(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 August 2019 (29.08.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/165050 A1

(51) International Patent Classification:
C12N 15/86 (2006.01) *C12N 15/87* (2006.01)
C12N 15/864 (2006.01)

(21) International Application Number:
PCT/US2019/018927

(22) International Filing Date:
21 February 2019 (21.02.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

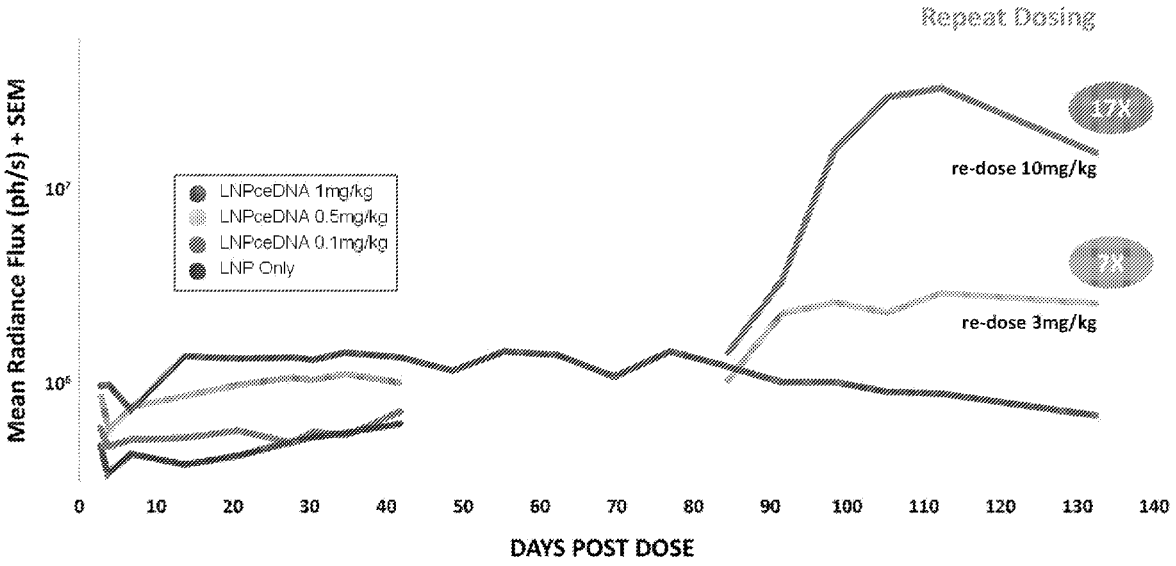
62/633,757	22 February 2018 (22.02.2018)	US
62/633,882	22 February 2018 (22.02.2018)	US
62/633,795	22 February 2018 (22.02.2018)	US
62/746,762	17 October 2018 (17.10.2018)	US

(71) Applicant: **GENERATION BIO CO.** [US/US]; 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US).

(72) Inventors: **KERR, Douglas A.**; c/o Generation Bio Co., 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US). **STANTON, Matthew G.**; c/o Generation Bio Co., 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US). **CHIOCCO, Matt**; c/o Generation Bio Co., 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US). **ANGELINO, Mark D.**; c/o Generation Bio Co., 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US). **KOTIN, Robert M.**; c/o Generation Bio Co., 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US). **SAMAYOA, Phillip**; c/o Generation Bio Co., 215 First Street, Suite 150, Cambridge, Massachusetts 02142 (US).

(54) Title: CONTROLLED EXPRESSION OF TRANSGENES USING CLOSE-ENDED DNA (CEDNA) VECTORS

FIG. 6



(57) Abstract: Provided herein are methods and constructs comprising close-ended DNA (ceDNA vectors) for maintaining or sustaining a level of transgene expression at a predetermined level or range for a predefined time, or increasing the level of transgene expression in a cell or a subject, where the transgene expression level can be modulated (e.g., increased) with one or more subsequent administrations (e.g., a re-dose or a booster administration) after an initial priming administration. Provided are methods for personalizing gene therapy throughout an individuals' lifespan to express a transgene at a level that meets an individual's needs, by modulating expression levels of a transgene expressed by ceDNA vector incrementally, or in a step-by-step manner, with one or more administrations after an initial priming administration (e.g., at time 0), thereby enabling titration of the level of expression of the transgene to a desired predetermined expression level or to a desired expression level range.

WO 2019/165050 A1

[Continued on next page]

(74) **Agent: BENN, Susanna C.** et al.; Nixon Peabody LLP, Exchange Place, 53 State Street, Boston, Massachusetts 02109 (US).

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

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

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

— *with sequence listing part of description (Rule 5.2(a))*

CONTROLLED EXPRESSION OF TRANSGENES USING CLOSE-ENDED DNA (CEDNA) VECTORS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application Nos. 62/633,882 filed February 22, 2018; 62/633,757 filed February 22, 2018; 62/633,795 filed February 22, 2018; and 62/746,762, filed October 17, 2018, the contents of each are incorporated herein by reference in their entireties.

[0002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on February 21, 2019, is named 080170-091190-WOPT_SL.txt and is 116,872 bytes in size.

TECHNICAL FIELD

[0003] The present invention relates to the field of gene therapy, including capsid-free vectors for controlled expression of a transgene or isolated polynucleotides in a subject or cell. The technology described herein relates to methods of controlled expression of a transgene *in vivo* from a capsid-free DNA vectors with closed ends (ceDNA) vector where the expression level can be sustained at desired level for a predetermined time or increased with one or more subsequent administrations (e.g., a booster administration, or re-dose).

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, by altering or silencing a defective gene, or delivering a therapeutic antibody, e.g., 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. 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 (rAVV) 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-pathologic 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.

[0009] Additionally, traditional gene therapy vectors for delivery of transgenes (e.g. adeno-associated virus (AAV), adenovirus, lentivirus vectors, etc.) are typically limited to a single administration of the vector to patients, due in part to the patient's immune response to viral proteins. Additionally, for sustained long term expression of the transgene, it is typically is required to administer a high titer on the initial administration, which can lead to deleterious side effects. As such, traditional viral vectors for gene therapy lack utility due to lack of sustained, long term transgene expression. In addition, the range of transgene genetic material suitable for delivery in such viral vectors is limited by the viral packaging capacity of the viral capsid proteins (e.g. about 4.5kb for AAV), thereby excluding delivery of larger transgenes for therapy. With respect to conventional adeno-associated virus (AAV) vectors for gene therapy, their use 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.

[0010] Additionally, there have been a number of reports raising concern about using too high a dose of a viral vector. Most viral vectors also suffer the immunogenicity concerns raised about AAV.

[0011] Accordingly, there is need in the field for a technology that allows multiple doses of a vector for gene therapy. In addition, there is a need in the art for methods of controlling gene expression from a gene therapy vector with minimal off-target effects, and there remains an important unmet need for controllable recombinant DNA vectors with improved production and/or expression properties. Further, as will be appreciated by a skilled physician, the ability to titrate the expression/dose of a transgene is desired to customize a gene therapy treatment based on a subject's particular set of symptoms and/or severity of disease and further to minimize side effects or toxicity.

BRIEF DESCRIPTION OF THE INVENTION

[0012] The invention described herein is a capsid-free DNA vector with covalently-closed ends (referred to herein as a "closed-ended DNA vector" or a "ceDNA vector") for controlled expression of a transgene in a cell, e.g., to treat a disease. In particular, the technology described herein relates to capsid-free close-ended DNA (ceDNA) vectors for controlled expression of a transgene, including but not limited to any of, sustained expression of a transgene, long-term controlled expression of a transgene, dose-dependent and/or tritratable expression of a transgene, and repeat dosing of a transgene using the vectors described herein. Accordingly, the methods disclosed herein enable one to personalize gene therapy throughout an individual's lifespan to express a transgene at a level that meets the individual's needs, by sustaining the transgene expression level at a predetermined level for a pre-determined time, or alternatively, increasing the expression of the transgene in a dose-dependent manner, by one or more administration after the initial priming administration, thereby controlling the transgene expression level

to a desired expression level or desired expression level range based on the concentration of ceDNA vector in the re-dose administration, allowing for a controlled and specific increases in the expression of the transgene in the cell or subject.

[0013] Accordingly, in some embodiments a ceDNA vector as disclosed herein can be re-administered (also referred to herein as a "redose" or "booster" administration) to continue transgene expression level at a predetermined level for a predetermined time, or to increase the expression level of the transgene above a prior expression level which was achieved on a first, or prior ceDNA vector administration, where the second administration or booster administration does not generate an immune reaction that prevents expression of the transgene by not generating an immune response to the vector itself that impacts expression of the transgene, or where the immune reaction is less as compared to a re-administration of a viral vector comprising viral proteins, including but not limited to a viral vector comprising a capsid, such as a parvovirus or a lentivirus.

[0014] Without wishing to be limited to theory, research shows that long-term transgene expression using conventional AAV viral vectors wanes over time. One way to ensure long-term transgene expression using conventional AAV vectors has traditionally been to increase the titre and dose of the delivered AAV vector at the initial administration. However, it is known that too high a titre or dose of the AAV vector can result in a variety of side-effects. Furthermore, as discussed above, re-administration of an AAV viral vector is traditionally not possible due to immune responses. Ways to circumvent causing an immune response to re-administration of AAV vectors typically require re-administration of an AAV vector with a different capsid configuration as compared to the AAV vector administered on any previous administrations. Such strategies however, can still pose significant risk to the subject by inducing an immune response to the AAV vector and potentially deleterious effects.

[0015] Herein, the invention provides a method for controlled expression of a transgene, including long term expression using a close-ended DNA (ceDNA) vector. It is demonstrated herein that a ceDNA vector can titrated to increase the transgene expression levels by, e.g., repeat or re-administration of the ceDNA vector. Additionally, the level of transgene expression can be maintained over a long term, and if any drop in expression levels is observed, a redose administration of the ceDNA vector can be used to maintain the desired level or even to increase the level of transgene expression if desirable for the subject and/or the disease or disorder to be treated.

[0016] Accordingly, provided herein are methods of administering ceDNA vectors for sustaining and/or increasing the level of expression of a transgene from a ceDNA vector in a cell or a subject, and where the expression level can be sustained or increased with one or more subsequent administrations (e.g., a re-dose or a booster administration). In the event that ceDNA-delivered transgene expression decreases for any reason, re-dosing of the ceDNA vector can re-establish or maintain the desired

expression of the transgene at a desired level. Also provided herein are methods for personalized gene therapy, such that the level of expression of a transgene expressed by the ceDNA vector can be increased incrementally, or in a step-by-step manner, with one or more administrations after an initial priming administration (e.g., at time 0), thereby tailoring (e.g., titration) of the level of expression of the transgene to a desired expression level or within a desired expression level range, as needed by the subject.

[0017] Accordingly, in some embodiments, the ceDNA vectors and methods disclosed herein can be used to titrate or effectuate an increase in the level of expression of a transgene by a ceDNA vector, and where the expression level of the transgene can be titrated in a dose-dependent manner with one or more subsequent administrations (e.g., a dose-dependent re-dose or booster administration). As such, the methods disclosed herein enable one to personalize gene therapy throughout an individual's lifespan to express a transgene at a level that meets the individual's needs, by sustaining the transgene expression level at a predetermined level, or alternatively, increasing the expression of the transgene in a dose-dependent manner, by one or more administration after the initial priming administration (e.g., at time 0), thereby controlling the transgene expression level to a desired expression level or desired expression level range based on the concentration of ceDNA vector in the re-dose administration, allowing for a controlled and specific increases in the expression of the transgene in the cell or subject.

[0018] Accordingly, in some embodiments a ceDNA vector as disclosed herein can be re-administered (also referred to herein as a "redose" or "booster" administration) to continue transgene expression level at a predetermined level for a predetermined time, or to increase the expression level of the transgene above a prior expression level which was achieved on a first, or prior ceDNA vector administration.

[0019] Accordingly, one aspect of the technology described herein relates to the use of ceDNA vector in methods for controlled transgene expression, for example, in a method for modulating expression levels of a transgene, or for a controlled increase in the transgene expression level, or for a dose-dependent expression of a transgene level in a cell or a subject, wherein the ceDNA vector comprises at least one heterologous nucleotide sequence (e.g., a transgene) operatively linked to a promoter and positioned between two inverted terminal repeat sequences, where the ITR sequences can be asymmetric, or symmetric, or substantially symmetrical as these terms are defined herein, wherein at least one of the ITRs comprises a functional terminal resolution site and a Rep binding site, and optionally the heterologous nucleic acid sequence encodes a transgene, and wherein the vector is not in a viral capsid.

[0020] In some embodiments, a ceDNA vector as described herein are capsid-free, linear duplex DNA molecules formed from a continuous strand of complementary DNA with covalently-closed ends (linear, continuous and non-encapsidated structure), which comprises two inverted terminal repeat (ITR) sequences flanking a transgene which is operatively linked to a promoter or other regulatory switch as

described herein. The 5' ITR and the 3' ITR can have the same symmetrical three-dimensional organization with respect to each other, (i.e., symmetrical or substantially symmetrical), or alternatively, the 5' ITR and the 3' ITR can have different three-dimensional organization with respect to each other (i.e., asymmetrical ITRs), as these terms are defined herein. In addition, the ITRs can be from the same or different serotypes. In some embodiments, 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 (i.e., they are the same or are mirror images with respect to each other). In some embodiments, one ITR can be from one AAV serotype, and the other ITR can be from a different AAV serotype.

[0021] Accordingly, some aspects of the technology described herein relate to a ceDNA vector for controlled transgene expression, including but not limited to, sustained or long-term expression of a transgene, dose-dependent or tritratable expression of a transgene, or repeated dosing of a transgene, where the ceDNA vector comprises ITR sequences selected from any of: (i) at least one WT ITR and at least one modified AAV inverted terminal repeat (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. The ceDNA vectors disclosed herein can be produced in eukaryotic cells, thus devoid of prokaryotic DNA modifications and bacterial endotoxin contamination in insect cells.

[0022] In some embodiments, the methods and ceDNA vectors as described herein allow a personalized genetic medicine approach, i.e., titrating an increase in the level of the transgene expression by re-dose administrations in a concentration-dependent manner. It is envisioned that increases in the transgene expression can be achieved in a dose-dependent, step-by-step manner using re-dose administrations, thus increasing the expression level of the transgene by a defined or certain amount by each re-dose administration. This enables controlled increases in the level of the transgene expression in a dose-dependent manner, and can be done incrementally. Accordingly, 1, 2, 3, 4, 5 or 6 or more than 6 re-doses of a define amount of ceDNA can be administered in order to increase the level of expression the transgene by a defined amount each time, to achieve a desired level, or to a desired expression level range, which is higher than the expression level achieved with the prior administration, or prior to this re-dose administration.

[0023] Accordingly, in some embodiments, A method of regulating expression of a transgene in a host comprising: (i) administering a sufficient amount of a ceDNA vector as disclosed herein comprising

a nucleic acid cassette containing at least one transgene operably linked to a promoter between flanking inverted terminal repeats (ITRs), to the host to express measurable levels of the transgene, wherein the transgene encodes a desired protein; and (ii) administering at least a second dose of the ceDNA vector comprising the at least one transgene or a modified transgene between flanking ITRs to (i) continue expression of the desired protein at a predetermined level for a predetermined time or (ii) modulate expression of the desired protein to a predetermined level, wherein the second administration of the ceDNA vector does not generate an immune reaction that prevents expression of the desired protein.

[0024] One aspect of the technology described herein relates to use of a ceDNA vector in a method for sustaining a desired level of expression of a transgene in a cell, method comprising: (a) administering to a cell at a first time point a first dose of a ceDNA vector to achieve expression of a transgene from the ceDNA vector, and (b) administering to the cell at a second time point another dose of the same or a different ceDNA vector to increase the level of expression of the transgene to a desired level, or to compensate for any decrease in expression level of the transgene after the initial ceDNA vector administration. It will be appreciated that such incremental increases in transgene expression permits titration of dosing in a subject to a desired level for such subject. In some embodiments, use of a ceDNA vector in a method for sustaining the level of expression of a transgene in a cell expresses the transgene at a desired expression level for at least 42 days. In some embodiments, the ceDNA vector expresses the transgene at a desired expression level for at least 84 days. In some embodiments, the ceDNA vector expresses the transgene at a desired expression level for at least 132 days.

[0025] In some embodiments, the ceDNA vector used in the methods described herein, e.g., in a method for sustaining expression of a transgene in a cell and/or for treating a subject with a disease, is administered in combination with a pharmaceutically acceptable carrier and/or excipient. In some embodiments, a ceDNA vector is administered at a second time point is administered at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months, or between 6-12 months, or between 1-2 years, or 2-3 years after the first time point.

[0026] In addition to a re-dose administration of a ceDNA vector to simply increase the level of transgene expression if expression levels have decreased over time (e.g., to continue transgene expression at a desired pre-determined level), in some embodiments, the methods and compositions of re-administration of a ceDNA vector can increase the level of transgene in a dose-dependent manner - that is, a re-dose administration of a defined amount of a ceDNA vector can effect a defined increase in expression level of a transgene. Stated differently and using arbitrary units for illustrative purposes only, a 1unit dose of the ceDNA in a re-dose administration will achieve a 10% increase in the level of transgene expression from a prior level, and a 2 unit dose of the cDNA vector will achieve a 20% increase in the

level of the transgene from a prior level, and a 0.5 unit dose of the ceDNA will achieve a 5% increase in the level of expression of the transgene from a prior level.

[0027] Accordingly, in one embodiment, a ceDNA vector as disclosed herein for controlled transgene expression can be used for increasing the level of expression of a transgene in a cell or a subject in a controlled manner. For example, the expression level of the transgene can be increased with one or more subsequent administrations (e.g., a re-dose or a booster administration) of the ceDNA vector.

[0028] Another aspect of the technology herein relates to a method for increasing expression of a transgene in a cell, e.g., to increase the expression level of a transgene above a prior expression level that was achieved with a prior ceDNA administration, the method comprising: (a) administering to a cell at a first time point, a priming dose of a ceDNA vector to achieve expression of a transgene, and (b) administering to the cell at a second time point, a dose of a ceDNA vector to increase the expression level of the transgene as compared to the level of expression of the transgene achieved after administration of the ceDNA vector at the first time point, or to increase the expression level of the transgene to achieve a desired expression level.

