein contains at least one functional ITR that comprises a Rep-binding site (RBS; *e.g.* 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60) for AAV2) and a functional terminal resolution site (*trs*; *e.g.* 5'-AGTT (SEQ ID NO: 62)).

[00293] In some embodiments, the ceDNA can be obtained from a vector polynucleotide that encodes a heterologous nucleic acid operatively positioned between two inverted terminal repeat sequences (ITRs) (*e.g.* AAV ITRs). In some embodiments, at least one of the ITRs comprises a functional terminal resolution site and a replicative protein binding site (RPS), *e.g.* a Rep binding site (*e.g.* wt AAV ITR SEQ ID NO: 1 or SEQ ID NO: 2 for AAV2), and one of the ITRs comprises a deletion, insertion, or substitution with respect to the other ITR, *e.g.* functional ITR.

[00294] As discussed above, any ITR can be used. For exemplary purposes, the ITRs in the ceDNA constructs are disclosed in Tables 7, 9A-9B and 10 herein, and can be symmetric, or asymmetric with respect to each other, as disclosed and defined herein. However, encompassed herein are ceDNA vectors that contain a heterologous nucleic acid sequence (*e.g.*, a transgene) positioned between two inverted terminal repeat (ITR) sequences, where the ITR sequences can be an asymmetrical ITR pair or a symmetrical- or substantially symmetrical ITR pair, as these terms are defined herein. A ceDNA vector as disclosed herein can comprise ITR sequences that are selected from any of: (i) at least one WT ITR and at least one modified AAV inverted terminal repeat (mod-ITR) (*e.g.*, asymmetric modified ITRs); (ii) two modified ITRs where the mod-ITR pair have a different three-dimensional spatial organization with respect to each other (*e.g.*, asymmetric modified ITRs), or (iii) symmetrical or substantially symmetrical WT-WT ITR pair, where each WT-ITR has the same three-dimensional spatial organization, or (iv) symmetrical or substantially symmetrical modified ITR pair, where each mod-ITR has the same three-dimensional spatial organization, where

the methods of the present disclosure may further include a delivery system, such as but not limited to a liposome nanoparticle delivery system.

[00295] In some embodiments, the methods and compositions described herein relate to the use of a ceDNA vector with a non-fusogenic LNP and an endosomolytic agent where the ceDNA vector is, but is not limited to, a ceDNA vector comprising asymmetric ITRS as disclosed in International Patent Application PCT/US18/49996, filed on September 7, 2018 (see, *e.g.*, Examples 1-4); a ceDNA vector for gene editing as disclosed on the International Patent Application PCT/US18/64242 filed on December 6, 2018 (see, *e.g.*, Examples 1-7), or a ceDNA vector for production of antibodies or fusion proteins, as disclosed in the International Patent Application PCT/US19/18016, filed on February 14, 2019, (*e.g.*, see Examples 1-4), or a ceDNA vector for controlled transgene expression, as disclosed in International Patent Application PCT/US19/18927 filed on February 22, 2019, each of which are incorporated herein in their entireties by reference. In some embodiments, it is also envisioned that the methods and compositions described herein can be used with a synthetically produced ceDNA vector, *e.g.*, a ceDNA vector produced in a cell free or insect-free system of ceDNA production, as disclosed in International Application PCT/US19/14122, filed on January 18, 2019, incorporated by reference in its entirety herein.

[00296] The non-viral capsid-free DNA vector with covalently-closed ends can be obtained by a process comprising the steps of: a) incubating a population of host cells (*e.g.* insect cells or mammalian cells, *e.g.*, 293 cells etc.) harboring the vector polynucleotide, which is devoid of viral capsid coding sequences, in the presence of a Rep protein under conditions effective and for a time sufficient to induce production of the capsid-free, non-viral DNA within the host cells, wherein the host cells do not comprise viral capsid coding sequences; and b) harvesting and isolating the capsid-free, non-viral DNA from the host cells. The presence of Rep protein induces replication of the vector polynucleotide with the modified ITR to produce the ceDNA vector in a host cell. In one embodiment, the presence of the capsid-free, non-viral close-ended vector isolated from the host cells can be confirmed, for example by digesting DNA isolated from the host cell with a restriction enzyme having a single recognition site on the DNA vector and analyzing the digested DNA material on a non-denaturing gel to confirm the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA. An exemplary method for preparing the ceDNA is disclosed in **Example 1**.

[00297] The host cells do not express viral capsid proteins and the polynucleotide vector template is devoid of any viral capsid coding sequences. In one embodiment, the polynucleotide vector template is devoid of AAV capsid genes but also of capsid genes of other viruses). In addition, in a particular embodiment, the nucleic acid molecule is also devoid of AAV Rep protein coding sequences. Accordingly, in a preferred embodiment, the nucleic acid molecule of the invention is devoid of both functional AAV cap and AAV rep genes.

V. Pharmaceutical Compositions

[00298] In another aspect, pharmaceutical compositions are provided. The pharmaceutical composition comprises a non-fusogenic LNP and/or an endosomolytic agent as disclosed herein and a ceDNA vector for expression of a desired transgene or therapeutic protein as described herein and a pharmaceutically acceptable carrier or diluent.

[00299] The ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein can be incorporated into pharmaceutical compositions suitable for administration to a subject for *in vivo* delivery to cells, tissues, or organs of the subject. Typically, the pharmaceutical composition comprises a ceDNA-vector as disclosed herein and a pharmaceutically acceptable carrier. For example, the ceDNA vectors for expression of a desired transgene or therapeutic protein as described herein can be incorporated into a pharmaceutical composition suitable for a desired route of therapeutic administration (*e.g.*, parenteral administration). Passive tissue transduction via high pressure intravenous or intra-arterial infusion, as well as intracellular injection, such as intranuclear microinjection or intracytoplasmic injection, are also contemplated. Pharmaceutical compositions for therapeutic purposes can be formulated as a solution, microemulsion, dispersion, liposomes, or other ordered structure suitable to high ceDNA vector concentration. Sterile injectable solutions can be prepared by incorporating the ceDNA vector compound in the required amount in an appropriate buffer with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization including a ceDNA vector can be formulated to deliver a transgene in the nucleic acid to the cells of a recipient, resulting in the therapeutic expression of the transgene or donor sequence therein. The composition can also include a pharmaceutically acceptable carrier.

[00300] Pharmaceutically active compositions comprising a ceDNA vector for expression of a desired transgene or therapeutic protein can be formulated to deliver a transgene for various purposes to the cell, *e.g.*, cells of a subject.

[00301] Pharmaceutical compositions for therapeutic purposes typically must be sterile and stable under the conditions of manufacture and storage. The composition can be formulated as a solution, microemulsion, dispersion, liposomes, or other ordered structure suitable to high ceDNA vector concentration. Sterile injectable solutions can be prepared by incorporating the ceDNA vector compound in the required amount in an appropriate buffer with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization.

[00302] A ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can be incorporated into a pharmaceutical composition suitable for topical, systemic, intra-amniotic, intrathecal, intracranial, intra-arterial, intravenous, intralymphatic, intraperitoneal, subcutaneous, tracheal, intra-tissue (*e.g.*, intramuscular, intracardiac, intrahepatic, intrarenal, intracerebral), intrathecal, intravesical, conjunctival (*e.g.*, extra-orbital, intraorbital, retroorbital, intraretinal, subretinal, choroidal, sub-choroidal, intrastromal, intracameral and intravitreal), intracochlear, and mucosal (*e.g.*, oral, rectal, nasal) administration. Passive tissue transduction via

high pressure intravenous or intraarterial infusion, as well as intracellular injection, such as intranuclear microinjection or intracytoplasmic injection, are also contemplated.

[00303] In some aspects, the methods provided herein comprise delivering one or more ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein to a host cell. Also provided herein are cells produced by such methods, and organisms (such as animals, plants, or fungi) comprising or produced from such cells. Methods of delivery of nucleic acids can include lipofection, nucleofection, microinjection, biolistics, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, and agent-enhanced uptake of DNA. Lipofection is described in *e.g.*, U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355, incorporated by reference in their entireties herein) and lipofection reagents are sold commercially (*e.g.*, TRANSFECTAM™ and LIPOFECTIN™). Delivery can be to cells (*e.g.*, *in vitro* or *ex vivo* administration) or target tissues (*e.g.*, *in vivo* administration).

VI. Methods of Use

[00304] A ceDNA vector for expression of a desired transgene or therapeutic proteins disclosed herein can also be used in a method for the delivery of a nucleotide sequence of interest (*e.g.*, encoding transgene or therapeutic protein) to a target cell (*e.g.*, a host cell). The method may in particular be a method for delivering a desired transgene or therapeutic protein to a cell of a subject in need thereof and treating a disease. The invention allows for the *in vivo* expression of a desired transgene or therapeutic protein encoded in the ceDNA vector in a cell in a subject such that therapeutic effect of the expression of a desired transgene or therapeutic protein occurs. These results are seen with both *in vivo* and *in vitro* modes of ceDNA vector delivery.

[00305] In addition, the invention provides a method for the delivery of a desired transgene or therapeutic protein in a cell of a subject in need thereof, comprising multiple administrations of the ceDNA vector of the invention encoding said transgene or therapeutic protein. Since the ceDNA vector of the invention does not induce an immune response like that typically observed against encapsidated viral vectors, such a multiple administration strategy will likely have greater success in a ceDNA-based system. The ceDNA vector are administered in sufficient amounts to transfect the cells of a desired tissue and to provide sufficient levels of gene transfer and expression of the a desired transgene or therapeutic protein without undue adverse effects. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, retinal administration (*e.g.*, subretinal injection, suprachoroidal injection or intravitreal injection), intravenous (*e.g.*, in a liposome formulation), direct delivery to the selected organ (*e.g.*, any one or more tissues selected from: liver, kidneys, gallbladder, prostate, adrenal gland, heart, intestine, lung, and stomach), intramuscular, and other parental routes of administration. Routes of administration may be combined, if desired.

[00306] Delivery of a ceDNA vector for expression of a desired transgene or therapeutic proteins as described herein is not limited to delivery of the expressed transgene or therapeutic protein. For example, conventionally produced (*e.g.*, using a cell-based production method (*e.g.*, insect-cell

production methods) or synthetically produced ceDNA vectors as described herein may be used with other delivery systems provided to provide a portion of the gene therapy. One non-limiting example of a system that may be combined with the ceDNA vectors in accordance with the present disclosure includes systems which separately deliver one or more co-factors or immune suppressors for effective gene expression of the ceDNA vector expressing the transgene or therapeutic protein.

[00307] The invention also provides for a method of treating a disease in a subject comprising introducing into a target cell in need thereof (in particular a muscle cell or tissue) of the subject a therapeutically effective amount of a ceDNA vector, optionally with a pharmaceutically acceptable carrier. While the ceDNA vector can be introduced in the presence of a carrier, such a carrier is not required. The ceDNA vector selected comprises a nucleotide sequence encoding a desired transgene or therapeutic protein useful for treating a disease. In particular, the ceDNA vector may comprise a desired transgene or therapeutic protein sequence operably linked to control elements capable of directing transcription of a desired transgene or therapeutic protein encoded by the exogenous DNA sequence when introduced into the subject. The ceDNA vector can be administered via any suitable route as provided above, and elsewhere herein.

[00308] The compositions and vectors provided herein can be used to deliver a desired transgene or therapeutic protein for various purposes. In some embodiments, the transgene encodes a therapeutic protein or transgene that is intended to be used for research purposes, *e.g.*, to create a somatic transgenic animal model harboring the transgene, *e.g.*, to study the function of the transgene or therapeutic protein product. In another example, the transgene encodes a transgene or therapeutic protein that is intended to be used to create an animal model of a disease. In some embodiments, the encoded transgene or therapeutic protein is useful for the treatment or prevention of a disease states in a mammalian subject. The transgene or therapeutic protein can be transferred (*e.g.*, expressed in) to a patient in a sufficient amount to treat a disease associated with reduced expression, lack of expression or dysfunction of the gene.

[00309] In principle, the expression cassette can include a nucleic acid or any transgene that encodes an transgene or therapeutic protein that is either reduced or absent due to a mutation or which conveys a therapeutic benefit when overexpressed is considered to be within the scope of the invention. Preferably, noninserted bacterial DNA is not present and preferably no bacterial DNA is present in the ceDNA compositions provided herein.

[00310] A ceDNA vector is not limited to one species of ceDNA vector. As such, in another aspect, multiple ceDNA vectors expressing different proteins or the same transgene or therapeutic protein but operatively linked to different promoters or cis-regulatory elements can be delivered simultaneously or sequentially to the target cell, tissue, organ, or subject. Therefore, this strategy can allow for the gene therapy or gene delivery of multiple proteins simultaneously. It is also possible to separate different portions of a transgene or therapeutic protein into separate ceDNA vectors (*e.g.*, different domains and/or co-factors required for functionality of a transgene or therapeutic protein) which can

be administered simultaneously or at different times, and can be separately regulatable, thereby adding an additional level of control of expression of a transgene or therapeutic protein. Delivery can also be performed multiple times and, importantly for gene therapy in the clinical setting, in subsequent increasing or decreasing doses, given the lack of an anti-capsid host immune response due to the absence of a viral capsid. It is anticipated that no anti-capsid response will occur as there is no capsid.

[00311] The invention also provides for a method of treating a disease in a subject comprising introducing into a target cell in need thereof (in particular a muscle cell or tissue) of the subject a therapeutically effective amount of a ceDNA vector as disclosed herein, optionally with a pharmaceutically acceptable carrier. While the ceDNA vector can be introduced in the presence of a carrier, such a carrier is not required. The ceDNA vector implemented comprises a nucleotide sequence of interest useful for treating the a disease. In particular, the ceDNA vector may comprise a desired exogenous DNA sequence operably linked to control elements capable of directing transcription of the desired polypeptide, protein, or oligonucleotide encoded by the exogenous DNA sequence when introduced into the subject. The ceDNA vector can be administered via any suitable route as provided above, and elsewhere herein.

VII. Methods of delivering ceDNA vectors for transgene or therapeutic protein production

[00312] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein can be delivered to a target cell *in vitro* or *in vivo* by various suitable methods. ceDNA vectors alone can be applied or injected. CeDNA vectors can be delivered to a cell without the help of a transfection reagent or other physical means. Alternatively, ceDNA vectors for expression of transgene or therapeutic protein can be delivered using any art-known transfection reagent or other art-known physical means that facilitates entry of DNA into a cell, *e.g.*, liposomes, alcohols, polylysine- rich compounds, arginine-rich compounds, calcium phosphate, microvesicles, microinjection, electroporation and the like.

[00313] The ceDNA vectors for expression of transgene or therapeutic protein as disclosed herein can efficiently target cell and tissue-types that are normally difficult to transduce with conventional AAV virions using various delivery reagent.

[00314] One aspect of the technology described herein relates to a method of delivering a transgene or therapeutic protein to a cell. Typically, for *in vivo* and *in vitro* methods, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein may be introduced into the cell using the methods as disclosed herein, as well as other methods known in the art. A ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein are preferably administered to the cell in a biologically-effective amount. If the ceDNA vector is administered to a cell *in vivo* (*e.g.*, to a subject), a biologically-effective amount of the ceDNA vector is an amount that is sufficient to result in transduction and expression of the desired transgene or therapeutic protein in a target cell.

[00315] Exemplary modes of administration of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein includes oral, rectal, transmucosal, intranasal, inhalation (*e.g.*, via an aerosol), buccal (*e.g.*, sublingual), vaginal, intrathecal, intraocular, transdermal, intraendothelial, in utero (or in ovo), parenteral (*e.g.*, intravenous, subcutaneous, intradermal, intracranial, intramuscular [including administration to skeletal, diaphragm and/or cardiac muscle], intrapleural, intracerebral, and intraarticular). Administration can be systemically or direct delivery to the liver or elsewhere (*e.g.*, any kidneys, gallbladder, prostate, adrenal gland, heart, intestine, lung, and stomach).

[00316] Administration can be topical (*e.g.*, to both skin and mucosal surfaces, including airway surfaces, and transdermal administration), intralymphatic, and the like, as well as direct tissue or organ injection (*e.g.*, but not limited to, liver, but also to eye, muscles, including skeletal muscle, cardiac muscle, diaphragm muscle, or brain).

[00317] Administration of the ceDNA vector can be to any site in a subject, including, without limitation, a site selected from the group consisting of the liver and/or also eyes, brain, a skeletal muscle, a smooth muscle, the heart, the diaphragm, the airway epithelium, the kidney, the spleen, the pancreas, the skin.

[00318] The most suitable route in any given case will depend on the nature and severity of the condition being treated, ameliorated, and/or prevented and on the nature of the particular ceDNA vector that is being used. Additionally, ceDNA permits one to administer more than one a transgene or therapeutic protein in a single vector, or multiple ceDNA vectors (*e.g.* a ceDNA cocktail).

A. Intramuscular Administration of a ceDNA vector

[00319] In some embodiments, a method of treating a disease in a subject comprises introducing into a target cell in need thereof (in particular a muscle cell or tissue) of the subject a therapeutically effective amount of a ceDNA vector encoding a desired transgene or therapeutic protein, optionally with a pharmaceutically acceptable carrier. In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein is administered to a muscle tissue of a subject.

[00320] In some embodiments, administration of the ceDNA vector can be to any site in a subject, including, without limitation, a site selected from the group consisting of a skeletal muscle, a smooth muscle, the heart, the diaphragm, or muscles of the eye.

[00321] Administration of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein to a skeletal muscle according to the present invention includes but is not limited to administration to the skeletal muscle in the limbs (*e.g.*, upper arm, lower arm, upper leg, and/or lower leg), back, neck, head (*e.g.*, tongue), thorax, abdomen, pelvis/perineum, and/or digits. The ceDNA as disclosed herein vector can be delivered to skeletal muscle by intravenous administration, intra-arterial administration, intraperitoneal administration, limb perfusion, (optionally, isolated limb perfusion of a leg and/or arm; see, *e.g.* Arruda *et al.*, (2005) Blood 105: 3458-3464), and/or direct intramuscular injection. In particular embodiments, the ceDNA vector as

disclosed herein is administered to the liver, eye, a limb (arm and/or leg) of a subject (*e.g.*, a subject with muscular dystrophy such as DMD) by limb perfusion, optionally isolated limb perfusion (*e.g.*, by intravenous or intra-articular administration. In embodiments, the ceDNA vector as disclosed herein can be administered without employing "hydrodynamic" techniques.

[00322] For instance, tissue delivery (*e.g.*, to retina) of conventional viral vectors is often enhanced by hydrodynamic techniques (*e.g.*, intravenous/intravenous administration in a large volume), which increase pressure in the vasculature and facilitate the ability of the viral vector to cross the endothelial cell barrier. In particular embodiments, the ceDNA vectors described herein can be administered in the absence of hydrodynamic techniques such as high volume infusions and/or elevated intravascular pressure (*e.g.*, greater than normal systolic pressure, for example, less than or equal to a 5%, 10%, 15%, 20%, 25% increase in intravascular pressure over normal systolic pressure). Such methods may reduce or avoid the side effects associated with hydrodynamic techniques such as edema, nerve damage and/or compartment syndrome.

[00323] Furthermore, a composition comprising a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein that is administered to a skeletal muscle can be administered to a skeletal muscle in the limbs (*e.g.*, upper arm, lower arm, upper leg, and/or lower leg), back, neck, head (*e.g.*, tongue), thorax, abdomen, pelvis/perineum, and/or digits. Suitable skeletal muscles include but are not limited to abductor digiti minimi (in the hand), abductor digiti minimi (in the foot), abductor hallucis, abductor ossis metatarsi quinti, abductor pollicis brevis, abductor pollicis longus, adductor brevis, adductor hallucis, adductor longus, adductor magnus, adductor pollicis, anconeus, anterior scalene, articularis genus, biceps brachii, biceps femoris, brachialis, brachioradialis, buccinator, coracobrachialis, corrugator supercilii, deltoid, depressor anguli oris, depressor labii inferioris, digastric, dorsal interossei (in the hand), dorsal interossei (in the foot), extensor carpi radialis brevis, extensor carpi radialis longus, extensor carpi ulnaris, extensor digiti minimi, extensor digitorum, extensor digitorum brevis, extensor digitorum longus, extensor hallucis brevis, extensor hallucis longus, extensor indicis, extensor pollicis brevis, extensor pollicis longus, flexor carpi radialis, flexor carpi ulnaris, flexor digiti minimi brevis (in the hand), flexor digiti minimi brevis (in the foot), flexor digitorum brevis, flexor digitorum longus, flexor digitorum profundus, flexor digitorum superficialis, flexor hallucis brevis, flexor hallucis longus, flexor pollicis brevis, flexor pollicis longus, frontalis, gastrocnemius, geniohyoid, gluteus maximus, gluteus medius, gluteus minimus, gracilis, iliocostalis cervicis, iliocostalis lumborum, iliocostalis thoracis, iliacus, inferior gemellus, inferior oblique, inferior rectus, infrapinnatus, interspinalis, intertransversi, lateral pterygoid, lateral rectus, latissimus dorsi, levator anguli oris, levator labii superioris, levator labii superioris alaeque nasi, levator palpebrae superioris, levator scapulae, long rotators, longissimus capitis, longissimus cervicis, longissimus thoracis, longus capitis, longus colli, lumbricals (in the hand), lumbricals (in the foot), masseter, medial pterygoid, medial rectus, middle scalene, multifidus, mylohyoid, obliquus capitis inferior, obliquus capitis superior, obturator externus, obturator internus,

occipitalis, omohyoid, opponens digiti minimi, opponens pollicis, orbicularis oculi, orbicularis oris, palmar interossei, palmaris brevis, palmaris longus, pectineus, pectoralis major, pectoralis minor, peroneus brevis, peroneus longus, peroneus tertius, piriformis, plantar interossei, plantaris, platysma, popliteus, posterior scalene, pronator quadratus, pronator teres, psoas major, quadratus femoris, quadratus plantae, rectus capitis anterior, rectus capitis lateralis, rectus capitis posterior major, rectus capitis posterior minor, rectus femoris, rhomboid major, rhomboid minor, risorius, sartorius, scalenus minimus, semimembranosus, semispinalis capitis, semispinalis cervicis, semispinalis thoracis, semitendinosus, serratus anterior, short rotators, soleus, spinalis capitis, spinalis cervicis, spinalis thoracis, splenius capitis, splenius cervicis, sternocleidomastoid, sternohyoid, sternothyroid, stylohyoid, subclavius, subscapularis, superior gemellus, superior oblique, superior rectus, supinator, supraspinatus, temporalis, tensor fascia lata, teres major, teres minor, thoracis, thyrohyoid, tibialis anterior, tibialis posterior, trapezius, triceps brachii, vastus intermedius, vastus lateralis, vastus medialis, zygomaticus major, and zygomaticus minor, and any other suitable skeletal muscle as known in the art.

[00324] Administration of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein to diaphragm muscle can be by any suitable method including intravenous administration, intra-arterial administration, and/or intra-peritoneal administration. In some embodiments, delivery of an expressed transgene from the ceDNA vector to a target tissue can also be achieved by delivering a synthetic depot comprising the ceDNA vector, where a depot comprising the ceDNA vector is implanted into skeletal, smooth, cardiac and/or diaphragm muscle tissue or the muscle tissue can be contacted with a film or other matrix comprising the ceDNA vector as described herein. Such implantable matrices or substrates are described in U.S. Pat. No. 7,201,898.

[00325] Administration of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein to cardiac muscle includes administration to the left atrium, right atrium, left ventricle, right ventricle and/or septum. The ceDNA vector as described herein can be delivered to cardiac muscle by intravenous administration, intra-arterial administration such as intra-aortic administration, direct cardiac injection (*e.g.*, into left atrium, right atrium, left ventricle, right ventricle), and/or coronary artery perfusion.

[00326] Administration of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein to smooth muscle can be by any suitable method including intravenous administration, intra-arterial administration, and/or intra-peritoneal administration. In one embodiment, administration can be to endothelial cells present in, near, and/or on smooth muscle. Non-limiting examples of smooth muscles include the iris of the eye, bronchioles of the lung, laryngeal muscles (vocal cords), muscular layers of the stomach, esophagus, small and large intestine of the gastrointestinal tract, ureter, detrusor muscle of the urinary bladder, uterine myometrium, penis, or prostate gland.

[00327] In some embodiments, of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is administered to skeletal muscle, diaphragm muscle and/or cardiac muscle. In representative embodiments, a ceDNA vector according to the present invention is used to treat and/or prevent disorders of skeletal, cardiac and/or diaphragm muscle.

[00328] Specifically, it is contemplated that a composition comprising a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can be delivered to one or more muscles of the eye (*e.g.*, Lateral rectus, Medial rectus, Superior rectus, Inferior rectus, Superior oblique, Inferior oblique), facial muscles (*e.g.*, Occipitofrontalis muscle, Temporoparietalis muscle, Procerus muscle, Nasalis muscle, Depressor septi nasi muscle, Orbicularis oculi muscle, Corrugator supercilii muscle, Depressor supercilii muscle, Auricular muscles, Orbicularis oris muscle, Depressor anguli oris muscle, Risorius, Zygomaticus major muscle, Zygomaticus minor muscle, Levator labii superioris, Levator labii superioris alaeque nasi muscle, Depressor labii inferioris muscle, Levator anguli oris, Buccinator muscle, Mentalis) or tongue muscles (*e.g.*, genioglossus, hyoglossus, chondroglossus, styloglossus, palatoglossus, superior longitudinal muscle, inferior longitudinal muscle, the vertical muscle, and the transverse muscle).

(i) *Intramuscular injection:*

[00329] In some embodiments, a composition comprising a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can be injected into one or more sites of a given muscle, for example, skeletal muscle (*e.g.*, deltoid, vastus lateralis, ventrogluteal muscle of dorsogluteal muscle, or anterolateral thigh for infants) in a subject using a needle. The composition comprising ceDNA can be introduced to other subtypes of muscle cells. Non-limiting examples of muscle cell subtypes include skeletal muscle cells, cardiac muscle cells, smooth muscle cells and/or diaphragm muscle cells.

[00330] Methods for intramuscular injection are known to those of skill in the art and as such are not described in detail herein. However, when performing an intramuscular injection, an appropriate needle size should be determined based on the age and size of the patient, the viscosity of the composition, as well as the site of injection. **Table 12** provides guidelines for exemplary sites of injection and corresponding needle size:

Table 12: Guidelines for intramuscular injection in human patients

<i>Injection Site</i>	<i>Needle Gauge</i>	<i>Needle Size</i>	<i>Maximum volume of composition</i>
Ventrogluteal site (gluteus medius and gluteus minimus)	<i>Aqueous solutions:</i> 20-25 gauge	Thin adult: 15 to 25 mm	3mL
		Average adult: 25 mm	
	<i>Viscous or oil-based solution:</i> 18-21 gauge	Larger adult (over 150 lbs): 25 to 38 mm.	
		Children and infants: will require	

		a smaller needle	
Vastus lateralis	<i>Aqueous solutions:</i> 20-25 gauge <i>Viscous or oil-based solution:</i> 18-21 gauge <u>Children/infants:</u> 22 to 25 gauge	Adult: 25 mm to 38 mm	3mL
Deltoid	22 to 25 gauge	Males: 130-260lbs: 25 mm Females: <130 lbs: 16 mm 130-200 lbs: 25mm >200lbs: 38mm	1mL

[00331] In certain embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is formulated in a small volume, for example, an exemplary volume as outlined in **Table 12** for a given subject. In some embodiments, the subject can be administered a general or local anesthetic prior to the injection, if desired. This is particularly desirable if multiple injections are required or if a deeper muscle is injected, rather than the common injection sites noted above.