[0029] In all aspects described herein, a ceDNA vector is administered at any time point (e.g. a first, second, third time point etc.) is administered in combination with a pharmaceutically acceptable carrier, and can be optionally administered with a carrier, for example, a particle, liposome or lipid nanoparticle (LNP). In all aspects herein, a ceDNA vector administered at any of: the first, second or any subsequent time point, is administered in combination with a pharmaceutically acceptable carrier.

[0030] In all aspects described herein, a ceDNA vector used in the methods for controlled transgene expression as described herein, e.g., in a method for sustaining expression of a transgene, or for a controlled increase in the expression of the transgene, or for a dose-dependent expression of the transgene, and/or for treating a subject with a disease, the ceDNA vector administered at any of: the first, second or any subsequent time point, is administered in combination with a pharmaceutically acceptable carrier and/or excipient.

[0031] In some embodiments, where more than one administration of the ceDNA is administered (e.g., at a second or any subsequent time point), the second time point, or any subsequent time point is at least 10 days or between 10-30 days, or at least 30 days, or between 30-60 days, at least 60 days, or between 60-90 days, or between 90-120 days, or between about 3-6 months, or between 6-12 months, or at least a year, after the ceDNA vector administration at the first time point, or the previous time point.

[0032] In some embodiments, a ceDNA vector administered at the first, second or any subsequent time point is the same ceDNA vector comprising the same transgene, or a modified transgene, and in alternative embodiments, a ceDNA vector administered at the first, second or any subsequent time point

is a different ceDNA vector comprising the same transgene, or a modified transgene, e.g., a different ceDNA vector with a different promoter operatively linked to the same transgene, or a modified transgene. In some embodiments, the promoter is an inducible or repressible promoter. The transgene can also be part of a regulatory switch, as disclosed herein.

[0033] In some embodiments, a ceDNA vector used in the methods described herein for controlled transgene expression, the ceDNA vector administered at the first, second or any subsequent time point is the same ceDNA vector comprising the same transgene, or a modified transgene. In alternative embodiments, the ceDNA vector administered at the first, second or any subsequent time point is a different ceDNA vector comprising the same transgene, or a modified transgene, for example, but not limited to, where the different ceDNA vector has a different promoter operatively linked to the same transgene, or to a modified transgene, or a different transgene. For illustrative purposes only, a ceDNA vector administered at the first timepoint can comprise a transgene and a first promoter or regulatory switch, and a ceDNA administered at a second or subsequent timepoint can comprise the same or a modified transgene and a second promoter or regulatory switch, where the first and second promoter (or regulator switch) are different promoters or different regulatory switches. Exemplary regulatory switches are defined herein.

[0034] In some embodiments, use of ceDNA vector in the methods for controlled transgene expression as described herein, e.g., in a method for sustaining expression of a transgene, or for a controlled increase in the expression of the transgene, or for a dose-dependent expression of the transgene, and/or for treating a subject with a disease, can optionally comprise a step of administering to the cell, at one or more time points after the second time point, a further dose of the ceDNA vector to increase the expression level of the transgene as compared to the level of expression of the transgene achieved after administration of the ceDNA vector at the second time point or previous time point, or to increase the expression level of the transgene to maintain a desired sustained expression level, wherein the composition administered at the one or more time points after the second time point comprises a ceDNA vector as described herein.

[0035] In some embodiments, the ceDNA vector useful in the methods disclosed herein for controlled transgene expression allows for expression of the transgene at a therapeutically effective amount.

[0036] In some embodiments, increasing the predetermined dose of the ceDNA vector administered at a second time point, or any subsequent time point, increases the expression level of the transgene in the cell and/or subject. In some embodiments, a predetermined dose of a ceDNA vector administered to the cell or subject at second time point, or subsequent time point, is determined using a dose-dependent relationship for the ceDNA vector to achieve the desired level of expression of the transgene in the cell or subject.

[0037] In some embodiments, a predetermined dose of the ceDNA vector administered at the second or any subsequent time point, is in an amount that is between 2-fold and 10-fold the dose of the ceDNA

vector administered at the first time point. In some embodiments, a predetermined dose of the ceDNA vector administered at the second or any subsequent time point, is in an amount that increases the expression of the transgene by at least 3-fold, or at least 5-fold, or least 10-fold, or between 2-15 fold or between 2-20 fold, or more than 20-fold as compared the expression of the transgene achieved after administration of the ceDNA at the first time point or previous time point. In some embodiments, the desired expression level of transgene achieved after the administration of the composition at one or more time points after the second time point is a therapeutically effective amount of the transgene.

[0038] Aspects of the invention relate to methods to produce a ceDNA vector used in the methods for controlled transgene expression as described herein, e.g., in a method for sustaining expression of a transgene, or for a controlled increase in the expression of the transgene, or for a dose-dependent expression of the transgene, and/or for treating a subject with a disease. In all aspects, the capsid free, non-viral DNA vector (ceDNA vector) for controlled transgene expression is obtained from a plasmid (referred to herein as a "ceDNA-plasmid") comprising a polynucleotide expression construct template comprising in this order: a first 5' inverted terminal repeat (e.g. AAV ITR); a heterologous nucleic acid sequence; and a 3' ITR (e.g. AAV ITR), where the 5' ITR and 3'ITR can be asymmetric relative to each other, or symmetric (e.g., WT-ITRs or modified symmetric ITRs) as defined herein.

[0039] A ceDNA vector useful in the methods for controlled transgene expression as described herein, (e.g., in a method for sustaining expression levels of a transgene, and/or for a controlled increase in the transgene expression level, or for a dose-dependent transgene expression level) is obtainable by a number of means that would be known to the ordinarily skilled artisan after reading this disclosure. For example, a polynucleotide expression construct template used for generating the ceDNA vectors of the present invention can be a ceDNA-plasmid (e.g. see **FIG. 4B),** a ceDNA-bacmid, and/or a ceDNA-baculovirus. In one embodiment, the ceDNA-plasmid comprises a restriction cloning site (e.g. SEQ ID NO: 123 and/or 124 operably positioned between the ITRs where an expression cassette comprising e.g., a promoter operatively linked to a transgene, e.g., a reporter gene and/or a therapeutic gene) can be inserted. In some embodiments, ceDNA vectors are produced from a polynucleotide template (e.g., ceDNA-plasmid, ceDNA-bacmid, ceDNA-baculovirus) containing symmetric or asymmetric ITRs (modified or WT ITRs).

[0040] In a permissive host cell, in the presence of e.g., Rep, the polynucleotide template having at least two ITRs 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. Rep proteins and Rep binding sites of the various AAV serotypes are well known to those of ordinary skill in the art. One of ordinary skill understands to choose a Rep protein from a serotype that

binds to and replicates the nucleic acid sequence based upon at least one functional ITR. For example, if the replication competent ITR is from AAV serotype 2, the corresponding Rep would be from an AAV serotype that works with that serotype such as AAV2 ITR with AAV2 or AAV4 Rep but not AAV5 Rep, which does not. Upon replication, the covalently-closed ended ceDNA vector continues to accumulate in permissive cells and ceDNA vector is preferably sufficiently stable over time in the presence of Rep protein under standard replication conditions, e.g. to accumulate in an amount that is at least 1 pg/cell, preferably at least 2 pg/cell, preferably at least 3 pg/cell, more preferably at least 4 pg/cell, even more preferably at least 5 pg/cell.

[0041] Accordingly, one aspect of the invention relates to a process of producing a ceDNA vector useful in the methods for controlled transgene expression as described herein, 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 virion-enforced size limitation.

[0042] The presence of the ceDNA vector useful for controlled expression of the transgene as described herein is 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 denaturing and non-denaturing gels to confirm the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA.

[0043] In another embodiment of this aspect and all other aspects provided herein, the transgene expressed in a controlled manner from the ceDNA vector is therapeutic transgene, e.g., a protein of interest, including but not limited to, a receptor, a toxin, a hormone, an enzyme, or a cell surface protein. In another embodiment of this aspect and all other aspects provided herein, the protein of interest is a receptor. In another embodiment of this aspect and all other aspects provided herein, the protein of interest is an enzyme. Exemplary genes to be targeted and proteins of interest are described in detail in the methods of use and methods of treatment sections herein.

[0044] In some embodiments, the present application may be defined in any of the following paragraphs:

[0045] A method of regulating expression of a transgene in a subject comprising: (a) administering to a subject a sufficient amount of a non-viral, capsid-free close-ended DNA (ceDNA) vector comprising

a nucleic acid cassette containing at least one transgene operably linked to a promoter between flanking inverted terminal repeats (ITRs), to express a measurable level of the transgene, wherein the transgene encodes a desired protein to treat a disease; and (b) titrating the ceDNA vector by administering to the subject at least a second dose of the ceDNA vector comprising the at least one transgene between flanking ITRs to obtain the transgene expression of the desired protein at a predetermined level for a predetermined time or to increase the transgene expression of the desired protein to a predetermined level.

[0046] In some embodiments, the subject is assessed at a predetermined time after step (a), e.g., at least 30 days, or at least 60 days, or between 60-90 days or longer than 90 days after step (a), to determine the titrating dose. For example, in some embodiments, the subject is assessed to determine the disease state in the subject after step (a) and/or the level of desired protein expressed by the ceDNA vector in the subject. In some embodiments, assessment of the disease state is an assessment of at least one symptom of the disease in the subject. In some embodiments, if the disease state of the subject has remained at a steady state, or has not improved, or where the disease state has declined in the subject, for example, as compared to the disease state at the time of the first administration of the ceDNA vector or any time before step (a), the subject is administered a second dose of the ceDNA vector according to step (b). The disease state for any given disease can be determined by a physician or person of skill in the art, and includes assessing one or more clinical symptoms and/or biomarkers of the disease, including protein biomarkers, miRNA and mRNA biomarkers and the like. In some embodiments, if the level of transgene expression in the subject has declined from a predetermined level or declined from a therapeutically effective amount, the subject is administered a second dose of the ceDNA vector according to step (b). In some embodiments, the level of the transgene expression is determined by measuring the level of the transgene (e.g., measuring protein level or mRNA levels) expressed from the ceDNA vector in a biological sample obtained from the subject. In some embodiments, the biological sample is selected from a blood sample, plasma, synovial fluid, CSF, saliva, or tissue biopsy sample. In some embodiments, where the ceDNA vector expresses a transgene encoding a desired protein or therapeutic gene and a reporter protein, the level of the transgene can be determined by measuring the desired reporter protein expressed from the ceDNA vector *in vivo,* using methods commonly known to persons of ordinary skill in the art. In some embodiments, the titrating the ceDNA vector is determining the level of transgene expressed from the ceDNA vector and administering a second dose of the ceDNA vector to the subject to adjust or modulate the transgene expression to a predetermined desired level.

[0047] Another aspect of the technology described herein relates to a method of regulating expression of a transgene in a subject comprising: (a) administering a sufficient amount of a non-viral capsid-free close-ended DNA (ceDNA) vector comprising a nucleic acid cassette containing at least one transgene operably linked to a promoter between flanking inverted terminal repeats (ITRs), to the subject

to express a measurable level of the transgene, wherein the transgene encodes a desired protein; and (b) administering to the subject at least a second dose of the ceDNA vector comprising the at least one transgene, or a modified transgene, between flanking ITRs to (i) continue expression of the desired protein at a predetermined level for a predetermined time or (ii) modulate expression of the desired protein to a predetermined level.

[0048] In all aspects herein, the second administration of the ceDNA vector to the subject does not generate an immune reaction sufficient to prevent obtaining the predetermined level of expression of the desired protein.

[0049] In some embodiments, the ceDNA vector is administered to the subject at first administration, or a second administration or any subsequent administration in combination with a pharmaceutically acceptable carrier.

[0050] In some embodiments, the second administration of the ceDNA vector is at a time when the level of the expression of the transgene decreases from a desired predetermined level, for example, in some embodiments, the second administration is at least about 30 days, or at least about 60 days, or at least about 90 days after the first administration. When more than two doses of the ceDNA vector are administered to the subject, each redose (e.g., 3rd, 4th, 5th, 6th and subsequent redoses) are administered at a time when the level of the expression of the transgene decreases or drops from a desired predetermined level achieved from the previous administration, for example, in some embodiments, each re-administration is at least about 30 days, or at least about 60 days, or at least about 90 days after the previous ceDNA vector administration.

[0051] In some embodiments, the method comprises administering at least three or more administrations of the ceDNA vector to the subject, and where at least three administrations of the ceDNA vector are administered, none of the administrations generate an immune response to the ceDNA vector that prevents the achieving the predetermined level of expression of the desired protein.

[0052] In some embodiments, the ceDNA vector is administered to the subject on a periodic schedule, e.g., every 2-months, every 3 months, every 6 months, every 12 months, every 18 months and the like.

[0053] In some embodiments, the second administration is to increase the level of expression of the desired protein, e.g., to prolong the expression of the desired protein at a predetermined level of expression.

[0054] In all aspects herein, the transgene encodes a therapeutic protein and the desired level of expression of the transgene is a therapeutically effective amount of the therapeutic protein. In some embodiments, the transgene is a genetic medicine selected from any of: a nucleic acid, an inhibitor, peptide or polypeptide, antibody or antibody fragment, fusion protein, antigen, antagonist, agonist, RNAi

molecule, etc. In some embodiments, the desired protein or therapeutic protein is an inhibitor protein, for example, but not limited to, an antibody or antigen-binding fragment, or a fusion protein. In some embodiments, the desired protein or therapeutic protein replaces a defective protein or a protein that is not being expressed or being expressed at low levels. In some embodiments, the transgene is under the control of a regulatory switch, as defined herein.

[0055] In some embodiments, the ceDNA vector comprises a promoter which is an inducible or repressible promoter.

[0056] In some embodiments, the ceDNA vector administered at the first, second or any subsequent time point is the same type of ceDNA vector comprising the same transgene, or a modified transgene. For example, stated differently, the same ceDNA vector is administered to the subject multiple times and is comparable to administering the same serotype of viral vector to a subject multiple times.

[0057] In some embodiments, a ceDNA vector administered to the subject at second administration or any subsequent administration thereafter (e.g., a redose administration) has a different promoter operatively linked to the same transgene, or a modified transgene, as compared to the promoter in the ceDNA vector administered at an earlier timepoint or administration.

[0058] In some embodiments, a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter comprises two inverted terminal repeat sequences (ITRs) that are AAV ITRs, and can be, e.g., AAV-2, or any ITR selected from Table 1, or AAV1, AAV3, AAV4, AAV5, AAV 5, AAV7, AAV8, AAV9, AAV10, AAV 11, AAV12, AAVrh8, AAVrh10, AAV-DJ, and AAV-DJ8. In some embodiments, at least one ITR comprises a functional terminal resolution site and a Rep binding site. In some embodiments, the flanking ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter are symmetric or substantially symmetrical or asymmetric, as defined herein. In some embodiments, one or both of the ITRs are wild type, or wherein both of the ITRs are wild-type. In some embodiments, the flanking ITRs are from different viral serotypes. In some embodiments, where the flanking ITRs are both wild type, they can be selected from any AAV serotype as shown in Table 1.

[0059] In some embodiments, the flanking ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter can comprise a sequence selected from the sequences in Tables 2, 4A, 4B or 5 herein.

[0060] In some embodiments, at least one of the ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter is altered from a wild-type AAV ITR sequence by a deletion, addition, or substitution that affects the overall three-dimensional conformation of the ITR. In some embodiments, one or both of the ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration

thereafter is derived from an AAV serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, and AAV12.

[0061] In some embodiments, one or both of the ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter are synthetic. In some embodiments, one or both of the ITRs is not a wild type ITR, or wherein both of the ITRs are not wild-type.

[0062] In some embodiments, one or both of the ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter is modified by a deletion, insertion, and/or substitution in at least one of the ITR regions selected from A, A', B, B', C, C', D, and D'. In some embodiments, a deletion, insertion, and/or substitution results in the deletion of all or part of a stem-loop structure normally formed by the A, A', B, B' C, or C' regions. In some embodiments, one or both of the ITRs are modified by a deletion, insertion, and/or substitution that results in the deletion of all or part of a stem-loop structure normally formed by the B and B' regions. In some embodiments, one or both of the ITRs are modified by a deletion, insertion, and/or substitution that results in the deletion of all or part of a stem-loop structure normally formed by the C and C' regions. In some embodiments, one or both of the ITRs are modified by a deletion, insertion, and/or substitution that results in the deletion of part of a stem-loop structure normally formed by the B and B' regions and/or part of a stem-loop structure normally formed by the C and C' regions. In some embodiments, one or both of the ITRs comprise a single stem-loop structure in the region that normally comprises 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. In some embodiments, one or both of the ITRs comprise a single stem and two loops in the region that normally comprises 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.

[0063] In some embodiments, both ITRs in a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter are altered in a manner that results in an overall three-dimensional symmetry when the ITRs are inverted relative to each other.

[0064] In some embodiments, a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter comprises at least one heterologous nucleotide sequence under the control of at least one regulatory switch, for example, at least one regulatory switch is selected from a binary regulatory switch, a small molecule regulatory switch, a passcode regulatory switch, a nucleic acid-based regulatory switch, a post-transcriptional regulatory switch, a radiation-controlled or ultrasound controlled regulatory switch, a hypoxia-mediated regulatory switch, an inflammatory response regulatory switch, a shear-activated regulatory switch, and a kill switch. Regulatory switches are disclosed herein in more detail below.

[0065] In some embodiments, a ceDNA vector administered at the first administration, or second administration or any subsequent administration thereafter is administered to a subject that has a disease or disorder selected from, e.g., cancer, autoimmune disease, a neurodegenerative disorder, hypercholesterolemia, acute organ rejection, multiple sclerosis, post-menopausal osteoporosis, skin conditions, asthma, or hemophilia. In some embodiments, a subject with cancer has a solid tumor, soft tissue sarcoma, lymphoma, and leukemia. In some embodiments, the subject has an autoimmune disease, e.g., selected from rheumatoid arthritis and Crohn's disease. In some embodiments, the subject has a skin condition, e.g., is selected from psoriasis and atopic dermatitis. In some embodiments, the subject has a neurodegenerative disorder, e.g., Alzheimer's disease, ALS, Parkinson's Disease, Huntington's Disease.

[0066] In some embodiments, the method disclosed herein further comprise administering to the subject, at one or more time points after the second time point, a dose of the ceDNA vector to increase the expression level of the heterologous nucleic acid sequence (e.g., the transgene) as compared to the level of expression of the transgene achieved after administration of the ceDNA vector at the second time point or previous time point, or to increase the expression level of the transgene to achieve a desired expression level.

[0067] In some embodiments, a predetermined dose of the ceDNA vector administered to the subject at a second or any subsequent time point, is in an amount that is between 2-fold and 10-fold the dose of the ceDNA vector composition administered at the first time point. In some embodiments, a predetermined dose of the ceDNA vector composition administered at the second or any subsequent time point, is in an amount that increases the expression of the transgene by at least 3-fold, or at least 5-fold, or least 10-fold, or between 2-15 fold or 2-20 fold as compared the level of expression of the transgene after administration of the ceDNA vector at the first time point or previous administration. In some embodiments, a predetermined dose of the ceDNA vector administered at the second administration, or second time point, is determined using a dose-dependent relationship for the ceDNA vector to achieve the desired level of expression of the transgene in the cell.

[0068] These and other aspects of the invention are described in further detail below.

DESCRIPTION OF DRAWINGS

[0069] 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.

[0070] **FIG. 1A** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression 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.

[0071] **FIG. 1B** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression as disclosed herein comprising asymmetric ITRs with an expression cassette containing CAG promoter, WPRE, and BGHpA. An open reading frame (ORF) encoding a 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.

[0072] **FIG. 1C** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression as disclosed herein comprising asymmetric ITRs, 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 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).

[0073] **FIG. 1D** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression 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 a 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.

[0074] **FIG. 1E** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression 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.

[0075] **FIG. 1F** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression 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.

[0076] **FIG. 1G** illustrates an exemplary structure of a ceDNA vector for controlled transgene expression 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.

[0077] **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.

[0078] **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.

[0079] **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 controlled transgene expression 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.

[0080] **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; BamH1 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.

[0081] **FIG. 6** is a graph showing the effect of a re-dose (i.e., a booster administration) for increasing the level of expression of a transgene from a ceDNA vector expressing luciferase present in a composition comprising a liposome. Expression of luciferase was measured following administration of a ceDNA vector as described in Example 6, and then later re-administration of a ceDNA vector produced from the ceDNA vector at day 84 or 87. Luciferase expression was assessed and detected in all three groups until at least 132 days (the longest time period assessed). **FIG. 6** shows that at or about day 80, the level of the expression of the transgene in mice administered 1mg/kg ceDNA vector in the presence of a liposome (LNPceDNA) decreases slightly. Re-administration of a ceDNA vector in the presence of a lipsome at day 84 or day 87 can be used to continue transgene expression at a desired pre-determined level (data not shown), or increase the transgene expression level from the ceDNA to a level above that achieved from the prior ceDNA vector administration. Shown here is an increase in expression by 7-fold above the previous transgene expression level by administering 3mg/kg LNPceDNA vector, or a 17-fold increase in expression level above the previous transgene expression level by administering a 10 mg/kg LNPceDNA vector composition.

[0082] **FIG. 7** 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.

[0083] **FIG. 8** 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.

[0084] **FIGS. 9A-9B** depict the results of the ocular studies set forth in Example 9. **FIG. 9A** 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. 9B** 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.

[0085] **FIGS. 10 and 10B** depict the results of the ceDNA persistence and redosing study in Rag2 mice described in Example 10. **FIG. 10A** shows a graph of total flux over time observed in LNP-ceDNA-Luc-treated wild-type c57bl/6 mice or Rag2 mice. **FIG. 10B** 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).

[0086] **FIG. 11** 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.

DETAILED DESCRIPTION

[0087] Described herein are methods and compositions comprising novel capsid-free DNA vectors with covalently-closed ends (ceDNA) for controlled expression of a transgene, e.g., to enable sustained expression of the desired transgene at a desired and for a predetermined time, or to modulate expression of the transgene level (including increasing the expression level) in a cell, either *in vivo* or *in vitro* and where the expression level of the transgene can be increased with at least one (i.e., one or more) subsequent administrations (e.g., a booster administration, or re-dose).

[0088] A ceDNA vector and methods as disclosed herein enable one to sustain the expression level of a transgene in vitro and in vivo in a host cell or subject, i.e., to maintain expression to a desired level, or to stop any deterioration in the expression level by at least one re-administration (herein also referred to as a re-dose or booster administration) at a time point after the initial administration.

[0089] A ceDNA vector and methods as disclosed herein enable one to increase the expression level of a transgene from a prior level *in vitro* and *in vivo*, i.e., to increase expression to, or above a desired level, or to increase the expression level to within a desired expression range, by at least one re-administration (herein also referred to as "re-dose" or "booster" administration) at a time point after the initial administration.

[0090] That is, expression of the transgene expressed by the ceDNA can be increased above a level from the prior administration. If the prior administration was an initial dose (i.e., a priming dose), then a

re-dose administration at a second time point can be used to increase the expression level of the transgene. Similarly, if the prior administration was a second administration (i.e., a re-dose administration), then an additional re-dose administration can be used to increase the level of expression of the transgene to a level higher than, or a to a desired expression level or range, than the prior re-dose administration. Therefore, the technology, methods and ceDNA vector as disclosed herein can be used to incrementally, in a controlled manner, increase the expression level of the transgene to a desired expression level. Such a step-wise and incremental increases in expression level of the transgene is advantageous for treatment of a subject, as it allows one to titrate the level of expression of the transgene to a particular individual, based on the subject's need and/or efficacy of the ceDNA vector and/or expressed transgene (e.g., genetic medicine) in the subject, without the risk of having to have a high initial dose administered in excess of what is actually needed and/or without the immune complications associated with other AAV-based vectors.

Definitions

[0091] 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.

[0092] 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 ceDNA vector as disclosed herein. Transgenes of interest include, but are not limited to, nucleic acids encoding polypeptides, preferably therapeutic (e.g., for medical, diagnostic, or veterinary uses) or immunogenic polypeptides (e.g., for vaccines). In some embodiments, nucleic acids of interest include nucleic acids that are transcribed into therapeutic RNA. Transgenes included for use in the ceDNA vectors of the invention include, but are not limited to, those that express or encode one or more polypeptides, peptides, ribozymes, aptamers, peptide nucleic acids, siRNAs, RNAis, miRNAs, lncRNAs, antisense oligo- or polynucleotides, antibodies, antigen binding fragments, or any combination thereof. A transgene can be a "genetic medicine" and encompasses any of: an inhibitor, nucleic acid, oligonucleotide, silencing nucleic acid, miRNA, RNAi, antagonist, agonist, polypeptide, peptide, antibody or antibody fragments, fusion proteins, or variants thereof, epitopes, antigens, aptamers, ribosomes, and the like. A transgene used herein in the ceDNA vector is not limited in size.

[0093] The term "genetic medicine" as disclosed herein relates to any DNA structure or nucleic acid sequence that can be used to treat or prevent a disease or disorder in a subject.

[0094] 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.

[0095] 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.

[0096] 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.

[0097] 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.

[0098] 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.

[0099] 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.

[00100] A DNA sequence that "encodes" a particular RNA or protein gene product 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").

[00101] 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.

[00102] As used herein, the term “gene delivery” means a process by which foreign DNA is transferred to host cells for applications of gene therapy.

[00103] 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".

[00104] 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).

[00105] 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.

[00106] 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.

[00107] As used herein, the term "asymmetric ITRs" also referred to herein 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. The difference in sequence between the two ITRs may be due to nucleotide addition, deletion, truncation, or point mutation. In one embodiment, one ITR of the pair may be a wild-type AAV sequence and the other a non-wild-type or synthetic sequence. In another embodiment, neither ITR of the pair is a wild-type AAV sequence and the two ITRs differ in sequence from one another. 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". 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.

[00108] As used herein, the term "symmetric ITRs" refers to a pair of ITRs within a single ceDNA genome or ceDNA vector that are mutated or modified relative to wild-type dependoviral ITR sequences and are inverse complements across their full length. 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".

[00109] 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.

[00110] 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.

[00111] 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.

[00112] 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.

[00113] 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 he 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.

[00114] 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).

[00115] As used herein, the term "ceDNA-plasmid" refers to a plasmid that comprises a ceDNA genome as an intermolecular duplex.

[00116] 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.

[00117] As used herein, the term "ceDNA-baculovirus" refers to a baculovirus that comprises a ceDNA genome as an intermolecular duplex within the baculovirus genome.

[00118] 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.

[00119] 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.

[00120] 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.

[00121] As defined 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.

[00122] 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.

[00123] Transcriptional regulators refer to transcriptional activators and repressors that either activate or repress transcription of a gene of interest. 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.

[00124] 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.

[00125] 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.

[00126] 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.

[00127] 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.

[00128] 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, or that of another promoter used in another modular component of the synthetic biological circuits described herein. 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 ceDNA 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.

[00129] 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 pars 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.

[00130] 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.

[00131] 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.

[00132] 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.

[00133] 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 or 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.

[00134] 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.

[00135] The term "operably linked" 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.

[00136] 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, cynomologous 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.

[00137] 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.

[00138] 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.

[00139] 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.

[00140] 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.

[00141] 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.

[00142] 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 ceDNA 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.

[00143] 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).

[00144] 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.

[00145] 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.

[00146] The term "biomarker" as used herein is meant any assayable characteristic or composition that can be used to identify a condition (e.g., a disease) or the status of a condition (e.g., disease state) of the subject or a sample. A biomarker can, in some examples disclosed herein, be a gene whose expression characteristics can be used to identify a condition or status of a condition in a subject or sample. In other examples, a biomarker can be a gene product. In some embodiments, the term "biomarker" refers to a polypeptide expressed endogenously in an individual or found or sequestered in a biological sample from an individual.

[00147] By "gene product" is meant a transcript (e.g., mRNA), nucleic acid (e.g., miRNA), or protein. Thus, disclosed herein are biomarkers whose presence, absence, or relative amount can be used to identify a condition or status of a condition in a subject or sample. In one particular example, a biomarker can be a gene product whose presence or absence in a subject is characteristic of a subject having or not having a particular neurodegenerative disease, having a particular risk for developing a disease, (e.g., a neurodegenerative disease), or being at a particular stage of disease. In still another example, a biomarker can be a gene product whose increase or decrease indicates a particular disease state, a particular risk for developing a disease, or a particular stage of disease. In another example, a biomarker can be a group of various gene products, the presence or absence of which is indicative of a subject having or not having a particular disease, having a particular risk for developing a disease, or being at a particular stage of disease. In a further example, a biomarker can be a group of gene products whose pattern of increasing and decreasing expression characterizes a particular disease or lack thereof. Still further, a biomarker can be a gene product or group of gene products whose pattern of expression is characteristic of the presence or absence of a disease, or a particular prognosis or outcome of a disease. As used herein, a biomarker can be a surrogate for other clinical tests. Biomarkers identified herein can be measured to determine levels, expression, activity, or to detect variants. As used throughout when detecting levels of expression or activity are discussed, it is understood that this could reflect variants of a given biomarker. Variants include amino acid or nucleic acid variants or post translationally modified variants.

[00148] The term "biological sample" as used herein refers to a cell or population of cells or a quantity of tissue or fluid from a subject. Most often, the sample has been removed from a subject, but the term "biological sample" can also refer to cells or tissue analyzed *in vivo*, i.e. without removal from the

subject. Often, a "biological sample" will contain cells from the subject, but the term can also refer to non-cellular biological material, such as non-cellular fractions of blood, saliva, or urine, that can be used to measure gene expression or protein expression levels. Biological samples include, but are not limited to, tissue biopsies, scrapes (e.g. buccal scrapes), whole blood, plasma, serum, urine, saliva, cell culture, or cerebrospinal fluid. Biological samples also include tissue biopsies, cell culture. A biological sample or tissue sample can refer to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, tumor biopsy, urine, stool, sputum, spinal fluid, pleural fluid, nipple aspirates, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, cells (including but not limited to blood cells), tumors, organs, and also samples of in vitro cell culture constituent. In some embodiments, the sample is from a resection, bronchoscopic biopsy, or core needle biopsy of a primary or metastatic tumor, or a cellblock from pleural fluid. In addition, fine needle aspirate samples are used. Samples can be either paraffin-embedded or frozen tissue. The sample can be obtained by removing a sample of cells from a subject, but can also be accomplished by using previously isolated cells (e.g. isolated by another person), or by performing analysis of the level of transgene expression from the ceDNA vector *in vivo*. Biological sample also refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, tumor biopsy, urine, stool, sputum, spinal fluid, pleural fluid, nipple aspirates, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, cells (including but not limited to blood cells), tumors, organs, and also samples of in vitro cell culture constituent. In some embodiments, the biological samples can be prepared, for example biological samples can be fresh, fixed, frozen, or embedded in paraffin.

[00149] The term “blood sample” or “blood” as used herein include, but are not limited to, whole blood, serum or plasma. In some embodiments, the whole blood sample is further processed into serum or plasma samples. The term also includes a mixture of the above-mentioned samples.

[00150] The term “inhibitor” as used herein refers to any agent or entity which results in the inhibition of a proteins biological activity. By a “decrease” or “inhibition” used in the context of the level of activity of a gene refers to a reduction in protein or nucleic acid level or biological activity in a cell, a cell extract, or a cell supernatant. For example, such inhibition may be due to decreased binding of the polypeptide to its endogenous ligand, or by non-completive binding of an inhibitor to a polypeptide to reduce catalytic activity or affinity for target ligand etc., or alternatively to reduced RNA stability, transcription, or translation, increased protein degradation, or RNA interference. Preferably, a decrease is at least about 5%, at least about 10%, at least about 25%, at least about 50%, at least about 75%, at least about 80%, or even at least about 90% of the level of expression or activity under control conditions. The term “inhibiting” as used herein as it pertains to the inhibition of the activity of topoisomerase I protein or

variants thereof does not necessarily mean complete inhibition of expression and/or activity. Rather, expression or activity of the protein, polypeptide or polynucleotide is inhibited to an extent, and/or for a time, sufficient to produce the desired effect.

[00151] The terms "lower", "reduced", "reduction" or "decrease" or "inhibit" are all used herein generally to mean a decrease by a statistically significant amount. However, for avoidance of doubt, "lower", "reduced", "reduction" or "decrease" or "inhibit" means a decrease by at least 10% as compared to a reference level, for example a decrease by at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% decrease (i.e. absent level as compared to a reference sample), or any decrease between 10-100% as compared to a reference level.

[00152] The terms "increased" ,"increase" or "enhance" or "higher" are all used herein to generally mean an increase by a statically significant amount; for the avoidance of any doubt, the terms "increased", "increase" or "enhance" or "higher" means an increase of at least 10% as compared to a reference level, for example an increase of at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% increase or any increase between 10-100% as compared to a reference level, or at least about a 2-fold, or at least about a 3-fold, or at least about a 4-fold, or at least about a 5-fold or at least about a 10-fold increase, or any increase between 2-fold and 10-fold or greater as compared to a reference level.

[00153] By an "increase" in the expression or activity of a gene or protein is meant a positive change in protein or polypeptide or nucleic acid level or activity in a cell, a cell extract, or a cell supernatant. For example, such an increase may be due to increased RNA stability, transcription, or translation, or decreased protein degradation. Preferably, this increase is at least 5%, at least about 10%, at least about 25%, at least about 50%, at least about 75%, at least about 80%, at least about 100%, at least about 200%, or even about 500% or more over the level of expression or activity under control conditions.

[00154] 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.

[00155] 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.

[00156] 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.

[00157] As used herein the term "consisting essentially of" refers to those elements required for a given embodiment. The term permits the presence of additional elements that do not materially affect the basic and novel or functional characteristic(s) of that embodiment of the invention.

[00158] 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."

[00159] 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 ±1%. The present invention is further explained in detail by the following examples, but the scope of the invention should not be limited thereto.

[00160] 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.

[00161] 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.

[00162] Other terms are defined herein within the description of the various aspects of the invention.

[00163] 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.

[00164] 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.

[00165] 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.

[00166] 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.

I. Method for controlled expression of a transgene using a ceDNA vector

[00167] Described herein are methods of administration of ceDNA vectors *in vivo* or *in vitro*, that enable either (i) sustained expression of level of a transgene for a predetermined period of time (i.e., maintaining transgene expression levels), or (ii) to increase the expression level of the transgene in a dose-dependent manner, the method comprising at least one (e.g., one or more) subsequent administrations (e.g., a "booster" or "re-dose" administration) of the ceDNA vector.

[00168] Accordingly, the technology described herein relates administration of a ceDNA vector *in vivo*, where the level of transgene expression is sustained at a desired level or increased in level with one or more subsequent administrations (e.g., a booster administration, or re-dose). The ceDNA vectors disclosed herein enable an increase in the level of transgene expression from a prior level *in vitro* and *in vivo*. The increase in the level of transgene can be to, or above a desired level, or to increase the expression level to within a desired expression range, by at least one re-administration (herein also referred to as "re-dose") at any time point after the initial administration. In some cases, the desired increase is to correct a naturally-occurring decrease in those levels such that the final dose is a return to the previously desired level, e.g., a sustained dose level. That such re-dose administration or repeated doses of a ceDNA vector as disclosed herein can be used to increase the level of transgene expression is possible because the ceDNA vector has no capsid to provoke a host immune response (e.g., the vector has non-immunogenic properties) and thus offers significant advantages over existing AAV vector technology, where re-doses are not possible due to immune responses or AAV immunity in the host from a prior AAV exposure, or e.g. from the initial AAV vector administration. Thus, to achieve a high level of expression using rAAV vectors, an initial high dose is typically required and re-doses are either not possible and/or not effective due to immune response issues. In contrast, one or more re-dose administrations of a ceDNA vector as disclosed herein can be administered to increase the level of expression of a transgene, e.g., to increase the level of expression to a desired expression level, or above a desired threshold level or within a desired expression level range.

[00169] Accordingly, the present invention relates to methods and ceDNA vectors for controlled transgene expression for any of: (i) maintain or sustain the level of transgene expression at a desired level for a predetermined time by re-administration of ceDNA vector at one or more time points, (ii) increasing transgene expression level by re-administration of ceDNA vector at one or more time points or (iii) dose-dependent transgene expression by administration of the ceDNA vector in a dose-dependent manner (i.e., titratable expression of the transgene). As discussed herein, the cDNA vectors are unique as compared to other viral vectors in that the ceDNA vector can be administered to the subject multiple times, e.g., over a short period of time (e.g., several months), or over a long period of time (e.g., over several years) to allow the expression of the transgene to be controlled, thereby enabling one to customize or tailor the gene therapy to the needs of the subject.