[00332] In some embodiments, intramuscular injection can be combined with electroporation, delivery pressure or the use of transfection reagents to enhance cellular uptake of the ceDNA vector.

(ii) Transfection Reagents

[00333] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is formulated in compositions comprising one or more transfection reagents to facilitate uptake of the vectors into myotubes or muscle tissue. Thus, in one embodiment, the nucleic acids described herein are administered to a muscle cell, myotube or muscle tissue by transfection using methods described elsewhere herein.

(iii) Electroporation

[00334] In certain embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is administered in the absence of a carrier to facilitate entry of ceDNA into the cells, or in a physiologically inert pharmaceutically acceptable carrier (*i.e.*, any carrier that does not improve or enhance uptake of the capsid free, non-viral vectors into the myotubes). In such embodiments, the uptake of the capsid free, non-viral vector can be facilitated by electroporation of the cell or tissue.

[00335] Cell membranes naturally resist the passage of extracellular into the cell cytoplasm. One method for temporarily reducing this resistance is “electroporation”, where electrical fields are used to create pores in cells without causing permanent damage to the cells. These pores are large enough to

allow DNA vectors, pharmaceutical drugs, DNA, and other polar compounds to gain access to the interior of the cell. With time, the pores in the cell membrane close and the cell once again becomes impermeable.

[00336] Electroporation can be used in both *in vitro* and *in vivo* applications to introduce *e.g.*, exogenous DNA into living cells. *In vitro* applications typically mix a sample of live cells with the composition comprising *e.g.*, DNA. The cells are then placed between electrodes such as parallel plates and an electrical field is applied to the cell/composition mixture.

[00337] There are a number of methods for *in vivo* electroporation; electrodes can be provided in various configurations such as, for example, a caliper that grips the epidermis overlying a region of cells to be treated. Alternatively, needle-shaped electrodes may be inserted into the tissue, to access more deeply located cells. In either case, after the composition comprising *e.g.*, nucleic acids are injected into the treatment region, the electrodes apply an electrical field to the region. In some electroporation applications, this electric field comprises a single square wave pulse on the order of 100 to 500 V/cm. of about 10 to 60 ms duration. Such a pulse may be generated, for example, in known applications of the Electro Square Porator T820, made by the BTX Division of Genetronics, Inc.

[00338] Typically, successful uptake of *e.g.*, nucleic acids occurs only if the muscle is electrically stimulated immediately, or shortly after administration of the composition, for example, by injection into the muscle.

[00339] In certain embodiments, electroporation is achieved using pulses of electric fields or using low voltage/long pulse treatment regimens (*e.g.*, using a square wave pulse electroporation system). Exemplary pulse generators capable of generating a pulsed electric field include, for example, the ECM600, which can generate an exponential wave form, and the ElectroSquarePorator (T820), which can generate a square wave form, both of which are available from BTX, a division of Genetronics, Inc. (San Diego, Calif.). Square wave electroporation systems deliver controlled electric pulses that rise quickly to a set voltage, stay at that level for a set length of time (pulse length), and then quickly drop to zero.

[00340] In some embodiments, a local anesthetic is administered, for example, by injection at the site of treatment to reduce pain that may be associated with electroporation of the tissue in the presence of a composition comprising a capsid free, non-viral vector as described herein. In addition, one of skill in the art will appreciate that a dose of the composition should be chosen that minimizes and/or prevents excessive tissue damage resulting in fibrosis, necrosis or inflammation of the muscle.

(iv) Delivery Pressure

[00341] In some embodiments, delivery of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein to muscle tissue is facilitated by delivery pressure, which uses a combination of large volumes and rapid injection into an artery supplying a limb (*e.g.*, iliac artery). This mode of administration can be achieved through a variety of methods that involve infusing limb

vasculature with a composition comprising a ceDNA vector, typically while the muscle is isolated from the systemic circulation using a tourniquet of vessel clamps. In one method, the composition is circulated through the limb vasculature to permit extravasation into the cells. In another method, the intravascular hydrodynamic pressure is increased to expand vascular beds and increase uptake of the ceDNA vector into the muscle cells or tissue. In one embodiment, the ceDNA composition is administered into an artery.

(v) Lipid Nanoparticle Compositions

[00342] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein for intramuscular delivery are formulated in a composition comprising a non-fusogenic LNP and/or endosomolytic agent as described elsewhere herein.

(vi) Systemic Administration of a ceDNA Vector targeted to Muscle Tissue

[00343] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is formulated to be targeted to the muscle via indirect delivery administration, where the ceDNA is transported to the muscle as opposed to the liver. Accordingly, the technology described herein encompasses indirect administration of compositions comprising a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein to muscle tissue, for example, by systemic administration. Such compositions can be administered topically, intravenously (by bolus or continuous infusion), intracellular injection, intratissue injection, orally, by inhalation, intraperitoneally, subcutaneously, intracavity, and can be delivered by peristaltic means, if desired, or by other means known by those skilled in the art. The agent can be administered systemically, for example, by intravenous infusion, if so desired.

[00344] In some embodiments, uptake of a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein into muscle cells/tissue is increased by using a targeting agent or moiety that preferentially directs the vector to muscle tissue. Thus, in some embodiments, a capsid free, ceDNA vector can be concentrated in muscle tissue as compared to the amount of capsid free ceDNA vectors present in other cells or tissues of the body.

[00345] In some embodiments, the composition comprising a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein further comprises a targeting moiety to muscle cells. In other embodiments, the expressed gene product comprises a targeting moiety specific to the tissue in which it is desired to act. The targeting moiety can include any molecule, or complex of molecules, which is/are capable of targeting, interacting with, coupling with, and/or binding to an intracellular, cell surface, or extracellular biomarker of a cell or tissue. The biomarker can include, for example, a cellular protease, a kinase, a protein, a cell surface receptor, a lipid, and/or fatty acid. Other examples of biomarkers that the targeting moieties can target, interact with, couple with, and/or bind to include molecules associated with a particular disease. For example, the biomarkers can include cell surface receptors implicated in cancer development, such as epidermal growth factor receptor and transferrin receptor. The targeting moieties can include, but are not limited to, synthetic

compounds, natural compounds or products, macromolecular entities, bioengineered molecules (*e.g.*, polypeptides, lipids, polynucleotides, antibodies, antibody fragments), and small entities (*e.g.*, small molecules, neurotransmitters, substrates, ligands, hormones and elemental compounds) that bind to molecules expressed in the target muscle tissue.

[00346] In certain embodiments, the targeting moiety may further comprise a receptor molecule, including, for example, receptors, which naturally recognize a specific desired molecule of a target cell. Such receptor molecules include receptors that have been modified to increase their specificity of interaction with a target molecule, receptors that have been modified to interact with a desired target molecule not naturally recognized by the receptor, and fragments of such receptors (see, *e.g.*, Skerra, 2000, *J. Molecular Recognition*, 13:167-187). A preferred receptor is a chemokine receptor. Exemplary chemokine receptors have been described in, for example, Lapidot *et al.*, 2002, *Exp Hematol*, 30:973-81 and Onuffer *et al.*, 2002, *Trends Pharmacol Sci*, 23:459-67.

[00347] In other embodiments, the additional targeting moiety may comprise a ligand molecule, including, for example, ligands which naturally recognize a specific desired receptor of a target cell, such as a Transferrin (Tf) ligand. Such ligand molecules include ligands that have been modified to increase their specificity of interaction with a target receptor, ligands that have been modified to interact with a desired receptor not naturally recognized by the ligand, and fragments of such ligands.

[00348] In still other embodiments, the targeting moiety may comprise an aptamer. Aptamers are oligonucleotides that are selected to bind specifically to a desired molecular structure of the target cell. Aptamers typically are the products of an affinity selection process similar to the affinity selection of phage display (also known as *in vitro* molecular evolution). The process involves performing several tandem iterations of affinity separation, *e.g.*, using a solid support to which the diseased immunogen is bound, followed by polymerase chain reaction (PCR) to amplify nucleic acids that bound to the immunogens. Each round of affinity separation thus enriches the nucleic acid population for molecules that successfully bind the desired immunogen. In this manner, a random pool of nucleic acids may be “educated” to yield aptamers that specifically bind target molecules. Aptamers typically are RNA, but may be DNA or analogs or derivatives thereof, such as, without limitation, peptide nucleic acids (PNAs) and phosphorothioate nucleic acids.

[00349] In some embodiments, the targeting moiety can comprise a photo-degradable ligand (*i.e.*, a ‘caged’ ligand) that is released, for example, from a focused beam of light such that the capsid free, non-viral vectors or the gene product are targeted to a specific tissue.

[00350] It is also contemplated herein that the compositions be delivered to multiple sites in one or more muscles of the subject. That is, injections can be in at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 70, at least 75, at least 80, at least 85, at least 90, at least 95, at least 100 injections sites. Such sites can be spread over the area of a single muscle or can be distributed among multiple muscles.

B. Administration of the ceDNA vector for expression of a therapeutic protein to non-muscle locations

[00351] In another embodiment, a ceDNA vector for expression of a desired transgene or therapeutic protein is administered to the liver. The ceDNA vector may also be administered to different regions of the eye such as the cornea and/or optic nerve. The ceDNA vector may also be introduced into the spinal cord, brainstem (medulla oblongata, pons), midbrain (hypothalamus, thalamus, epithalamus, pituitary gland, substantia nigra, pineal gland), cerebellum, telencephalon (corpus striatum, cerebrum including the occipital, temporal, parietal and frontal lobes, cortex, basal ganglia, hippocampus and portaamygdala), limbic system, neocortex, corpus striatum, cerebrum, and inferior colliculus.. The ceDNA vector may be delivered into the cerebrospinal fluid (*e.g.*, by lumbar puncture). The ceDNA vector for expression of a desired transgene or therapeutic protein may further be administered intravascularly to the CNS in situations in which the blood-brain barrier has been perturbed (*e.g.*, brain tumor or cerebral infarct).

[00352] In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein can be administered to the desired region(s) of the eye by any route known in the art, including but not limited to, intrathecal, intra-ocular, intracerebral, intraventricular, intravenous (*e.g.*, in the presence of a sugar such as mannitol), intranasal, intra-aural, intra-ocular (*e.g.*, intra-vitreous, sub-retinal, anterior chamber) and peri-ocular (*e.g.*, sub-Tenon's region) delivery as well as intramuscular delivery with retrograde delivery to motor neurons.

[00353] In some embodiments, the ceDNA vector for expression of a desired transgene or therapeutic protein is administered in a liquid formulation by direct injection (*e.g.*, stereotactic injection) to the desired region or compartment in the CNS. In other embodiments, the ceDNA vector can be provided by topical application to the desired region or by intra-nasal administration of an aerosol formulation. Administration to the eye may be by topical application of liquid droplets. As a further alternative, the ceDNA vector can be administered as a solid, slow-release formulation (see, *e.g.*, U.S. Pat. No. 7,201,898). In yet additional embodiments, the ceDNA vector can be used for retrograde transport to treat, ameliorate, and/or prevent diseases and disorders involving motor neurons (*e.g.*, amyotrophic lateral sclerosis (ALS); spinal muscular atrophy (SMA), etc.). For example, the ceDNA vector can be delivered to muscle tissue from which it can migrate into neurons.

C. Ex vivo treatment

[00354] In some embodiments, cells are removed from a subject, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is introduced therein, and the cells are then replaced back into the subject. Methods of removing cells from subject for treatment *ex vivo*, followed by introduction back into the subject are known in the art (see, *e.g.*, U.S. Pat. No. 5,399,346; the disclosure of which is incorporated herein in its entirety). Alternatively, a ceDNA vector is introduced into cells from another subject, into cultured cells, or into cells from any other suitable source, and the cells are administered to a subject in need thereof.

[00355] Cells transduced with a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein are preferably administered to the subject in a "therapeutically-effective amount" in combination with a pharmaceutical carrier. Those skilled in the art will appreciate that the therapeutic effects need not be complete or curative, as long as some benefit is provided to the subject.

[00356] In some embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can encode a desired transgene or therapeutic protein as described herein (sometimes called a transgene or heterologous nucleotide sequence) that is to be produced in a cell *in vitro*, *ex vivo*, or *in vivo*. For example, in contrast to the use of the ceDNA vectors described herein in a method of treatment as discussed herein, in some embodiments a ceDNA vector for expression of a desired transgene or therapeutic protein may be introduced into cultured cells and the expressed a desired transgene or therapeutic protein isolated from the cells, *e.g.*, for the production of antibodies and fusion proteins. In some embodiments, the cultured cells comprising a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can be used for commercial production of antibodies or fusion proteins, *e.g.*, serving as a cell source for small or large scale biomanufacturing of antibodies or fusion proteins. In alternative embodiments, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein is introduced into cells in a host non-human subject, for *in vivo* production of antibodies or fusion proteins, including small scale production as well as for commercial large scale a desired transgene or therapeutic protein production.

[00357] The ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein can be used in both veterinary and medical applications. Suitable subjects for *ex vivo* gene delivery methods as described above include both avians (*e.g.*, chickens, ducks, geese, quail, turkeys and pheasants) and mammals (*e.g.*, humans, bovines, ovines, caprines, equines, felines, canines, and lagomorphs), with mammals being preferred. Human subjects are most preferred. Human subjects include neonates, infants, juveniles, and adults.

D. Dose ranges

[00358] Provided herein are methods of treatment comprising administering to the subject an effective amount of a composition comprising a ceDNA vector encoding a desired transgene or therapeutic protein as described herein. As will be appreciated by a skilled practitioner, the term "effective amount" refers to the amount of the ceDNA composition administered that results in expression of the transgene or therapeutic protein in a "therapeutically effective amount" for the treatment of a disease.

[00359] *In vivo* and/or *in vitro* assays can optionally be employed to help identify optimal dosage ranges for use. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the condition, and should be decided according to the judgment

of the person of ordinary skill in the art and each subject's circumstances. Effective doses can be extrapolated from dose-response curves derived from *in vitro* or animal model test systems, *e.g.*, .

[00360] A ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein is administered in sufficient amounts to transfect the cells of a desired tissue and to provide sufficient levels of gene transfer and expression without undue adverse effects. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, those described above in the "Administration" section, such as direct delivery to the selected organ (*e.g.*, intraportal delivery to the liver), oral, inhalation (including intranasal and intratracheal delivery), intraocular, intravenous, intramuscular, subcutaneous, intradermal, intratumoral, and other parental routes of administration. Routes of administration can be combined, if desired.

[00361] The dose of the amount of a ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein required to achieve a particular "therapeutic effect," will vary based on several factors including, but not limited to: the route of nucleic acid administration, the level of gene or RNA expression required to achieve a therapeutic effect, the specific disease or disorder being treated, and the stability of the gene(s), RNA product(s), or resulting expressed protein(s). One of skill in the art can readily determine a ceDNA vector dose range to treat a patient having a particular disease or disorder based on the aforementioned factors, as well as other factors that are well known in the art.

[00362] Dosage regime can be adjusted to provide the optimum therapeutic response. For example, the oligonucleotide can be repeatedly administered, *e.g.*, several doses can be administered daily or the dose can be proportionally reduced as indicated by the exigencies of the therapeutic situation. One of ordinary skill in the art will readily be able to determine appropriate doses and schedules of administration of the subject oligonucleotides, whether the oligonucleotides are to be administered to cells or to subjects.

[00363] A "therapeutically effective dose" will fall in a relatively broad range that can be determined through clinical trials and will depend on the particular application (neural cells will require very small amounts, while systemic injection would require large amounts). For example, for direct *in vivo* injection into skeletal or cardiac muscle of a human subject, a therapeutically effective dose will be on the order of from about 1 µg to 100 g of the ceDNA vector. If exosomes or microparticles are used to deliver the ceDNA vector, then a therapeutically effective dose can be determined experimentally, but is expected to deliver from 1 µg to about 100 g of vector. Moreover, a therapeutically effective dose is an amount ceDNA vector that expresses a sufficient amount of the transgene to have an effect on the subject that results in a reduction in one or more symptoms of the disease, but does not result in significant off-target or significant adverse side effects. In one embodiment, a "therapeutically effective amount" is an amount of an expressed a desired transgene or therapeutic protein that is sufficient to produce a statistically significant, measurable change in

expression of a disease biomarker or reduction of a given disease symptom. Such effective amounts can be gauged in clinical trials as well as animal studies for a given ceDNA vector composition.

[00364] Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens.

[00365] For *in vitro* transfection, an effective amount of a ceDNA vectors for expression of a desired transgene or therapeutic protein as disclosed herein to be delivered to cells (1×10^6 cells) will be on the order of 0.1 to 100 μg ceDNA vector, preferably 1 to 20 μg , and more preferably 1 to 15 μg or 8 to 10 μg . Larger ceDNA vectors will require higher doses. If exosomes or microparticles are used, an effective *in vitro* dose can be determined experimentally but would be intended to deliver generally the same amount of the ceDNA vector.

[00366] For the treatment of a disease, the appropriate dosage of a ceDNA vector that expresses a desired transgene or therapeutic protein as disclosed herein will depend on the specific type of disease to be treated, the type of therapeutic protein, the severity and course of the a disease, previous therapy, the patient's clinical history and response to the antibody, and the discretion of the attending physician. The ceDNA vector encoding a desired transgene or therapeutic protein is suitably administered to the patient at one time or over a series of treatments. Various dosing schedules including, but not limited to, single or multiple administrations over various time-points, bolus administration, and pulse infusion are contemplated herein.

[00367] Depending on the type and severity of the disease, a ceDNA vector is administered in an amount that the transgene or therapeutic protein is expressed at about 0.3 mg/kg to 100 mg/kg (*e.g.* 15 mg/kg-100 mg/kg, or any dosage within that range), by one or more separate administrations, or by continuous infusion. One typical daily dosage of the ceDNA vector is sufficient to result in the expression of the encoded transgene or therapeutic protein at a range from about 15 mg/kg to 100 mg/kg or more, depending on the factors mentioned above. One exemplary dose of the ceDNA vector is an amount sufficient to result in the expression of the encoded transgene or therapeutic protein as disclosed herein in a range from about 10 mg/kg to about 50 mg/kg. Thus, one or more doses of a ceDNA vector in an amount sufficient to result in the expression of the encoded transgene or therapeutic protein at about 0.5 mg/kg, 1 mg/kg, 1.5 mg/kg, 2.0 mg/kg, 3 mg/kg, 4.0 mg/kg, 5 mg/kg, 10 mg/kg, 15 mg/kg, 20 mg/kg, 25 mg/kg, 30 mg/kg, 35 mg/kg, 40 mg/kg, 50 mg/kg, 60 mg/kg, 70 mg/kg, 80 mg/kg, 90 mg/kg, or 100 mg/kg (or any combination thereof) may be administered to the patient. In some embodiments, the ceDNA vector is an amount sufficient to result in the expression of the encoded transgene or therapeutic protein for a total dose in the range of 50 mg to 2500 mg. An exemplary dose of a ceDNA vector is an amount sufficient to result in the total expression of the encoded transgene or therapeutic protein at about 50 mg, about 100 mg, 200 mg, 300 mg, 400 mg, about 500 mg, about 600 mg, about 700 mg, about 720 mg, about 1000 mg, about 1050 mg, about 1100 mg, about 1200 mg, about 1300 mg, about 1400 mg, about 1500 mg, about 1600 mg, about 1700

mg, about 1800 mg, about 1900 mg, about 2000 mg, about 2050 mg, about 2100 mg, about 2200 mg, about 2300 mg, about 2400 mg, or about 2500 mg (or any combination thereof). As the expression of the transgene or therapeutic protein from ceDNA vector can be carefully controlled by regulatory switches herein, or alternatively multiple dose of the ceDNA vector administered to the subject, the expression of the transgene or therapeutic protein from the ceDNA vector can be controlled in such a way that the doses of the expressed transgene or therapeutic protein may be administered intermittently, *e.g.* every week, every two weeks, every three weeks, every four weeks, every month, every two months, every three months, or every six months from the ceDNA vector. The progress of this therapy can be monitored by conventional techniques and assays.

[00368] In certain embodiments, a ceDNA vector is administered an amount sufficient to result in the expression of the encoded transgene or therapeutic protein at a dose of 15 mg/kg, 30 mg/kg, 40 mg/kg, 45 mg/kg, 50 mg/kg, 60 mg/kg or a flat dose, *e.g.*, 300 mg, 500 mg, 700 mg, 800 mg, or higher. In some embodiments, the expression of the transgene or therapeutic protein from the ceDNA vector is controlled such that the transgene or therapeutic protein is expressed every day, every other day, every week, every 2 weeks or every 4 weeks for a period of time. In some embodiments, the expression of the transgene or therapeutic protein from the ceDNA vector is controlled such that the transgene or therapeutic protein is expressed every 2 weeks or every 4 weeks for a period of time. In certain embodiments, the period of time is 6 months, one year, eighteen months, two years, five years, ten years, 15 years, 20 years, or the lifetime of the patient.

[00369] Treatment can involve administration of a single dose or multiple doses. In some embodiments, more than one dose can be administered to a subject; in fact, multiple doses can be administered as needed, because the ceDNA vector elicits does not elicit an anti-capsid host immune response due to the absence of a viral capsid. As such, one of skill in the art can readily determine an appropriate number of doses. The number of doses administered can, for example, be on the order of 1-100, preferably 2-20 doses.

[00370] Without wishing to be bound by any particular theory, the lack of typical anti-viral immune response elicited by administration of a ceDNA vector as described by the disclosure (*i.e.*, the absence of capsid components) allows the ceDNA vector for expression of transgene or therapeutic protein to be administered to a host on multiple occasions. In some embodiments, the number of occasions in which a heterologous nucleic acid is delivered to a subject is in a range of 2 to 10 times (*e.g.*, 2, 3, 4, 5, 6, 7, 8, 9, or 10 times). In some embodiments, a ceDNA vector is delivered to a subject more than 10 times.

[00371] In some embodiments, a dose of a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein is administered to a subject no more than once per calendar day (*e.g.*, a 24-hour period). In some embodiments, a dose of a ceDNA vector is administered to a subject no more than once per 2, 3, 4, 5, 6, or 7 calendar days. In some embodiments, a dose of a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein is administered to a subject no more

than once per calendar week (*e.g.*, 7 calendar days). In some embodiments, a dose of a ceDNA vector is administered to a subject no more than bi-weekly (*e.g.*, once in a two calendar week period). In some embodiments, a dose of a ceDNA vector is administered to a subject no more than once per calendar month (*e.g.*, once in 30 calendar days). In some embodiments, a dose of a ceDNA vector is administered to a subject no more than once per six calendar months. In some embodiments, a dose of a ceDNA vector is administered to a subject no more than once per calendar year (*e.g.*, 365 days or 366 days in a leap year).

[00372] In particular embodiments, more than one administration (*e.g.*, two, three, four or more administrations) of a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein may be employed to achieve the desired level of gene expression over a period of various intervals, *e.g.*, daily, weekly, monthly, yearly, etc.

[00373] In some embodiments, a therapeutic a transgene or therapeutic protein encoded by a ceDNA vector as disclosed herein can be regulated by a regulatory switch, inducible or repressible promotor so that it is expressed in a subject for at least 1 hour, at least 2 hours, at least 5 hours, at least 10 hours, at least 12 hours, at least 18 hours, at least 24 hours, at least 36 hours, at least 48 hours, at least 72 hours, at least 1 week, at least 2 weeks, at least 1 month, at least 2 months, at least 6 months, at least 12 months/one year, at least 2 years, at least 5 years, at least 10 years, at least 15 years, at least 20 years, at least 30 years, at least 40 years, at least 50 years or more. In one embodiment, the expression can be achieved by repeated administration of the ceDNA vectors described herein at predetermined or desired intervals. Alternatively, a ceDNA vector for expression of a desired transgene or therapeutic protein as disclosed herein can further comprise components of a gene editing system (*e.g.*, CRISPR/Cas, TALENs, zinc finger endonucleases etc) to permit insertion of the one or more nucleic acid sequences encoding the transgene or therapeutic protein for substantially permanent treatment or “curing” the disease. Such ceDNA vectors comprising gene editing components are disclosed in International Application PCT/US18/64242, and can include the 5’ and 3’ homology arms (*e.g.*, SEQ ID NO: 151-154, or sequences with at least 40%, 50%, 60%, 70% or 80% homology thereto) for insertion of the nucleic acid encoding a transgene or therapeutic protein into safe harbor regions, such as, but not including albumin gene or CCR5 gene. By way of example, a ceDNA vector expressing a transgene or therapeutic protein can comprise at least one genomic safe harbor (GSH)-specific homology arms for insertion of the a transgene into a genomic safe harbor is disclosed in International Patent Application PCT/US2019/020225, filed on March 1, 2019, which is incorporated herein in its entirety by reference.

[00374] The duration of treatment depends upon the subject’s clinical progress and responsiveness to therapy. Continuous, relatively low maintenance doses are contemplated after an initial higher therapeutic dose.

E. Unit dosage forms

[00375] In some embodiments, the pharmaceutical compositions comprising a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can conveniently be presented in unit dosage form. A unit dosage form will typically be adapted to one or more specific routes of administration of the pharmaceutical composition. In some embodiments, the unit dosage form is adapted for droplets to be administered directly to the eye. In some embodiments, the unit dosage form is adapted for administration by inhalation. In some embodiments, the unit dosage form is adapted for administration by a vaporizer. In some embodiments, the unit dosage form is adapted for administration by a nebulizer. In some embodiments, the unit dosage form is adapted for administration by an aerosolizer. In some embodiments, the unit dosage form is adapted for oral administration, for buccal administration, or for sublingual administration. In some embodiments, the unit dosage form is adapted for intravenous, intramuscular, or subcutaneous administration. In some embodiments, the unit dosage form is adapted for subretinal injection, suprachoroidal injection or intravitreal injection.

[00376] In some embodiments, the unit dosage form is adapted for intrathecal or intracerebroventricular administration. In some embodiments, the pharmaceutical composition is formulated for topical administration. The amount of active ingredient which can be combined with a carrier material to produce a single dosage form will generally be that amount of the compound which produces a therapeutic effect.

VIII. Methods of Treatment

[00377] The technology described herein also demonstrates methods for making, as well as methods of using the disclosed ceDNA vectors for expression of a desired transgene or therapeutic protein in a variety of ways, including, for example, *ex vivo*, *ex situ*, *in vitro* and *in vivo* applications, methodologies, diagnostic procedures, and/or gene therapy regimens.

[00378] In one embodiment, the expressed therapeutic transgene or therapeutic protein expressed from a ceDNA vector as disclosed herein is functional for the treatment of disease. In a preferred embodiment, the therapeutic transgene or therapeutic protein does not cause an immune system reaction, unless so desired.