A. Increased or dose-dependent controlled expression from the transgene

(i) Controlled transgene expression: sustained transgene expression by re-administration of ceDNA vector

[00170] One aspect of the technology described herein relates to use of a ceDNA vector as disclosed herein in a method for maintaining or sustaining the level of expression of a transgene in a cell, the method comprising: (a) administering to a cell at a first time point, a priming dose of a ceDNA vector to achieve expression of a transgene, and (b) administering to the cell at a second time point, a dose of a ceDNA vector to compensate for any decrease in expression level of the transgene after the ceDNA vector administration at the first time point.

[00171] As shown in **FIG. 6,** the expression level of a transgene expressed by the ceDNA can be achieved and sustained (i.e., maintained) for at least 42 days, or at least 60 days or at least 80 days, and can be increased to the required pre-determined expression level with a re-dose administration at a second or more subsequent time points after an initial priming administration (e.g., at time 0). **FIG. 6** also demonstrates that the expression level of a transgene expressed by the ceDNA can be achieved and sustained for at least 42 days, or at least 60 days, and can be increased in a dose-dependent manner with a re-dose administration comprising a defined amount of the ceDNA vector, at a second or more subsequent time points after an initial priming administration (e.g., at time 0).

[00172] Expression of the transgene typically occurs within about 7-20 days after the re-dose administration, and maintains that level of expression (i.e., is sustained expression) of the transgene from after the re-dose administration for at least about 30 days, or about 60 days, or about 90 days, or about 120 days or longer than 120 days after administration of the second dose (i.e., the re-dose, or administration at the second or subsequent time point).

[00173] Another aspect of the technology described herein relates to a method for treating a disease in a subject, or alternatively, a method for controlled transgene expression in a subject, the method comprising: (a) administering to the subject at a first time point, a priming dose of a composition comprising a ceDNA vector as described herein, to achieve expression of the transgene at a first level (or first amount), and (b) administering to the subject at a second time point, a dose or amount of a ceDNA vector to maintain the expression level of the transgene at a desired sustained expression level, where the sustained expression level is a higher level of expression as compared to the level of transgene expression which achieved if only the initial (i.e.., priming) administration of the ceDNA vector at the first time point was given, thereby treating the disease in the subject.

[00174] Accordingly, in some embodiments a ceDNA vector as disclosed herein can be re-administered (also referred to herein as "re-dosed") to maintain a desired transgene expression level,

where the ceDNA vector comprises two ITR sequences (e.g., a symmetric ITR pair or asymmetric ITR pair as described herein) flanking a transgene polynucleotide sequence operatively linked to a promoter.

[00175] In some embodiments, use of a ceDNA vector in a method for sustaining the level of expression of a transgene in a cell expresses the transgene at a desired expression level for at least 42 days. In some embodiments, the ceDNA vector expresses the transgene at a desired expression level for at least 84 days. In some embodiments, the ceDNA vector expresses the transgene at a desired expression level for at least 132 days. For illustrative purposes only, a ceDNA vector produced by the methods disclosed in Example 1 can sustain the expression level of a transgene in a murine model of CD-1 IGS mice at a higher level from days 7-28 post-administration as compared to similar close-ended DNA vectors produced by other methods and tested in the same mouse model.

[00176] In some embodiments, the increased transgene expression is not dose-dependent, or not entirely dose dependent, but it is sustained expression, in that the increased expression of the transgene occurs typically within 7-20 days after the re-dose administration, and maintains that level of increased expression of the transgene from the re-dose for at least about 30days, or at least about 60 days, or at least about 90 days, or at least about 120 days or longer than about 120 days after administration of the second dose (i.e., the re-dose, or administration at the second or subsequent time point).

[00177] The specific dose response relationship for a given ceDNA vector or composition disclosed herein can be determined by means well known to those of skill in the art, including e.g., as described in Example 6. The ceDNA vector as disclosed herein can be titrated by administration of additional doses of the ceDNA vector at one or more times following the initial administration as required to achieve the desired level of expression. More than one, 2, 3, 4, 5 or 6 or more repeat administrations can be administered in order to titrate the expression levels of a transgene at or near a desired level. Typically, each repeat administration is administered approximately 7-days before a decrease in the expression of the transgene from the previous administration is observed. In effect, the repeated administrations serve to titrate the expression level of the transgene to a desired level, or stated another way, the repeated administrations enable the expression level of the transgene to be maintained within a desired expression level range, e.g., at a desired range which is therapeutic to treat a disease or disorder or within the therapeutic window of the composition.

[00178] In some embodiments, a ceDNA vector described herein can be used to tailor the level of transgene expression *in vivo,* where the level of expression is increased from a prior level (where the prior level is the level after a prior priming or re-dose administration) to a desired expression level, or within a desired expression level range, or above a desired threshold level *in vivo,* where the increased expression level is sustained or maintained between about 30-60 days, or between about 40-70 days, or between about 50-80 days, or between about 60-90 days, or between about 70-100 days or between about 80-110

days, or between about 40-120 days, or longer than 120 days after the re-dose administration of the ceDNA vector.

[00179] In some embodiments, the ceDNA vector used in the methods described herein, e.g., in a method for sustaining expression of a transgene in a cell and/or for treating a subject with a disease, is administered in combination with a pharmaceutically acceptable carrier and/or excipient. In some embodiments, a ceDNA vector is administered at a second time point is administered at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months, or between 6-12 months, or between 1-2 years, or 2-3 years after the first time point.

(ii) Controlled transgene expression: increasing transgene expression by re-administration of ceDNA vector

[00180] In addition to a re-dose administration of a ceDNA vector to simply increase the level of transgene expression if expression levels have decreased over time (e.g., to continue or maintain the transgene expression at a desired pre-determined level), in some embodiments, the methods and compositions of re-administration of a ceDNA vector can increase the level of transgene in a dose-dependent manner - that is, a re-dose administration of a defined amount of a ceDNA vector can effect a defined increase in expression level of a transgene. Stated differently and using arbitrary units for illustrative purposes only, a 1unit dose of the ceDNA in a re-dose administration will achieve a 10% increase in the level of transgene expression from a prior level, and a 2 unit dose of the cDNA vector will achieve a 20% increase in the level of the transgene from a prior level, and a 0.5 unit dose of the ceDNA will achieve a 5% increase in the level of expression of the transgene from a prior level.

[00181] Accordingly, in one embodiment, a ceDNA vector as disclosed herein for controlled transgene expression can be used for increasing the level of expression of a transgene in a cell or a subject in a controlled manner. For example, the expression level of the transgene can be increased with one or more subsequent administrations (e.g., a re-dose or a booster administration) of the ceDNA vector.

[00182] The ceDNA vectors as described herein enable a dose- or concentration dependent re-dose administration of the ceDNA at one or more points after the initial priming administration to increase the expression level of the transgene by a defined amount *in vivo*. In some embodiments, the increase in transgene expression by a defined amount can bring the transgene expression level *in vivo* at, or above a desired threshold (or predetermined level), or within a desired expression level range, where the desired threshold or desired expression level range is above the transgene expression level of the prior administration (i.e., the initial priming administration or a prior re-dose administration).

[00183] **FIG. 6** illustrates that one can readily increase in the expression level of the transgene after the subject is administered a re-dose (i.e., a re-administration or booster) of the ceDNA vector *in vivo*. In

particular, **FIG. 6** shows different increases in the expression level of the transgene after the subject is administered different concentrations of the ceDNA vector in re-dose administrations *in vivo*. In particular, a re-dose concentration of 3mg/ml of ceDNA achieved a 7-fold increase in the expression of the transgene, and a re-dose concentration of 10mg/kg of ceDNA resulted in a 17-fold increase in the expression of the transgene as compared to the expression level without the dose-dependent re-dose administration. Accordingly, the technology described herein relates to at least two administrations of a ceDNA to a subject *in vivo*, where the second or subsequent administrations results in a dose dependent increase in the level of the expression of the transgene by a desired amount, and in some embodiments, to achieve a desired expression level range, or a desired expression level, or to a threshold expression level, as compared to the expression level of the transgene achieved with the prior administration of the ceDNA vector , or without the dose-dependent re-dose administration. That is, an increase in the level of expression of the transgene is achieved by one or more dose-dependent re-dose administration to increase the expression level of the transgene in a controlled manner, that is, to titrate the expression of the transgene based on the dose (or amount) of ceDNA in the re-dose administration.

[00184] Stated differently, the dose-dependent re-dose administration disclosed herein adds to the transgene expression level. The increased transgene expression is dose-dependent, and is a sustainable expression – that is, the expression of the transgene at the higher level (due to the dose-dependent re-dose administration) is maintained for a defined period of time, or does not decrease, or drop below the level of expression observed without the re-dose administration.

[00185] Typically, each dose-dependent re-dose is administered approximately 7-days before the desired increase in the expression of the transgene is desired. In effect, the dose-dependent re-doses serve to titrate the expression level of the transgene to a desired level, or desired expression level range, or stated another way, the dose-dependent re-doses enable the expression level of the transgene to be increased by a defined amount above the prior expression level, and that the increases can be in at least one dose-dependent re-dose administration, or alternatively, in incremental increases with more than one dose-dependent re-dose administration, such that the transgene is increased in a controlled, titrated manner to be expressed at a level that is within a desired expression level range, e.g., at a desired range which is therapeutic to treat a disease or disorder.

[00186] In some embodiments, a desired range of expression level, or desired expression level range of the transgene may be a therapeutically effective amount of transgene to effectively treat or reduce a symptom of a disease. Accordingly, in some embodiments, to achieve such a therapeutically effective amount of transgene, one can increase the level of expression of the transgene using one or more re-dose administrations as described herein, to incrementally increase the levels to the therapeutically effective amount of the transgene. Therefore, in some embodiments, a subject can be administered a priming dose

of ceDNA vector that expresses the transgene at a low expression level (i.e., a sub-therapeutically effective amount), and one or more re-dose administrations can be administered to the subject over a period of time to increase the expression level until a desired therapeutic effect is achieved. Such a strategy allows the subject's body to adjust to the level of the expressed transgene, and effectively allows titration or adjusting (in this case, increasing) the level of the expression of the transgene in increments to reach a desired therapeutic goal or effect and/or prevent over medication and/or side effects due to over expression of the transgene. Alternatively, in some embodiments, a subject can be administered a priming dose of ceDNA vector that expresses the transgene at a desired expression level when the subject is an infant or child, typically a low expression level, and one or more dose-dependent re-dose administrations can be administered to the subject over a period of time as the subject grows to increase the expression level so that the therapeutic effect is maintained.

B. Timing and amount of ceDNA vector in re-administration

[00187] The level of expression of a transgene from a ceDNA vector as disclosed herein can be increased from a prior level (i.e.., expression level achieved from a prior priming administration at day 0, or a prior re-dose) by re-administration (i.e., re-dose) of the ceDNA vector at one or more times following the initial administration. Typically, the re-dose administration to increase the level of expression to a desired level or a desired expression level range is administered about 7 days, or more than 7 days before the increase in expression is desired. As an illustrative example, if an increase in the level of transgene expression is desired at about 30-days after the prior administration (i.e., a priming or prior re-dose administration), then a re-dose can be given at 28 days or earlier. Similarly, if an increase in the transgene expression is desired at about 90 days (or about 3-months) after a prior administration (i.e., a priming or prior re-dose administration), then a re-dose can be administered at about 83 or 84 days or earlier in order to increase the level of the transgene expression at or around 90 days to a desired level or within a desired expression level range, where the desired level or desired expression level range is above the transgene expression level achieved with the prior administration.

[00188] As discussed herein, as the methods and ceDNA vectors described herein allow a personalized genetic medicine approach, i.e., titrating the level of the transgene expression in a step-by-step manner with re-dose administrations to increase the expression levels incrementally, it is envisioned that 1, 2, 3, 4, 5 or 6 or more than 6 re-doses can be administered over time in order to increase the level of expression the transgene to a desired level, or to a desired expression level range, which is higher than the expression level achieved with the prior administration, or prior to this re-dose administration. The incremental increases in expression level of the transgene by each re-dose administration can be the same, i.e., each re-dose administration can increase the expression level of the transgene by about 10% from the

prior expression level, or can be different, i.e., a first re-dose administration can increase the expression about 10% from the prior expression level, and a second re-dose administration can increase the expression about 20% from the prior expression level. Typically, each re-dose is administered approximately 7-days before the desired increase in the expression of the transgene is desired. In effect, the re-doses serve to titrate the expression level of the transgene to a desired level or desired expression level range, or stated another way, the re-doses enable the expression level of the transgene to be increased above the prior expression level, and that the increases can be incremental with one or more re-dose administration such that the transgene is expressed at a level that is within a desired expression level range, e.g., at a desired range which is therapeutic to treat a disease or disorder.

[00189] In some embodiments, a re-dose to increase the transgene expression level is at least about 20 days, or at least about 30 days, or at least about 40 days, or at least about 50 days, or at least about 60 days, or between about 60-90 days, or between about 90-120 days, or between about 120-150 days after a prior administration (e.g., an initial priming administration at day 0, or a prior re-dose administration) of the ceDNA composition.

[00190] In some embodiments, the technology described herein relates to a re-dose of ceDNA for increasing transgene expression *in vivo*, where expression of the transgene can be increased by one or more re-doses (i.e., re-administration or booster administrations) of the ceDNA composition. In some embodiments, the dose (or amount) of ceDNA vector in the re-dose administration at a second or subsequent time point is the same, or a different amount to the dose (i.e., amount) of ceDNA in the administration prior to, or proceeding this re-dose administration (an initial priming administration at day 0, or a prior re-dose administration). As one example, if the amount of ceDNA administered at the first time point is 1mg/kg, the amount in the re-dose administration at a second or subsequent time point can be 1mg/kg, or less than 1mg/kg or more than 1mg/kg. In some embodiments, the re-dose can be at an amount selected from any of: about 2mg/kg, about 3mg/kg, about 4mg/kg, about 5mg/kg, about 6mg/kg, about 7mg/kg, about 8mg/kg, about 9mg/kg, about 10mg/kg, or between about 2-5mg/kg, or between 5-10mg/kg, or between 10-15mg/kg or greater than 15mg/kg.

[00191] In some embodiments, to increase the expression level of the transgene by one or more re-dose administrations (i.e., sequential re-dose administrations to incrementally increase the transgene expression level), each re-dose administration can be the same, i.e., each re-dose administration can increase the expression level of the transgene by about 10% from the prior expression level, or can be different, i.e., a first re-dose administration can increase the expression about 10% from the prior expression level, and a second re-dose administration can increase the expression about 20% from the prior expression level.

[00192] In some embodiments, where more than one re-dose is administered, the amount of increase in the expression level of the transgene by each re-dose administration can be the same, i.e., each re-dose administered can increase the expression level of the transgene by about 10% from the prior expression level, or can be different, i.e., a first re-dose administration can increase the expression about 10% from the prior expression level, and a second re-dose administration can increase the expression about 20% from the prior expression level. In some embodiments, where more than one re-dose is administered, the amount of increase in the expression level of the transgene by each re-dose administration can be the same, i.e., each re-dose administered can increase the expression level of the transgene by about 1-fold, or 2-fold, or 3-fold etc. from the prior expression level, or can be different, i.e., a first re-dose administration can increase the expression about 2-fold from the prior expression level, and a second re-dose administration can increase the expression about 6-fold from the prior expression level, or about 6-fold from the expression level achieved from the initial priming administration.

[00193] In some embodiments, the ceDNA vector is the same ceDNA vector administered at the prime administration (i.e., first administration at time 0) as that administered to the cell or subject at a second or any subsequent administration (e.g., re-dose administrations). In some embodiments, the ceDNA vector can be the same ceDNA vector. For illustrative purposes only, re-administration of viral vectors, e.g., AAV vectors usually are a different serotype to that administered previously. In contrast, a ceDNA vector in a re-administration is the same as the ceDNA vector administered previously, - that is, the ceDNA vector has not changed such that it is equivalent to administering the same serotype of AAV multiple times.

[00194] That being said, while it is the equivalent of the same ceDNA vector serotype being re-administered, in some embodiments, the ceDNA vector administered at a second or any subsequent administration (e.g., re-dose administrations) after the initial prime administration is different, such, as, e.g., different ITR-pair, a different promoter operatively linked to the transgene, a different transgene or modified transgene or the like. In some embodiments, the transgene gene is the same, or can be a modified transgene.

[00195] In some embodiments, the intervals between the first administration (i.e., priming administration) of a ceDNA vector and a re-dose administration (i.e., at a second time point, or any subsequent time point after that, e.g., 3rd, 4th, 5th, 6th, 7th, 8th, 9th, 10th etc.) can be at least 30 days, or at least 60 days, or at least 80 days, or between 60-90 days, or between 90-120 days, or between about 2-3 months, or between about 3-6 months, or between about 6-12 months, or at about 1 year or between 1-2 years, or at about 2 years or between 2-3 years, or at about 3 years or between 3-4 years, or at about 5 years or between 5-6 years, or between 5-10 years, or between 10-20 years etc. after the prior administration (i.e., the priming dose or a prior-re-dose administration). As discussed herein, in one

embodiment a re-dose administration to increase the level of expression of a transgene is administered about 7 days, or more than 7 days (e.g., between 8-10, or 14 days) before an increase in expression is desired.

[00196] For illustration purposes only, if the amount of the ceDNA vector in the initial administration at day 0 is set at an arbitrary unit of 1, then the amount of ceDNA vector in a re-dose administration, at a second, third, fourth, fifth, sixth time point can be selected from any of 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, or between 15-20-fold, or between 20-50-fold greater (or more) than the amount of ceDNA vector in the initial administration at day 0.

[00197] In some embodiments, if the amount of the ceDNA vector in the initial administration at day 0 results in a level of expression that is given an arbitrary unit of 1-fold, then the amount of ceDNA vector in a re-dose administration, at a second, third, fourth, fifth, sixth time point can be an amount of a ceDNA vector that results in an increase in the level of expression of the transgene by at least 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, or between 15-20-fold, or between 20-50-fold greater (or more) as compared to the level of transgene expression from the initial administration of the ceDNA vector at day 0.

[00198] In some embodiments, a re-dose is administered in the same manner, or same route of administration as the initial administration of the ceDNA at day 0. In some embodiments, a re-dose is administered in a different manner or by a different route of administration as the initial administration of the ceDNA at day 0. In some embodiments, when the initial administration (i.e., prime administration at time 0) is followed by one or more re-doses (i.e., booster administrations), the administration can be by intravenous administration, intranasal or intramuscular administration – or any other medically appropriate route of administration of the composition comprising ceDNA vector.

[00199] A ceDNA vector as disclosed herein can be administered to a subject at a first time point (e.g., the initial administration, e.g. at day 0), and at a time after the first time point if necessary or desired. A ceDNA vector as disclosed herein can be administered at a second time point to titrate the levels of the transgene at a desired level (e.g., above a threshold value for efficacy) or within a desired expression level range (e.g., within the therapeutic window of the composition). It is envisioned that more than one dose of ceDNA vector as disclosed herein, can be administered to the subject, e.g., a repeated administration can be given at any one or more of: a second time point, a third time point, a fourth time point, etc. It is encompassed that additional doses can be administered to maintain the desired level of transgene expression (i.e., to maintain or sustain the same level of transgene expression). The intervals between the first and second or any two successive re-doses do not need to be the same.

C. Pre-determined transgene expression levels

[00200] In some embodiments, a predetermined transgene expression level (also referred to as desired range of expression level, or desired expression level range of the transgene) may be a therapeutically effective amount of transgene to effectively treat or reduce a symptom of a disease. Accordingly, in some embodiments, to achieve such a therapeutically effective amount of transgene, one can maintain the level of transgene expression and/or increase the level of expression of the transgene, as disclosed herein, using one or more re-dose administrations as described herein, to incrementally increase the levels to the therapeutically effective amount of the transgene. Therefore, in some embodiments, a subject can be administered a priming dose of ceDNA vector that expresses the transgene at a desired expression level, typically a low expression level (i.e., a sub-therapeutically effective amount), and one or more dose-dependent re-dose administrations can be administered to the subject over a period of time to increase the expression level until a desired therapeutic effect is achieved. Such a strategy allows the subject's body to adjust to the level of the expressed transgene, and effectively allows titration or adjusting (in this case, increasing) the level of the expression of the transgene in at least one dose-dependent re-dose administration or more (i.e., in at least two or more increments) to reach a desired therapeutic goal or effect and/or prevent over medication and/or side effects due to over expression of the transgene. Alternatively, in some embodiments, a subject can be administered a priming dose of ceDNA vector that expresses the transgene at a desired expression level when the subject is an infant or child, typically a low expression level, and one or more dose-dependent re-dose administrations can be administered to the subject over a period of time as the subject grows to increase the expression level so that the therapeutic effect is maintained. Similarly, in some embodiments, a subject can be administered a priming dose of ceDNA vector that expresses the transgene at a desired expression level in the subject, and one or more dose-dependent re-dose administrations can be administered to the subject over a period of time as the subject gains weight etc. to increase the expression level so that the therapeutic effect is maintained.

[00201] That is, a ceDNA vector as disclosed herein is administered to a subject at a first time point (e.g., the initial administration, e.g. at day 0), and at time after the first time point, a ceDNA vector as disclosed herein is administered at a second time point to increase the level of expression of the transgene to a predetermined transgene expression level (e.g., to a desired level or within a desired expression level range), where the predetermined transgene expression level is above the level of the expression of transgene prior to the re-dose administration. In some embodiments, a predetermined transgene expression level is not necessarily the therapeutically effective amount of the transgene, it is envisioned that more than one dose-dependent re-doses of a ceDNA vector as disclosed herein, can be administered to the subject at any one or more of: a second time point, a third time point, a fourth time point, etc. where the transgene is increased to a predetermined transgene expression level. It is encompassed that dose-

dependent re-doses can be administered to increase the level of expression of the transgene by a defined amount which is dependent on the dose of the ceDNA in the re-dose administration, and in some embodiments, can be used to increase the expression levels in a step-by-step manner to achieve an expression level that is a therapeutically effective amount of the transgene (e.g., is a level that produces a desired therapeutic effect or reduces one or more symptoms of the disease or disorder).

[00202] While the re-dose administration results in a dose-dependent increase in expression of the transgene, in some embodiments the effect of the re-dose administration is synergistic, - that is, the increase in transgene expression is greater than the sum of the single administrations. For example, as demonstrated in **FIG. 6**, a 3-fold increase in amount of ceDNA vector in a re-dose administration resulted in an increase in transgene expression greater than 3-fold; and a 10-fold increase in the amount of ceDNA vector in a re-dose administration resulted in a greater than 10-fold increase in transgene expression.

[00203] In particular embodiments described herein, it may be desirable to select a dose low on the dose-response curve of the ceDNA for the initial administration (i.e., at a first time point, i.e., day 0) and optionally to increase the dose (i.e., level of transgene expression) in increments with one or more dose-dependent re-dose administrations at a second, third, fourth etc., administrations to induce a therapeutic effect while also preventing onset of untoward side effects or intolerance to the composition. In one embodiment, the dose response relationship for a given ceDNA vector is used to determine and/or estimate an optimal dose for effective treatment of a given disease that is well within the bounds of the therapeutic window of the composition. That is, titration of the dose of ceDNA vectors as described herein using an initial priming administration (e.g., at day 0) and incremental dose-dependent increase in re-dose administrations at subsequent time points can be achieved maximize the therapeutic effect of the expressed transgene while also minimizing side effects or unwanted toxicity.

[00204] *In vivo* and/or *in vitro* assays can optionally be employed to help identify optimal dosage ranges of the ceDNA vector in re-administrations to achieve a predetermined transgene expression level. The precise dose of ceDNA vector in the initial priming administration (e.g., at time 0) and each re-administration thereafter 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.

[00205] A ceDNA vector for controlled transgene expression is administered in sufficient amounts to transfect the cells of a desired tissue and to provide sufficient levels of transgene 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.

[00206] The dose of the amount of a ceDNA vector in the initial priming administration (e.g., at time 0) and each re-administration thereafter for controlled transgene expression 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.

[00207] 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.

[00208] 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.

[00209] 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.

[00210] For *in vitro* transfection, an effective amount of a ceDNA vector to be delivered to cells ($1x10^6$ 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.

[00211] 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 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.

[00212] In particular embodiments, more than one administration (e.g., two, three, four or more administrations) may be employed to achieve the desired level of gene expression over a period of various intervals, e.g., daily, weekly, monthly, yearly, etc.

[00213] In some embodiments, a transgene encoded by a ceDNA vector for controlled transgene expression 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, controlled expression from a ceDNA vector 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 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 the transgene into safe harbor regions, such as, but not including albumin gene or CCR5 gene.

[00214] 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.

II. Personalized gene therapy

[00215] As discussed herein, the methods and ceDNA vectors as described herein allow a personalized genetic medicine approach, i.e., dose-dependent titration of the level of the transgene expression using re-doses, e.g., in a step-by-step manner with one or more re-dose administrations, in order to increase the transgene expression levels by a certain amount with each re-dose administration. As

such, the re-doses can be used to titrate the expression level of the transgene in increments. Accordingly, in some embodiments 1, 2, 3, 4, 5 or 6 or more than 6 dose-dependent re-doses can be administered in order to either maintain and/or or increase the level of expression the transgene by a defined amount each time (i.e., each re-dose), to increase the expression level to a desired level, or to a desired expression level range, which is higher than the expression level achieved with the prior administration, or prior to this re-dose administration.

[00216] The ability of a skilled artisan to titrate the dose of a composition comprising a ceDNA vectors as described herein based on the dose-response relationship for the vector is beneficial for the treatment of disease in a variety of ways. At a minimum, a skilled artisan can increase the dose of the ceDNA vector when an increase in the effect is desired (e.g., expression of a transgene). Alternatively, a skilled artisan can select a dose that is known to achieve a level of expression that is therapeutic based on prior knowledge of the dose-response relationship for the ceDNA composition. In some embodiments, it may be desirable to select a dose relatively high on the dose-response curve, for example, to accelerate treatment for a subject having severe symptoms of a given disease. In contrast, it is generally desirable to select a dose low on the dose-response curve and optionally to increase the dose in increments to induce a therapeutic effect while also preventing onset of untoward side effects or intolerance to the composition. In one embodiment, the dose response relationship for a given ceDNA vector is used to determine and/or estimate an optimal dose for effective treatment of a given disease that is well within the bounds of the therapeutic window of the composition. That is, titration of the dose of ceDNA vectors as described herein maximize the therapeutic effect of the expressed transgene while also minimizing side effects or unwanted toxicity.

[00217] As an illustrative example only, subjects with cystic fibrosis can have differing severity of disease, and/or respond differently to the same level of transgene expression of the CFTR1 gene, and/or have a lower or higher than normal drug clearance, and thus, by increasing the expression of the CFTR1 transgene in a step-by-step manner with one or more re-dose administrations of a ceDNA vector comprising a CFTR1 transgene allows the level of expression of the CFTR1 transgene to be increased in incremental steps, i.e., to a level of expression that is effective at reducing one or more symptoms of the cystic fibrosis disease in that particular subject. Previously, such a personalized approach, or titration method to increase the expression level of a transgene was either not effective, and/or not possible due to the immune responses associated with other viral-based vectors such as AAV vectors.

[00218] In some embodiments, a dose-dependent re-dose administration allows for a controlled increase in the level of transgene expression, and therefore, the methods and compositions as disclosed herein allows a personalized medicine approach to gene therapy. As an illustrative example only, subjects with cystic fibrosis can have different severity of disease, and/or respond differently to the same level of

transgene expression of the CFTR1 gene, and/or have a lower or higher than normal drug clearance, and thus, a controlled increase of the expression of the CFTR1 transgene by one or more dose-dependent re-dose administrations of a ceDNA vector comprising a CFTR1 transgene allows the level of expression of the CFTR1 transgene to be increased in a controlled manner, and in some embodiments, the controlled increased in expression can be in increased to a level of expression that is effective at reducing one or more symptoms of the cystic fibrosis disease in that particular subject. Previously, such a personalized approach, or titration method to dose-dependently increase the expression level of a transgene was either not effective, and/or not possible due to the immune responses associated with other viral- or AAV-based vectors.

[00219] As discussed herein, in some embodiments, a subject is assessed at a predetermined time after a first administration of a ceDNA vector, e.g., at any timepoint at least 30 days, or at least 60 days, or between 60-90 days or longer than 90 days after the first administration of a ceDNA vector to determine the titrating dose. For example, in some embodiments, the subject is assessed to determine the disease state of the subject after a first administration of a ceDNA vector and/or the level of transgene expressed by the ceDNA vector in the subject.

[00220] In some embodiments, assessment of the disease state is an assessment of at least one symptom of the disease in the subject. The disease state for any given disease can be determined by a physician or person of skill in the art, and includes assessing one or more clinical symptoms and/or biomarkers of the disease, including protein biomarkers, miRNA and mRNA biomarkers and other molecular profiling systems.

[00221] In some embodiments, assessment of the disease state in a subject can be determined using molecular profiling in combination with clinical characterization of a patient such as observations made by a physician (such as a code from the International Classification of Diseases, for example, and the dates such codes were determined), laboratory test results, x-rays, biopsy results, statements made by the patient, and any other medical information typically relied upon by a physician to make a diagnosis in a specific disease. Methods to determine a disease state based on molecular profiling in a subject are disclosed in Patents and US Patent Applications, 7,167,734, 9,372,193, 9,383,365, 2006/0224191, 2011/0172501, 2009/0104596, 2009/0023149 which are each incorporated herein in their entireties.

[00222] In some embodiments, the methods described herein can be used to titrate the ceDNA vector to the subject for individualized medical intervention for a particular disease state.

[00223] Examining the changes of a variety of biomarkers provides information about the status of the subject from which the biomarkers are obtained. Understanding how biomarkers change (e.g., increase, decrease, no change) with disease progression so that by measuring a single biological sample at

a single point in time permits verification (e.g., disease or no disease), disease typing, and characterization of a disease state (e.g., early or "onset" versus late or "recovery" phase).

[00224] In some embodiments, biomarkers to assess the state of disease progression, such as onset or recovery, based on the level of each of the biomarkers as well as their trend (increase, decrease or constant) with time can be assessed. Profiling biomarkers for a subject disease state can also be combined with other techniques, such as stable isotope ratios naturally occurring in breath (e.g., U.S. Pat. No. 5,912,178), for assessing whether an individual is healthy or in a disease state. Disease states are detected by measuring changes in biomarker levels, and particularly, a plurality of biomarkers interrelated within a biological pathway associated with the disease state. In some embodiments, a particular disease state can be characterized by detecting and analyzing complex signals from NMR spectra to determine biomarkers whose levels are changing as the disease progresses. This initial disease state assessment allows for "fingerprinting" the dynamic changes associated with disease progression and assists in assessing current status of the disease progression and process. When the disease state is identified, the administration of the ceDNA vector can be tailored and/or titrated to the subject so as to reduce the disease time course.

[00225] In some embodiments, one can assess a disease state by assessing the biomarker profile within a biological sample obtained from a subject. The specific biomarkers that are measured are determined from an analysis of the key biochemical pathways underlying the disease and the associated host immune response. In an embodiment, a standard biomarker profile is obtained from a healthy individual and from an individual with the disease. Comparing the biomarker profile from the biological sample to the standard biomarker profile (healthy and disease) permits a disease state to be positively identified. Optionally, a second biological sample is isolated from the patient at a second time point or disease progression time point to obtain a biomarker profile trend (e.g., which biomarkers are changing between the first and second samples), thereby providing further information about the disease status or state of the patient. In some embodiments, a standard biomarker profile is assessed at one or more of the following times; before the first administration of the ceDNA vector, after at least 30 days, or at least 60 days, or between 60-90 days or longer than 90 days after the first administration of a ceDNA vector, or after at least 30 days, or at least 60 days, or between 60-90 days or longer than 90 days after the second (e.g., redose) administration of a ceDNA vector, or after at least 30 days, or at least 60 days, or between 60-90 days or longer than 90 days after any subsequent administrations (e.g., redose administrations) of a ceDNA vector.

[00226] Protein biomarkers have been identified for diabetes, Alzheimer's Disease, and cancer. (See, for Example, U.S. Pat. Nos. 7,125,663; 7,097,989; 7,074,576; and 6,925,389, which are incorporated herein in their entirety). Methods for detection of protein biomarkers, such as mass spectrometry and specific binding to antibodies, can also be used. High throughput expression analysis methods using

microarrays can be used for mRNA biomarkers, as well as focused arrays and qPCR for multiple relevant genes to identify stress related genes see. e.g., WO2007106685A2. DNA microarrays have been used to measure gene expression in tumor samples from patients and to facilitate diagnosis. Gene expression can reveal the presence of cancer in a patient, its type, stage, and origin, and whether genetic mutations are involved. Gene expression may even have a role in predicting the efficacy of chemotherapy. Over recent decades, the National Cancer Institute (NCI) has tested compounds, including chemotherapy agents, for their effect in limiting the growth of 60 human cancer cell lines. The NCI has also measured gene expression in these 60 cancer cell lines using DNA microarrays. Various studies have explored the relationship between gene expression and compound effect using the NCI datasets. Critical time is often lost due to a trial and error approach to finding an effective chemotherapy for patients with cancer. In addition, cancer cells often develop resistance to a previously effective therapy. In such situations, patient outcome could be greatly improved by early detection of such resistance.

[0001] In some embodiments, the level of biomarker expression of a disease state is determined by measuring the level of mRNA transcribed from the gene(s), by detecting the level of a protein product of the gene(s), or by detecting the level of the biological activity of a protein product of the gene(s). In some embodiments, the level of a biomarker (including miRNA biomarkers) of a disease state is measured using a quantitative reverse transcription-polymerase chain reaction (qRT-PCR). Such methods to measure gene expression products, *e.g.,* protein level, include ELISA (enzyme linked immunosorbent assay), western blot, and immunoprecipitation, immunofluorescence using detection reagents such as an antibody or protein binding agents. Alternatively, a peptide can be detected in a subject by introducing into a subject a labeled anti-peptide antibody and other types of detection agent. For example, the antibody can be labeled with a radioactive marker whose presence and location in the subject is detected by standard imaging techniques.

[0002] In certain embodiments, the gene expression products can be determined by measuring the level of messenger RNA (mRNA) expression of a disease biomarker. Such molecules can be isolated, derived, or amplified from a biological sample, such as a whole blood or plasma, e.g., platelet rich plasma. Detection of mRNA expression is known by persons skilled in the art, and comprise, for example but not limited to, PCR procedures, RT-PCR, Northern blot analysis, differential gene expression, RNA protection assay, microarray analysis, hybridization methods etc. In some embodiments, the level of the mRNAs can be measured using quantitative RT-PCR. In general, the PCR procedure describes a method of gene amplification which is comprised of (i) sequence-specific hybridization of primers to specific genes or sequences within a nucleic acid sample or library, (ii) subsequent amplification involving multiple rounds of annealing, elongation, and denaturation using a thermostable DNA polymerase, and (iii) screening the PCR products for a band of the correct size. The primers used are oligonucleotides of

sufficient length and appropriate sequence to provide initiation of polymerization, *i.e.* each primer is specifically designed to be complementary to a strand of the genomic locus to be amplified. In an alternative embodiment, mRNA level of gene expression products described herein can be determined by reverse-transcription (RT) PCR and by quantitative RT-PCR (QRT-PCR) or real-time PCR methods. Methods of RT-PCR and QRT-PCR are well known in the art.

[00227] In some embodiments, methods to measure one or more biomarkers for a disease state, or the level of the expression of the transgene from the ceDNA vector can an assay selected from any of: immunohistochemical (IHC) analysis and/or a micro array analysis, a comparative genomic hybridization (CGH) micro array, a single nucleotide polymorphism (SNP) micro array, a fluorescent in-situ hybridization (ISH), an in-situ hybridization (ISH), and a proteomic array.