[00379] Provided herein is a method of treating a disease in a subject comprising introducing into a target cell in need thereof (for example, a muscle cell or tissue, or other affected cell type) of the subject a therapeutically effective amount of a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein, optionally with a pharmaceutically acceptable carrier. While the ceDNA vector can be introduced in the presence of a carrier, such a carrier is not required. The ceDNA vector implemented comprises a nucleotide sequence encoding an transgene or therapeutic protein as described herein useful for treating the disease. In particular, a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein may comprise a desired transgene or therapeutic protein DNA sequence operably linked to control elements capable of directing

transcription of the desired transgene or therapeutic protein encoded by the exogenous DNA sequence when introduced into the subject. The ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can be administered via any suitable route as provided above, and elsewhere herein.

[00380] Disclosed herein are ceDNA vector compositions and formulations for expression of transgene or therapeutic protein as disclosed herein that include one or more of the ceDNA vectors of the present invention together with one or more pharmaceutically-acceptable buffers, diluents, or excipients. Such compositions may be included in one or more diagnostic or therapeutic kits, for diagnosing, preventing, treating or ameliorating one or more symptoms of a disease. In one aspect the disease, injury, disorder, trauma or dysfunction is a human disease, injury, disorder, trauma or dysfunction.

[00381] Another aspect of the technology described herein provides a method for providing a subject in need thereof with a diagnostically- or therapeutically-effective amount of a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein, the method comprising providing to a cell, tissue or organ of a subject in need thereof, an amount of the ceDNA vector as disclosed herein; and for a time effective to enable expression of the transgene or therapeutic protein from the ceDNA vector thereby providing the subject with a diagnostically- or a therapeutically-effective amount of the transgene or therapeutic protein expressed by the ceDNA vector. In a further aspect, the subject is human.

[00382] Another aspect of the technology described herein provides a method for diagnosing, preventing, treating, or ameliorating at least one or more symptoms of a disease, a disorder, a dysfunction, an injury, an abnormal condition, or trauma in a subject. In an overall and general sense, the method includes at least the step of administering to a subject in need thereof one or more of the disclosed ceDNA vector for transgene or therapeutic protein production, in an amount and for a time sufficient to diagnose, prevent, treat or ameliorate the one or more symptoms of the disease, disorder, dysfunction, injury, abnormal condition, or trauma in the subject. In such an embodiment, the subject can be evaluated for efficacy of the transgene or therapeutic protein, or alternatively, detection of the transgene or therapeutic protein or tissue location (including cellular and subcellular location) of the transgene or therapeutic protein in the subject. As such, the ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can be used as an *in vivo* diagnostic tool, *e.g.*, for the detection of cancer or other indications. In a further aspect, the subject is human.

[00383] Another aspect is use of a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein as a tool for treating or reducing one or more symptoms of a disease or disease states. There are a number of inherited diseases in which defective genes are known, and typically fall into two classes: deficiency states, usually of enzymes, which are generally inherited in a recessive manner, and unbalanced states, which may involve regulatory or structural proteins, and which are typically but not always inherited in a dominant manner. For unbalanced disease states, a ceDNA

vector for expression of transgene or therapeutic protein as disclosed herein can be used to create a disease state in a model system, which could then be used in efforts to counteract the disease state. Thus the ceDNA vector for expression of transgene or therapeutic protein as disclosed herein permit the treatment of genetic diseases. As used herein, a disease state is treated by partially or wholly remedying the deficiency or imbalance that causes the disease or makes it more severe.

A. Host cells

[00384] In some embodiments, a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein delivers the transgene or therapeutic protein transgene into a subject host cell. In some embodiments, the cells are photoreceptor cells. In some embodiments, the cells are RPE cells. In some embodiments, the subject host cell is a human host cell, including, for example blood cells, stem cells, hematopoietic cells, CD34⁺ cells, liver cells, cancer cells, vascular cells, muscle cells, pancreatic cells, neural cells, ocular or retinal cells, epithelial or endothelial cells, dendritic cells, fibroblasts, or any other cell of mammalian origin, including, without limitation, hepatic (*i.e.*, liver) cells, lung cells, cardiac cells, pancreatic cells, intestinal cells, diaphragmatic cells, renal (*i.e.*, kidney) cells, neural cells, blood cells, bone marrow cells, or any one or more selected tissues of a subject for which gene therapy is contemplated. In one aspect, the subject host cell is a human host cell.

[00385] The present disclosure also relates to recombinant host cells as mentioned above, including a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein. Thus, one can use multiple host cells depending on the purpose as is obvious to the skilled artisan. A construct or a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein including donor sequence is introduced into a host cell so that the donor sequence is maintained as a chromosomal integrant as described earlier. The term host cell encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the donor sequence and its source.

[00386] The host cell may also be a eukaryote, such as a mammalian, insect, plant, or fungal cell. In one embodiment, the host cell is a human cell (*e.g.*, a primary cell, a stem cell, or an immortalized cell line). In some embodiments, the host cell can be administered a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein *ex vivo* and then delivered to the subject after the gene therapy event. A host cell can be any cell type, *e.g.*, a somatic cell or a stem cell, an induced pluripotent stem cell, or a blood cell, *e.g.*, T-cell or B-cell, or bone marrow cell. In certain embodiments, the host cell is an allogenic cell. For example, T-cell genome engineering is useful for cancer immunotherapies, disease modulation such as HIV therapy (*e.g.*, receptor knock out, such as CXCR4 and CCR5) and immunodeficiency therapies. MHC receptors on B-cells can be targeted for immunotherapy. In some embodiments, gene modified host cells, *e.g.*, bone marrow stem cells, *e.g.*, CD34⁺ cells, or induced pluripotent stem cells can be transplanted back into a patient for expression of a therapeutic protein.

B. Additional diseases for gene therapy

[00387] In general, a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can be used to deliver any transgene or therapeutic protein in accordance with the description above to treat, prevent, or ameliorate the symptoms associated with a disease related to an aberrant protein expression or gene expression in a subject.

[00388] In some embodiments, a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can be used to deliver an transgene or therapeutic protein to skeletal, cardiac or diaphragm muscle, for production of an transgene or therapeutic protein for secretion and circulation in the blood or for systemic delivery to other tissues to treat, ameliorate, and/or a disease.

[00389] The a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can be administered to the lungs of a subject by any suitable means, optionally by administering an aerosol suspension of respirable particles comprising the ceDNA vectors, which the subject inhales. The respirable particles can be liquid or solid. Aerosols of liquid particles comprising the ceDNA vectors may be produced by any suitable means, such as with a pressure-driven aerosol nebulizer or an ultrasonic nebulizer, as is known to those of skill in the art. See, *e.g.*, U.S. Pat. No. 4,501,729. Aerosols of solid particles comprising the ceDNA vectors may likewise be produced with any solid particulate medicament aerosol generator, by techniques known in the pharmaceutical art.

[00390] In some embodiments, a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can be administered to tissues of the CNS (*e.g.*, brain, eye).

[00391] Ocular disorders that may be treated, ameliorated, or prevented with a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein include ophthalmic disorders involving the retina, posterior tract, and optic nerve (*e.g.*, retinitis pigmentosa, diabetic retinopathy and other retinal degenerative diseases, uveitis, age-related macular degeneration, glaucoma). Many ophthalmic diseases and disorders are associated with one or more of three types of indications: (1) angiogenesis, (2) inflammation, and (3) degeneration. In some embodiments, the ceDNA vector as disclosed herein can be employed to deliver anti-angiogenic factors; anti-inflammatory factors; factors that retard cell degeneration, promote cell sparing, or promote cell growth and combinations of the foregoing. Diabetic retinopathy, for example, is characterized by angiogenesis. Diabetic retinopathy can be treated by delivering one or more anti-angiogenic antibodies or fusion proteins either intraocularly (*e.g.*, in the vitreous) or periocularly (*e.g.*, in the sub-Tenon's region). Additional ocular diseases that may be treated, ameliorated, or prevented with the ceDNA vectors of the invention include geographic atrophy, vascular or "wet" macular degeneration, PKU, Leber Congenital Amaurosis (LCA), Usher syndrome, pseudoxanthoma elasticum (PXE), x-linked retinitis pigmentosa (XLRP), x-linked retinoschisis (XLRs), Choroideremia, Leber hereditary optic neuropathy (LHON), Archomatopsia, cone-rod dystrophy, Fuchs endothelial corneal dystrophy, diabetic macular edema and ocular cancer and tumors.

[00392] In some embodiments, inflammatory ocular diseases or disorders (*e.g.*, uveitis) can be treated, ameliorated, or prevented by a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein. One or more anti-inflammatory antibodies or fusion proteins can be expressed by intraocular (*e.g.*, vitreous or anterior chamber) administration of the ceDNA vector as disclosed herein.

[00393] In some embodiments, a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein can encode an transgene or therapeutic protein that is associated with transgene encoding a reporter polypeptide (*e.g.*, an enzyme such as Green Fluorescent Protein, or alkaline phosphatase). In some embodiments, a transgene that encodes a reporter protein useful for experimental or diagnostic purposes, is selected from any of: β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, and others well known in the art. In some aspects, ceDNA vectors expressing a transgene or therapeutic protein linked to a reporter polypeptide may be used for diagnostic purposes, as well as to determine efficacy or as markers of the ceDNA vector's activity in the subject to which they are administered.

C. Testing for successful gene expression using a ceDNA vector

[00394] Assays well known in the art can be used to test the efficiency of gene delivery of a transgene or therapeutic protein by a ceDNA vector can be performed in both *in vitro* and *in vivo* models. Levels of the expression of the transgene or therapeutic protein by ceDNA can be assessed by one skilled in the art by measuring mRNA and protein levels of the transgene or therapeutic protein (*e.g.*, reverse transcription PCR, western blot analysis, and enzyme-linked immunosorbent assay (ELISA)). In one embodiment, ceDNA comprises a reporter protein that can be used to assess the expression of the transgene or therapeutic protein, for example by examining the expression of the reporter protein by fluorescence microscopy or a luminescence plate reader. For *in vivo* applications, protein function assays can be used to test the functionality of a given transgene or therapeutic protein to determine if gene expression has successfully occurred. One skilled will be able to determine the best test for measuring functionality of an transgene or therapeutic protein expressed by the ceDNA vector *in vitro* or *in vivo*.

[00395] It is contemplated herein that the effects of gene expression of an transgene or therapeutic protein from the ceDNA vector in a cell or subject can last for at least 1 month, at least 2 months, at least 3 months, at least four months, at least 5 months, at least six months, at least 10 months, at least 12 months, at least 18 months, at least 2 years, at least 5 years, at least 10 years, at least 20 years, or can be permanent.

[00396] In some embodiments, a transgene or therapeutic protein in the expression cassette, expression construct, or ceDNA vector described herein can be codon optimized for the host cell. As used herein, the term "codon optimized" or "codon optimization" refers to the process of modifying a nucleic acid sequence for enhanced expression in the cells of the vertebrate of interest, *e.g.*, mouse or

human (*e.g.*, humanized), by replacing at least one, more than one, or a significant number of codons of the native sequence (*e.g.*, a prokaryotic sequence) with codons that are more frequently or most frequently used in the genes of that vertebrate. Various species exhibit particular bias for certain codons of a particular amino acid. Typically, codon optimization does not alter the amino acid sequence of the original translated protein. Optimized codons can be determined using *e.g.*, Aptagen's Gene Forge® codon optimization and custom gene synthesis platform (Aptagen, Inc.) or another publicly available database.

D. Determining Efficacy by Assessing transgene or therapeutic protein Expression from the ceDNA vector

[00397] Essentially any method known in the art for determining protein expression can be used to analyze expression of a transgene or therapeutic protein from a ceDNA vector. Non-limiting examples of such methods/assays include enzyme-linked immunoassay (ELISA), affinity ELISA, ELISPOT, serial dilution, flow cytometry, surface plasmon resonance analysis, kinetic exclusion assay, mass spectrometry, Western blot, immunoprecipitation, and PCR.

[00398] For assessing transgene or therapeutic protein expression *in vivo*, a biological sample can be obtained from a subject for analysis. Exemplary biological samples include a biofluid sample, a body fluid sample, blood (including whole blood), serum, plasma, urine, saliva, a biopsy and/or tissue sample *etc.* A biological sample or tissue sample can also refer to a sample of tissue or fluid isolated from an individual including, but not limited to, tumor biopsy, stool, spinal fluid, pleural fluid, nipple aspirates, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, breast milk, cells (including, but not limited to, blood cells), tumors, organs, and also samples of *in vitro* cell culture constituent. The term also includes a mixture of the above-mentioned samples. The term "sample" also includes untreated or pretreated (or pre-processed) biological samples. In some embodiments, the sample used for the assays and methods described herein comprises a serum sample collected from a subject to be tested.

E. Determining Efficacy of the expressed transgene or therapeutic protein by Clinical Parameters

[00399] The efficacy of a given transgene or therapeutic protein expressed by a ceDNA vector for treatment of a disease (*i.e.*, functional expression) can be determined by the skilled clinician. However, a treatment is considered "effective treatment," as the term is used herein, if any one or all of the signs or symptoms of a disease is/are altered in a beneficial manner, or other clinically accepted symptoms or markers of disease are improved, or ameliorated, *e.g.*, by at least 10% following treatment with a ceDNA vector encoding a therapeutic transgene or therapeutic protein as described herein. Efficacy can also be measured by failure of an individual to worsen as assessed by stabilization of a disease, or the need for medical interventions (*i.e.*, progression of the disease is halted or at least slowed). Methods of measuring these indicators are known to those of skill in the art and/or described herein. Treatment includes any treatment of a disease in an individual or an animal (some non-limiting examples include a human, or a mammal) and includes: (1) inhibiting, *e.g.*,

arresting, or slowing progression of a disease; or (2) relieving a symptom of the disease being treated disease, *e.g.*, causing regression of a disease symptoms; and (3) preventing or reducing the likelihood of the development of the disease, or preventing secondary diseases/disorders associated with the disease. An effective amount for the treatment of a disease means that amount which, when administered to a mammal in need thereof, is sufficient to result in effective treatment as that term is defined herein, for that disease. Efficacy of an agent can be determined by assessing physical indicators that are particular to the disease being treated.

[00400] The efficacy of a ceDNA vector expressing a therapeutic protein as disclosed herein can be determined by assessing physical indicators that are particular to a given A disease. Standard methods of analysis of disease indicators are known in the art.

IX. Various applications of ceDNA vectors expressing antibodies or fusion proteins

[00401] As disclosed herein, the compositions and ceDNA vectors for expression of transgene or therapeutic protein as described herein can be used to express a transgene or therapeutic protein for a range of purposes. In one embodiment, the ceDNA vector expressing a transgene or therapeutic protein can be used to create a somatic transgenic animal model harboring the transgene, *e.g.*, to study the function or disease progression of a disease. In some embodiments, a ceDNA vector expressing a transgene or therapeutic protein is useful for the treatment, prevention, or amelioration of a disease state or disorders in a mammalian subject.

[00402] In some embodiments the transgene or therapeutic protein can be expressed from the ceDNA vector in a subject in a sufficient amount to treat a disease associated with increased expression, increased activity of the gene product, or inappropriate upregulation of a gene.

[00403] In some embodiments the transgene or therapeutic protein can be expressed from the ceDNA vector in a subject in a sufficient amount to treat a with a reduced expression, lack of expression or dysfunction of a protein.

[00404] It will be appreciated by one of ordinary skill in the art that the transgene may not be an open reading frame of a gene to be transcribed itself; instead it may be a promoter region or repressor region of a target gene, and the ceDNA vector may modify such region with the outcome of so modulating the expression of the gene.

[00405] The compositions and ceDNA vectors for expression of transgene or therapeutic protein as disclosed herein can be used to deliver a transgene or therapeutic protein for various purposes as described above.

[00406] In some embodiments, the transgene encodes one or more transgene or therapeutic proteins which are useful for the treatment, amelioration, or prevention of a disease state in a mammalian subject. The transgene or therapeutic protein expressed by the ceDNA vector is administered to a patient in a sufficient amount to treat a disease associated with an abnormal gene sequence, which can result in any one or more of the following: increased protein expression, over activity of the protein, reduced expression, lack of expression or dysfunction of the target gene or protein.

[00407] In some embodiments, the ceDNA vectors for expression of transgene or therapeutic protein as disclosed herein are envisioned for use in diagnostic and screening methods, whereby a transgene or therapeutic protein is transiently or stably expressed in a cell culture system, or alternatively, a transgenic animal model.

[00408] Another aspect of the technology described herein provides a method of transducing a population of mammalian cells with a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein. In an overall and general sense, the method includes at least the step of introducing into one or more cells of the population, a composition that comprises an effective amount of one or more of the ceDNA vectors for expression of transgene or therapeutic protein as disclosed herein.

[00409] Additionally, the present invention provides compositions, as well as therapeutic and/or diagnostic kits that include one or more of the disclosed ceDNA vectors for expression of transgene or therapeutic protein as disclosed herein or ceDNA compositions, formulated with one or more additional ingredients, or prepared with one or more instructions for their use.

[00410] A cell to be administered a ceDNA vector for expression of transgene or therapeutic protein as disclosed herein may be of any type, including but not limited to neural cells (including cells of the peripheral and central nervous systems, in particular, brain cells), lung cells, retinal cells, epithelial cells (*e.g.*, gut and respiratory epithelial cells), muscle cells, dendritic cells, pancreatic cells (including islet cells), hepatic cells, myocardial cells, bone cells (*e.g.*, bone marrow stem cells), hematopoietic stem cells, spleen cells, keratinocytes, fibroblasts, endothelial cells, prostate cells, germ cells, and the like. Alternatively, the cell may be any progenitor cell. As a further alternative, the cell can be a stem cell (*e.g.*, neural stem cell, liver stem cell). As still a further alternative, the cell may be a cancer or tumor cell. Moreover, the cells can be from any species of origin, as indicated above.

A. Production and Purification of ceDNA vectors expressing a transgene or therapeutic protein

[00411] The ceDNA vectors disclosed herein are to be used to produce transgene or therapeutic protein either *in vitro* or *in vivo*. The transgene or therapeutic proteins produced in this manner can be isolated, tested for a desired function, and purified for further use in research or as a therapeutic treatment. Each system of protein production has its own advantages/disadvantages. While proteins produced *in vitro* can be easily purified and can proteins in a short time, proteins produced *in vivo* can have post-translational modifications, such as glycosylation.

[00412] A transgene or therapeutic protein produced using ceDNA vectors can be purified using any method known to those of skill in the art, for example, ion exchange chromatography, affinity chromatography, precipitation, or electrophoresis.

An transgene or therapeutic protein produced by the methods and compositions described herein can be tested for binding to the desired target protein.

X. Definitions

[00413] Unless otherwise defined herein, scientific and technical terms used in connection with the present application shall have the meanings that are commonly understood by those of ordinary skill in the art to which this disclosure belongs. It should be understood that this invention is not limited to the particular methodology, protocols, and reagents, etc., described herein and as such can vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims. Definitions of common terms in immunology and molecular biology can be found in The Merck Manual of Diagnosis and Therapy, 19th Edition, published by Merck Sharp & Dohme Corp., 2011 (ISBN 978-0-911910-19-3); Robert S. Porter *et al.* (eds.), Fields Virology, 6th Edition, published by Lippincott Williams & Wilkins, Philadelphia, PA, USA (2013), Knipe, D.M. and Howley, P.M. (ed.), The Encyclopedia of Molecular Cell Biology and Molecular Medicine, published by Blackwell Science Ltd., 1999-2012 (ISBN 9783527600908); and Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8); Immunology by Werner Luttmann, published by Elsevier, 2006; Janeway's Immunobiology, Kenneth Murphy, Allan Mowat, Casey Weaver (eds.), Taylor & Francis Limited, 2014 (ISBN 0815345305, 9780815345305); Lewin's Genes XI, published by Jones & Bartlett Publishers, 2014 (ISBN-1449659055); Michael Richard Green and Joseph Sambrook, Molecular Cloning: A Laboratory Manual, 4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (2012) (ISBN 1936113414); Davis *et al.*, Basic Methods in Molecular Biology, Elsevier Science Publishing, Inc., New York, USA (2012) (ISBN 044460149X); Laboratory Methods in Enzymology: DNA, Jon Lorsch (ed.) Elsevier, 2013 (ISBN 0124199542); Current Protocols in Molecular Biology (CPMB), Frederick M. Ausubel (ed.), John Wiley and Sons, 2014 (ISBN 047150338X, 9780471503385), Current Protocols in Protein Science (CPPS), John E. Coligan (ed.), John Wiley and Sons, Inc., 2005; and Current Protocols in Immunology (CPI) (John E. Coligan, ADA M Kruisbeek, David H Margulies, Ethan M Shevach, Warren Strobe, (eds.) John Wiley and Sons, Inc., 2003 (ISBN 0471142735, 9780471142737), the contents of which are all incorporated by reference herein in their entireties.

[00414] As used herein, the terms, "administration," "administering" and variants thereof refers to introducing a composition or agent (*e.g.*, a therapeutic nucleic acid or an immunosuppressant as described herein) into a subject and includes concurrent and sequential introduction of one or more compositions or agents. "Administration" can refer, *e.g.*, to therapeutic, pharmacokinetic, diagnostic, research, placebo, and experimental methods. "Administration" also encompasses *in vitro* and *ex vivo* treatments. The introduction of a composition or agent into a subject is by any suitable route, including orally, pulmonarily, intranasally, parenterally (intravenously, intramuscularly, intraperitoneally, or subcutaneously), rectally, intralymphatically, intratumorally, or topically. The introduction of a composition or agent into a subject is by electroporation. Administration includes

self-administration and the administration by another. Administration can be carried out by any suitable route. A suitable route of administration allows the composition or the agent to perform its intended function. For example, if a suitable route is intravenous, the composition is administered by introducing the composition or agent into a vein of the subject.

[00415] As used herein, the phrases “nucleic acid therapeutic”, “therapeutic nucleic acid” and “TNA” are used interchangeably and refer to any modality of therapeutic using nucleic acids as an active component of therapeutic agent to treat a disease or disorder. As used herein, these phrases refer to RNA-based therapeutics and DNA-based therapeutics. Non-limiting examples of RNA-based therapeutics include mRNA, antisense RNA and oligonucleotides, ribozymes, aptamers, interfering RNAs (RNAi), Dicer-substrate dsRNA, small hairpin RNA (shRNA), asymmetrical interfering RNA (aiRNA), microRNA (miRNA). Non-limiting examples of DNA-based therapeutics include minicircle DNA, minigene, viral DNA (*e.g.*, Lentiviral or AAV genome) or non-viral synthetic DNA vectors, closed-ended linear duplex DNA (ceDNA / CELiD), plasmids, bacmids, doggybone (dbDNA™) DNA vectors, minimalistic immunological-defined gene expression (MIDGE)-vector, nonviral ministring DNA vector (linear-covalently closed DNA vector), or dumbbell-shaped DNA minimal vector (“dumbbell DNA”).

[00416] As used herein, an “effective amount” or “therapeutically effective amount” of an active agent or therapeutic agent, such as an immunosuppressant and/or therapeutic nucleic acid, is an amount sufficient to produce the desired effect, *e.g.*, a normalization or reduction of immune response (*e.g.*, innate immune response) and expression or inhibition of expression of a target sequence in comparison to the expression level detected in the absence of a therapeutic nucleic acid and/or immunosuppressant. Suitable assays for measuring expression of a target gene or target sequence include, *e.g.*, examination of protein or RNA levels using techniques known to those of skill in the art such as dot blots, northern blots, *in situ* hybridization, ELISA, immunoprecipitation, enzyme function, as well as phenotypic assays known to those of skill in the art. However, dosage levels are based on a variety of factors, including the type of injury, the age, weight, sex, medical condition of the patient, the severity of the condition, the route of administration, and the particular active agent employed. Thus, the dosage regimen may vary widely, but can be determined routinely by a physician using standard methods. Additionally, the terms “therapeutic amount”, “therapeutically effective amounts” and “pharmaceutically effective amounts” include prophylactic or preventative amounts of the compositions of the described invention. In prophylactic or preventative applications of the described invention, pharmaceutical compositions or medicaments are administered to a patient susceptible to, or otherwise at risk of, a disease, disorder or condition in an amount sufficient to eliminate or reduce the risk, lessen the severity, or delay the onset of the disease, disorder or condition, including biochemical, histologic and/or behavioral symptoms of the disease, disorder or condition, its complications, and intermediate pathological phenotypes presenting during development of the disease, disorder or condition. It is generally preferred that a maximum dose be used, that is, the

highest safe dose according to some medical judgment. The terms “dose” and “dosage” are used interchangeably herein.

[00417] As used herein the term “therapeutic effect” refers to a consequence of treatment, the results of which are judged to be desirable and beneficial. A therapeutic effect can include, directly or indirectly, the arrest, reduction, or elimination of a disease manifestation. A therapeutic effect can also include, directly or indirectly, the arrest reduction or elimination of the progression of a disease manifestation.

[00418] For any therapeutic agent described herein therapeutically effective amount may be initially determined from preliminary in vitro studies and/or animal models. A therapeutically effective dose may also be determined from human data. The applied dose may be adjusted based on the relative bioavailability and potency of the administered compound. Adjusting the dose to achieve maximal efficacy based on the methods described above and other well-known methods is within the capabilities of the ordinarily skilled artisan. General principles for determining therapeutic effectiveness, which may be found in Chapter 1 of Goodman and Gilman’s *The Pharmacological Basis of Therapeutics*, 10th Edition, McGraw-Hill (New York) (2001), incorporated herein by reference, are summarized below.

[00419] Pharmacokinetic principles provide a basis for modifying a dosage regimen to obtain a desired degree of therapeutic efficacy with a minimum of unacceptable adverse effects. In situations where the drug’s plasma concentration can be measured and related to therapeutic window, additional guidance for dosage modification can be obtained.

[00420] As used herein, the terms “heterologous nucleotide sequence” and “transgene” are used interchangeably and refer to a nucleic acid of interest (other than a nucleic acid encoding a capsid polypeptide) that is incorporated into and may be delivered and expressed by a ccdNA vector as disclosed herein.

[00421] As used herein, the terms “expression cassette” and “transcription cassette” are used interchangeably and refer to a linear stretch of nucleic acids that includes a transgene that is operably linked to one or more promoters or other regulatory sequences sufficient to direct transcription of the transgene, but which does not comprise capsid-encoding sequences, other vector sequences or inverted terminal repeat regions. An expression cassette may additionally comprise one or more *cis*-acting sequences (*e.g.*, promoters, enhancers, or repressors), one or more introns, and one or more post-transcriptional regulatory elements.

[00422] The terms “polynucleotide” and “nucleic acid,” used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, this term includes single, double, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer including purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases. “Oligonucleotide” generally refers to polynucleotides of between about 5 and about 100 nucleotides of single- or double-stranded

DNA. However, for the purposes of this disclosure, there is no upper limit to the length of an oligonucleotide. Oligonucleotides are also known as “oligomers” or “oligos” and may be isolated from genes, or chemically synthesized by methods known in the art. The terms “polynucleotide” and “nucleic acid” should be understood to include, as applicable to the embodiments being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

[00423] DNA may be in the form of, *e.g.*, antisense molecules, plasmid DNA, DNA-DNA duplexes, pre-condensed DNA, PCR products, vectors (P1, PAC, BAC, YAC, artificial chromosomes), expression cassettes, chimeric sequences, chromosomal DNA, or derivatives and combinations of these groups. DNA may be in the form of minicircle, plasmid, bacmid, minigene, ministring DNA (linear covalently closed DNA vector), closed-ended linear duplex DNA (CELiD or ceDNA), doggybone (dbDNATM) DNA, dumbbell shaped DNA, minimalistic immunological-defined gene expression (MIDGE)-vector, viral vector or nonviral vectors. RNA may be in the form of small interfering RNA (siRNA), Dicer-substrate dsRNA, small hairpin RNA (shRNA), asymmetrical interfering RNA (aiRNA), microRNA (miRNA), mRNA, rRNA, tRNA, viral RNA (vRNA), and combinations thereof. Nucleic acids include nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, and which have similar binding properties as the reference nucleic acid. Examples of such analogs and/or modified residues include, without limitation, phosphorothioates, phosphorodiamidate morpholino oligomer (morpholino), phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2'-O-methyl ribonucleotides, locked nucleic acid (LNATM), and peptide nucleic acids (PNAs). Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated.