[00228] The term "Microarray" as used herein means a device employed by any method that quantifies one or more subject oligonucleotides, e.g., DNA or RNA, or analogues thereof, at a time. One exemplary class of microarrays consists of DNA probes attached to a glass or quartz surface. Many microarrays, e.g., those made by Affymetrix, use several probes for determining the expression of a single gene. The DNA microarray can contain oligonucleotide probes that may be, e.g., full-length cDNAs complementary to an RNA or cDNA fragments that hybridize to part of an RNA. Exemplary RNAs include mRNA, miRNA, and miRNA precursors. Exemplary microarrays also include a "nucleic acid microarray" having a substrate-bound plurality of nucleic acids, hybridization to each of the plurality of bound nucleic acids being separately detectable. The substrate can be solid or porous, planar or non-planar, unitary or distributed. Exemplary nucleic acid microarrays include all of the devices so called in Schena (ed.), DNA Microarrays: A Practical Approach (Practical Approach Series), Oxford University Press (1999); Nature Genet. 21(1)(suppl.):1-60 (1999); and Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000). Additionally, exemplary nucleic acid microarrays can include a substrate-bound plurality of nucleic acids in which the plurality of nucleic acids is disposed on a plurality of beads, rather than on a unitary planar substrate, as is described, inter alia, in Brenner et al., Proc. Natl. Acad. Sci. USA 97(4):1665-1670 (2000). Examples of nucleic acid microarrays may be found in U.S. Pat. Nos. 6,391,623, 6,383,754, 6,383,749, 6,380,377, 6,379,897, 6,376,191, 6,372,431, 6,351,712 6,344,316, 6,316,193, 6,312,906, 6,309,828, 6,309,824, 6,306,643, 6,300,063, 6,287,850, 6,284,497, 6,284,465, 6,280,954, 6,262,216, 6,251,601, 6,245,518, 6,263,287, 6,251,601, 6,238,866, 6,228,575, 6,214,587, 6,203,989, 6,171,797, 6,103,474, 6,083,726, 6,054,274, 6,040,138, 6,083,726, 6,004,755, 6,001,309, 5,958,342, 5,952,180, 5,936,731, 5,843,655, 5,814,454, 5,837,196, 5,436,327, 5,412,087, and 5,405,783, herein incorporated by reference.

[00229] Exemplary microarrays can also include "peptide microarrays" or "protein microarrays" having a substrate-bound plurality of polypeptides, the binding of a oligonucleotide, a peptide, or a

protein to the plurality of bound polypeptides being separately detectable. Alternatively, the peptide microarray, can have a plurality of binders, including, but not limited to, monoclonal antibodies, polyclonal antibodies, phage display binders, yeast 2 hybrid binders, aptamers, that can specifically detect the binding of specific oligonucleotides, peptides, or proteins. Examples of peptide arrays may be found in International Patent Publication Nos. WO 02/31463, WO 02/25288, WO 01/94946, WO 01/88162, WO 01/68671, WO 01/57259, WO 00/61806, WO 00/54046, WO 00/47774, WO 99/40434, WO 99/39210, and WO 97/42507, and in U.S. Pat. Nos. 6,268,210, 5,766,960, and 5,143,854, herein incorporated by reference.

[00230] In some embodiments, if the disease state of the subject has remained at a steady state, or has not improved, or where the disease state has declined in the subject, for example, as compared to the disease state at the time of the first administration of the ceDNA vector or any time before administration of the ceDNA vector, the subject is administered a second dose of the ceDNA vector, e.g., wherein in some embodiments, the amount of ceDNA vector administered is a titrated dose.

[00231] In alternative embodiments, if the level of transgene expression in the subject has declined from a predetermined level or declined from a therapeutically effective amount, e.g., dropped from the initial transgene expression level after the first administration of the ceDNA vector, the subject is administered a second dose of the ceDNA vector e.g., wherein in some embodiments, the amount of ceDNA vector administered is a titrated dose. In some embodiments, the level of the transgene expression is determined by measuring the level of the transgene (e.g., measuring protein level and/or mRNA levels) expressed from the ceDNA vector in a biological sample obtained from the subject. In some embodiments, the biological sample is selected from any of: blood, plasma, synovial fluid, CSF, saliva, or tissue biopsy sample.

[00232] In some embodiments, where the ceDNA vector expresses a reporter protein in addition to a transgene encoding a desired protein or therapeutic gene, the level of the transgene can be determined by measuring the level of reporter protein expressed from the ceDNA vector *in vivo,* using methods commonly known to persons of ordinary skill in the art. In some embodiments, the titrating the ceDNA vector is determining the level of transgene expressed from the ceDNA vector and administering a second dose of the ceDNA vector to the subject to adjust or modulate the transgene expression to a predetermined desired level.

III. ceDNA vector in general

[00233] Embodiments of the invention are based on methods and compositions comprising close ended linear duplexed (ceDNA) vectors that can express a transgene, as defined herein. The ceDNA vectors 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 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. In some embodiments, a ceDNA vector as disclosed herein is translocated to the nucleus where expression of the transgene in the ceDNA vector, e.g., genetic medicine transgene can occur. In some embodiments, a ceDNA vector as disclosed herein translocated to the nucleus where expression of the transgene, e.g., genetic medicine transgene located between the two ITRs can occur.

[00234] In general, a ceDNA vector 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.

[00235] Encompassed herein are methods and compositions comprising the ceDNA vector, which may further include a delivery system, such as but not limited to, a liposome nanoparticle delivery system. Nonlimiting 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, which is incorporated herein.

[00236] The ceDNA vectors 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.

[00237] **FIG. 1A-1E** show schematics of nonlimiting, exemplary ceDNA vectors, or the corresponding sequence of ceDNA plasmids. ceDNA vectors 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).

[00238] The expression cassette can also comprise an internal ribosome entry site (IRES) (e.g., SEQ ID NO: 190) 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. 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 the transgene, 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.

[00239] 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. In some embodiments, the ceDNA vector is devoid of prokaryote-specific methylation.

[00240] 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.

[00241] The expression cassette can comprise any transgene useful for treating a disease or disorder in a subject. A ceDNA vector can be used to deliver and express any gene of interest in the subject, which includes but are not limited to, 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, antigen binding fragments, or any combination thereof.

[00242] 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.

[00243] Sequences provided in the expression cassette, expression construct of a ceDNA vector 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.

[00244] In some embodiments, a transgene expressed by the ceDNA vector for controlled expression as disclosed herein is a therapeutic gene. In some embodiments, a therapeutic gene is an antibody, or antibody fragment, or antigen-binding fragment thereof, or a fusion protein. In some embodiments, the antibody or fusion protein thereof is an activating antibody or a neutralizing antibody or antibody fragment and the like. In some embodiments, a ceDNA vector for controlled gene expression comprises an antibody or fusion protein as disclosed in International patent PCT/US19/18016, filed on February 14, 2019, which is incorporated herein in its entirity by reference.

[00245] In particular, a therapeutic gene is one or more therapeutic agent(s), including, but not limited to, for example, protein(s), polypeptide(s), peptide(s), enzyme(s), antibodies, antigen binding fragments, as well as variants, and/or active fragments thereof, for use in the treatment, prophylaxis, and/or

amelioration of one or more symptoms of a disease, dysfunction, injury, and/or disorder. Exemplary therapeutic genes are described herein in the section entitled "Method of Treatment".

[00246] There are many structural features of ceDNA vectors 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 nonlimiting 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.

[00247] ceDNA vectors 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.

[00248] There are several advantages of using a ceDNA vector 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.

IV. ITRs

[00249] As disclosed herein, ceDNA vectors for controlled transgene expressioncontain 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.

[00250] 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).

[00251] 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.

[00252] 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

[00253] In some embodiments, a ceDNA vector 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.

[00254] *(i) Wildtype ITRs*

[00255] 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.

[00256] Accordingly, as disclosed herein, ceDNA vectors for controlled transgene expressioncontain 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).

[00257] In one aspect, ceDNA vectors for controlled transgene expressionare 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.

[00258] 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.

[00259] In some embodiments, one aspect of the technology described herein relates to a ceDNA vector, wherein the ceDNA vector comprises at least one heterologous nucleotide sequence, 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.

[00260] 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.

[00261] Exemplary WT-ITR sequences for use in the ceDNA vectors for controlled transgene expressioncomprising WT-ITRs are shown in **Table 2** herein, which shows pairs of WT-ITRs (5' WT-ITR and the 3' WT-ITR).

[00262] As an exemplary example, the present disclosure provides a ceDNA vector 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.

[00263] 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.

[00264] 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 sire (trs).

[00265] By way of example only, **Table 1** indicates exemplary combinations of WT-ITRs.

[00266] **Table 1:** 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 (AAAV), bovine AAV (BAAV), canine, equine, and ovine AAV), ITRs from B19 parvoviris (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).

[00267] **Table 1:**

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,AVVDJ8	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,PORCINE
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				

[00268] By way of example only, **Table 2** shows the sequences of exemplary WT-ITRs from some different AAV serotypes.

[00269] **TABLE 2**

AAV serotype	5' WT-ITR (LEFT)	3' WT-ITR (RIGHT)

AAV1	5'- TTGCCCACTCCCTCTCTGCGCGCTCGC TCGCTCGGTGGGGCCTGCGGACCAAA GGTCCGCAGACGGCAGAGGTCTCCTC TGCCGGCCCCACCGAGCGAGCGACGC GCGCAGAGAGGGAGTGGGCAACTCCA TCACTAGGGTAA-3' (SEQ ID NO: 5)	5'- TTACCCTAGTGATGGAGTTGCCCACTC CCTCTCTGCGCGCGTCGCTCGCTCGGT GGGGCCGGCAGAGGAGACCTCTGCCG TCTGCGGACCTTTGGTCCGCAGGCCCC ACCGAGCGAGCGAGCGCGCAGAGAGG GAGTGGGCAA-3' (SEQ ID NO: 10)
AAV2	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCC CGGGCGTCGGGCGACCTTTGGTCGCC CGGCCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTGGCCAACTCCATCAC TAGGGGTTCCT (SEQ ID NO: 2)	AGGAACCCCTAGTGATGGAGTTGGCCA CTCCCTCTCTGCGCGCTCGCTCGCTCAC TGAGGCCGGGCGACCAAAGGTCGCCC GACGCCCGGGCTTTGCCCGGGCGGCCT CAGTGAGCGAGCGAGCGCGCAGCTGC CTGCAGG (SEQ ID NO: 1)
AAV3	5'- TTGGCCACTCCCTCTATGCGCACTCGC TCGCTCGGTGGGGCCTGGCGACCAAA GGTCGCCAGACGGACGTGGGTTTCCA CGTCCGGCCCCACCGAGCGAGCGAGT GCGCATAGAGGGAGTGGCCAACTCCA TCACTAGAGGTAT-3' (SEQ ID NO: 6)	5'- ATACCTCTAGTGATGGAGTTGGCCACT CCCTCTATGCGCACTCGCTCGCTCGGT GGGGCCGGACGTGGAAACCCACGTCC GTCTGGCGACCTTTGGTCGCCAGGCCC CACCGAGCGAGCGAGTGCGCATAGAG GGAGTGGCCAA-3' (SEQ ID NO: 11)
AAV4	5'- TTGGCCACTCCCTCTATGCGCGCTCGC TCACTCACTCGGCCCTGGAGACCAAA GGTCTCCAGACTGCCGGCCTCTGGCC GGCAGGGCCGAGTGAGTGAGCGAGC GCGCATAGAGGGAGTGGCCAACT-3' (SEQ ID NO: 7)	5'- AGTTGGCCACATTAGCTATGCGCGCTC GCTCACTCACTCGGCCCTGGAGACCAA AGGTCTCCAGACTGCCGGCCTCTGGCC GGCAGGGCCGAGTGAGTGAGCGAGCG CGCATAGAGGGAGTGGCCAA-3' (SEQ ID NO: 12)

AAV5	5'- TCCCCCCTGTCGCGTTCGCTCGCTCGC TGGCTCGTTTGGGGGGGCGACGGCCA GAGGGCCGTCGTCTGGCAGCTCTTTG AGCTGCCACCCCCCCAAACGAGCCAG CGAGCGAGCGAACGCGACAGGGGGG AGAGTGCCACACTCTCAAGCAAGGGG GTTTTGTAAG -3' (SEQ ID NO: 8)	5'- CTTACAAAACCCCCTTGCTTGAGAGTG TGGCACTCTCCCCCCTGTCGCGTTCGCT CGCTCGCTGGCTCGTTTGGGGGGGTGG CAGCTCAAAGAGCTGCCAGACGACGG CCCTCTGGCCGTCGCCCCCCCAAACGA 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)

[00270] 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.

[00271] In certain embodiments of the present invention, the synthetically produced ceDNA vector 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 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.

[00272] The ceDNA vector 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 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 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 for controlled transgene expressioncomprising asymmetric ITR pairs or symmetric ITR pairs

[00273] As discussed herein, a ceDNA vector 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).

[00274] 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.

[00275] 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.

[00276] 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.

[00277] 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 for controlled transgene expressioncomprising 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).

[00278] 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.

[00279] 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.

[00280] By way of example only, **Table 3** 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.

[00281] **Table 3:** 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).

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

[00282] In some embodiments, mod-ITR for use in a ceDNA vector comprising 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 3**, 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 3, 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 3**, 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 3**, 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 3**, 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 3**, and also a modification of at least one nucleotide (e.g., a deletion, insertion and/ or substitution) in the D region.

[00283] 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). 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.

[00284] 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). 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.

[00285] 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.

[00286] 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.

[00287] In some embodiments, a synthetically produced ceDNA vector 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.

[00288] 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.

[00289] 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.

[00290] 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.

[00291] The ceDNA vector 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. In some embodiments, the ceDNA vector 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 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.

[00292] In some embodiments, the modified ITR (e.g., the left or right ITR) of the synthetically produced ceDNA vector 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.

[00293] In some embodiments, the modified ITR for use in a ceDNA vector 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.

[00294] Additional exemplary modified ITRs for use in a ceDNA vector comprising an asymmetric ITR pair, or symmetric mod-ITR pair in each of the above classes are provided in Tables 4A and 4B. The predicted secondary structure of the Right modified ITRs in Table 4A are shown in FIG. 7A of Internatioanl Application PCT/US2018/064242, filed December 6, 2018, and the predicted secondary structure of the Left modified ITRs in Table 4B are shown in FIG. 7B of Internatioanl Application PCT/US2018/064242, filed December 6, 2018, which is incorporated herein in its entirety by reference.

[00295] **Table 4A and Table 4B** show exemplary right and left modified ITRs.

[00296] **Table 4A:** 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 4A: Exemplary Right modified ITRs

ITR Construct	Sequence	SEQ ID NO:
ITR-18 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCGCACGCCCGGGTTTCCCGGGCGGCCTCAGTG AGCGAGCGAGCGCGCAGCTGCCTGCAGG	15
ITR-19 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGACGCCCGGGCTTTGCCCGGGCGGCCTCA GTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	16
ITR-20 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG CGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	17
ITR-21 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCTTTGCCTCAGTGAGCGAGCGAGCGCGCAGC TGCCTGCAGG	18
ITR-22 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACAAAGTCGCCCGACGCCCGGGCT TTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC AGG	19

ITR-23 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGAAAATCGCCCGACGCCCGGGCTTT GCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAG G	20
ITR-24 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGAAACGCCCGACGCCCGGGCTTTGC CCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	21
ITR-25 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCAAAGCCCGACGCCCGGGCTTTGCCC GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	22
ITR-26 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG TTTCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC AGG	23
ITR-27 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGT TTCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAG G	24
ITR-28 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGTT TCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	25
ITR-29 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCTTT GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	26
ITR-30 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCTTTG GCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	27
ITR-31 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCTTTGC GGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	28
ITR-32 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGTTTCGG CCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	29

ITR-49 Right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGGCCTCA GTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	30
ITR-50 right	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG CGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	31

[00297] **TABLE 4B**: 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 4B: Exemplary modified left ITRs

ITR-33 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG AAACCCGGGCGTGCGCCTCAGTGAGCGAGCGAGCGCGCAGAGAG GGAGTGGCCAACTCCATCACTAGGGGTTCCT	32
ITR-34 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGTCGGGC GACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACTCCATCACTAGGGGTTCCT	33
ITR-35 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG CAAAGCCCGGGCGTCGGCCTCAGTGAGCGAGCGAGCGCGCAGAG AGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	34
ITR-36 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCGCCCGGGC GTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGC GCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	35
ITR-37 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCAAAGCCTC AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA CTAGGGGTTCCT	36
ITR-38 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG CAAAGCCCGGGCGTCGGGCGACTTTGTCGCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGT TCCT	37

ITR-39 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG CAAAGCCCGGGCGTCGGGCGATTTTCGCCCGGCCTCAGTGAGCGA GCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTC CT	38
ITR-40 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG CAAAGCCCGGGCGTCGGGCGTTTCGCCCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	39
ITR-41 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG CAAAGCCCGGGCGTCGGGCTTTGCCCGGCCTCAGTGAGCGAGCGA GCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	40
ITR-42 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG AAACCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGT TCCT	41
ITR-43 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGA AACCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGA GCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTC CT	42
ITR-44 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGAA ACGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	43
ITR-45 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCAAA GGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGA GCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	44
ITR-46 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCAAAG GCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGC GCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	45
ITR-47 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCAAAGC GTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGC GCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	46
ITR-48 Left	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGAAACGT CGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	47

[00298] In one embodiment, a ceDNA vector 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 the 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 and for use to generate a ceDNA-plasmid are shown in **Table 4A and 4B.**

[00299] In an alternative embodiment, a synthetically produced ceDNA vector 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 ATC**GA**TCG with an addition of *AACG* between the G and A to result in the sequence ATC**G***AACG***A**TCG (SEQ ID NO: 51). The corresponding 3' ITR sense strand is CGA**TC**GAT (the reverse complement of ATC**GA**TCG) with an addition of *CGTT* (i.e. the reverse complement of AACG) between the T and C to result in the sequence CGA**T***CGTT***C**GAT (SEQ ID NO: 49) (the reverse complement of ATC**G***AACG***A**TCG) (SEQ ID NO: 51).

[00300] 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.

[00301] 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 ATCG*AACG***A**TCG (SEQ ID NO: 51), and modified 3' ITR as CGA**T***CGTT***C**GAT (SEQ ID NO: 49) (i.e., the reverse complement of ATCG*AACG***A**TCG (SEQ ID NO: 51)), these modified ITRs would still be symmetrical if, for example, the 5' ITR had the sequence of ATCG*AACC***A**TCG (SEQ ID NO: 50), where G in the addition is modified to C, and the substantially symmetrical 3' ITR has the sequence of CGA**T***CGTT***C**GAT (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.