[00424] “Nucleotides” contain a sugar deoxyribose (DNA) or ribose (RNA), a base, and a phosphate group. Nucleotides are linked together through the phosphate groups.

[00425] “Bases” include purines and pyrimidines, which further include natural compounds adenine, thymine, guanine, cytosine, uracil, inosine, and natural analogs, and synthetic derivatives of purines and pyrimidines, which include, but are not limited to, modifications which place new reactive groups such as, but not limited to, amines, alcohols, thiols, carboxylates, and alkylhalides.

[00426] As used herein, the term “interfering RNA” or “RNAi” or “interfering RNA sequence” includes single-stranded RNA (*e.g.*, mature miRNA, ssRNAi oligonucleotides, ssDNAi oligonucleotides), double-stranded RNA (*i.e.*, duplex RNA such as siRNA, Dicer-substrate dsRNA, shRNA, aiRNA, or pre-miRNA), a DNA-RNA hybrid (see, *e.g.*, PCT Publication No. WO 2004/078941), or a DNA-DNA hybrid (see, *e.g.*, PCT Publication No. WO 2004/104199) that is capable of reducing or inhibiting the expression of a target gene or sequence (*e.g.*, by mediating the

degradation or inhibiting the translation of mRNAs which are complementary to the interfering RNA sequence) when the interfering RNA is in the same cell as the target gene or sequence. Interfering RNA thus refers to the single-stranded RNA that is complementary to a target mRNA sequence or to the double-stranded RNA formed by two complementary strands or by a single, self-complementary strand. Interfering RNA may have substantial or complete identity to the target gene or sequence, or may comprise a region of mismatch (*i.e.*, a mismatch motif). The sequence of the interfering RNA can correspond to the full-length target gene, or a subsequence thereof. Preferably, the interfering RNA molecules are chemically synthesized. The disclosures of each of the above patent documents are herein incorporated by reference in their entirety for all purposes.

[00427] Interfering RNA includes “small-interfering RNA” or “siRNA,” *e.g.*, interfering RNA of about 15-60, 15-50, or 15-40 (duplex) nucleotides in length, more typically about 15-30, 15-25, or 19-25 (duplex) nucleotides in length, and is preferably about 20-24, 21-22, or 21-23 (duplex) nucleotides in length (*e.g.*, each complementary sequence of the double-stranded siRNA is 15-60, 15-50, 15-40, 15-30, 15-25, or 19-25 nucleotides in length, preferably about 20-24, 21-22, or 21-23 nucleotides in length, and the double-stranded siRNA is about 15-60, 15-50, 15-40, 15-30, 15-25, or 19-25 base pairs in length, preferably about 18-22, 19-20, or 19-21 base pairs in length). siRNA duplexes may comprise 3' overhangs of about 1 to about 4 nucleotides or about 2 to about 3 nucleotides and 5' phosphate termini. Examples of siRNA include, without limitation, a double-stranded polynucleotide molecule assembled from two separate stranded molecules, wherein one strand is the sense strand and the other is the complementary antisense strand; a double-stranded polynucleotide molecule assembled from a single stranded molecule, where the sense and antisense regions are linked by a nucleic acid-based or non-nucleic acid-based linker; a double-stranded polynucleotide molecule with a hairpin secondary structure having self-complementary sense and antisense regions; and a circular single-stranded polynucleotide molecule with two or more loop structures and a stem having self-complementary sense and antisense regions, where the circular polynucleotide can be processed *in vivo* or *in vitro* to generate an active double-stranded siRNA molecule. As used herein, the term “siRNA” includes RNA-RNA duplexes as well as DNA-RNA hybrids (see, *e.g.*, PCT Publication No. WO 2004/078941)

[00428] The term “nucleic acid construct” as used herein refers to a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic. The term nucleic acid construct is synonymous with the term “expression cassette” when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present disclosure. An “expression cassette” includes a DNA coding sequence operably linked to a promoter.

[00429] By “hybridizable” or “complementary” or “substantially complementary” it is meant that a nucleic acid (*e.g.*, RNA) includes a sequence of nucleotides that enables it to non-covalently bind, *i.e.*

form Watson-Crick base pairs and/or G/U base pairs, “anneal”, or “hybridize,” to another nucleic acid in a sequence-specific, antiparallel, manner (*i.e.*, a nucleic acid specifically binds to a complementary nucleic acid) under the appropriate *in vitro* and/or *in vivo* conditions of temperature and solution ionic strength. As is known in the art, standard Watson-Crick base-pairing includes: adenine (A) pairing with thymidine (T), adenine (A) pairing with uracil (U), and guanine (G) pairing with cytosine (C). In addition, it is also known in the art that for hybridization between two RNA molecules (*e.g.*, dsRNA), guanine (G) base pairs with uracil (U). For example, G/U base-pairing is partially responsible for the degeneracy (*i.e.*, redundancy) of the genetic code in the context of tRNA anti-codon base-pairing with codons in mRNA. In the context of this disclosure, a guanine (G) of a protein-binding segment (dsRNA duplex) of a subject DNA-targeting RNA molecule is considered complementary to a uracil (U), and vice versa. As such, when a G/U base-pair can be made at a given nucleotide position a protein-binding segment (dsRNA duplex) of a subject DNA-targeting RNA molecule, the position is not considered to be non-complementary, but is instead considered to be complementary.

[00430] The terms “peptide,” “polypeptide,” and “protein” are used interchangeably herein, and refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones.

[00431] A DNA sequence that “encodes” a particular a transgene or therapeutic protein (*e.g.*, ATP8B1, ABCB11, ABCB4 and TJP2) is a DNA nucleic acid sequence that is transcribed into the particular RNA and/or protein. A DNA polynucleotide may encode an RNA (mRNA) that is translated into protein, or a DNA polynucleotide may encode an RNA that is not translated into protein (*e.g.*, tRNA, rRNA, or a DNA-targeting RNA; also called “non-coding” RNA or “ncRNA”).

[0001] As used herein, the term “fusion protein” as used herein refers to a polypeptide which comprises protein domains from at least two different proteins. For example, a fusion protein may comprise (i) a therapeutic protein or fragment thereof and (ii) at least one non-GOI protein. Fusion proteins encompassed herein include, but are not limited to, an antibody, or Fc or antigen-binding fragment of an antibody fused to a therapeutic protein, *e.g.*, an extracellular domain of a receptor, ligand, enzyme or peptide. The therapeutic protein or fragment thereof that is part of a fusion protein can be a monospecific antibody or a bispecific or multispecific antibody.

[00432] As used herein, the term “genomic safe harbor gene” or “safe harbor gene” refers to a gene or loci that a nucleic acid sequence can be inserted such that the sequence can integrate and function in a predictable manner (*e.g.*, express a protein of interest) without significant negative consequences to endogenous gene activity, or the promotion of cancer. In some embodiments, a safe harbor gene is also a loci or gene where an inserted nucleic acid sequence can be expressed efficiently and at higher levels than a non-safe harbor site.

[00433] As used herein, the term “gene delivery” means a process by which foreign DNA is transferred to host cells for applications of gene therapy.

[00434] As used herein, the term “terminal repeat” or “TR” includes any viral terminal repeat or synthetic sequence that comprises at least one minimal required origin of replication and a region comprising a palindrome hairpin structure. A Rep-binding sequence (“RBS”) (also referred to as RBE (Rep-binding element)) and a terminal resolution site (“TRS”) together constitute a “minimal required origin of replication” and thus the TR comprises at least one RBS and at least one TRS. TRs that are the inverse complement of one another within a given stretch of polynucleotide sequence are typically each referred to as an “inverted terminal repeat” or “ITR”. In the context of a virus, ITRs mediate replication, virus packaging, integration and provirus rescue. As was unexpectedly found in the invention herein, TRs that are not inverse complements across their full length can still perform the traditional functions of ITRs, and thus the term ITR is used herein to refer to a TR in a ceDNA genome or ceDNA vector that is capable of mediating replication of ceDNA vector. It will be understood by one of ordinary skill in the art that in complex ceDNA vector configurations more than two ITRs or asymmetric ITR pairs may be present. The ITR can be an AAV ITR or a non-AAV ITR, or can be derived from an AAV ITR or a non-AAV ITR. For example, the ITR can be derived from the family Parvoviridae, which encompasses parvoviruses and dependoviruses (*e.g.*, canine parvovirus, bovine parvovirus, mouse parvovirus, porcine parvovirus, human parvovirus B-19), or the SV40 hairpin that serves as the origin of SV40 replication can be used as an ITR, which can further be modified by truncation, substitution, deletion, insertion and/or addition. Parvoviridae family viruses consist of two subfamilies: Parvovirinae, which infect vertebrates, and Densovirinae, which infect invertebrates. Dependoparvoviruses include the viral family of the adeno-associated viruses (AAV) which are capable of replication in vertebrate hosts including, but not limited to, human, primate, bovine, canine, equine and ovine species. For convenience herein, an ITR located 5’ to (upstream of) an expression cassette in a ceDNA vector is referred to as a “5’ ITR” or a “left ITR”, and an ITR located 3’ to (downstream of) an expression cassette in a ceDNA vector is referred to as a “3’ ITR” or a “right ITR”.

[00435] A “wild-type ITR” or “WT-ITR” refers to the sequence of a naturally occurring ITR sequence in an AAV or other dependovirus that retains, *e.g.*, Rep binding activity and Rep nicking ability. The nucleotide sequence of a WT-ITR from any AAV serotype may slightly vary from the canonical naturally occurring sequence due to degeneracy of the genetic code or drift, and therefore WT-ITR sequences encompassed for use herein include WT-ITR sequences as result of naturally occurring changes taking place during the production process (*e.g.*, a replication error).

[00436] As used herein, the term “substantially symmetrical WT-ITRs” or a “substantially symmetrical WT-ITR pair” refers to a pair of WT-ITRs within a single ceDNA genome or ceDNA vector that are both wild type ITRs that have an inverse complement sequence across their entire length. For example, an ITR can be considered to be a wild-type sequence, even if it has one or more nucleotides that deviate from the canonical naturally occurring sequence, so long as the changes do not affect the properties and overall three-dimensional structure of the sequence. In some aspects, the

deviating nucleotides represent conservative sequence changes. As one non-limiting example, a sequence that has at least 95%, 96%, 97%, 98%, or 99% sequence identity to the canonical sequence (as measured, *e.g.*, using BLAST at default settings), and also has a symmetrical three-dimensional spatial organization to the other WT-ITR such that their 3D structures are the same shape in geometrical space. The substantially symmetrical WT-ITR has the same A, C-C' and B-B' loops in 3D space. A substantially symmetrical WT-ITR can be functionally confirmed as WT by determining that it has an operable Rep binding site (RBE or RBE') and terminal resolution site (TRS) that pairs with the appropriate Rep protein. One can optionally test other functions, including transgene expression under permissive conditions.

[00437] As used herein, the phrases of “modified ITR” or “mod-ITR” or “mutant ITR” are used interchangeably herein and refer to an ITR that has a mutation in at least one or more nucleotides as compared to the WT-ITR from the same serotype. The mutation can result in a change in one or more of A, C, C', B, B' regions in the ITR, and can result in a change in the three-dimensional spatial organization (*i.e.* its 3D structure in geometric space) as compared to the 3D spatial organization of a WT-ITR of the same serotype.

[00438] As used herein, the term “asymmetric ITRs” also referred to as “asymmetric ITR pairs” refers to a pair of ITRs within a single ceDNA genome or ceDNA vector that are not inverse complements across their full length. As one non-limiting example, an asymmetric ITR pair does not have a symmetrical three-dimensional spatial organization to their cognate ITR such that their 3D structures are different shapes in geometrical space. Stated differently, an asymmetrical ITR pair have the different overall geometric structure, *i.e.*, they have different organization of their A, C-C' and B-B' loops in 3D space (*e.g.*, one ITR may have a short C-C' arm and/or short B-B' arm as compared to the cognate ITR). The difference in sequence between the two ITRs may be due to one or more nucleotide addition, deletion, truncation, or point mutation. In one embodiment, one ITR of the asymmetric ITR pair may be a wild-type AAV ITR sequence and the other ITR a modified ITR as defined herein (*e.g.*, a non-wild-type or synthetic ITR sequence). In another embodiment, neither ITRs of the asymmetric ITR pair is a wild-type AAV sequence and the two ITRs are modified ITRs that have different shapes in geometrical space (*i.e.*, a different overall geometric structure). In some embodiments, one mod-ITRs of an asymmetric ITR pair can have a short C-C' arm and the other ITR can have a different modification (*e.g.*, a single arm, or a short B-B' arm etc.) such that they have different three-dimensional spatial organization as compared to the cognate asymmetric mod-ITR.

[00439] As used herein, the term “symmetric ITRs” refers to a pair of ITRs within a single ceDNA genome or ceDNA vector that are wild-type or mutated (*e.g.*, modified relative to wild-type) dependoviral ITR sequences and are inverse complements across their full length. In one non-limiting example, both ITRs are wild type ITRs sequences from AAV2. In another example, neither ITRs are wild type ITR AAV2 sequences (*i.e.*, they are a modified ITR, also referred to as a mutant ITR), and can have a difference in sequence from the wild type ITR due to nucleotide addition,

deletion, substitution, truncation, or point mutation. For convenience herein, an ITR located 5' to (upstream of) an expression cassette in a ceDNA vector is referred to as a "5' ITR" or a "left ITR", and an ITR located 3' to (downstream of) an expression cassette in a ceDNA vector is referred to as a "3' ITR" or a "right ITR".

[00440] As used herein, the terms "substantially symmetrical modified-ITRs" or a "substantially symmetrical mod-ITR pair" refers to a pair of modified-ITRs within a single ceDNA genome or ceDNA vector that are both that have an inverse complement sequence across their entire length. For example, the a modified ITR can be considered substantially symmetrical, even if it has some nucleotide sequences that deviate from the inverse complement sequence so long as the changes do not affect the properties and overall shape. As one non-limiting example, a sequence that has at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to the canonical sequence (as measured using BLAST at default settings), and also has a symmetrical three-dimensional spatial organization to their cognate modified ITR such that their 3D structures are the same shape in geometrical space. Stated differently, a substantially symmetrical modified-ITR pair have the same A, C-C' and B-B' loops organized in 3D space. In some embodiments, the ITRs from a mod-ITR pair may have different reverse complement nucleotide sequences but still have the same symmetrical three-dimensional spatial organization – that is both ITRs have mutations that result in the same overall 3D shape. For example, one ITR (*e.g.*, 5' ITR) in a mod-ITR pair can be from one serotype, and the other ITR (*e.g.*, 3' ITR) can be from a different serotype, however, both can have the same corresponding mutation (*e.g.*, if the 5' ITR has a deletion in the C region, the cognate modified 3' ITR from a different serotype has a deletion at the corresponding position in the C' region), such that the modified ITR pair has the same symmetrical three-dimensional spatial organization. In such embodiments, each ITR in a modified ITR pair can be from different serotypes (*e.g.* AAV1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12) such as the combination of AAV2 and AAV6, with the modification in one ITR reflected in the corresponding position in the cognate ITR from a different serotype. In one embodiment, a substantially symmetrical modified ITR pair refers to a pair of modified ITRs (mod-ITRs) so long as the difference in nucleotide sequences between the ITRs does not affect the properties or overall shape and they have substantially the same shape in 3D space. As a non-limiting example, a mod-ITR that has at least 95%, 96%, 97%, 98% or 99% sequence identity to the canonical mod-ITR as determined by standard means well known in the art such as BLAST (Basic Local Alignment Search Tool), or BLASTN at default settings, and also has a symmetrical three-dimensional spatial organization such that their 3D structure is the same shape in geometric space. A substantially symmetrical mod-ITR pair has the same A, C-C' and B-B' loops in 3D space, *e.g.*, if a modified ITR in a substantially symmetrical mod-ITR pair has a deletion of a C-C' arm, then the cognate mod-ITR has the corresponding deletion of the C-C' loop and also has a similar 3D structure of the remaining A and B-B' loops in the same shape in geometric space of its cognate mod-ITR.

[00441] The term “flanking” refers to a relative position of one nucleic acid sequence with respect to another nucleic acid sequence. Generally, in the sequence ABC, B is flanked by A and C. The same is true for the arrangement AxBxC. Thus, a flanking sequence precedes or follows a flanked sequence but need not be contiguous with, or immediately adjacent to the flanked sequence. In one embodiment, the term flanking refers to terminal repeats at each end of the linear duplex ceDNA vector.

[00442] As used herein, the terms “treat,” “treating,” and/or “treatment” include abrogating, substantially inhibiting, slowing or reversing the progression of a condition, substantially ameliorating clinical symptoms of a condition, or substantially preventing the appearance of clinical symptoms of a condition, obtaining beneficial or desired clinical results. Treating further refers to accomplishing one or more of the following: (a) reducing the severity of the disorder; (b) limiting development of symptoms characteristic of the disorder(s) being treated; (c) limiting worsening of symptoms characteristic of the disorder(s) being treated; (d) limiting recurrence of the disorder(s) in patients that have previously had the disorder(s); and (e) limiting recurrence of symptoms in patients that were previously asymptomatic for the disorder(s). Beneficial or desired clinical results, such as pharmacologic and/or physiologic effects include, but are not limited to, preventing the disease, disorder or condition from occurring in a subject that may be predisposed to the disease, disorder or condition but does not yet experience or exhibit symptoms of the disease (prophylactic treatment), alleviation of symptoms of the disease, disorder or condition, diminishment of extent of the disease, disorder or condition, stabilization (*i.e.*, not worsening) of the disease, disorder or condition, preventing spread of the disease, disorder or condition, delaying or slowing of the disease, disorder or condition progression, amelioration or palliation of the disease, disorder or condition, and combinations thereof, as well as prolonging survival as compared to expected survival if not receiving treatment.

[00443] As used herein, the term “increase,” “enhance,” “raise” (and like terms) generally refers to the act of increasing, either directly or indirectly, a concentration, level, function, activity, or behavior relative to the natural, expected, or average, or relative to a control condition.

[00444] As used herein, the term “suppress,” “decrease,” “interfere,” “inhibit” and/or “reduce” (and like terms) generally refers to the act of reducing, either directly or indirectly, a concentration, level, function, activity, or behavior relative to the natural, expected, or average, or relative to a control condition. By “decrease,” “decreasing,” “reduce,” or “reducing” of an immune response (*e.g.*, an immune response (*e.g.*, innate immune response)) by an immunosuppressant is intended to mean a detectable decrease of an immune response to a given immunosuppressant. The amount of decrease of an immune response by the immunosuppressant may be determined relative to the level of an immune response in the presence of an immunosuppressant. A detectable decrease can be about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100%, or more lower than the immune response detected in the presence of the immunosuppressant.

[00445] As used herein, the term “lipid” refers to a group of organic compounds that include, but are not limited to, esters of fatty acids and are characterized by being insoluble in water, but soluble in many organic solvents. They are usually divided into at least three classes: (1) “simple lipids,” which include fats and oils as well as waxes; (2) “compound lipids,” which include phospholipids and glycolipids; and (3) “derived lipids” such as steroids.

[00446] As used herein, the term “lipid particle” includes a lipid formulation that can be used to deliver a therapeutic agent such as nucleic acid therapeutics and/or an immunosuppressant to a target site of interest (*e.g.*, cell, tissue, organ, and the like). In preferred embodiments, the lipid particle of the invention is a nucleic acid containing lipid particle, which is typically formed from a cationic lipid, a non-cationic lipid, and optionally a conjugated lipid that prevents aggregation of the particle. In other preferred embodiments, a therapeutic agent such as a therapeutic nucleic acid may be encapsulated in the lipid portion of the particle, thereby protecting it from enzymatic degradation. In other preferred embodiments, an immunosuppressant can be optionally included in the nucleic acid containing lipid particles.

[00447] As used herein, the term “lipid encapsulated” can refer to a lipid particle that provides an active agent or therapeutic agent, such as a nucleic acid (*e.g.*, a ceDNA), with full encapsulation, partial encapsulation, or both. In a preferred embodiment, the nucleic acid is fully encapsulated in the lipid particle (*e.g.*, to form a nucleic acid containing lipid particle).

[00448] As used herein, the term “lipid conjugate” refers to a conjugated lipid that inhibits aggregation of lipid particles. Such lipid conjugates include, but are not limited to, PEG-lipid conjugates such as, *e.g.*, PEG coupled to dialkyloxypropyls (*e.g.*, PEG-DAA conjugates), PEG coupled to diacylglycerols (*e.g.*, PEG-DAG conjugates), PEG coupled to cholesterol, PEG coupled to phosphatidylethanolamines, and PEG conjugated to ceramides (see, *e.g.*, U.S. Pat. No. 5,885,613), cationic PEG lipids, polyoxazoline (POZ)-lipid conjugates (*e.g.*, POZ-DAA conjugates; see, *e.g.*, U.S. Provisional Application No. 61/294,828, filed Jan. 13, 2010, and U.S. Provisional Application No. 61/295,140, filed Jan. 14, 2010), polyamide oligomers (*e.g.*, ATTA-lipid conjugates), and mixtures thereof. Additional examples of POZ-lipid conjugates are described in PCT Publication No. WO 2010/006282. PEG or POZ can be conjugated directly to the lipid or may be linked to the lipid *via* a linker moiety. Any linker moiety suitable for coupling the PEG or the POZ to a lipid can be used including, *e.g.*, non-ester containing linker moieties and ester-containing linker moieties. In certain preferred embodiments, non-ester containing linker moieties, such as amides or carbamates, are used. The disclosures of each of the above patent documents are herein incorporated by reference in their entirety for all purposes.

[00449] Representative examples of phospholipids include, but are not limited to, phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, phosphatidic acid, palmitoyloleoyl phosphatidylcholine, lysophosphatidylcholine, lysophosphatidylethanolamine, dipalmitoylphosphatidylcholine, dioleoylphosphatidylcholine,

distearoylphosphatidylcholine, and dilinoleoylphosphatidylcholine. Other compounds lacking in phosphorus, such as sphingolipid, glycosphingolipid families, diacylglycerols, and β -acyloxyacids, are also within the group designated as amphipathic lipids. Additionally, the amphipathic lipids described above can be mixed with other lipids including triglycerides and sterols.

[00450] As used herein, the term “neutral lipid” refers to any of a number of lipid species that exist either in an uncharged or neutral zwitterionic form at a selected pH. At physiological pH, such lipids include, for example, diacylphosphatidylcholine, diacylphosphatidylethanolamine, ceramide, sphingomyelin, cephalin, cholesterol, cerebroside, and diacylglycerols.

[00451] As used herein, the term “non-cationic lipid” refers to any amphipathic lipid as well as any other neutral lipid or anionic lipid.

[00452] As used herein, the term “anionic lipid” refers to any lipid that is negatively charged at physiological pH. These lipids include, but are not limited to, phosphatidylglycerols, cardiolipins, diacylphosphatidylserines, diacylphosphatidic acids, N-dodecanoyl phosphatidylethanolamines, N-succinyl phosphatidylethanolamines, N-glutarylphosphatidylethanolamines, lysylphosphatidylglycerols, palmitoyloleyolphosphatidylglycerol (POPG), and other anionic modifying groups joined to neutral lipids.

[00453] As used herein, the term “hydrophobic lipid” refers to compounds having apolar groups that include, but are not limited to, long-chain saturated and unsaturated aliphatic hydrocarbon groups and such groups optionally substituted by one or more aromatic, cycloaliphatic, or heterocyclic group(s). Suitable examples include, but are not limited to, diacylglycerol, dialkylglycerol, N—N-dialkylamino, 1,2-diacyloxy-3-aminopropane, and 1,2-dialkyl-3-aminopropane.

[00454] As used herein, the term “aqueous solution” refers to a composition comprising in whole, or in part, water.

[00455] As used herein, the term “organic lipid solution” refers to a composition comprising in whole, or in part, an organic solvent having a lipid.

[00456] As used herein, the term “systemic delivery” refers to delivery of lipid particles that leads to a broad biodistribution of an active agent such as an interfering RNA (*e.g.*, siRNA) within an organism. Some techniques of administration can lead to the systemic delivery of certain agents, but not others. Systemic delivery means that a useful, preferably therapeutic, amount of an agent is exposed to most parts of the body. To obtain broad biodistribution generally requires a blood lifetime such that the agent is not rapidly degraded or cleared (such as by first pass organs (liver, lung, etc.) or by rapid, nonspecific cell binding) before reaching a disease site distal to the site of administration. Systemic delivery of lipid particles can be by any means known in the art including, for example, intravenous, subcutaneous, and intraperitoneal. In a preferred embodiment, systemic delivery of lipid particles is by intravenous delivery.

[00457] As used herein, the term “local delivery” refers to delivery of an active agent such as an interfering RNA (*e.g.*, siRNA) directly to a target site within an organism. For example, an agent can

be locally delivered by direct injection into a disease site such as a tumor or other target site such as a site of inflammation or a target organ such as the liver, heart, pancreas, kidney, and the like.

[00458] As used herein, the term “ceDNA genome” refers to an expression cassette that further incorporates at least one inverted terminal repeat region. A ceDNA genome may further comprise one or more spacer regions. In some embodiments the ceDNA genome is incorporated as an intermolecular duplex polynucleotide of DNA into a plasmid or viral genome.

[00459] As used herein, the term “ceDNA spacer region” refers to an intervening sequence that separates functional elements in the ceDNA vector or ceDNA genome. In some embodiments, ceDNA spacer regions keep two functional elements at a desired distance for optimal functionality. In some embodiments, ceDNA spacer regions provide or add to the genetic stability of the ceDNA genome within *e.g.*, a plasmid or baculovirus. In some embodiments, ceDNA spacer regions facilitate ready genetic manipulation of the ceDNA genome by providing a convenient location for cloning sites and the like. For example, in certain aspects, an oligonucleotide “polylinker” containing several restriction endonuclease sites, or a non-open reading frame sequence designed to have no known protein (*e.g.*, transcription factor) binding sites can be positioned in the ceDNA genome to separate the *cis* – acting factors, *e.g.*, inserting a 6mer, 12mer, 18mer, 24mer, 48mer, 86mer, 176mer, etc. between the terminal resolution site and the upstream transcriptional regulatory element. Similarly, the spacer may be incorporated between the polyadenylation signal sequence and the 3'-terminal resolution site.

[00460] As used herein, the terms “Rep binding site,” “Rep binding element,” “RBE” and “RBS” are used interchangeably and refer to a binding site for Rep protein (*e.g.*, AAV Rep 78 or AAV Rep 68) which upon binding by a Rep protein permits the Rep protein to perform its site-specific endonuclease activity on the sequence incorporating the RBS. An RBS sequence and its inverse complement together form a single RBS. RBS sequences are known in the art, and include, for example, 5'-GCGCGCTCGCTCGCTC-3' (SEQ ID NO: 60), an RBS sequence identified in AAV2. Any known RBS sequence may be used in the embodiments of the invention, including other known AAV RBS sequences and other naturally known or synthetic RBS sequences. Without being bound by theory it is thought that the nuclease domain of a Rep protein binds to the duplex nucleotide sequence GCTC, and thus the two known AAV Rep proteins bind directly to and stably assemble on the duplex oligonucleotide, 5'-(GCGC)(GCTC)(GCTC)(GCTC)-3' (SEQ ID NO: 60). In addition, soluble aggregated conformers (*i.e.*, undefined number of inter-associated Rep proteins) dissociate and bind to oligonucleotides that contain Rep binding sites. Each Rep protein interacts with both the nitrogenous bases and phosphodiester backbone on each strand. The interactions with the nitrogenous bases provide sequence specificity whereas the interactions with the phosphodiester backbone are non- or less- sequence specific and stabilize the protein-DNA complex.

[00461] As used herein, the terms “terminal resolution site” and “TRS” are used interchangeably herein and refer to a region at which Rep forms a tyrosine-phosphodiester bond with the 5' thymidine

generating a 3' OH that serves as a substrate for DNA extension via a cellular DNA polymerase, *e.g.*, DNA pol delta or DNA pol epsilon. Alternatively, the Rep-thymidine complex may participate in a coordinated ligation reaction. In some embodiments, a TRS minimally encompasses a non-base-paired thymidine. In some embodiments, the nicking efficiency of the TRS can be controlled at least in part by its distance within the same molecule from the RBS. When the acceptor substrate is the complementary ITR, then the resulting product is an intramolecular duplex. TRS sequences are known in the art, and include, for example, 5'-GGTTGA-3' (SEQ ID NO: 61), the hexanucleotide sequence identified in AAV2. Any known TRS sequence may be used in the embodiments of the invention, including other known AAV TRS sequences and other naturally known or synthetic TRS sequences such as AGTT (SEQ ID NO: 62), GGTTGG (SEQ ID NO: 63), AGTTGG (SEQ ID NO: 64), AGTTGA (SEQ ID NO: 65), and other motifs such as RRTTRR (SEQ ID NO: 66).

[00462] As used herein, the term “ceDNA-plasmid” refers to a plasmid that comprises a ceDNA genome as an intermolecular duplex.

[00463] As used herein, the term “ceDNA-bacmid” refers to an infectious baculovirus genome comprising a ceDNA genome as an intermolecular duplex that is capable of propagating in *E. coli* as a plasmid, and so can operate as a shuttle vector for baculovirus.

[00464] As used herein, the term “ceDNA-baculovirus” refers to a baculovirus that comprises a ceDNA genome as an intermolecular duplex within the baculovirus genome.

[00465] As used herein, the terms “ceDNA-baculovirus infected insect cell” and “ceDNA-BIIC” are used interchangeably, and refer to an invertebrate host cell (including, but not limited to an insect cell (*e.g.*, an Sf9 cell)) infected with a ceDNA-baculovirus.

[00466] As used herein, the term “closed-ended DNA vector” refers to a capsid-free DNA vector with at least one covalently closed end and where at least part of the vector has an intramolecular duplex structure.

[00467] As used herein, the term “ceDNA” refers to capsid-free closed-ended linear double stranded (*ds*) duplex DNA for non-viral gene transfer, synthetic or otherwise. Detailed description of ceDNA is described in International application of PCT/US2017/020828, filed March 3, 2017, the entire contents of which are expressly incorporated herein by reference. Certain methods for the production of ceDNA comprising various inverted terminal repeat (ITR) sequences and configurations using cell-based methods are described in Example 1 of International applications PCT/US18/49996, filed September 7, 2018, and PCT/US2018/064242, filed December 6, 2018 each of which is incorporated herein in its entirety by reference. Certain methods for the production of synthetic ceDNA vectors comprising various ITR sequences and configurations are described, *e.g.*, in International application PCT/US2019/14122, filed January 18, 2019, the entire content of which is incorporated herein by reference.

[00468] As used herein, the terms “ceDNA vector” and “ceDNA” are used interchangeably and refer to a closed-ended DNA vector comprising at least one terminal palindrome. In some embodiments, the ceDNA comprises two covalently-closed ends.

[00469] As used herein, the term “neDNA” or “nicked ceDNA” refers to a closed-ended DNA having a nick or a gap of 1-100 base pairs in a stem region or spacer region 5' upstream of an open reading frame (*e.g.*, a promoter and transgene to be expressed).

[00470] As used herein, the terms “gap” and “nick” are used interchangeably and refer to a discontinued portion of synthetic DNA vector of the present invention, creating a stretch of single stranded DNA portion in otherwise double stranded ceDNA. The gap can be 1 base-pair to 100 base-pair long in length in one strand of a duplex DNA. Typical gaps, designed and created by the methods described herein and synthetic vectors generated by the methods can be, for example, 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 or 60 bp long in length. Exemplified gaps in the present disclosure can be 1 bp to 10 bp long, 1 to 20 bp long, 1 to 30 bp long in length.

[00471] As used herein, the terms “sense” and “antisense” refer to the orientation of the structural element on the polynucleotide. The sense and antisense versions of an element are the reverse complement of each other.

[00472] As used herein, the term “synthetic AAV vector” and “synthetic production of AAV vector” refers to an AAV vector and synthetic production methods thereof in an entirely cell-free environment.

[00473] As used herein, “reporters” refer to proteins that can be used to provide detectable read-outs. Reporters generally produce a measurable signal such as fluorescence, color, or luminescence. Reporter protein coding sequences encode proteins whose presence in the cell or organism is readily observed. For example, fluorescent proteins cause a cell to fluoresce when excited with light of a particular wavelength, luciferases cause a cell to catalyze a reaction that produces light, and enzymes such as β -galactosidase convert a substrate to a colored product. Exemplary reporter polypeptides useful for experimental or diagnostic purposes include, but are not limited to β -lactamase, β -galactosidase (LacZ), alkaline phosphatase (AP), thymidine kinase (TK), green fluorescent protein (GFP) and other fluorescent proteins, chloramphenicol acetyltransferase (CAT), luciferase, and others well known in the art.

[00474] As used herein, the term “effector protein” refers to a polypeptide that provides a detectable read-out, either as, for example, a reporter polypeptide, or more appropriately, as a polypeptide that kills a cell, *e.g.*, a toxin, or an agent that renders a cell susceptible to killing with a chosen agent or lack thereof. Effector proteins include any protein or peptide that directly targets or damages the host cell's DNA and/or RNA. For example, effector proteins can include, but are not limited to, a restriction endonuclease that targets a host cell DNA sequence (whether genomic or on

an extrachromosomal element), a protease that degrades a polypeptide target necessary for cell survival, a DNA gyrase inhibitor, and a ribonuclease-type toxin. In some embodiments, the expression of an effector protein controlled by a synthetic biological circuit as described herein can participate as a factor in another synthetic biological circuit to thereby expand the range and complexity of a biological circuit system's responsiveness.

[00475] Transcriptional regulators refer to transcriptional activators and repressors that either activate or repress transcription of a gene of interest, such as transgene or therapeutic protein. Promoters are regions of nucleic acid that initiate transcription of a particular gene. Transcriptional activators typically bind nearby to transcriptional promoters and recruit RNA polymerase to directly initiate transcription. Repressors bind to transcriptional promoters and sterically hinder transcriptional initiation by RNA polymerase. Other transcriptional regulators may serve as either an activator or a repressor depending on where they bind and cellular and environmental conditions. Non-limiting examples of transcriptional regulator classes include, but are not limited to homeodomain proteins, zinc-finger proteins, winged-helix (forkhead) proteins, and leucine-zipper proteins.

[00476] As used herein, a "repressor protein" or "inducer protein" is a protein that binds to a regulatory sequence element and represses or activates, respectively, the transcription of sequences operatively linked to the regulatory sequence element. Preferred repressor and inducer proteins as described herein are sensitive to the presence or absence of at least one input agent or environmental input. Preferred proteins as described herein are modular in form, comprising, for example, separable DNA-binding and input agent-binding or responsive elements or domains.

[00477] As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutically active substances is well known in the art. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce a toxic, an allergic, or similar untoward reaction when administered to a host.

[00478] As used herein, an "input agent responsive domain" is a domain of a transcription factor that binds to or otherwise responds to a condition or input agent in a manner that renders a linked DNA binding fusion domain responsive to the presence of that condition or input. In one embodiment, the presence of the condition or input results in a conformational change in the input agent responsive domain, or in a protein to which it is fused, that modifies the transcription-modulating activity of the transcription factor.

[00479] The term "*in vivo*" refers to assays or processes that occur in or within an organism, such as a multicellular animal. In some of the aspects described herein, a method or use can be said to occur "*in vivo*" when a unicellular organism, such as a bacterium, is used. The term "*ex vivo*" refers to methods and uses that are performed using a living cell with an intact membrane that is outside of the

body of a multicellular animal or plant, *e.g.*, explants, cultured cells, including primary cells and cell lines, transformed cell lines, and extracted tissue or cells, including blood cells, among others. The term "*in vitro*" refers to assays and methods that do not require the presence of a cell with an intact membrane, such as cellular extracts, and can refer to the introducing of a programmable synthetic biological circuit in a non-cellular system, such as a medium not comprising cells or cellular systems, such as cellular extracts.

[00480] The term “promoter,” as used herein, refers to any nucleic acid sequence that regulates the expression of another nucleic acid sequence by driving transcription of the nucleic acid sequence, which can be a heterologous target gene encoding a protein or an RNA. Promoters can be constitutive, inducible, repressible, tissue-specific, or any combination thereof. A promoter is a control region of a nucleic acid sequence at which initiation and rate of transcription of the remainder of a nucleic acid sequence are controlled. A promoter can also contain genetic elements at which regulatory proteins and molecules can bind, such as RNA polymerase and other transcription factors. In some embodiments of the aspects described herein, a promoter can drive the expression of a transcription factor that regulates the expression of the promoter itself. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain “TATA” boxes and “CAT” boxes. Various promoters, including inducible promoters, may be used to drive the expression of transgenes in the ccdNA vectors disclosed herein. A promoter sequence may be bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background.

[00481] The term “enhancer” as used herein refers to a cis-acting regulatory sequence (*e.g.*, 50-1,500 base pairs) that binds one or more proteins (*e.g.*, activator proteins, or transcription factor) to increase transcriptional activation of a nucleic acid sequence. Enhancers can be positioned up to 1,000,000 base pairs upstream of the gene start site or downstream of the gene start site that they regulate. An enhancer can be positioned within an intronic region, or in the exonic region of an unrelated gene.

[00482] A promoter can be said to drive expression or drive transcription of the nucleic acid sequence that it regulates. The phrases “operably linked,” “operatively positioned,” “operatively linked,” “under control,” and “under transcriptional control” indicate that a promoter is in a correct functional location and/or orientation in relation to a nucleic acid sequence it regulates to control transcriptional initiation and/or expression of that sequence. An “inverted promoter,” as used herein, refers to a promoter in which the nucleic acid sequence is in the reverse orientation, such that what was the coding strand is now the non-coding strand, and vice versa. Inverted promoter sequences can be used in various embodiments to regulate the state of a switch. In addition, in various embodiments, a promoter can be used in conjunction with an enhancer.

[00483] A promoter can be one naturally associated with a gene or sequence, as can be obtained by isolating the 5' non-coding sequences located upstream of the coding segment and/or exon of a given gene or sequence. Such a promoter can be referred to as “endogenous.” Similarly, in some embodiments, an enhancer can be one naturally associated with a nucleic acid sequence, located either downstream or upstream of that sequence.

[00484] In some embodiments, a coding nucleic acid segment is positioned under the control of a “recombinant promoter” or “heterologous promoter,” both of which refer to a promoter that is not normally associated with the encoded nucleic acid sequence it is operably linked to in its natural environment. A recombinant or heterologous enhancer refers to an enhancer not normally associated with a given nucleic acid sequence in its natural environment. Such promoters or enhancers can include promoters or enhancers of other genes; promoters or enhancers isolated from any other prokaryotic, viral, or eukaryotic cell; and synthetic promoters or enhancers that are not “naturally occurring,” *i.e.*, comprise different elements of different transcriptional regulatory regions, and/or mutations that alter expression through methods of genetic engineering that are known in the art. In addition to producing nucleic acid sequences of promoters and enhancers synthetically, promoter sequences can be produced using recombinant cloning and/or nucleic acid amplification technology, including PCR, in connection with the synthetic biological circuits and modules disclosed herein (see, *e.g.*, U.S. Pat. No. 4,683,202, U.S. Pat. No. 5,928,906, each incorporated herein by reference). Furthermore, it is contemplated that control sequences that direct transcription and/or expression of sequences within non-nuclear organelles such as mitochondria, chloroplasts, and the like, can be employed as well.

[00485] As described herein, an “inducible promoter” is one that is characterized by initiating or enhancing transcriptional activity when in the presence of, influenced by, or contacted by an inducer or inducing agent. An “inducer” or “inducing agent,” as defined herein, can be endogenous, or a normally exogenous compound or protein that is administered in such a way as to be active in inducing transcriptional activity from the inducible promoter. In some embodiments, the inducer or inducing agent, *i.e.*, a chemical, a compound or a protein, can itself be the result of transcription or expression of a nucleic acid sequence (*i.e.*, an inducer can be an inducer protein expressed by another component or module), which itself can be under the control of an inducible promoter. In some embodiments, an inducible promoter is induced in the absence of certain agents, such as a repressor. Examples of inducible promoters include but are not limited to, tetracycline, metallothionine, ecdysone, mammalian viruses (*e.g.*, the adenovirus late promoter; and the mouse mammary tumor virus long terminal repeat (MMTV-LTR)) and other steroid-responsive promoters, rapamycin responsive promoters and the like.

[00486] The terms “DNA regulatory sequences,” “control elements,” and “regulatory elements,” used interchangeably herein, refer to transcriptional and translational control sequences, such as promoters, enhancers, polyadenylation signals, terminators, protein degradation signals, and the like,

that provide for and/or regulate transcription of a non-coding sequence (*e.g.*, DNA-targeting RNA) or a coding sequence (*e.g.*, site-directed modifying polypeptide, or Cas9/Csn1 polypeptide) and/or regulate translation of an encoded polypeptide.

[00487] The term “Operably linked” as used herein refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression. An “expression cassette” includes a heterologous DNA sequence that is operably linked to a promoter or other regulatory sequence sufficient to direct transcription of the transgene in the ceDNA vector. Suitable promoters include, for example, tissue specific promoters. Promoters can also be of AAV origin.

[00488] The term “subject” as used herein refers to a human or animal, to whom treatment, including prophylactic treatment, with the ceDNA vector according to the present invention, is provided. Usually the animal is a vertebrate such as, but not limited to a primate, rodent, domestic animal or game animal. Primates include but are not limited to, chimpanzees, cynomolgous monkeys, spider monkeys, and macaques, *e.g.*, Rhesus. Rodents include mice, rats, woodchucks, ferrets, rabbits and hamsters. Domestic and game animals include, but are not limited to, cows, horses, pigs, deer, bison, buffalo, feline species, *e.g.*, domestic cat, canine species, *e.g.*, dog, fox, wolf, avian species, *e.g.*, chicken, emu, ostrich, and fish, *e.g.*, trout, catfish and salmon. In certain embodiments of the aspects described herein, the subject is a mammal, *e.g.*, a primate or a human. A subject can be male or female. Additionally, a subject can be an infant or a child. In some embodiments, the subject can be a neonate or an unborn subject, *e.g.*, the subject is *in utero*. Preferably, the subject is a mammal. The mammal can be a human, non-human primate, mouse, rat, dog, cat, horse, or cow, but is not limited to these examples. Mammals other than humans can be advantageously used as subjects that represent animal models of diseases and disorders. In addition, the methods and compositions described herein can be used for domesticated animals and/or pets. A human subject can be of any age, gender, race or ethnic group, *e.g.*, Caucasian (white), Asian, African, black, African American, African European, Hispanic, Mideastern, etc. In some embodiments, the subject can be a patient or other subject in a clinical setting. In some embodiments, the subject is already undergoing treatment. In some embodiments, the subject is an embryo, a fetus, neonate, infant, child, adolescent, or adult. In some embodiments, the subject is a human fetus, human neonate, human infant, human child, human adolescent, or human adult. In some embodiments, the subject is an animal embryo, or non-human embryo or non-human primate embryo. In some embodiments, the subject is a human embryo.

[00489] As used herein, the term “host cell”, includes any cell type that is susceptible to transformation, transfection, transduction, and the like with a nucleic acid construct or ceDNA expression vector of the present disclosure. As non-limiting examples, a host cell can be an isolated primary cell, pluripotent stem cells, CD34⁺ cells), induced pluripotent stem cells, or any of a number

of immortalized cell lines (*e.g.*, HepG2 cells). Alternatively, a host cell can be an *in situ* or *in vivo* cell in a tissue, organ or organism.

[00490] The term “exogenous” refers to a substance present in a cell other than its native source. The term “exogenous” when used herein can refer to a nucleic acid (*e.g.*, a nucleic acid encoding a polypeptide) or a polypeptide that has been introduced by a process involving the hand of man into a biological system such as a cell or organism in which it is not normally found and one wishes to introduce the nucleic acid or polypeptide into such a cell or organism. Alternatively, “exogenous” can refer to a nucleic acid or a polypeptide that has been introduced by a process involving the hand of man into a biological system such as a cell or organism in which it is found in relatively low amounts and one wishes to increase the amount of the nucleic acid or polypeptide in the cell or organism, *e.g.*, to create ectopic expression or levels. In contrast, the term “endogenous” refers to a substance that is native to the biological system or cell.

[00491] The term “sequence identity” refers to the relatedness between two nucleotide sequences. For purposes of the present disclosure, the degree of sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled “longest identity” (obtained using the -nobrief option) is used as the percent identity and is calculated as follows: (Identical Deoxyribonucleotides.times.100)/(Length of Alignment-Total Number of Gaps in Alignment). The length of the alignment is preferably at least 10 nucleotides, preferably at least 25 nucleotides more preferred at least 50 nucleotides and most preferred at least 100 nucleotides.

[00492] The term “homology” or “homologous” as used herein is defined as the percentage of nucleotide residues that are identical to the nucleotide residues in the corresponding sequence on the target chromosome, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleotide sequence homology can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN, ClustalW2 or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. In some embodiments, a nucleic acid sequence (*e.g.*, DNA sequence), for example of a homology arm, is considered “homologous” when the sequence is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more, identical to the corresponding native or unedited nucleic acid sequence (*e.g.*, genomic sequence) of the host cell.

[00493] The term “heterologous,” as used herein, means a nucleotide or polypeptide sequence that is not found in the native nucleic acid or protein, respectively. A heterologous nucleic acid sequence may be linked to a naturally-occurring nucleic acid sequence (or a variant thereof) (*e.g.*, by genetic engineering) to generate a chimeric nucleotide sequence encoding a chimeric polypeptide. A heterologous nucleic acid sequence may be linked to a variant polypeptide (*e.g.*, by genetic engineering) to generate a nucleotide sequence encoding a fusion variant polypeptide.

[00494] A “vector” or “expression vector” is a replicon, such as plasmid, bacmid, phage, virus, virion, or cosmid, to which another DNA segment, *i.e.* an “insert”, may be attached so as to bring about the replication of the attached segment in a cell. A vector can be a nucleic acid construct designed for delivery to a host cell or for transfer between different host cells. As used herein, a vector can be viral or non-viral in origin and/or in final form, however for the purpose of the present disclosure, a “vector” generally refers to a cDNA vector, as that term is used herein. The term “vector” encompasses any genetic element that is capable of replication when associated with the proper control elements and that can transfer gene sequences to cells. In some embodiments, a vector can be an expression vector or recombinant vector.

[00495] As used herein, the term “expression vector” refers to a vector that directs expression of an RNA or polypeptide from sequences linked to transcriptional regulatory sequences on the vector. The sequences expressed will often, but not necessarily, be heterologous to the cell. An expression vector may comprise additional elements, for example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in human cells for expression and in a prokaryotic host for cloning and amplification. The term “expression” refers to the cellular processes involved in producing RNA and proteins and as appropriate, secreting proteins, including where applicable, but not limited to, for example, transcription, transcript processing, translation and protein folding, modification and processing. “Expression products” include RNA transcribed from a gene, and polypeptides obtained by translation of mRNA transcribed from a gene. The term “gene” means the nucleic acid sequence which is transcribed (DNA) to RNA *in vitro* or *in vivo* when operably linked to appropriate regulatory sequences. The gene may or may not include regions preceding and following the coding region, *e.g.*, 5' untranslated (5'UTR) or “leader” sequences and 3' UTR or “trailer” sequences, as well as intervening sequences (introns) between individual coding segments (exons).

[00496] By “recombinant vector” is meant a vector that includes a heterologous nucleic acid sequence, or “transgene” that is capable of expression *in vivo*. It should be understood that the vectors described herein can, in some embodiments, be combined with other suitable compositions and therapies. In some embodiments, the vector is episomal. The use of a suitable episomal vector provides a means of maintaining the nucleotide of interest in the subject in high copy number extra chromosomal DNA thereby eliminating potential effects of chromosomal integration.

[00497] The phrase “genetic disease” as used herein refers to a disease, partially or completely, directly or indirectly, caused by one or more abnormalities in the genome, especially a condition that is present from birth. The abnormality may be a mutation, an insertion or a deletion. The abnormality may affect the coding sequence of the gene or its regulatory sequence. The genetic disease may be, but not limited to DMD, hemophilia, cystic fibrosis, Huntington's chorea, familial hypercholesterolemia (LDL receptor defect), hepatoblastoma, Wilson's disease, congenital hepatic porphyria, inherited disorders of hepatic metabolism, Lesch Nyhan syndrome, sickle cell anemia, thalassaemias, xeroderma pigmentosum, Fanconi's anemia, retinitis pigmentosa, ataxia telangiectasia, Bloom's syndrome, retinoblastoma, and Tay-Sachs disease.

[00498] As used herein the term “comprising” or “comprises” is used in reference to compositions, methods, and respective component(s) thereof, that are essential to the method or composition, yet open to the inclusion of unspecified elements, whether essential or not.

[00499] As used herein the term “consisting essentially of” refers to those elements required for a given embodiment. The term permits the presence of elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment. The use of “comprising” indicates inclusion rather than limitation.

[00500] The term “consisting of” refers to compositions, methods, and respective components thereof as described herein, which are exclusive of any element not recited in that description of the embodiment.

[00501] As used in this specification and the appended claims, the singular forms “a,” “an,” and “the” include plural references unless the context clearly dictates otherwise. Thus for example, references to “the method” includes one or more methods, and/or steps of the type described herein and/or which will become apparent to those persons skilled in the art upon reading this disclosure and so forth. Similarly, the word “or” is intended to include “and” unless the context clearly indicates otherwise. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of this disclosure, suitable methods and materials are described below. The abbreviation, “*e.g.*” is derived from the Latin *exempli gratia*, and is used herein to indicate a non-limiting example. Thus, the abbreviation “*e.g.*” is synonymous with the term “for example.”

[00502] Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction conditions used herein should be understood as modified in all instances by the term “about.” The term “about” when used in connection with percentages can mean $\pm 1\%$. The present invention is further explained in detail by the following examples, but the scope of the invention should not be limited thereto.

[00503] Groupings of alternative elements or embodiments of the invention disclosed herein are not to be construed as limitations. Each group member can be referred to and claimed individually or in any combination with other members of the group or other elements found herein. One or more members of a group can be included in, or deleted from, a group for reasons of convenience and/or

patentability. When any such inclusion or deletion occurs, the specification is herein deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

[00504] In some embodiments of any of the aspects, the disclosure described herein does not concern a process for cloning human beings, processes for modifying the germ line genetic identity of human beings, uses of human embryos for industrial or commercial purposes or processes for modifying the genetic identity of animals which are likely to cause them suffering without any substantial medical benefit to man or animal, and also animals resulting from such processes.

[00505] Other terms are defined herein within the description of the various aspects of the invention.

[00506] All patents and other publications; including literature references, issued patents, published patent applications, and co-pending patent applications; cited throughout this application are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the technology described herein. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

[00507] The description of embodiments of the disclosure is not intended to be exhaustive or to limit the disclosure to the precise form disclosed. While specific embodiments of, and examples for, the disclosure are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the disclosure, as those skilled in the relevant art will recognize. For example, while method steps or functions are presented in a given order, alternative embodiments may perform functions in a different order, or functions may be performed substantially concurrently. The teachings of the disclosure provided herein can be applied to other procedures or methods as appropriate. The various embodiments described herein can be combined to provide further embodiments. Aspects of the disclosure can be modified, if necessary, to employ the compositions, functions and concepts of the above references and application to provide yet further embodiments of the disclosure. Moreover, due to biological functional equivalency considerations, some changes can be made in protein structure without affecting the biological or chemical action in kind or amount. These and other changes can be made to the disclosure in light of the detailed description. All such modifications are intended to be included within the scope of the appended claims.

[00508] Specific elements of any of the foregoing embodiments can be combined or substituted for elements in other embodiments. Furthermore, while advantages associated with certain embodiments of the disclosure have been described in the context of these embodiments, other embodiments may

also exhibit such advantages, and not all embodiments need necessarily exhibit such advantages to fall within the scope of the disclosure.

[00509] The technology described herein is further illustrated by the following examples which in no way should be construed as being further limiting. It should be understood that this invention is not limited to the particular methodology, protocols, and reagents, etc., described herein and as such can vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims.

EXAMPLES

[00510] The following examples are provided by way of illustration not limitation. It will be appreciated by one of ordinary skill in the art that ceDNA vectors can be constructed from any of the wild-type or modified ITRs described herein, and that the following exemplary methods can be used to construct and assess the activity of such ceDNA vectors. While the methods are exemplified with certain ceDNA vectors, they are applicable to any ceDNA vector in keeping with the description.

EXAMPLE 1: Constructing ceDNA Vectors Using an Insect Cell-Based Method

[00511] Production of the ceDNA vectors using a polynucleotide construct template is described in Example 1 of PCT/US18/49996, which is incorporated herein in its entirety by reference. For example, a polynucleotide construct template used for generating the ceDNA vectors of the present invention can be a ceDNA-plasmid, a ceDNA-Bacmid, and/or a ceDNA-baculovirus. Without being limited to theory, in a permissive host cell, in the presence of *e.g.*, Rep, the polynucleotide construct template having two symmetric ITRs and an expression construct, where at least one of the ITRs is modified relative to a wild-type ITR sequence, replicates to produce ceDNA vectors. ceDNA vector production undergoes two steps: first, excision (“rescue”) of template from the template backbone (*e.g.* ceDNA-plasmid, ceDNA-bacmid, ceDNA-baculovirus genome etc.) via Rep proteins, and second, Rep mediated replication of the excised ceDNA vector.