[00302] **Table 5** shows exemplary symmetric modified ITR pairs (i.e. a left modified ITRs and the symmetric right modified ITR). 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).

Table 5: exemplary symmetric modified ITR pairs

LEFT modified ITR (modified 5′ ITR)		Symmetric RIGHT modified ITR (modified 3′ ITR)	
SEQ ID NO:32 (ITR-33 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGAAACCCG GGCGTGCGCCTCAGTGAGCGAGCGAG**	SEQ ID NO: 15 (ITR-18, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCGCAC**

	CGCGCAGAGAGGGAGTGGCCAACTCC ATCACTAGGGGTTCCT		GCCCGGGTTTCCCGGGCGGCC TCAGTGAGCGAGCGAGCGCGC AGCTGCCTGCAGG
SEQ ID NO: 33 (ITR-34 left)	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCCGTCGGGCGACCTTTG GTCGCCCGGCCTCAGTGAGCGAGCGA GCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCCT	SEQ ID NO: 48 (ITR-51, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCTGCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG GCCTCAGTGAGCGAGCGAGCG CGCAGCTGCCTGCAGG
SEQ ID NO: 34 (ITR-35 left)	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCCC GGGCGTCGGCCTCAGTGAGCGAGCGA GCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCCT	SEQ ID NO: 16 (ITR-19, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCTGCGCGC TCGCTCGCTCACTGAGGCCGAC GCCCGGGCTTTGCCCGGGCGG CCTCAGTGAGCGAGCGAGCGC GCAGCTGCCTGCAGG
SEQ ID NO: 35 (ITR-36 left)	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCGCCCGGGCGTCGGGC GACCTTTGGTCGCCCGGCCTCAGTGAG CGAGCGAGCGCGCAGAGAGGGAGTG GCCAACTCCATCACTAGGGGTTCCT	SEQ ID NO: 17 (ITR-20, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCTGCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG CCCGGGCGCCTCAGTGAGCGA GCGAGCGCGCAGCTGCCTGCA GG
SEQ ID NO: 36 (ITR-37 left)	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCAAAGCCTCAGTGAGCG AGCGAGCGCGCAGAGAGGGAGTGGC CAACTCCATCACTAGGGGTTCCT	SEQ ID NO: 18 (ITR-21, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCTGCGCGC TCGCTCGCTCACTGAGGCTTTG CCTCAGTGAGCGAGCGAGCGC GCAGCTGCCTGCAGG
SEQ ID NO: 37 (ITR-38 left)	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCCC GGGCGTCGGGCGACTTTGTCGCCCGG CCTCAGTGAGCGAGCGAGCGCGCAGA	SEQ ID NO: 19 (ITR-22 right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCTGCGCGC TCGCTCGCTCACTGAGGCCGG GCGACAAAGTCGCCCGACGCC

	GAGGGAGTGGCCAACTCCATCACTAG GGGTTCCT		**CGGGCTTTGCCCGGGCGGCCT CAGTGAGCGAGCGAGCGCGC**A GCTGCCTGCAGG
SEQ ID NO: 38 (ITR-39 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCCC GGGCGTCGGGCGATTTTCGCCCGGCCT CAGTGAGCGAGCGAGCGCGC**AGAGAG GGAGTGGCCAACTCCATCACTAGGGG TTCCT	SEQ ID NO: 20 (ITR-23, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGAAAATCGCCCGACGCCCG GGCTTTGCCCGGGCGGCCTCA GTGAGCGAGCGAGCGCGC**AGC TGCCTGCAGG
SEQ ID NO: 39 (ITR-40 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCCC GGGCGTCGGGCGTTTCGCCCGGCCTCA GTGAGCGAGCGAGCGCGC**AGAGAGG GAGTGGCCAACTCCATCACTAGGGGT TCCT	SEQ ID NO: 21 (ITR-24, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGAAACGCCCGACGCCCGGG CTTTGCCCGGGCGGCCTCAGTG AGCGAGCGAGCGCGC**AGCTGC CTGCAGG
SEQ ID NO: 40 (ITR-41 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCCC GGGCGTCGGGCTTTGCCCGGCCTCAGT GAGCGAGCGAGCGCGC**AGAGAGGGA GTGGCCAACTCCATCACTAGGGGTTCC T	SEQ ID NO: 22 (ITR-25 right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCAAAGCCCGACGCCCGGGCT TTGCCCGGGCGGCCTCAGTGA GCGAGCGAGCGCGC**AGCTGCC TGCAGG
SEQ ID NO: 41 (ITR-42 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGAAACCCG GGCGTCGGGCGACCTTTGGTCGCCCG GCCTCAGTGAGCGAGCGAGCGCGC**AG AGAGGGAGTGGCCAACTCCATCACTA GGGGTTCCT	SEQ ID NO: 23 (ITR-26 right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG CCCGGGTTTCCCGGGCGGCCTC AGTGAGCGAGCGAGCGCGC**AG CTGCCTGCAGG

SEQ ID NO: 42(ITR-43 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGGAAACCGGG CGTCGGGCGACCTTTGGTCGCCCGGCC TCAGTGAGCGAGCGAGCGCGC**AGAGA GGGAGTGGCCAACTCCATCACTAGGG GTTCCT	SEQ ID NO: 24 (ITR-27 right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG CCCGGTTTCCGGGCGGCCTCAG TGAGCGAGCGAGCGCGC**AGCT GCCTGCAGG
SEQ ID NO: 43 (ITR-44 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCGAAACGGGCG TCGGGCGACCTTTGGTCGCCCGGCCTC AGTGAGCGAGCGAGCGCGC**AGAGAG GGAGTGGCCAACTCCATCACTAGGGG TTCCT	SEQ ID NO: 25 (ITR-28 right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG CCCGTTTCGGGCGGCCTCAGTG AGCGAGCGAGCGCGC**AGCTGC CTGCAGG
SEQ ID NO:44 (ITR-45 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCCAAAGGGCGTC GGGCGACCTTTGGTCGCCCGGCCTCAG TGAGCGAGCGAGCGCGC**AGAGAGGG AGTGGCCAACTCCATCACTAGGGGTTC CT	SEQ ID NO:26 (ITR-29, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG CCCTTTGGGCGGCCTCAGTGAG CGAGCGAGCGCGC**AGCTGCCT GCAGG
SEQ ID NO:45 (ITR-46 left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCCAAAGGCGTCGG GCGACCTTTGGTCGCCCGGCCTCAGTG AGCGAGCGAGCGCGC**AGAGAGGGAG TGGCCAACTCCATCACTAGGGGTTCCT	SEQ ID NO: 27(ITR-30, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG CCTTTGGCGGCCTCAGTGAGCG AGCGAGCGCGC**AGCTGCCTGC AGG
SEQ ID NO: 46 (ITR-47, left)	CCTGCAGGCAGCT**GCGCGCTCGCTCG CTCACTGAGGCCGCAAAGCGTCGGGC GACCTTTGGTCGCCCGGCCTCAGTGAG**	SEQ ID NO: 28 (ITR-31, right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCT**GCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG**

	CGAGCGAGCGCGCAGAGAGGGAGTG GCCAACTCCATCACTAGGGGTTCCT		CTTTGCGGCCTCAGTGAGCGA GCGAGCGCGCAGCTGCCTGCA GG
SEQ ID NO: 47 (ITR-48, left)	CCTGCAGGCAGCTGCGCGCTCGCTCG CTCACTGAGGCCGAAACGTCGGGCGA CCTTTGGTCGCCCGGCCTCAGTGAGCG AGCGAGCGCGCAGAGAGGGAGTGGC CAACTCCATCACTAGGGGTTCCT	SEQ ID NO: 29 (ITR-32 right)	AGGAACCCCTAGTGATGGAGT TGGCCACTCCCTCTCTGCGCGC TCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACG TTTCGGCCTCAGTGAGCGAGC GAGCGCGCAGCTGCCTGCAGG

[00303] In some embodiments, a ceDNA vector 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 4A-4B 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.

V. Exemplary ceDNA vectors for controlled transgene expression

[00304] As described above, the present disclosure relates to recombinant ceDNA expression vectors and ceDNA vectors for controlled transgene expressionthat encode a transgene 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 controlled transgene expressionhaving 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.

[00305] The ceDNA vector for controlled transgene expression 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 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.

[00306] Referring now to **FIGS 1A-1G,** schematics of the functional components of two non-limiting plasmids useful in making the ceDNA vectors of the present disclosure are shown. **FIG. 1A, 1B, 1D, 1F** show the construct of ceDNA vectors or the corresponding sequences of ceDNA plasmids. 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 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).

[00307] **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.

[00308] The ceDNA vectors 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. In some embodiments, the ITR can act as the promoter for the transgene. In some embodiments, the ceDNA vector for controlled transgene expressioncomprises 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. 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.

[00309] 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.

[00310] The ceDNA vectors 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:

[00311] It will be appreciated by one of ordinary skill in the art that promoters used in the ceDNA vectors of the invention should be tailored as appropriate for the specific sequences they are promoting. For example, a guide RNA may not require a promoter at all, since its function is to form a duplex with a specific target sequence on the native DNA to effect a recombination event. In contrast, a nuclease encoded by the ceDNA vector would benefit from a promoter so that it can be efficiently expressed from the vector – and, optionally, in a regulatable fashion.

[00312] Expression cassettes of the present invention include a promoter, which can influence overall expression levels as well as cell-specificity. For transgene expression, 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 preferred embodiments, an expression cassette can contain a 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 (EF1a) 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.

[00313] Suitable promoters, including those described 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), a CAG promoter (SEQ ID NO: 72), a human alpha 1-antitypsin (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.

[00314] 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).

[00315] Non-limiting examples of suitable promoters for use in accordance with the present invention include 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 EF1a promoter (SEQ ID NO: 78), IE2 promoter (e.g., SEQ ID NO: 84) and the rat EF1-α promoter (SEQ ID NO: 85), or 1E1 promoter fragment (SEQ ID NO: 125).

(ii). Polyadenylation Sequences:

[00316] A sequence encoding a polyadenylation sequence can be included in the ceDNA vector for controlled transgene expressionto 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 vector 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. In some embodiments, where the ceDNA vector for controlled transgene expression can comprises two transgenes, e.g., in the case of controlled expression of an antibody, a ceDNA vector can comprise a nucleic acid encoding an antibody heavy chain (e.g., an exemplary heavy chain is SEQ ID NO: 57) and a nucleic acid encoding an antibody light chain (e.g., an exemplary light chain is SEQ ID NO: 58), and there can be a polyadenylation 3' of the first transgene, and an IRES (e.g., SEQ ID NO: 190) located between the first and second transgene (e.g., between the nucleic acid encoding an antibody heavy chain and the nucleic acid encoding an antibody light chain). In such embodiments, a ceDNA vector for controlled transgene expressionthat encodes more than one transgene (e.g., 2, or 3 or more) can comprise an IRES (internal ribosome entry site) sequence (SEQ ID NO: 190), e.g., where the IRES sequence is located 3' of a polyadenylation sequence, such that a second transgene (e.g., antibody or antigen-binding fragment) that is located 3' of a first transgene, is translated and expressed by the same ceDNA vector, such that the ceDNA vector can express two or more transgenes encoded by the ceDNA vector.

[00317] The expression cassettes can include a poly-adenylation sequence known in the art or a variation thereof, such as a 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, the, USE can be used in combination with SV40pA or heterologous poly-A signal.

[00318] 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 (SEQ ID NO: 88) and VK-A26 sequences (SEQ ID NO: 89), or IgK signal sequence (SEQ ID NO: 128), Glu secretory signal sequence (SEQ ID NO: 188) or TND secretory signal sequence (SEQ ID NO: 189).

(iii). Nuclear Localization Sequences

[00319] In some embodiments, the vector encoding an RNA guided endonuclease 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 6.**

[00320] Table 6: Nuclear Localization Signals

SOURCE	***SEQUENCE***	***SEQ ID NO.***
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 importin-alpha	RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV	96
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	KRKGDEVDGVDEVAKKKSKK	121
steroid hormone receptors (human) glucocorticoid	RKCLQAGMNLEARKTKK	122

B. Additional Components of ceDNA vectors

[00321] The ceDNA vectors 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 selections 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.

[00322] In embodiments, the ceDNA vector for controlled transgene expressionproduced using the synthetic process as described herein can be used for gene editing, for example, as disclosed in International Application PCT/US2018/064242, filed on December 6, 2018, which is incorporated herein in its entirety by reference, and may include one or more of: a 5' homology arm, a 3' homology arm, a polyadenylation site upstream and proximate to the 5' homology arm. Exemplary homology arms are 5' and 3' albumin homology arms (SEQ ID NO: 151 and 152) or CCR5 5'- and 3' homology arms (e.g., SEQ ID NO: 153, 154).

C. Regulatory Switches

[00323] 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 described herein to control the output of expression of the transgene from the ceDNA vector. In some embodiments, the ceDNA vector for controlled transgene expressioncomprises a regulatory switch that serves to fine tune expression of the transgene. 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 of the gene of interest in the ceDNA 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 controlled transgene expressioncan 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

[00324] In some embodiments, the ceDNA vector for controlled transgene expressioncomprises a regulatory switch that can serve to controllably modulate expression of the transgene. For example, the

expression cassette located between the ITRs of the ceDNA vector for controlled transgene expressionmay additionally comprise a regulatory region, e.g., a promoter, cis-element, repressor, enhancer etc., that is operatively linked to the gene of interest, 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. Nonlimiting 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

[00325] A variety of art-known small-molecule based regulatory switches are known in the art and can be combined with the ceDNA vectors 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 for controlled transgene expressionis a pro-drug activation switch, such as that disclosed in US patents 8,771,679, and 6,339,070.

(iii) "Passcode" Regulatory Switches

[00326] 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 for controlled transgene expressionwhen 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.

[00327] In some embodiments, a passcode regulatory switch or "Passcode circuit" encompassed for use in the ceDNA vector for controlled transgene expressioncomprises 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.

[00328] 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 in Table 11.

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

[00329] In some embodiments, the regulatory switch to control the transgene expressed 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 for controlled transgene expressioncan comprise a regulatory switch that encodes a RNAi molecule that is complementary to the transgene expressed by the ceDNA vector. When such RNAi is expressed even if the transgene 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 is not silenced by the RNAi.

[00330] 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 expression 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.

[00331] In some embodiments, the regulatory switch to control the transgene or gene of interest expressed by the ceDNA vector for controlled transgene expressionis 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

[00332] Any known regulatory switch can be used in the ceDNA vector to control the gene expression of the transgene expressed 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.

[00333] In some embodiments, a regulatory switch envisioned for use in the ceDNA vector for controlled transgene expressionis 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 for controlled transgene expressionafter ischemia or in ischemic tissues, and/or tumors.

(iv). Kill Switches

[00334] Other embodiments of the invention relate to a ceDNA vector for controlled transgene expressioncomprising 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 of the invention 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 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 from a subject or to ensure that it will not express the encoded transgene.

VI. Detailed method of Production of a ceDNA Vector

A. Production in General

[00335] Certain methods for the production of a ceDNA vector for controlled transgene expressioncomprising 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 controlled transgene expressionfor use in the methods and compositions as disclosed herein can be produced using insect cells, as described herein. In anterlative embodiments, a for use in the methods and compositions 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.

[00336] As described herein, in one embodiment, a ceDNA vector for controlled transgene expressioncan 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.

[00337] 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.

[00338] 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.

[00339] In one embodiment, the host cells used to make the ceDNA vectors 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.

[00340] 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.

[00341] 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.

[00342] The presence of the ceDNA vector 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.

[00343]

B. ceDNA Plasmid

[00344] A ceDNA-plasmid is a plasmid used for later production of a ceDNA vector. 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.

[00345] In one aspect, a ceDNA vector for controlled transgene expressionis 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 are have the same modifications (i.e., they are inverse complement or symmetric relative to each other).

[00346] 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.

[00347] 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 web address: 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.

[00348] 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.

[00349] An Exemplary ceDNA (e.g., rAAV0) 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

[00350] Methods for making capsid-less ceDNA vectors are also provided herein, notably a method with a sufficiently high yield to provide sufficient vector for *in vivo* experiments.

[00351] In some embodiments, a method for the production of a ceDNA vector for controlled transgene expressioncomprises 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 for controlled transgene expressioncan 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*:**

[00352] Host cell lines used in the production of a ceDNA vector for controlled transgene expressioncan 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, HeplA, 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.

[00353] 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*:**

[00354] Examples of the process for obtaining and isolating ceDNA vectors are described in **FIGS. 4A-4E** and the specific examples below. ceDNA-vectors 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 incorporating one or more Rep protein(s) and plasmids used to obtain a ceDNA vector. Exemplary plasmids for production of ceDNA vector to for controlled expression of a transgene is a plasmid as shown in FIG. 6B of International application PCT/US2018/064242, filed December 6, 2018, which is incorporataed herein in its entirety. A ceDNA plasmid for production of a ceDNA vector for controlled expression of an antibody is disclosed in FIG.6A and is SEQ ID NO: 56 of International Application PCT/US19/18016 filed on February 14, 2019, which discloses an exemplary ceDNA plasmid for production of aducanmab.

[00355] In one aspect, a polynucleotide encodes the AAV Rep protein (Rep 78 or Rep68) is 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.

[00356] Methods to produce a ceDNA-vector, which is an exemplary ceDNA vector, are described herein. Expression constructs used for generating a ceDNA vectors of the present invention 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 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.

[00357] The bacmid (e.g., ceDNA-bacmid) can be transfected into a 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.

[00358] 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. 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.

[00359] 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 nonlimiting 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.

[00360] In some embodiments, ceDNA vectors 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.

[00361] 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)

[00362] 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.

[00363] **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, as discussed with respect to **FIG. 4D** in the Examples.

VII. Pharmaceutical Compositions

[00364] In another aspect, pharmaceutical compositions are provided. The pharmaceutical composition comprises a closed-ended DNA vector, e.g., ceDNA vector for controlled transgene expressionproduced using the synthetic process as described herein and a pharmaceutically acceptable carrier or diluent.

[00365] The ceDNA vectors 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 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.

[00366] Pharmaceutically active compositions comprising a ceDNA vector for controlled transgene expressioncan be formulated to deliver a transgene for various purposes to the cell, e.g., cells of a subject.