[00512] An exemplary method to produce ceDNA vectors is from a ceDNA-plasmid as described herein. Referring to **FIG. 1A and 1B**, the polynucleotide construct template of each of the ceDNA-plasmids includes both a left modified ITR and a right modified ITR with the following between the ITR sequences: (i) an enhancer/promoter; (ii) a cloning site for a transgene; (iii) a posttranscriptional response element (*e.g.* the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE)); and (iv) a poly-adenylation signal (*e.g.* from bovine growth hormone gene (BGHpA). Unique restriction endonuclease recognition sites (R1-R6) (shown in **FIG. 1A** and **FIG. 1B**) were also introduced between each component to facilitate the introduction of new genetic components into the specific sites in the construct. R3 (PmeI) GTTTAAAC (SEQ ID NO: 123) and R4 (PacI) TTAATTAA (SEQ ID NO: 124) enzyme sites are engineered into the cloning site to introduce an open reading frame of a transgene. These sequences were cloned into a pFastBac HT B plasmid obtained from ThermoFisher Scientific.

[00513] Production of ceDNA-bacmids:

[00514] DH10Bac competent cells (MAX EFFICIENCY® DH10Bac™ Competent Cells, Thermo Fisher) were transformed with either test or control plasmids following a protocol according to the manufacturer's instructions. Recombination between the plasmid and a baculovirus shuttle vector in the DH10Bac cells were induced to generate recombinant ceDNA-bacmids. The recombinant bacmids were selected by screening a positive selection based on blue-white screening in *E. coli* (Φ 80dlacZAM15 marker provides α -complementation of the β -galactosidase gene from the bacmid vector) on a bacterial agar plate containing X-gal and IPTG with antibiotics to select for transformants and maintenance of the bacmid and transposase plasmids. White colonies caused by transposition that disrupts the β -galactoside indicator gene were picked and cultured in 10 ml of media.

[00515] The recombinant ceDNA-bacmids were isolated from the *E. coli* and transfected into Sf9 or Sf21 insect cells using FugeneHD to produce infectious baculovirus. The adherent Sf9 or Sf21 insect cells were cultured in 50 ml of media in T25 flasks at 25°C. Four days later, culture medium (containing the P0 virus) was removed from the cells, filtered through a 0.45 μ m filter, separating the infectious baculovirus particles from cells or cell debris.

[00516] Optionally, the first generation of the baculovirus (P0) was amplified by infecting naïve Sf9 or Sf21 insect cells in 50 to 500 ml of media. Cells were maintained in suspension cultures in an orbital shaker incubator at 130 rpm at 25 °C, monitoring cell diameter and viability, until cells reach a diameter of 18-19 nm (from a naïve diameter of 14-15 nm), and a density of $\sim 4.0 \times 10^6$ cells/mL. Between 3 and 8 days post-infection, the P1 baculovirus particles in the medium were collected following centrifugation to remove cells and debris then filtration through a 0.45 μ m filter.

[00517] The ceDNA-baculovirus comprising the test constructs were collected and the infectious activity, or titer, of the baculovirus was determined. Specifically, four x 20 ml Sf9 cell cultures at 2.5×10^6 cells/ml were treated with P1 baculovirus at the following dilutions: 1/1000, 1/10,000, 1/50,000, 1/100,000, and incubated at 25-27°C. Infectivity was determined by the rate of cell diameter increase and cell cycle arrest, and change in cell viability every day for 4 to 5 days.

[00518] A "Rep-plasmid" as disclosed in FIG. 8A of PCT/US18/49996, which is incorporated herein in its entirety by reference, was produced in a pFASTBAC™-Dual expression vector (ThermoFisher) comprising both the Rep78 (SEQ ID NO: 131 or 133) and Rep52 (SEQ ID NO: 132) or Rep68 (SEQ ID NO: 130) and Rep40 (SEQ ID NO: 129). The Rep-plasmid was transformed into the DH10Bac competent cells (MAX EFFICIENCY® DH10Bac™ Competent Cells (Thermo Fisher) following a protocol provided by the manufacturer. Recombination between the Rep-plasmid and a baculovirus shuttle vector in the DH10Bac cells were induced to generate recombinant bacmids ("Rep-bacmids"). The recombinant bacmids were selected by a positive selection that included-blue-white screening in *E. coli* (Φ 80dlacZAM15 marker provides α -complementation of the β -galactosidase gene from the bacmid vector) on a bacterial agar plate containing X-gal and IPTG. Isolated white colonies were picked and inoculated in 10 ml of selection media (kanamycin,

gentamicin, tetracycline in LB broth). The recombinant bacmids (Rep-bacmids) were isolated from the *E. coli* and the Rep-bacmids were transfected into Sf9 or Sf21 insect cells to produce infectious baculovirus.

[00519] The Sf9 or Sf21 insect cells were cultured in 50 ml of media for 4 days, and infectious recombinant baculovirus (“Rep-baculovirus”) were isolated from the culture. Optionally, the first generation Rep-baculovirus (P0) were amplified by infecting naïve Sf9 or Sf21 insect cells and cultured in 50 to 500 ml of media. Between 3 and 8 days post-infection, the P1 baculovirus particles in the medium were collected either by separating cells by centrifugation or filtration or another fractionation process. The Rep-baculovirus were collected and the infectious activity of the baculovirus was determined. Specifically, four x 20 mL Sf9 cell cultures at 2.5×10^6 cells/mL were treated with P1 baculovirus at the following dilutions, 1/1000, 1/10,000, 1/50,000, 1/100,000, and incubated. Infectivity was determined by the rate of cell diameter increase and cell cycle arrest, and change in cell viability every day for 4 to 5 days.

[00520] ceDNA vector generation and characterization

[00521] With reference to **FIG. 4B**, Sf9 insect cell culture media containing either (1) a sample-containing a ceDNA-bacmid or a ceDNA-baculovirus, and (2) Rep-baculovirus described above were then added to a fresh culture of Sf9 cells (2.5×10^6 cells/mL, 20ml) at a ratio of 1:1000 and 1:10,000, respectively. The cells were then cultured at 130 rpm at 25°C. 4-5 days after the co-infection, cell diameter and viability are detected. When cell diameters reached 18-20nm with a viability of ~70-80%, the cell cultures were centrifuged, the medium was removed, and the cell pellets were collected. The cell pellets are first resuspended in an adequate volume of aqueous medium, either water or buffer. The ceDNA vector was isolated and purified from the cells using Qiagen MIDI PLUS™ purification protocol (Qiagen, 0.2mg of cell pellet mass processed per column).

[00522] Yields of ceDNA vectors produced and purified from the Sf9 insect cells were initially determined based on UV absorbance at 260nm.

[00523] ceDNA vectors can be assessed by identified by agarose gel electrophoresis under native or denaturing conditions as illustrated in **FIG. 4D**, where (a) the presence of characteristic bands migrating at twice the size on denaturing gels versus native gels after restriction endonuclease cleavage and gel electrophoretic analysis and (b) the presence of monomer and dimer (2x) bands on denaturing gels for uncleaved material is characteristic of the presence of ceDNA vector.

[00524] Structures of the isolated ceDNA vectors were further analyzed by digesting the DNA obtained from co-infected Sf9 cells (as described herein) with restriction endonucleases selected for a) the presence of only a single cut site within the ceDNA vectors, and b) resulting fragments that were large enough to be seen clearly when fractionated on a 0.8% denaturing agarose gel (>800 bp). As illustrated in **FIGS. 4D and 4E**, linear DNA vectors with a non-continuous structure and ceDNA vector with the linear and continuous structure can be distinguished by sizes of their reaction products— for example, a DNA vector with a non-continuous structure is expected to produce 1kb and

2kb fragments, while a non-encapsidated vector with the continuous structure is expected to produce 2kb and 4kb fragments.

[00525] Therefore, to demonstrate in a qualitative fashion that isolated ceDNA vectors are covalently closed-ended as is required by definition, the samples were digested with a restriction endonuclease identified in the context of the specific DNA vector sequence as having a single restriction site, preferably resulting in two cleavage products of unequal size (*e.g.*, 1000 bp and 2000 bp). Following digestion and electrophoresis on a denaturing gel (which separates the two complementary DNA strands), a linear, non-covalently closed DNA will resolve at sizes 1000 bp and 2000 bp, while a covalently closed DNA (*i.e.*, a ceDNA vector) will resolve at 2x sizes (2000 bp and 4000 bp), as the two DNA strands are linked and are now unfolded and twice the length (though single stranded). Furthermore, digestion of monomeric, dimeric, and *n*-meric forms of the DNA vectors will all resolve as the same size fragments due to the end-to-end linking of the multimeric DNA vectors (see **FIG. 4D**).

[00526] As used herein, the phrase “assay for the Identification of DNA vectors by agarose gel electrophoresis under native gel and denaturing conditions” refers to an assay to assess the close-endedness of the ceDNA by performing restriction endonuclease digestion followed by electrophoretic assessment of the digest products. One such exemplary assay follows, though one of ordinary skill in the art will appreciate that many art-known variations on this example are possible. The restriction endonuclease is selected to be a single cut enzyme for the ceDNA vector of interest that will generate products of approximately 1/3x and 2/3x of the DNA vector length. This resolves the bands on both native and denaturing gels. Before denaturation, it is important to remove the buffer from the sample. The Qiagen PCR clean-up kit or desalting “spin columns,” *e.g.* GE HEALTHCARE ILUSTRA™ MICROSPIN™ G-25 columns are some art-known options for the endonuclease digestion. The assay includes for example, i) digest DNA with appropriate restriction endonuclease(s), 2) apply to *e.g.*, a Qiagen PCR clean-up kit, elute with distilled water, iii) adding 10x denaturing solution (10x = 0.5 M NaOH, 10mM EDTA), add 10X dye, not buffered, and analyzing, together with DNA ladders prepared by adding 10X denaturing solution to 4x, on a 0.8 – 1.0 % gel previously incubated with 1mM EDTA and 200mM NaOH to ensure that the NaOH concentration is uniform in the gel and gel box, and running the gel in the presence of 1x denaturing solution (50 mM NaOH, 1mM EDTA). One of ordinary skill in the art will appreciate what voltage to use to run the electrophoresis based on size and desired timing of results. After electrophoresis, the gels are drained and neutralized in 1x TBE or TAE and transferred to distilled water or 1x TBE/TAE with 1x SYBR Gold. Bands can then be visualized with *e.g.* Thermo Fisher, SYBR® Gold Nucleic Acid Gel Stain (10,000X Concentrate in DMSO) and epifluorescent light (blue) or UV (312nm).

[00527] The purity of the generated ceDNA vector can be assessed using any art-known method. As one exemplary and non-limiting method, contribution of ceDNA-plasmid to the overall UV absorbance of a sample can be estimated by comparing the fluorescent intensity of ceDNA vector to a

standard. For example, if based on UV absorbance 4 μ g of ceDNA vector was loaded on the gel, and the ceDNA vector fluorescent intensity is equivalent to a 2kb band which is known to be 1 μ g, then there is 1 μ g of ceDNA vector, and the ceDNA vector is 25% of the total UV absorbing material. Band intensity on the gel is then plotted against the calculated input that band represents – for example, if the total ceDNA vector is 8kb, and the excised comparative band is 2kb, then the band intensity would be plotted as 25% of the total input, which in this case would be .25 μ g for 1.0 μ g input. Using the ceDNA vector plasmid titration to plot a standard curve, a regression line equation is then used to calculate the quantity of the ceDNA vector band, which can then be used to determine the percent of total input represented by the ceDNA vector, or percent purity.

[00528] For comparative purposes, Example 1 describes the production of ceDNA vectors using an insect cell based method and a polynucleotide construct template, and is also described in Example 1 of PCT/US18/49996, which is incorporated herein in its entirety by reference. For example, a polynucleotide construct template used for generating the ceDNA vectors of the present invention according to Example 1 can be a ceDNA-plasmid, a ceDNA-Bacmid, and/or a ceDNA-baculovirus. Without being limited to theory, in a permissive host cell, in the presence of *e.g.*, Rep, the polynucleotide construct template having two symmetric ITRs and an expression construct, where at least one of the ITRs is modified relative to a wild-type ITR sequence, replicates to produce ceDNA vectors. ceDNA vector production undergoes two steps: first, excision (“rescue”) of template from the template backbone (*e.g.* ceDNA-plasmid, ceDNA-bacmid, ceDNA-baculovirus genome etc.) via Rep proteins, and second, Rep mediated replication of the excised ceDNA vector.

[00529] An exemplary method to produce ceDNA vectors in a method using insect cell is from a ceDNA-plasmid as described herein. Referring to **FIG. 1A and 1B**, the polynucleotide construct template of each of the ceDNA-plasmids includes both a left modified ITR and a right modified ITR with the following between the ITR sequences: (i) an enhancer/promoter; (ii) a cloning site for a transgene; (iii) a posttranscriptional response element (*e.g.* the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE)); and (iv) a poly-adenylation signal (*e.g.* from bovine growth hormone gene (BGHpA). Unique restriction endonuclease recognition sites (R1-R6) (shown in **FIG. 1A** and **FIG. 1B**) were also introduced between each component to facilitate the introduction of new genetic components into the specific sites in the construct. R3 (PmeI) GTTTAAAC (SEQ ID NO: 123) and R4 (PacI) TTAATTAA (SEQ ID NO: 124) enzyme sites are engineered into the cloning site to introduce an open reading frame of a transgene. These sequences were cloned into a pFastBac HT B plasmid obtained from ThermoFisher Scientific.

[00530] Production of ceDNA-bacmids:

[00531] DH10Bac competent cells (MAX EFFICIENCY® DH10Bac™ Competent Cells, Thermo Fisher) were transformed with either test or control plasmids following a protocol according to the manufacturer’s instructions. Recombination between the plasmid and a baculovirus shuttle vector in the DH10Bac cells were induced to generate recombinant ceDNA-bacmids. The recombinant

bacmids were selected by screening a positive selection based on blue-white screening in *E. coli* (Φ 80dlacZ Δ M15 marker provides α -complementation of the β -galactosidase gene from the bacmid vector) on a bacterial agar plate containing X-gal and IPTG with antibiotics to select for transformants and maintenance of the bacmid and transposase plasmids. White colonies caused by transposition that disrupts the β -galactoside indicator gene were picked and cultured in 10 ml of media.

[00532] The recombinant ceDNA-bacmids were isolated from the *E. coli* and transfected into Sf9 or Sf21 insect cells using FugeneHD to produce infectious baculovirus. The adherent Sf9 or Sf21 insect cells were cultured in 50 ml of media in T25 flasks at 25°C. Four days later, culture medium (containing the P0 virus) was removed from the cells, filtered through a 0.45 μ m filter, separating the infectious baculovirus particles from cells or cell debris.

[00533] Optionally, the first generation of the baculovirus (P0) was amplified by infecting naïve Sf9 or Sf21 insect cells in 50 to 500 ml of media. Cells were maintained in suspension cultures in an orbital shaker incubator at 130 rpm at 25 °C, monitoring cell diameter and viability, until cells reach a diameter of 18-19 nm (from a naïve diameter of 14-15 nm), and a density of $\sim 4.0 \times 10^6$ cells/mL. Between 3 and 8 days post-infection, the P1 baculovirus particles in the medium were collected following centrifugation to remove cells and debris then filtration through a 0.45 μ m filter.

[00534] The ceDNA-baculovirus comprising the test constructs were collected and the infectious activity, or titer, of the baculovirus was determined. Specifically, four x 20 ml Sf9 cell cultures at 2.5×10^6 cells/ml were treated with P1 baculovirus at the following dilutions: 1/1000, 1/10,000, 1/50,000, 1/100,000, and incubated at 25-27°C. Infectivity was determined by the rate of cell diameter increase and cell cycle arrest, and change in cell viability every day for 4 to 5 days.

[00535] A “Rep-plasmid” was produced in a pFASTBACTM-Dual expression vector (ThermoFisher) comprising both the Rep78 (SEQ ID NO: 131 or 133) or Rep68 (SEQ ID NO: 130) and Rep52 (SEQ ID NO: 132) or Rep40 (SEQ ID NO: 129). The Rep-plasmid was transformed into the DH10Bac competent cells (MAX EFFICIENCY® DH10BacTM Competent Cells (Thermo Fisher) following a protocol provided by the manufacturer. Recombination between the Rep-plasmid and a baculovirus shuttle vector in the DH10Bac cells were induced to generate recombinant bacmids (“Rep-bacmids”). The recombinant bacmids were selected by a positive selection that included-blue-white screening in *E. coli* (Φ 80dlacZ Δ M15 marker provides α -complementation of the β -galactosidase gene from the bacmid vector) on a bacterial agar plate containing X-gal and IPTG. Isolated white colonies were picked and inoculated in 10 ml of selection media (kanamycin, gentamicin, tetracycline in LB broth). The recombinant bacmids (Rep-bacmids) were isolated from the *E. coli* and the Rep-bacmids were transfected into Sf9 or Sf21 insect cells to produce infectious baculovirus.

[00536] The Sf9 or Sf21 insect cells were cultured in 50 ml of media for 4 days, and infectious recombinant baculovirus (“Rep-baculovirus”) were isolated from the culture. Optionally, the first generation Rep-baculovirus (P0) were amplified by infecting naïve Sf9 or Sf21 insect cells and

cultured in 50 to 500 ml of media. Between 3 and 8 days post-infection, the P1 baculovirus particles in the medium were collected either by separating cells by centrifugation or filtration or another fractionation process. The Rep-baculovirus were collected and the infectious activity of the baculovirus was determined. Specifically, four x 20 mL Sf9 cell cultures at 2.5×10^6 cells/mL were treated with P1 baculovirus at the following dilutions, 1/1000, 1/10,000, 1/50,000, 1/100,000, and incubated. Infectivity was determined by the rate of cell diameter increase and cell cycle arrest, and change in cell viability every day for 4 to 5 days.

[00537] ceDNA vector generation and characterization

[00538] Sf9 insect cell culture media containing either (1) a sample-containing a ceDNA-bacmid or a ceDNA-baculovirus, and (2) Rep-baculovirus described above were then added to a fresh culture of Sf9 cells (2.5×10^6 cells/mL, 20ml) at a ratio of 1:1000 and 1:10,000, respectively. The cells were then cultured at 130 rpm at 25°C. 4-5 days after the co-infection, cell diameter and viability are detected. When cell diameters reached 18-20nm with a viability of ~70-80%, the cell cultures were centrifuged, the medium was removed, and the cell pellets were collected. The cell pellets are first resuspended in an adequate volume of aqueous medium, either water or buffer. The ceDNA vector was isolated and purified from the cells using Qiagen MIDI PLUS™ purification protocol (Qiagen, 0.2mg of cell pellet mass processed per column).

[00539] Yields of ceDNA vectors produced and purified from the Sf9 insect cells were initially determined based on UV absorbance at 260nm. The purified ceDNA vectors can be assessed for proper closed-ended configuration using the electrophoretic methodology described in Example 5.

EXAMPLE 2: Synthetic ceDNA production via excision from a double-stranded DNA molecule

[00540] Synthetic production of the ceDNA vectors is described in Examples 2-6 of International Application PCT/US19/14122, filed January 18, 2019, which is incorporated herein in its entirety by reference. One exemplary method of producing a ceDNA vector using a synthetic method that involves the excision of a double-stranded DNA molecule. In brief, a ceDNA vector can be generated using a double stranded DNA construct, *e.g.*, see FIGS. 7A-8E of PCT/US19/14122. In some embodiments, the double stranded DNA construct is a ceDNA plasmid, *e.g.*, see, *e.g.*, FIG. 6 in International patent application PCT/US2018/064242, filed December 6, 2018).

[00541] In some embodiments, a construct to make a ceDNA vector comprises a regulatory switch as described herein.

[00542] For illustrative purposes, Example 1 describes producing ceDNA vectors as exemplary closed-ended DNA vectors generated using this method. However, while ceDNA vectors are exemplified in this Example to illustrate *in vitro* synthetic production methods to generate a closed-ended DNA vector by excision of a double-stranded polynucleotide comprising the ITRs and expression cassette (*e.g.*, heterologous nucleic acid sequence) followed by ligation of the free 3' and 5' ends as described herein, one of ordinary skill in the art is aware that one can, as illustrated above,

modify the double stranded DNA polynucleotide molecule such that any desired closed-ended DNA vector is generated, including but not limited to, doggybone DNA, dumbbell DNA and the like.

Exemplary ceDNA vectors for production of transgenes and therapeutic proteins can be produced by the synthetic production method described in Example 2.

[00543] The method involves (i) excising a sequence encoding the expression cassette from a double-stranded DNA construct and (ii) forming hairpin structures at one or more of the ITRs and (iii) joining the free 5' and 3' ends by ligation, *e.g.*, by T4 DNA ligase.

[00544] The double-stranded DNA construct comprises, in 5' to 3' order: a first restriction endonuclease site; an upstream ITR; an expression cassette; a downstream ITR; and a second restriction endonuclease site. The double-stranded DNA construct is then contacted with one or more restriction endonucleases to generate double-stranded breaks at both of the restriction endonuclease sites. One endonuclease can target both sites, or each site can be targeted by a different endonuclease as long as the restriction sites are not present in the ceDNA vector template. This excises the sequence between the restriction endonuclease sites from the rest of the double-stranded DNA construct (see Fig. 9 of PCT/US19/14122). Upon ligation a closed-ended DNA vector is formed.

[00545] One or both of the ITRs used in the method may be wild-type ITRs. Modified ITRs may also be used, where the modification can include deletion, insertion, or substitution of one or more nucleotides from the wild-type ITR in the sequences forming B and B' arm and/or C and C' arm (see, *e.g.*, Figs. 6-8 and 10 FIG. 11B of PCT/US19/14122), and may have two or more hairpin loops (see, *e.g.*, Figs. 6-8 FIG. 11B of PCT/US19/14122) or a single hairpin loop (see, *e.g.*, Fig. 10A-10B FIG. 11B of PCT/US19/14122). The hairpin loop modified ITR can be generated by genetic modification of an existing oligo or by de novo biological and/or chemical synthesis.

[00546] In a non-limiting example, ITR-6 Left and Right (SEQ ID NOS: 111 and 112), include 40 nucleotide deletions in the B-B' and C-C' arms from the wild-type ITR of AAV2. Nucleotides remaining in the modified ITR are predicted to form a single hairpin structure. Gibbs free energy of unfolding the structure is about -54.4 kcal/mol. Other modifications to the ITR may also be made, including optional deletion of a functional Rep binding site or a Trs site.

EXAMPLE 3: ceDNA production via oligonucleotide construction

[00547] Another exemplary method of producing a ceDNA vector using a synthetic method that involves assembly of various oligonucleotides, is provided in Example 3 of PCT/US19/14122, where a ceDNA vector is produced by synthesizing a 5' oligonucleotide and a 3' ITR oligonucleotide and ligating the ITR oligonucleotides to a double-stranded polynucleotide comprising an expression cassette. FIG. 11B of PCT/US19/14122 shows an exemplary method of ligating a 5' ITR oligonucleotide and a 3' ITR oligonucleotide to a double stranded polynucleotide comprising an expression cassette.

[00548] As disclosed herein, the ITR oligonucleotides can comprise WT-ITRs (*e.g.*, see **FIG. 3A**, **FIG. 3C**), or modified ITRs (*e.g.*, see **FIG. 3B** and **FIG. 3D**). (See also, *e.g.*, FIGS. 6A, 6B, 7A and

7B of PCT/US19/14122, which is incorporated herein in its entirety). Exemplary ITR oligonucleotides include, but are not limited to SEQ ID NOS: 134-145 (*e.g.*, see Table 7 in of PCT/US19/14122).

Modified ITRs can include deletion, insertion, or substitution of one or more nucleotides from the wild-type ITR in the sequences forming B and B' arm and/or C and C' arm. ITR oligonucleotides, comprising WT-ITRs or mod-ITRs as described herein, to be used in the cell-free synthesis, can be generated by genetic modification or biological and/or chemical synthesis. As discussed herein, the ITR oligonucleotides in Examples 2 and 3 can comprise WT-ITRs, or modified ITRs (mod-ITRs) in symmetrical or asymmetrical configurations, as discussed herein.

EXAMPLE 4: ceDNA production via a single-stranded DNA molecule

[00549] Another exemplary method of producing a ceDNA vector using a synthetic method is provided in Example 4 of PCT/US19/14122, and uses a single-stranded linear DNA comprising two sense ITRs which flank a sense expression cassette sequence and are attached covalently to two antisense ITRs which flank an antisense expression cassette, the ends of which single stranded linear DNA are then ligated to form a closed-ended single-stranded molecule. One non-limiting example comprises synthesizing and/or producing a single-stranded DNA molecule, annealing portions of the molecule to form a single linear DNA molecule which has one or more base-paired regions of secondary structure, and then ligating the free 5' and 3' ends to each other to form a closed single-stranded molecule.

[00550] An exemplary single-stranded DNA molecule for production of a ceDNA vector comprises, from 5' to 3': a sense first ITR; a sense expression cassette sequence; a sense second ITR; an antisense second ITR; an antisense expression cassette sequence; and an antisense first ITR.

[00551] A single-stranded DNA molecule for use in the exemplary method of Example 4 can be formed by any DNA synthesis methodology described herein, *e.g.*, *in vitro* DNA synthesis, or provided by cleaving a DNA construct (*e.g.*, a plasmid) with nucleases and melting the resulting dsDNA fragments to provide ssDNA fragments.

[00552] Annealing can be accomplished by lowering the temperature below the calculated melting temperatures of the sense and antisense sequence pairs. The melting temperature is dependent upon the specific nucleotide base content and the characteristics of the solution being used, *e.g.*, the salt concentration. Melting temperatures for any given sequence and solution combination are readily calculated by one of ordinary skill in the art.

[00553] The free 5' and 3' ends of the annealed molecule can be ligated to each other, or ligated to a hairpin molecule to form the ceDNA vector. Suitable exemplary ligation methodologies and hairpin molecules are described in Examples 2 and 3.

EXAMPLE 5: Purifying and/or confirming production of ceDNA

[00554] Any of the DNA vector products produced by the methods described herein, *e.g.*, including the insect cell based production methods described in Example 1, or synthetic production methods described in Examples 2-4 can be purified, *e.g.*, to remove impurities, unused components, or

byproducts using methods commonly known by a skilled artisan; and/or can be analyzed to confirm that DNA vector produced, (in this instance, a ceDNA vector) is the desired molecule. An exemplary method for purification of the DNA vector, *e.g.*, ceDNA is using Qiagen Midi Plus purification protocol (Qiagen) and/or by gel purification,

[00555] The following is an exemplary method for confirming the identity of ceDNA vectors.

[00556] ceDNA vectors can be assessed by identified by agarose gel electrophoresis under native or denaturing conditions as illustrated in **FIG. 4D**, where (a) the presence of characteristic bands migrating at twice the size on denaturing gels versus native gels after restriction endonuclease cleavage and gel electrophoretic analysis and (b) the presence of monomer and dimer (2x) bands on denaturing gels for uncleaved material is characteristic of the presence of ceDNA vector.

[00557] Structures of the isolated ceDNA vectors were further analyzed by digesting the purified DNA with restriction endonucleases selected for a) the presence of only a single cut site within the ceDNA vectors, and b) resulting fragments that were large enough to be seen clearly when fractionated on a 0.8% denaturing agarose gel (>800 bp). As illustrated in **FIGS. 4C and 4D**, linear DNA vectors with a non-continuous structure and ceDNA vector with the linear and continuous structure can be distinguished by sizes of their reaction products— for example, a DNA vector with a non-continuous structure is expected to produce 1kb and 2kb fragments, while a ceDNA vector with the continuous structure is expected to produce 2kb and 4kb fragments.

[00558] Therefore, to demonstrate in a qualitative fashion that isolated ceDNA vectors are covalently closed-ended as is required by definition, the samples were digested with a restriction endonuclease identified in the context of the specific DNA vector sequence as having a single restriction site, preferably resulting in two cleavage products of unequal size (*e.g.*, 1000 bp and 2000 bp). Following digestion and electrophoresis on a denaturing gel (which separates the two complementary DNA strands), a linear, non-covalently closed DNA will resolve at sizes 1000 bp and 2000 bp, while a covalently closed DNA (*i.e.*, a ceDNA vector) will resolve at 2x sizes (2000 bp and 4000 bp), as the two DNA strands are linked and are now unfolded and twice the length (though single stranded). Furthermore, digestion of monomeric, dimeric, and *n*-meric forms of the DNA vectors will all resolve as the same size fragments due to the end-to-end linking of the multimeric DNA vectors (*see FIG. 4E*).

[00559] As used herein, the phrase “assay for the Identification of DNA vectors by agarose gel electrophoresis under native gel and denaturing conditions” refers to an assay to assess the close-endedness of the ceDNA by performing restriction endonuclease digestion followed by electrophoretic assessment of the digest products. One such exemplary assay follows, though one of ordinary skill in the art will appreciate that many art-known variations on this example are possible. The restriction endonuclease is selected to be a single cut enzyme for the ceDNA vector of interest that will generate products of approximately 1/3x and 2/3x of the DNA vector length. This resolves the bands on both native and denaturing gels. Before denaturation, it is important to remove

the buffer from the sample. The Qiagen PCR clean-up kit or desalting “spin columns,” *e.g.* GE HEALTHCARE ILUSTRA™ MICROSPIN™ G-25 columns are some art-known options for the endonuclease digestion. The assay includes for example, i) digest DNA with appropriate restriction endonuclease(s), 2) apply to *e.g.*, a Qiagen PCR clean-up kit, elute with distilled water, iii) adding 10x denaturing solution (10x = 0.5 M NaOH, 10mM EDTA), add 10X dye, not buffered, and analyzing, together with DNA ladders prepared by adding 10X denaturing solution to 4x, on a 0.8 – 1.0 % gel previously incubated with 1mM EDTA and 200mM NaOH to ensure that the NaOH concentration is uniform in the gel and gel box, and running the gel in the presence of 1x denaturing solution (50 mM NaOH, 1mM EDTA). One of ordinary skill in the art will appreciate what voltage to use to run the electrophoresis based on size and desired timing of results. After electrophoresis, the gels are drained and neutralized in 1x TBE or TAE and transferred to distilled water or 1x TBE/TAE with 1x SYBR Gold. Bands can then be visualized with *e.g.* Thermo Fisher, SYBR® Gold Nucleic Acid Gel Stain (10,000X Concentrate in DMSO) and epifluorescent light (blue) or UV (312nm). The foregoing gel-based method can be adapted to purification purposes by isolating the ceDNA vector from the gel band and permitting it to renature.

[00560] The purity of the generated ceDNA vector can be assessed using any art-known method. As one exemplary and non-limiting method, contribution of ceDNA-plasmid to the overall UV absorbance of a sample can be estimated by comparing the fluorescent intensity of ceDNA vector to a standard. For example, if based on UV absorbance 4μg of ceDNA vector was loaded on the gel, and the ceDNA vector fluorescent intensity is equivalent to a 2kb band which is known to be 1μg, then there is 1μg of ceDNA vector, and the ceDNA vector is 25% of the total UV absorbing material. Band intensity on the gel is then plotted against the calculated input that band represents – for example, if the total ceDNA vector is 8kb, and the excised comparative band is 2kb, then the band intensity would be plotted as 25% of the total input, which in this case would be .25μg for 1.0μg input. Using the ceDNA vector plasmid titration to plot a standard curve, a regression line equation is then used to calculate the quantity of the ceDNA vector band, which can then be used to determine the percent of total input represented by the ceDNA vector, or percent purity.

EXAMPLE 6: Controlled transgene expression from ceDNA: transgene expression from the ceDNA vector *in vivo* can be sustained and/or increased by re-dose administration.

[00561] A ceDNA vector was produced according to the methods described in Example 1 above, using a ceDNA plasmid comprising a CAG promoter (SEQ ID NO: 72) and a luciferase transgene (SEQ ID NO: 56) is used as an exemplary transgene, flanked between asymmetric ITRs (*e.g.*, a 5' WT-ITR (SEQ ID NO: 2) and a 3' mod-ITR (SEQ ID NO: 3) and was assessed in different treatment paradigms *in vivo*. This ceDNA vector was used in all subsequent experiments described in Examples 6-10. In Example 6, the ceDNA vector was purified and formulated with a lipid nanoparticle (LNP ceDNA) and injected into the tail vein of each CD-1® IGS mice. Liposomes were formulated with a

suitable lipid blend comprising four components to form lipid nanoparticles (LNP) liposomes, including cationic lipids, helper lipids, cholesterol and PEG-lipids.

[00562] To assess the sustained expression of the transgene *in vivo* from the ceDNA vector over a long time period, the LNP-ceDNA was administered in sterile PBS by tail vein intravenous injection to CD-1® IGS mice of approximately 5-7 weeks of age. Three different dosage groups were assessed: 0.1mg/kg, 0.5 mg/kg, and 1.0 mg/kg, ten mice per group (except 1.0 mg/kg which had 15 mice per group). Injections were administered on day 0. Five mice from each of the groups were injected with an additional identical dose on day 28. Luciferase expression was measured by IVIS imaging following intravenous administration into CD-1® IGS mice (Charles River Laboratories; WT mice). Luciferase expression was assessed by IVIS imaging following intraperitoneal injection of 150 mg/kg luciferin substrate on days 3, 4, 7, 14, 21, 28, 31, 35, and 42, and routinely (*e.g.*, weekly, biweekly or every 10-days or every 2 weeks), between days 42-110 days. Luciferase transgene expression as the exemplary a transgene as measured by IVIS imaging for at least 132 days after 3 different administration protocols (data not shown).

[00563] An extension study was performed to investigate the effect of a re-dose, *e.g.*, a re-administration of LNP-ceDNA expressing luciferase of the LNP-ceDNA treated subjects. In particular, it was assessed to determine if expression levels can be increased by one or more additional administrations of the ceDNA vector.

[00564] In this study, the biodistribution of luciferase expression from a ceDNA vector was assessed by IVIS in CD-1® IGS mice after an initial intravenous administration of 1.0mg/kg (*i.e.*, a priming dose) at days 0 and 28 (Group A). A second administration of a ceDNA vector was administered via tail vein injection of 3mg/kg (Group B) or 10mg/kg (Group C) in 1.2 mL in the tail vein at day 84. In this study, five (5) CD-1® mice were used in each of Groups A, B and C. IVIS imaging of the mice for luciferase expression was performed prior to the additional dosing at days 49, 56, 63, and 70 as described above, as well as post-redose on day 84 and on days 91, 98, 105, 112, and 132. Luciferase expression was assessed and detected in all three Groups A, B and C until at least 110 days (the longest time period assessed).

[00565] The level of expression of luciferase was shown to be increased by a re-dose (*i.e.*, re-administration of the ceDNA composition) of the LNP-ceDNA-Luc, as determined by assessment of luciferase activity in the presence of luciferin. Luciferase transgene expression as an exemplary a transgene as measured by IVIS imaging for at least 110 days after 3 different administration protocols (Groups A, B and C). The mice that had not been given any additional redose (1mg/kg priming dose (*i.e.*, Group A) treatment had stable luciferase expression observed over the duration of the study. The mice in Group B that had been administered a re-dose of 3mg/kg of the ceDNA vector showed an approximately seven-fold increase in observed radiance relative to the mice in Group C. Surprisingly, the mice re-dosed with 10 mg/kg of the ceDNA vector had a 17-fold increase in observed luciferase radiance over the mice not receiving any redose (Group A).

[00566] Group A shows luciferase expression in CD-1[®] IGS mice after intravenous administration of 1mg/kg of a ceDNA vector into the tail vein at days 0 and 28. Group B and C show luciferase expression in CD-1[®] IGS mice administered 1mg/kg of a ceDNA vector at a first time point (day 0) and re-dosed with administration of a ceDNA vector at a second time point of 84 days. The second administration (*i.e.*, re-dose) of the ceDNA vector increased expression by at least 7-fold, even up to 17-fold.

[00567] A 3-fold increase in the dose (*i.e.*, the amount) of ceDNA vector in a re-dose administration in Group B (*i.e.*, 3mg/kg administered at re-dose) resulted in a 7-fold increase in expression of the luciferase. Also unexpectedly, a 10-fold increase in the amount of ceDNA vector in a re-dose administration (*i.e.*, 10mg/kg re-dose administered) in Group C resulted in a 17-fold increase in expression of the luciferase. Thus, the second administration (*i.e.*, re-dose) of the ceDNA increased expression by at least 7-fold, even up to 17-fold. This shows that the increase in transgene expression from the re-dose is greater than expected and dependent on the dose or amount of the ceDNA vector in the re-dose administration, and appears to be synergistic to the initial transgene expression from the initial priming administration at day 0. That is, the dose-dependent increase in transgene expression is not additive, rather, the expression level of the transgene is dose-dependent and greater than the sum of the amount of the ceDNA vector administered at each time point.

[00568] Both Groups B and C showed significant dose-dependent increase in expression of luciferase as compared to control mice (Group A) that were not re-dosed with a ceDNA vector at the second time point. Taken together, these data show that the expression of a transgene from ceDNA vector can be increased in a dose-dependent manner by re-dose (*i.e.*, re-administration) of the ceDNA vector at least a second time point.

[00569] Taken together, these data demonstrate that the expression level of a transgene, *e.g.*, inflammasome antagonist from ceDNA vectors can be maintained at a sustained level for at least 84 days and can be increased *in vivo* after a redose of the ceDNA vector administered at least at a second time point.

EXAMPLE 7: Sustained transgene expression *in vivo* of LNP-Formulated ceDNA vectors

[00570] The reproducibility of the results in Example 6 with a different lipid nanoparticle was assessed *in vivo* in mice. Mice were dosed on day 0 with either ceDNA vector comprising a luciferase transgene driven by a CAG promoter that was encapsulated in an LNP different from that used in Example 6 or with that same LNP comprising polyC but lacking ceDNA or a luciferase gene. Specifically, male CD-1[®] mice of approximately 4 weeks of age were treated with a single injection of 0.5 mg/kg LNP-TTX-luciferase or control LNP-polyC, administered intravenously via lateral tail vein on day 0. At day 14 animals were dosed systemically with luciferin at 150 mg/kg via intraperitoneal injection at 2.5 mL/kg. At approximately 15 minutes after luciferin administration each animal was imaged using an In Vivo Imaging System (“IVIS”).

As shown in **FIG. 7**, significant fluorescence in the liver was observed in all four ceDNA-treated mice, and very little other fluorescence was observed in the animals other than at the injection site, indicating that the LNP mediated liver-specific delivery of the ceDNA construct and that the delivered ceDNA vector was capable of controlled sustained expression of its transgene for at least two weeks after administration.

EXAMPLE 8: Sustained transgene expression in the liver *in vivo* from ceDNA vector administration

[00571] In a separate experiment, the localization of LNP-delivered ceDNA within the liver of treated animals was assessed. A ceDNA vector comprising a functional transgene of interest was encapsulated in the same LNP as used in Example 7 and administered to mice *in vivo* at a dose level of 0.5 mg/kg by intravenous injection. After 6 hours the mice were terminated and liver samples taken, formalin fixed and paraffin-embedded using standard protocols. RNAscope® *in situ* hybridization assays were performed to visualize the ceDNA vectors within the tissue using a probe specific for the ceDNA transgene and detecting using chromogenic reaction and hematoxylin staining (Advanced Cell Diagnostics). **FIG. 8** shows the results, which indicate that ceDNA is present in hepatocytes. One of skill will appreciate that luciferase can be replaced in ceDNA vector for any nucleic acid sequence selected for any transgene of interest.

EXAMPLE 9: Sustained Ocular transgene Expression of ceDNA *in vivo*

[00572] The sustainability of ceDNA vector transgene expression in tissues other than the liver was assessed to determine tolerability and expression of a ceDNA vector after ocular administration *in vivo*. While luciferase was used as an exemplary transgene in Example 9, one of ordinary skill can readily substitute the luciferase transgene with any transgene of interest.

[00573] On day 0, male Sprague Dawley rats of approximately 9 weeks of age were injected sub-retinally with 5 µL of either ceDNA vector comprising a luciferase transgene formulated with jetPEI® transfection reagent (Polyplus) or plasmid DNA encoding luciferase formulated with jetPEI®, both at a concentration of 0.25 µg/µL. Four rats were tested in each group. Animals were sedated and injected sub-retinally in the right eye with the test article using a 33-gauge needle. The left eye of each animal was untreated. Immediately after injection eyes were checked with optical coherence tomography or fundus imaging in order to confirm the presence of a subretinal bleb. Rats were treated with buprenorphine and topical antibiotic ointment according to standard procedures. At days 7, 14, 21, 28, and 35, the animals in both groups were dosed systemically with freshly made luciferin at 150 mg/kg via intraperitoneal injection at 2.5mL/kg at 5-15 minutes post luciferin administration, all animals were imaged using IVIS while under isoflurane anesthesia. Total Flux [p/s] and average Flux (p/s/sr/cm²) in a region of interest encompassing the eye were obtained over 5 minutes of exposure. The results were graphed as average radiance of each treatment group in the treated eye (“injected”) relative to the average radiance of each treatment group in the untreated eye (“uninjected”) (**FIG. 9B**). Significant fluorescence was readily detectable in the ceDNA vector-

treated eyes but much weaker in the plasmid-treated eyes (**FIG. 9A**). After 35 days, the plasmid-injected rats were terminated, while the study continued for the ceDNA-treated rats, with luciferin injection and IVIS imaging at days 42, 49, 56, 63, 70, and 99. The results demonstrate that ceDNA vector introduced in a single injection to rat eye mediated transgene expression *in vivo* and that that expression was sustained at a high level at least through 99 days after injection.

EXAMPLE 10: Sustained dosing and redosing of ceDNA vector in Rag2 mice.

[00574] In situations where one or more of the transgenes encoded in the gene expression cassette of the ceDNA vector is expressed in a host environment (*e.g.*, cell or subject) where the expressed protein is recognized as foreign, the possibility exists that the host will mount an adaptive immune response that may result in undesired depletion of the expression product, which could potentially be confused for lack of expression. In some cases this may occur with a reporter molecule that is heterologous to the normal host environment. Accordingly, ceDNA vector transgene expression was assessed *in vivo* in the Rag2 mouse model which lacks B and T cells and therefore does not mount an adaptive immune response to non-native murine proteins such as luciferase. Briefly, c57bl/6 and Rag2 knockout mice were dosed intravenously via tail vein injection with 0.5 mg/kg of LNP-encapsulated ceDNA vector expressing luciferase or a polyC control at day 0, and at day 21 certain mice were redosed with the same LNP-encapsulated ceDNA vector at the same dose level. All testing groups consisted of 4 mice each. IVIS imaging was performed after luciferin injection as described in Example 9 at weekly intervals.

[00575] Comparing the total flux observed from the IVIS analyses, the fluorescence observed in the wild-type mice (an indirect measure of the presence of expressed luciferase) dosed with LNP-ceDNA vector-Luc decreased gradually after day 21 whereas the Rag2 mice administered the same treatment displayed relatively constant sustained expression of luciferase over the 42 day experiment (**FIG. 9A**). The approximately 21-day time point of the observed decrease in the wild-type mice corresponds to the timeframe in which an adaptive immune response might expect to be produced. Re-administration of the LNP-ceDNA vector in the Rag2 mice resulted in a marked increase in expression which was sustained over the at least 21 days it was tracked in this study (**FIG. 9B**). The results suggest that adaptive immunity may play a role when a non-native protein is expressed from a ceDNA vector in a host, and that observed decreases in expression in the 20+ day timeframe from initial administration may signal a confounding adaptive immune response to the expressed molecule rather than (or in addition to) a decline in expression. Of note, this response is expected to be low when expressing native proteins in a host where it is anticipated that the host will properly recognize the expressed molecules as self and will not develop such an immune response.

EXAMPLE 11: Impact of liver-specific expression and CpG modulation on sustained expression

[00576] As described in Example 10, undesired host immune response may in some cases artificially dampen what would otherwise be sustained expression of one or more desired transgenes from an introduced ceDNA vector. Two approaches were taken to assess the impact of avoiding

and/or dampening potential host immune response on sustained expression from a ceDNA vector. First, since the ceDNA-Luc vector used in the preceding examples was under the control of a constitutive CAG promoter, a similar construct was made using a liver-specific promoter (hAAT) or a different constitutive promoter (hEF-1) to see whether avoiding prolonged exposure to myeloid cells or non-liver tissue reduced any observed immune effects. Second, certain of the ceDNA-luciferase constructs were engineered to be reduced in CpG content, a known trigger for host immune reaction. ceDNA-encoded luciferase gene expression upon administration of such engineered and promoter-switched ceDNA vectors to mice was measured.

[00577] Three different ceDNA vectors were used, each encoding luciferase as the transgene. The first ceDNA vector had a high number of unmethylated CpG (~350) and comprised the constitutive CAG promoter (“ceDNA CAG”); the second had a moderate number of unmethylated CpG (~60) and comprised the liver-specific hAAT promoter (“ceDNA hAAT low CpG”); and the third was a methylated form of the second, such that it contained no unmethylated CpG and also comprised the hAAT promoter (“ceDNA hAAT No CpG”). The ceDNA vectors were otherwise identical. The vectors were prepared as described above.

[00578] Four groups of four male CD-1® mice, approximately 4 weeks old, were treated with one of the ceDNA vectors encapsulated in an LNP or a polyC control. On day 0 each mouse was administered a single intravenous tail vein injection of 0.5 mg/kg ceDNA vector in a volume of 5 mL/kg. Body weights were recorded on days -1, -, 1, 2, 3, 7, and weekly thereafter until the mice were terminated. Whole blood and serum samples were taken on days 0, 1, and 35. In-life imaging was performed on days 7, 14, 21, 28, and 35, and weekly thereafter using an in vivo imaging system (IVIS). For the imaging, each mouse was injected with luciferin at 150 mg/kg via intraperitoneal injection at 2.5 mL/kg. After 15 minutes, each mouse was anaesthetized and imaged. The mice were terminated at day 93 and terminal tissues collected, including liver and spleen. Cytokine measurements were taken 6 hours after dosing on day 0.

[00579] While all of the ceDNA-treated mice displayed significant fluorescence at days 7 and 14, the fluorescence decreased rapidly in the ceDNA CAG mice after day 14 and more gradually decreased for the remainder of the study. In contrast, the total flux for the ceDNA hAAT low CpG and No CpG-treated mice remained at a steady high level (**FIG. 10**). This suggested that directing the ceDNA vector delivery specifically to the liver resulted in sustained, durable transgene expression from the vector over at least 77 days after a single injection. Constructs that were CpG minimized or completely absent of CpG content had similar durable sustained expression profiles, while the high CpG constitutive promoter construct exhibited a decline in expression over time, suggesting that host immune activation by the ceDNA vector introduction may play a role in any decreased expression observed from such vector in a subject. These results provide alternative methods of tailoring the duration of the response to the desired level by selecting a tissue-restricted promoter and/or altering

the CpG content of the ceDNA vector in the event that a host immune response is observed – a potentially transgene-specific response.

EXAMPLE 12: Preparation of lipid formulations

[00580] Lipid nanoparticles (LNP) were prepared at a total lipid to ceDNA weight ratio of approximately 10:1 to 30:1. Briefly, a condensing agent (*e.g.*, a cationic lipid, such DOTAP), a non-cationic-lipid (*e.g.*, distearoylphosphatidylcholine (DSPC)), a component to provide membrane integrity (such as a sterol, *e.g.*, cholesterol) and a conjugated lipid molecule (such as a PEG-lipid, *e.g.*, 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol, with an average PEG molecular weight of 2000 (“PEG-DMG”)), were solubilized in alcohol (*e.g.*, ethanol) at a molar ratio of 50:10:38.5:1.5 or 20:40:38:1.5. The therapeutic nucleic acid like ceDNA was diluted to a desired concentration in buffer solution. For example, the ceDNA were diluted to a concentration of 0.1 mg/ml to 0.25 mg/ml in a buffer solution comprising sodium acetate, sodium acetate and magnesium chloride, citrate, malic acid, or malic acid and sodium chloride. In one example, the ceDNA was diluted to 0.2 mg/mL in 10 to 50 mM citrate buffer, pH 4. The alcoholic lipid solution was mixed with ceDNA aqueous solution using, for example, syringe pumps or an impinging jet mixer, at a ratio of about 1:5 to 1:3 (vol/vol) with total flow rates above 10 ml/min. In one example, the alcoholic lipid solution was mixed with ceDNA aqueous at a ratio of about 1:3 (vol/vol) with a flow rate of 12 ml/min. The alcohol was removed and the buffer is replaced with PBS by dialysis. Alternatively, the buffer was replaced with PBS using centrifugal tubes. Alcohol removal and simultaneous buffer exchange were accomplished by, for example, dialysis or tangential flow filtration. The obtained lipid nanoparticles are filtered through a 0.2 µm pore sterile filter.

[00581] In one study lipid nanoparticles comprising exemplary ceDNAs were prepared using a lipid solution comprising MC3, DSPC, Cholesterol and DMG-PEG2000 (mol ratio 50:10:38.5:1.5) or DOTAP, DSPC, Cholesterol and DMG-PEG2000 (mol ratio 20:40:38.5:1.5). Aqueous solutions of ceDNA in buffered solutions were prepared. The lipid solution and the ceDNA solution were mixed using an in-house procedure on a NanoAssembler at a total flow rate of 12 ml/min at a lipid to ceDNA ratio of 1:3 (v/v). **Table 13** shows exemplary LNPs prepared in this study, where DSPC/MC3/Chol/PEG is a fusogenic LNP control and the others were expected to form non-fusogenic LNPs.

Table 13: Exemplary LNPs

MLNP#	Lipid mix*	Lipid Molar Ratio	Lipid Feed [mg/ml]	ceDNA	Scale (mg ceDNA)
2	DSPC/MC3/Chol/PEG	10/50/38.5/1.5	8	TTX-A	1.7
21	DSPC/DOTAP/Chol/PEG	40/20/38.5/1.5	8	TTX-A	0.2
23	DSPC/DOTAP/Chol/PEG	40/20/38.5/1.5	8	TTX-B	0.2
24	DSPC/DOTAP/Chol/PEG	40/20/38.5/1.5	8	TTX-B	1.4
48	DSPC/DOTAP/Chol/PEG	40/20/38.5/1.5	8	TTX-C	1.23
49	CHEMS/DOTAP/Chol/PEG	53/25/20/2	16	TTX-C	1.23

50	DSPC/SS-E-P4C2/Chol/PEG2	20/40/38.5/1.5	8	TTX-C	1.23
----	--------------------------	----------------	---	-------	------

*DSPC = distearoylphosphatidylcholine; MC3 = heptatriaconta-6,9,28,31-tetraen-19-yl-4-(dimethylamino)butanoate; Chol = Cholesterol; PEG = 1-(monomethoxy-polyethyleneglycol)-2,3-dimyristoylglycerol (PEG₂₀₀₀-DMG); CHEMS = cholesteryl hemisuccinate; and SS-E-P4C2 = COATSOME® SS-E-P4C2/SS-33/4PE-15.

EXAMPLE 13: Analysis of lipid particle formulations

[00582] Lipid nanoparticle size and zeta potential, and encapsulation of ceDNA into the lipid nanoparticles were determined. Particle size was determined by dynamic light scattering and zeta potential was measured by electrophoretic light scattering (Zetasizer Nano ZS, Malvern Instruments). Results are shown in **FIGS. 11 and 12**.

[00583] Encapsulation of ceDNA in lipid particles was determined by an Oligreen® assay. Oligreen® is an ultra-sensitive fluorescent nucleic acid stain for quantitating oligonucleotides and single-stranded DNA or RNA in solution (available from Invitrogen Corporation; Carlsbad, Calif.). As an alternative, PicoGreen® were used. Briefly, encapsulation was determined by performing a membrane-impermeable fluorescent dye exclusion assay, which uses a dye that has enhanced fluorescence when associated with nucleic acid. Encapsulation was determined by adding the dye to the lipid particle formulation, measuring the resulting fluorescence, and comparing it to the fluorescence observed upon addition of a small amount of nonionic detergent. Detergent-mediated disruption of the lipid bilayer releases the encapsulated ceDNA, allowing it to interact with the membrane-impermeable dye. Encapsulation of ceDNA can be calculated as $E = (I_0 - I)/I_0$, where/and I_0 refers to the fluorescence intensities before and after the addition of detergent.

[00584] Here, encapsulation efficiency was calculated by determining unencapsulated ceDNA content by measuring the fluorescence upon the addition of PicoGreen (Thermo Scientific) to the LNP slurry (C_{free}) and comparing this value to the total ceDNA content that is obtained upon lysis of the LNPs by 1% Triton X-100 (C_{total}), where % encapsulation = $(C_{\text{total}} - C_{\text{free}})/C_{\text{total}} \times 100$. Results are shown in **FIG. 13**.

[00585] Next, release of ceDNA from LNPs was determined. Endosome mimicking anionic liposome was prepared by mixing DOPS:DOPC:DOPE (mol ratio 1:1:2) in chloroform, followed by solvent evaporation at vacuum. The dried lipid film was resuspended in DPBS with brief sonication, followed by filtration through 0.45 μm syringe filter to form anionic liposome.

[00586] Anionic liposome was added to LNP solution at desired anionic/cationic lipid mole ratio in DPBS at either pH 7.4 or 6.0. The resulting combination was then incubated at 37°C for another 15 min. Free ceDNA at pH 7.4 or pH 6.0 was calculated by determining unencapsulated ceDNA content by measuring the fluorescence upon the addition of PicoGreen (Thermo Scientific®) to the LNP slurry (C_{free}) and comparing this value to the total ceDNA content that is obtained upon lysis of the LNPs by 1% Triton X-100 (C_{total}), where % free = $C_{\text{free}}/C_{\text{total}} \times 100$. The % ceDNA released after incubation with anionic liposome was calculated based on the equation below:

% ceDNA released = % free ceDNA_{mixed with anionic liposome} - % free ceDNA_{mixed with DPBS} Results are shown in FIG. 14.

[00587] *In vivo* relative activity was determined by measuring luciferase expression in the liver 4 hours following administration via tail vein injection. The activity was compared at a dose of 0.3 and 1.0 mg ceDNA/kg and expressed as ng luciferase/g liver measured 4 hours after administration.

[00588] The pKa of formulated cationic lipids can be correlated with the effectiveness of the LNPs for delivery of nucleic acids (see Jayaraman *et al.*, Angewandte Chemie, International Edition (2012), 51(34), 8529-8533; Semple *et al.*, Nature Biotechnology 28, 172-176 (2010), both of which are incorporated by reference in their entirety). The preferred range of pKa is ~5 to ~7. The pKa of each cationic lipid was determined in lipid nanoparticles using an assay based on fluorescence of 2-(p-toluidino)-6-naphthalene sulfonic acid (TNS). Lipid nanoparticles comprising of cationic lipid/DSPC/cholesterol/PEG-lipid (50/10/38.5/1.5 mol %; or 50/10/37/3 mol %) in PBS at a concentration of 0.4 mM total lipid were prepared using the in-line process as described herein and elsewhere. TNS was prepared as a 100 μ M stock solution in distilled water. Vesicles were diluted to 24 μ M lipid in 2 mL of buffered solutions containing, 10 mM HEPES, 10 mM MES, 10 mM ammonium acetate, 130 mM NaCl, where the pH ranges from 2.5 to 11. An aliquot of the TNS solution was added to give a final concentration of 1 μ M, and following vortex mixing, fluorescence intensity was measured at room temperature in a SLM Aminco Series 2 Luminescence Spectrophotometer using excitation and emission wavelengths of 321 nm and 445 nm. A sigmoidal best fit analysis was applied to the fluorescence data and the pKa was measured as the pH giving rise to half-maximal fluorescence intensity.

[00589] Binding of the lipid nanoparticles to ApoE was determined as follows. LNP (10 μ g/mL) is incubated at 37°C for 20 min with equal volume of recombinant ApoE3 (500 μ g/mL) in 1x DPBS. After incubation, LNP samples were diluted 10-fold using 1x DPBS and analyzed by heparin sepharose chromatography on AKTA pure 150 (GE Healthcare) according to the conditions below:

HiTrap chromatographic conditions

Column	HiTrap Heparin Sepharose HP 1mL
Equilibration buffer	1x DPBS
Wash buffer	1x DPBS
Elution buffer	1 M NaCl in 10 mM sodium phosphate buffer, pH 7.0
Flow rate	1 mL/min
Injection volume	500 μ L
Detection	260 nm

	CV	A (%)	B (%)
Equilibration	1	100	0
Column wash	4	100	0
Elution (linear)	10	0	100
Equilibration	3	100	0

[00590] Expression of ceDNA encapsulated into the lipid nanoparticles was assayed as follows. HEK293 cells were maintained at 37°C with 5% CO₂ in DMEM + GlutaMAX™ culture medium (Thermo Scientific®) supplemented with 10% Fetal Bovine Serum and 1% Penicillin-Streptomycin. Cells were plated in 96-well plates at a density of 30,000 cells/well the day before transfection. Lipofectamine™ 3000 (Thermo Scientific) transfection reagent was used for transfecting 100 ng/well of control ceDNA according to the manufacturer's protocol. The control ceDNA was diluted in Opti-MEM™ (Thermo Scientific) and P3000™ Reagent was added. Subsequently, Lipofectamine™ 3000 was diluted to a final concentration of 3% in Opti-MEM™. Diluted Lipofectamine™ 3000 was added to diluted ceDNA at a 1:1 ratio and incubated for 15 minutes at room temperature. Desired amount of ceDNA-lipid complex or LNP was then directly added to each well containing cells. The cells were incubated at 37°C with 5% CO₂ for 72 hours. The expression levels of secreted protein encoded by ceDNA (e.g., Factor IX) in HEK293-conditioned media were determined by ELISA according to manufacturer's instructions.

EXAMPLE 14: Innate immunity by dose in mice

[00591] Exemplary lipid nanoparticles were tested *in vivo* to determine whether endosomally-restricted ceDNA is less inflammatory. Mice were dosed with LNP formulations as shown in **Table 14** on Day 0.

Table 14: Innate immunity study design

Group	Strain/#	Formulation ¹	CeDNA	Endpoints
1	C57Bl/8	Lipid 1 and buffer	N/A	Dosing on Day 0 Cytokines at 6 hours post dose Harvested liver at 24 hours post-dose for later qPCR analysis
2		Lipid 1 and polyC RNA	N/A	
3		Lipid 1	TTX-D (0.8 mg/kg)	
4			TTX-D (0.4 mg/kg)	
5		Inactive LNP - either DSPC liposomes or the DOTAP/CHEMS	none	
6			TTX-D (0.8 mg/kg)	
7			TTX-D (0.4 mg/kg)	

¹Lipid 1 is an ionizable/cationic lipid used in the art for nucleic acid delivery to cells. Inactive LNPs are exemplary non-fusogenic lipid nanoparticles of the invention, e.g., DSPC/DOTAP/Chol/PEG; or CHEMS/DOTAP/Chol/PEG, which may require one or more endosomolytic agents.

[00592] Cytokines levels were measured at 6 hours post dose using the ProcartaPlex Multiplex Immunoassay (Invitrogen), which is a relative quantitative multiplex bead-based immunoassay for measuring levels of various cytokines and chemokines in study samples, using the Luminex technology platform. A pre-mixed custom mouse cytokine 8-plex kit, with magnetic beads, was used

to assay the following cytokines: IFN- α , IFN- γ , IL-6, IP-10, IL-18, IL-1 β , MCP-1, and TNF-alpha. A pre-mixed custom mouse cytokine 1-plex kit was used for assaying INF-beta.

[00593] Livers were harvested 24 hours post-dose for qPCR analysis. Body weight change was determined at 24 hours post dose. Results are shown in FIGS. 15A-15C and FIG. 16.

REFERENCES

[00594] All publications and references, including but not limited to patents and patent applications, cited in this specification and Examples herein are incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in the manner described above for publications and references.

CLAIMS

What is claimed is:

1. A composition comprising:
 - a. a capsid free, non-viral vector encapsulated in a lipid nanoparticle (LNP), wherein the LNP lacks fusogenic activity; and
 - b. an endosomolytic agent.
2. The composition of claim 1, wherein the capsid free, non-viral vector when digested with a restriction enzyme having a single recognition site on the DNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA when analyzed on a non-denaturing gel.
3. The composition of claim 1, wherein the endosomolytic agent targets a target cell.
4. The composition of claim 1, wherein the endosomolytic agent is:
 - (a) a membrane-destabilizing polymer;
 - (b) a copolymer, a peptide, a membrane-destabilizing toxin, or a viral fusogenic peptide;
 - (c) a pH-sensitive polymer;
 - (d) polyanionic peptide, polycationic peptide, amphipathic peptide, hydrophobic peptide or a peptidomimetic; or
 - (e) a peptide selected from the group consisting of:
AALEALAEALAEALAEALAEAAAAGGC (SEQ ID NO: 530);
AALAEALAEALAEALAEALAEALAAAAGGC (SEQ ID NO: 531);
ALEALAEALAEALAEA (SEQ ID NO: 532);
GLFEAIEGFIENGWEGMIWDYG (SEQ ID NO: 533);
GLFGAIAAGFIENGWEGMIDGWYG (SEQ ID NO: 534);
GLFEAIEGFIENGWEGMIDGWYGC (SEQ ID NO: 535);
GLFEAIEGFIENGWEGMIDGWYGC (SEQ ID NO: 536);
GLFEAIEGFIENGWEGMIDGGC (SEQ ID NO: 537);
GLFEAIEGFIENGWEGMIDGGC (SEQ ID NO: 538);
CGLFGEIEELIEEGLENLIDWGNG (SEQ ID NO: 539);
GLFGALAEALAEALAEHLAEALAEALAEALAAGGSC (SEQ ID NO: 540);
GLFEAIEGFIENGWEGLAELAEALAEALAEALAAGGSC (SEQ ID NO: 541);
GLFEAIEGFIENGWEGnIDGK (n = norleucine) (SEQ ID NO: 542);
GLFEAIEGFIENGWEGnIDG (n = norleucine) (SEQ ID NO: 543);
GLFEALLESLESLWELLLEA (SEQ ID NO: 544);
GLFKALLKLLKSLWKLLLKA (SEQ ID NO: 545);
GLFRALLRLLRSLWRLLLR (SEQ ID NO: 546);
WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 547);

GLFFEIAIEFIEGGWEGLIEGC (SEQ ID NO: 548);
GIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 549);
H5WYG (SEQ ID NO: 550);
CHK₆HC (SEQ ID NO: 551);
RQIKIWFQNRRMKWKK (SEQ ID NO: 552);
GRKKRRQRRRPPQC (SEQ ID NO: 553);
GALFLGWLGAAGSTM (SEQ ID NO: 554);
GAWSQPKKKRKV (SEQ ID NO: 555);
LLIILRRRIRKQAHAAHSK (SEQ ID NO: 556);
GWTLSAGYLLKINLKALAALAKKIL (SEQ ID NO: 557);
KLALKLALKALKAALKLA (SEQ ID NO: 558);
RRRRRRRRR (SEQ ID NO: 559);
KFFKFFKFFK (SEQ ID NO: 560);
LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES (SEQ ID NO: 561);
SWLSKTAKKLENSAKKRISGIAIAIQGGPR (SEQ ID NO: 562);
ACYCRIPACIAGERRYGTCTIYQGRLWAFCC (SEQ ID NO: 563);
DHYNVCVSSGGQCLYSACPIFTKIQTCTYRGKAKCCK ((SEQ ID NO: 564);
RKCRIVVIRVCRRRRRPRPPYLPRPRPPFPFPRLLPRIIPGFPFRFPFRPGKR (SEQ ID NO: 565);
ILPWKWPWWPWRR (566);
WEAALAEALAEALAEHLAEALAEALEALAA (SEQ ID NO: 567);
CAEALAEALAEALAEALA (SEQ ID NO: 568);
GIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 569);
CGIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 570);
FIIDIIAFLLMGGFIVYVKNL (SEQ ID NO: 571);
CAAFIIDHAFLLMGGFIVYVKNL (SEQ ID NO: 572);
CARGWEVLKYWWNLLQY (SEQ ID NO: 573);
MVKSKIGSWILVLFVAMWSDVGLCKKRPKP (SEQ ID NO: 574);
KLALKLALKALKAALKLA (SEQ ID NO: 575);
YARAAARQARA (SEQ ID NO: 576);
GDCLPHLKLCKENKDCCSKKCKRRGTNIE (SEQ ID NO: 577);
RRLSYSRRRF (SEQ ID NO: 578);
RGGRLSYSRRRFSTSTGR (SEQ ID NO: 579);
IAWVKAFIRKLRKGPLG (SEQ ID NO: 580);
YTAIAWVKAFIRKLRK (SEQ ID NO: 581);
GLWRALWRLRLRWLLWRA (SEQ ID NO: 582);
KWFETWFTWPKKRK (SEQ ID NO: 583);

KETWWETWWTEWSQPKKRKV (SEQ ID NO: 584);
 AGYLLGK(eNH_a)INLKALAALAKKIL (SEQ ID NO: 585);
 AGYLLGKINLKALAALAKKIL (SEQ ID NO: 586);
 RQIKIVVFQNRMRKWKK (SEQ ID NO: 587);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 588);
 LLIILRRRIRKQAHASHK (SEQ ID NO: 589);
 YTIVVMPENPRPGTPCDIFTNSRGKRASNG (SEQ ID NO: 590);
 AAVALLPAVLLALLAK (SEQ ID NO: 591);
 GWTLNSAGYLLGKINLKALAALAKKIL (SEQ ID NO: 592);
 GRKKRRQRRPPQ (SEQ ID NO: 593);
 KMTRAQRRAAARRNRRTAR (SEQ ID NO: 594);
 KKRKAPKKRKFA-KFHTFPQTAIGVGAP (SEQ ID NOS: 595 and 600, respectively);
 MVTVLFRRLRIRRASGPPRV (SEQ ID NO: 596);
 LIRLWSHLIHIVVFQNRRLKWKKK (SEQ ID NO: 597);
 GALFLGFLGAAGSTMGAWSQPKKRKV (SEQ ID NO: 598); and
 GALFLAFLAAALSLMGLWSQPKKRKV (SEQ ID NO: 599).

5. The composition of claim 1 wherein the endosomolytic agent is in form of a nanoparticle.
6. The composition of claim 5, wherein the nanoparticle further comprises a cationic lipid, a non-cationic lipid, a sterol, a conjugated lipid, or any combination thereof.
7. The composition of claim 1, wherein the lipid nanoparticle comprises the endosomolytic agent.
8. The composition of claim 1, wherein the endosomolytic agent is preferentially or specifically taken up by the target cell relative to a non-target cell.
9. The composition of claim 1, wherein the endosomolytic agent comprises a targeting ligand.
10. The composition of claim 9, wherein the ligand binds to a cell surface molecule selected from the group consisting of a transferrin receptor type 1, a transferrin receptor type 2, the EGF receptor, HER2/Neu, a VEGF receptor, a PDGF receptor, an integrin, an NGF receptor, CD2, CD3, CD4, CD8, CD19, CD20, CD22, CD33, CD43, CD38, CD56, CD69, the asialoglycoprotein receptor (ASGPR), GalNAc receptor, prostate-specific membrane antigen (PSMA), a folate receptor, and a sigma receptor.
11. The composition of claim 9, wherein the targeting ligand is a monovalent or multivalent D-galactose or N-acetyl-D-galactose.
12. The composition of claim 1, further comprising an additional compound.
13. The composition of any of claims 1-12, wherein the capsid free, non-viral vector is a closed-ended DNA (ceDNA) vector comprising:

- at least one nucleotide sequence between flanking inverted terminal repeats (ITRs), wherein the at least one nucleotide sequence encodes at least one transgene or therapeutic protein of interest.
14. The composition of claim 13, wherein the ceDNA vector comprise a promoter selected from any of those in Table 7 operatively linked to the least one nucleotide sequence that encodes at least one transgene or therapeutic protein;
an enhancer selected from any of those in Table 8;
a 5' UTR and/or intron sequence selected from any of those in Table 9A;
a 3' UTR selected from any of those in Table 9B; and/or
at least one poly A sequence selected from any of those in Table 10.
 15. The composition of any one of claims 13 to 14, wherein the ceDNA vector comprises at least one promoter linked to the at least one nucleotide sequence.
 16. The composition of any one of claims 13 and 14, wherein at least one flanking ITR comprises a functional terminal resolution site and a Rep binding site;
one or both of the flanking ITRs are from a virus selected from the group consisting of : a parvovirus, a dependovirus, and an adeno-associated virus (AAV); one or both of the flanking ITRs are from a virus selected from the group consisting of: a parvovirus, a dependovirus, and an adeno-associated virus (AAV);
the flanking ITRs are symmetric or asymmetric;
the flanking ITRs are symmetrical or substantially symmetrical;
one of the ITRs is a wild type ITR, or wherein both of the flanking ITRs are wild-type ITRs;
the flanking ITRs are from different viral serotypes;
the flanking ITRs are from a pair of viral serotypes shown in Table 2;
one or both of the flanking ITRs comprises a sequence selected from the sequences in Table 3;
at least one of the flanking ITRs is altered from a wild-type AAV ITR sequence by a deletion, an addition, and/or a substitution that affects the overall three-dimensional conformation of the ITR;
one or both of the flanking ITRs are synthetic; and/or
one of the flanking ITRs is not a wild-type ITR, or wherein both of the flanking ITRs are not wild-type ITRs.
 17. The composition of any one of claims 13, 14 and 16, wherein one or both of the flanking ITRs is modified by a deletion, an insertion, and/or a substitution in at least one of the flanking ITR regions selected from the group consisting of: an A, A', B, B', C, C', D, and D' regions; one or both of the flanking ITRs is modified by a deletion, an insertion, and/or a substitution in at least one of the flanking ITR regions selected from the group consisting of: an A, A', B, B', C, C', D, and D' regions;

wherein the deletion, insertion, and/or substitution results in the deletion of all or part of a stem-loop structure formed by the A, A', B, B' C, or C' regions;

wherein one or both of the flanking ITRs are modified by a deletion, an insertion, and/or a substitution that results in the deletion of all or part of a stem-loop structure formed by the B and B' regions;

one or both of the flanking ITRs are modified by a deletion, an insertion, and/or a substitution that results in the deletion of all or part of a stem-loop structure formed by the C and C' regions;

one or both of the flanking ITRs are modified by a deletion, an insertion, and/or a substitution that results in the deletion of part of a stem-loop structure formed by the B and B' regions and/or part of a stem-loop structure formed by the C and C' regions;

one or both of the flanking ITRs comprise a single stem-loop structure in the region that, in a wild-type ITR, would comprise a first stem-loop structure formed by the B and B' regions and a second stem-loop structure formed by the C and C' regions;

one or both of the flanking ITRs comprise a single stem and two loops in the region that, in a wild-type ITR, would comprise a first stem-loop structure formed by the B and B' regions and a second stem-loop structure formed by the C and C' regions;

one or both of the flanking ITRs comprise a single stem and a single loop in the region that, in a wild-type ITR, would comprise a first stem-loop structure formed by the B and B' regions and a second stem-loop structure formed by the C and C' regions; or

both flanking ITRs are altered in a manner that results in an overall three-dimensional symmetry when the flanking ITRs are inverted relative to each other.

18. The composition of any one of claims 13, 14, 16, 17, wherein the ceDNA vector further comprises at least one regulatory switch.
19. Use of a capsid free, non-viral vector in the manufacture of a medicament for delivering the capsid free, non-viral vector to the cytosol of a target cell within a subject,
 - wherein the capsid free, non-viral vector is encapsulated in a lipid nanoparticle (LNP),
 - wherein the LNP lacks fusogenic activity; and
 - wherein the capsid free, non-viral vector and an endosomolytic agent are to be co-administered.
20. The use of claim 19, wherein the capsid free, non-viral vector when digested with a restriction enzyme having a single recognition site on the DNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA when analyzed on a non-denaturing gel.
21. The use of claim 19, wherein the endosomolytic agent targets the target cell.
22. The use of claim 19, wherein the capsid free, non-viral vector is translocated to nucleus of the cell after administration.

23. The use of claim 19, wherein the target cell is a cell that lacks or does not express a functional innate DNA-sensing pathway, has reduced innate DNA-sensing pathway activity or lacks or does not express functional cGAS and/or STING, or which has reduced cGAS and/or STING activity.
24. The use of claim 19, wherein the target cell is a hepatocyte.
25. The use of claim 19, wherein the endosomolytic agent comprises:
 - (a) a membrane-destabilizing polymer,
 - (b) a copolymer, a peptide, a membrane-destabilizing toxin, or a viral fusogenic peptide;
 - (c) a pH-sensitive polymer;
 - (d) a polyanionic peptide, a polycationic peptide, an amphipathic peptide, a hydrophobic peptide or a peptidomimetic; or
 - (e) a peptide selected from the group consisting of
 AALEALAEALAEALAEALAEAAAAGGC (SEQ ID NO: 530);
 AALAEALAEALAEALAEALAEALAAAAGGC (SEQ ID NO: 531);
 ALEALAEALAEALAEA (SEQ ID NO: 532);
 GLFEAIEGFIENGWEGMIWDYG (SEQ ID NO: 533);
 GLFGAIAAGFIENGWEGMIDGWYG (SEQ ID NO: 534);
 GLFEAIEGFIENGWEGMIDGWYGC (SEQ ID NO: 535);
 GLFEAIEGFIENGWEGMIDGWYGC (SEQ ID NO: 536);
 GLFEAIEGFIENGWEGMIDGGC (SEQ ID NO: 537);
 GLFEAIEGFIENGWEGMIDGGC (SEQ ID NO: 538);
 CGLFGEIEELIEEGLENLIDWGNG (SEQ ID NO: 539);
 GLFGALAEALAEALAEHLAEALAEALAEALAAGGSC (SEQ ID NO: 540);
 GLFEAIEGFIENGWEGLAELAEALAEALAEALAAGGSC (SEQ ID NO: 541);
 GLFEAIEGFIENGWEGnIDGK (n = norleucine) (SEQ ID NO: 542);
 GLFEAIEGFIENGWEGnIDG (n = norleucine) (SEQ ID NO: 543);
 GLFEALLELLESLWELLLEA (SEQ ID NO: 544);
 GLFKALLKLLKSLWKLLLLKA (SEQ ID NO: 545);
 GLFRALLRLLRSLWRLLLLRA (SEQ ID NO: 546);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 547);
 GLFFEIAIEFIEGGWEGGLIEGC (SEQ ID NO: 548);
 GIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 549);
 H5WYG (SEQ ID NO: 550);
 CHK₆HC (SEQ ID NO: 551);
 RQIKIWFQNRRMKWKK (SEQ ID NO: 552);
 GRKKRRQRRRPPQC (SEQ ID NO: 553);
 GALFLGWLGAAGSTM (SEQ ID NO: 554);

GAWSQPKKKRKV (SEQ ID NO: 555);
 LLIILRRRIRKQAHAAHSK (SEQ ID NO: 556);
 GWTLNSAGYLLKINLKALAALAKKIL (SEQ ID NO: 557);
 KLALKLALKALKAAALKLA (SEQ ID NO: 558);
 RRRRRRRRRR (SEQ ID NO: 559);
 KFFKFFKFFK (SEQ ID NO: 560);
 LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES (SEQ ID NO: 561);
 SWLSKTAKKLENSAKKRISGIAIAIQGGPR (SEQ ID NO: 562);
 ACYCRIPACIAGERRYGTCTIYQGRLWAFCC (SEQ ID NO: 563);
 DHYNCVSSGGQCLYSACPIFTKIQTCTYRGKAKCCK ((SEQ ID NO: 564);
 RKCRIVVIRVCRRRRRPRPPYLPRPRPPFFPPRLPPRIIPGFPFRFPFRPGKR (SEQ ID
 NO: 565);
 ILPWKWPWWPWRR (566);
 WEAALAEALAEALAEHLAEALAEALAEALAA (SEQ ID NO: 567);
 CAEALAEALAEALAEALA (SEQ ID NO: 568);
 GIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 569);
 CGIGAVLKVLTTGLPALISWIKRKRQQ (SEQ ID NO: 570);
 FIIDIIFLLMGGFIVYVKNL (SEQ ID NO: 571);
 CAAFIIDHAFLLMGGFIVYVKNL (SEQ ID NO: 572);
 CARGWEVLKYWWNLLQY (SEQ ID NO: 573);
 MVKSKIGSWILVLFVAMWSDVGLCKKRPKP (SEQ ID NO: 574);
 KLALKLALKALKAAALKLA (SEQ ID NO: 575);
 YARAAARQARA (SEQ ID NO: 576);
 GDCLPHLKLCKENKDCCSKKCKRRGTNIE (SEQ ID NO: 577);
 RRLSYSRRRF (SEQ ID NO: 578);
 RGGRLSYSRRRFSTSTGR (SEQ ID NO: 579);
 IAWVKAFIRKLRKGPLG (SEQ ID NO: 580);
 YTAIAWVKAFIRKLRK (SEQ ID NO: 581);
 GLWRALWRLRLWRLWRA (SEQ ID NO: 582);
 KWFETWFTWEPKKRK (SEQ ID NO: 583);
 KETWWETWTEWSQPKKKRKV (SEQ ID NO: 584);
 AGYLLGK(eNH_a)INLKALAALAKKIL (SEQ ID NO: 585);
 AGYLLGKINLKALAALAKKIL (SEQ ID NO: 586);
 RQIKIVVFQNRMRMKWKK (SEQ ID NO: 587);
 WEAKLAKALAKALAKHLAKALAKALKACEA (SEQ ID NO: 588);
 LLIILRRRIRKQAHAAHSK (SEQ ID NO: 589);
 YTIVVMPENPRPGTPCDIFTNSRGKRASNG (SEQ ID NO: 590);

AAVALLPAVLLALLAK (SEQ ID NO: 591);
 GWTLNSAGYLLGKINLKALAALAKKIL (SEQ ID NO: 592);
 GRKKRRQRRPPQ (SEQ ID NO: 593);
 KMTRAQRRAAARRNRRWTAR (SEQ ID NO: 594);
 KKRKAPKKRKFA-KFHTFPQTAIGVGAP (SEQ ID NOS: 595 and 600, respectively);
 MVTVLFRRLRIRRASGPPRVRV (SEQ ID NO: 596);
 LIRLWSHLIHIVVFQNRRLKWKKK (SEQ ID NO: 597);
 GALFLGFLGAAGSTMGAWSQPKKKRKV (SEQ ID NO: 598); and
 GALFLAFLAAALSLMGLWSQPKKKRKV (SEQ ID NO: 599).

26. The use of claim 19, wherein the endosomolytic agent is in the form of a nanoparticle; optionally wherein the nanoparticle further comprises a cationic lipid, a non-cationic lipid, a sterol, a conjugated lipid, or any combination thereof.
27. The use of claim 19, wherein the lipid nanoparticle and the endosomolytic agent are formulated into separate compositions for administration to the subject.
28. The use of claim 27, wherein the separate compositions are to be simultaneously administered are to be sequentially or subsequently administered.
29. The use of claim 19, wherein the lipid nanoparticle comprises the endosomolytic agent.
30. The use of claim 19, wherein at least one of the endosomolytic agent and the lipid nanoparticle includes a first targeting ligand that binds to a molecule on the surface of the target cell.
31. The use of claim 30, wherein one of the endosomolytic agent and the lipid nanoparticle includes the first targeting ligand, and the other of the endosomolytic agent and the lipid nanoparticle includes a second targeting ligand.
32. The use of claim 30, wherein the molecule on surface of the target cell is selected from the group consisting of a transferrin receptor type 1, a transferrin receptor type 2, the EGF receptor, HER2/Neu, a VEGF receptor, a PDGF receptor, an integrin, an NGF receptor, CD2, CD3, CD4, CD8, CD19, CD20, CD22, CD33, CD43, CD38, CD56, CD69, the asialoglycoprotein receptor (ASGPR), GalNAc receptor, prostate-specific membrane antigen (PSMA), a folate receptor, and a sigma receptor.
33. The use of claim 31, wherein the targeting ligand is a monovalent or multivalent D-galactose or N-acetyl-D-galactose.
34. The use of claim 19, further comprising that an additional compound is to be administered to the subject.
35. The use of claim 34, wherein said additional compound is a therapeutic agent, an immune modulating agent, or a second capsid free, non-viral vector, wherein the first and second capsid free, non-viral vectors are different.

36. Use of a capsid free, non-viral vector in the manufacture of a medicament for delivering the capsid free, non-viral vector to the cytosol of a target cell within a subject,
wherein the capsid free, non-viral vector is encapsulated in a lipid nanoparticle (LNP),
wherein the LNP lacks fusogenic activity;
wherein the capsid free, non-viral vector and an endosomolytic agent are to be co-administered, and
wherein at least one of the lipid nanoparticle and the endosomolytic agent includes a first targeting ligand that binds to a molecule on the surface of the target cell.
37. The use of claim 36, wherein the capsid free, non-viral vector when digested with a restriction enzyme having a single recognition site on the DNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA when analyzed on a non-denaturing gel.
38. Use of the composition of any one of claims 13, 14, 16, 17 and 18 in the manufacture of a medicament for expressing a desired transgene or therapeutic protein in a cell *in vivo* comprising contacting the cell with the composition of any one of claims 13, 14, 16, 17 and 18.
39. A cell containing a composition of any of claims 13, 14, 16, 17 and 18.
40. The use of any of claims 19-37, wherein the capsid free, non-viral vector is a close-ended DNA (ceDNA) vector comprising:
at least one nucleotide sequence between flanking inverted terminal repeats (ITRs), wherein
at least one nucleotide sequence encodes at least one transgene or therapeutic protein of interest.
41. An *in vitro* method of expressing a desired transgene or therapeutic protein in a cell, comprising contacting the cell with the composition of any one of claims 13, 14, 16, 17 and 18.