[00367] The ceDNA vectors 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 the DNA-vectors disclosed herein and a pharmaceutically acceptable carrier. For example, the ceDNA vectors of the invention 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 intraarterial 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.

[00368] Pharmaceutically active compositions comprising 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 therein. The composition can also optionally include a pharmaceutically acceptable carrier and/or excipient.

[00369] The compositions and vectors provided herein can be used to deliver a transgene for various purposes. In some embodiments, the transgene encodes a protein or functional RNA 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 product. In another example, the transgene encodes a protein or functional RNA that is intended to be used to create an animal model of disease. In some embodiments,

the transgene encodes one or more peptides, polypeptides, or proteins, which are useful for the treatment or prevention of disease states in a mammalian subject. The transgene 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. In some embodiments, the transgene is a gene editing molecule (e.g., nuclease). In certain embodiments, the nuclease is a CRISPR-associated nuclease (Cas nuclease).

[00370] Pharmaceutical compositions for therapeutic purposes typically must be sterile and stable under the conditions of manufacture and storage. 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.

[00371] In certain circumstances, it will be desirable to deliver a ceDNA composition or vector as disclosed herein in suitably formulated pharmaceutical compositions disclosed herein either subcutaneously, intraopancreatically, intranasally, parenterally, intravenously, intramuscularly, intrathecally, systemic administration, or orally, intraperitoneally, or by inhalation.

[00372] It is specifically contemplated herein that the compositions described herein comprise a ceDNA vector for controlled transgene expressionat a given dose that is determined by the dose-response relationship of the ceDNA vector, for example, a "unit dose" that, upon administration, can be reliably expected to produce a desired effect or level of expression of the genetic medicine in a typical subject.

[00373] 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.

[00374] A ceDNA vector for controlled transgene expressionas 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.

[00375] In some aspects, the methods provided herein comprise delivering one or more ceDNA vectors 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) 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).

[00376] Various techniques and methods are known in the art for delivering nucleic acids to cells. For example, nucleic acids, such as ceDNA can be formulated into lipid nanoparticles (LNPs), lipidoids, liposomes, lipid nanoparticles, lipoplexes, or core-shell nanoparticles. Typically, LNPs are composed of nucleic acid (*e.g.*, ceDNA) molecules, one or more ionizable or cationic lipids (or salts thereof), one or more non-ionic or neutral lipids (*e.g.*, a phospholipid), a molecule that prevents aggregation (*e.g.*, PEG or a PEG-lipid conjugate), and optionally a sterol (*e.g.*, cholesterol).

[00377] Another method for delivering nucleic acids, such as ceDNA to a cell is by conjugating the nucleic acid with a ligand that is internalized by the cell. For example, the ligand can bind a receptor on the cell surface and internalized via endocytosis. The ligand can be covalently linked to a nucleotide in the nucleic acid. Exemplary conjugates for delivering nucleic acids into a cell are described, example, in WO2015/006740, WO2014/025805, WO2012/037254, WO2009/082606, WO2009/073809, WO2009/018332, WO2006/112872, WO2004/090108, WO2004/091515 and WO2017/177326.

[00378] Nucleic acids, such as ceDNA, can also be delivered to a cell by transfection. Useful transfection methods include, but are not limited to, lipid-mediated transfection, cationic polymer-mediated transfection, or calcium phosphate precipitation. Transfection reagents are well known in the art and include, but are not limited to, TurboFect Transfection Reagent (Thermo Fisher Scientific), Pro-Ject Reagent (Thermo Fisher Scientific), TRANSPASS™ P Protein Transfection Reagent (New England Biolabs), CHARIOT™ Protein Delivery Reagent (Active Motif), PROTEOJUICE™ Protein Transfection Reagent (EMD Millipore), 293fectin, LIPOFECTAMINE™ 2000, LIPOFECTAMINE™ 3000 (Thermo Fisher Scientific), LIPOFECTAMINE™ (Thermo Fisher Scientific), LIPOFECTIN™ (Thermo Fisher Scientific), DMRIE-C, CELLFECTIN™ (Thermo Fisher Scientific), OLIGOFECTAMINE™ (Thermo Fisher Scientific), LIPOFECTACE™, FUGENE™ (Roche, Basel, Switzerland), FUGENE™ HD (Roche), TRANSFECTAM™(Transfectam, Promega, Madison, Wis.), TFX-10™ (Promega), TFX-20™ (Promega), TFX-50™ (Promega), TRANSFECTIN™ (BioRad, Hercules, Calif.), SILENTFECT™ (Bio-Rad), Effectene™ (Qiagen, Valencia, Calif.), DC-chol (Avanti Polar Lipids), GENEPORTER™ (Gene Therapy Systems, San Diego, Calif.), DHARMAFECT 1™ (Dharmacon, Lafayette, Colo.), DHARMAFECT 2™ (Dharmacon), DHARMAFECT 3™ (Dharmacon), DHARMAFECT 4™ (Dharmacon), ESCORT™ III (Sigma, St. Louis, Mo.), and ESCORT™ IV (Sigma Chemical Co.).

Nucleic acids, such as ceDNA, can also be delivered to a cell via microfluidics methods known to those of skill in the art.

[00379] ceDNA vectors as described herein can also be administered directly to an organism for transduction of cells *in vivo*. Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells including, but not limited to, injection, infusion, topical application and electroporation. Suitable methods of administering such nucleic acids are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

[00380] Methods for introduction of a nucleic acid vector ceDNA vector for controlled transgene expressionas disclosed herein can be delivered into hematopoietic stem cells, for example, by the methods as decribed, for example, in U.S. Pat. No. 5,928,638.

[00381] The ceDNA vectors in accordance with the present invention can be added to liposomes for delivery to a cell or target organ in a subject. Liposomes are vesicles that possess at least one lipid bilayer. Liposomes are typical used as carriers for drug/ therapeutic delivery in the context of pharmaceutical development. They work by fusing with a cellular membrane and repositioning its lipid structure to deliver a drug or active pharmaceutical ingredient (API). Liposome compositions for such delivery are composed of phospholipids, especially compounds having a phosphatidylcholine group, however these compositions may also include other lipids. Exemplary liposomes and liposome formulations, including but not limited to polyethylene glycol (PEG)-functional group containing compounds are disclosed in International Application PCT/US2018/050042, filed on September 7, 2018 and in International application PCT/US2018/064242, filed on December 6, 2018, e.g., see the section entitled "Pharmaceutical Formulations".]

[00382] Various delivery methods known in the art or modification thereof can be used to deliver ceDNA vectors *in vitro* or *in vivo.* For example, in some embodiments, ceDNA vectors are delivered by making transient penetration in cell membrane by mechanical, electrical, ultrasonic, hydrodynamic, or laser-based energy so that DNA entrance into the targeted cells is facilitated. For example, a ceDNA vector for controlled transgene expressioncan be delivered by transiently disrupting cell membrane by squeezing the cell through a size-restricted channel or by other means known in the art. In some cases, a ceDNA vector alone is directly injected as naked DNA into skin, thymus, cardiac muscle, skeletal muscle, or liver cells. In some cases, a ceDNA vector is delivered by gene gun. Gold or tungsten spherical particles (1–3 μm diameter) coated with capsid-free AAV vectors can be accelerated to high speed by pressurized gas to penetrate into target tissue cells.

[00383] Compositions comprising a ceDNA vector for controlled transgene expressionand a pharmaceutically acceptable carrier are specifically contemplated herein. In some embodiments, the ceDNA vector for controlled transgene expressionis formulated with a lipid delivery system, for example, liposomes as described herein. In some embodiments, such compositions are administered by any route desired by a skilled practitioner. The compositions may be administered to a subject by different routes including orally, parenterally, sublingually, transdermally, rectally, transmucosally, topically, via inhalation, via buccal administration, intrapleurally, intravenous, intra-arterial, intraperitoneal, subcutaneous, intramuscular, intranasal intrathecal, and intraarticular or combinations thereof. For veterinary use, the composition may be administered as a suitably acceptable formulation in accordance with normal veterinary practice. The veterinarian may readily determine the dosing regimen and route of administration that is most appropriate for a particular animal. The compositions may be administered by traditional syringes, needleless injection devices, "microprojectile bombardment gene guns", or other physical methods such as electroporation ("EP"), hydrodynamic methods, or ultrasound.

[00384] In some cases, a ceDNA vector for controlled transgene expressionis delivered by hydrodynamic injection, which is a simple and highly efficient method for direct intracellular delivery of any water-soluble compounds and particles into internal organs and skeletal muscle in an entire limb.

[00385] In some cases, ceDNA vectors are delivered by ultrasound by making nanoscopic pores in membrane to facilitate intracellular delivery of DNA particles into cells of internal organs or tumors, so the size and concentration of plasmid DNA have great role in efficiency of the system. In some cases, ceDNA vectors are delivered by magnetofection by using magnetic fields to concentrate particles containing nucleic acid into the target cells.

[00386] In some cases, chemical delivery systems can be used, for example, by using nanomeric complexes, which include compaction of negatively charged nucleic acid by polycationic nanomeric particles, belonging to cationic liposome/micelle or cationic polymers. Cationic lipids used for the delivery method includes, but not limited to monovalent cationic lipids, polyvalent cationic lipids, guanidine containing compounds, cholesterol derivative compounds, cationic polymers, (*e.g.,* poly(ethylenimine), poly-L-lysine, protamine, other cationic polymers), and lipid-polymer hybrid.

A. Exosomes:

[00387] In some embodiments, a ceDNA vector for controlled transgene expressionas disclosed herein is delivered by being packaged in an exosome. 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 1μm, 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.

B. Microparticle/Nanoparticles:

[00388] In some embodiments, a ceDNA vector for controlled transgene expressionas disclosed herein is delivered by a lipid nanoparticle. Generally, lipid nanoparticles comprise an ionizable amino lipid (*e.g.*, heptatriaconta-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.

[00389] 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 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.

[00390] Various lipid nanoparticles known in the art can be used to deliver ceDNA vector for controlled transgene expressiondisclosed 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.

[00391] In some embodiments, a ceDNA vector for controlled transgene expressiondisclosed 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.

C. Conjugates

[00392] In some embodiments, a ceDNA vector for controlled transgene expressionas disclosed herein is conjugated (*e.g.*, covalently bound to an agent that increases cellular uptake. An "agent that increases cellular uptake" is a molecule that facilitates transport of a nucleic acid across a lipid membrane. For example, a nucleic acid can be conjugated to a lipophilic compound (*e.g.*, cholesterol, tocopherol, *etc.*), a

cell penetrating peptide (CPP) (*e.g.*, penetratin, TAT, Syn1B, *etc.*), and polyamines (*e.g.*, spermine). Further examples of agents that increase cellular uptake are disclosed, for example, in Winkler (2013). *Oligonucleotide conjugates for therapeutic applications*. Ther. Deliv. 4(7); 791-809.

[00393] In some embodiments, a ceDNA vector for controlled transgene expressionas disclosed herein is conjugated to a polymer (*e.g.*, a polymeric molecule) or a folate molecule (*e.g.*, folic acid molecule). Generally, delivery of nucleic acids conjugated to polymers is known in the art, for example as described in WO2000/34343 and WO2008/022309. In some embodiments, a ceDNA vector for controlled transgene expressionas disclosed herein is conjugated to a poly(amide) polymer, for example as described by U.S. Patent No. 8,987,377. In some embodiments, a nucleic acid described by the disclosure is conjugated to a folic acid molecule as described in U.S. Patent No. 8,507,455.

[00394] In some embodiments, a ceDNA vector for controlled transgene expressionas disclosed herein is conjugated to a carbohydrate, for example as described in U.S. Patent No. 8,450,467.

D. Nanocapsule

[00395] Alternatively, nanocapsule formulations of a ceDNA vector for controlled transgene expressionas disclosed herein can be used. Nanocapsules can generally entrap substances in a stable and reproducible way. To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use.

E. Liposomes

[00396] The ceDNA vectors in accordance with the present invention can be added to liposomes for delivery to a cell or target organ in a subject. Liposomes are vesicles that possess at least one lipid bilayer. Liposomes are typical used as carriers for drug/ therapeutic delivery in the context of pharmaceutical development. They work by fusing with a cellular membrane and repositioning its lipid structure to deliver a drug or active pharmaceutical ingredient (API). Liposome compositions for such delivery are composed of phospholipids, especially compounds having a phosphatidylcholine group, however these compositions may also include other lipids.

[00397] The formation and use of liposomes is generally known to those of skill in the art. Liposomes have been developed with improved serum stability and circulation half-times (U.S. Pat. No. 5,741,516). Further, various methods of liposome and liposome like preparations as potential drug carriers have been described (U.S. Pat. Nos. 5,567,434; 5,552,157; 5,565,213; 5,738,868 and 5,795,587).

F. Exemplary liposome and Lipid Nanoparticle (LNP) Compositions

[00398] The ceDNA vectors in accordance with the present invention can be added to liposomes for delivery to a cell, e.g., a cell in need of expression of the transgene. Liposomes are vesicles that possess

at least one lipid bilayer. Liposomes are typical used as carriers for drug/ therapeutic delivery in the context of pharmaceutical development. They work by fusing with a cellular membrane and repositioning its lipid structure to deliver a drug or active pharmaceutical ingredient (API). Liposome compositions for such delivery are composed of phospholipids, especially compounds having a phosphatidylcholine group, however these compositions may also include other lipids.

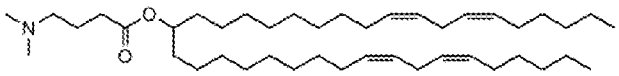
[00399] Lipid nanoparticles (LNPs) comprising ceDNA are disclosed in International Application PCT/US2018/050042, filed on September 7, 2018, and International Application PCT/US2018/064242, filed on December 6, 2018 which are incorporated herein in their entirety and envisioned for use in the methods and compostions as disclosed herein.

[00400] In some aspects, a lipid nanoparticle comprising a ceDNA is an ionizable lipid.

[00401] 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. Ionizable lipids are also referred to as cationic lipids herein. Exemplary ionizable lipids are described in International PCT patent publications WO2015/095340, WO2015/199952, WO2018/011633, WO2017/049245, WO2015/061467, WO2012/040184, WO2012/000104, WO2015/074085, WO2016/081029, WO2017/004143, WO2017/075531, WO2017/117528, WO2011/022460, WO2013/148541, WO2013/116126, WO2011/153120, WO2012/044638, WO2012/054365, WO2011/090965, WO2013/016058, WO2012/162210, WO2008/042973, WO2010/129709, WO2010/144740 , WO2012/099755, WO2013/049328, WO2013/086322, WO2013/086373, WO2011/071860, WO2009/132131, WO2010/048536, WO2010/088537, WO2010/054401, WO2010/054406 , WO2010/054405, WO2010/054384, WO2012/016184, WO2009/086558, WO2010/042877, WO2011/000106, WO2011/000107, WO2005/120152, WO2011/141705, WO2013/126803, WO2006/007712, WO2011/038160, WO2005/121348, WO2011/066651, WO2009/127060, WO2011/141704, WO2006/069782, WO2012/031043, WO2013/006825, WO2013/033563, WO2013/089151, WO2017/099823, WO2015/095346, and WO2013/086354, and US patent publications US2016/0311759, US2015/0376115, US2016/0151284, US2017/0210697, US2015/0140070, US2013/0178541, US2013/0303587, US2015/0141678, US2015/0239926, US2016/0376224, US2017/0119904, US2012/0149894, US2015/0057373, US2013/0090372, US2013/0274523, US2013/0274504, US2013/0274504, US2009/0023673, US2012/0128760, US2010/0324120, US2014/0200257, US2015/0203446,

US2018/0005363, US2014/0308304, US2013/0338210, US2012/0101148, US2012/0027796, US2012/0058144, US2013/0323269, US2011/0117125, US2011/0256175, US2012/0202871, US2011/0076335, US2006/0083780, US2013/0123338, US2015/0064242, US2006/0051405, US2013/0065939, US2006/0008910, US2003/0022649, US2010/0130588, US2013/0116307, US2010/0062967, US2013/0202684, US2014/0141070, US2014/0255472, US2014/0039032, US2018/0028664, US2016/0317458, and US2013/0195920, the contents of all of which are incorporated herein by reference in their entirety.

[00402] In some embodiments, the ionizable lipid is MC3 (6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl-4-(dimethylamino) butanoate (DLin-MC3-DMA or MC3) having the following structure:



DLin-M-C3-DMA ("MC3")

VIII. Methods of delivering ceDNA vectors

[00403] In some embodiments, a ceDNA vector for controlled transgene expressioncan 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 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.

[00404] 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.

[00405] In another embodiment, a ceDNA vector for controlled transgene expressionis administered to the CNS (e.g., to the brain or to the eye). The ceDNA vector for controlled transgene expressionmay 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 also be administered to different regions of the eye such as the retina, cornea and/or optic nerve. The ceDNA vector may be delivered into the cerebrospinal fluid (e.g., by lumbar puncture).

The ceDNA vector 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).

[00406] In some embodiments, the ceDNA vector for controlled transgene expressioncan be administered to the desired region(s) of the CNS 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.

[00407] In some embodiments, the ceDNA vector for controlled transgene expressionis 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 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.

IX. Additional uses of the ceDNA vectors

[00408] The compositions and ceDNA vectors as described herein can be used to express a target gene or transgene for various purposes. In some embodiments, the resulting transgene encodes a protein or functional RNA 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 product. In another example, the transgene encodes a protein or functional RNA that is intended to be used to create an animal model of disease. In some embodiments, the resulting transgene encodes one or more peptides, polypeptides, or proteins, which are useful for the treatment, prevention, or amelioration of disease states or disorders in a mammalian subject. The resulting transgene can be transferred (e.g., expressed in) to a subject in a sufficient amount to treat a disease associated with reduced expression, lack of expression or dysfunction of the gene. In some embodiments the resulting transgene can be expressed in a subject in a sufficient amount to treat a disease associated with increased expression, activity of the gene product, or inappropriate upregulation of a gene that the resulting transgene suppresses or otherwise causes the expression of which to be reduced. In yet other embodiments, the resulting transgene replaces or supplements a defective copy of the native gene. 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 a gene of interest.

[00409] In some embodiments, the transgene encodes a protein or functional RNA that is intended to be used to create an animal model of disease. In some embodiments, the transgene encodes one or more peptides, polypeptides, or proteins, which are useful for the treatment or prevention of disease states in a mammalian subject. The transgene 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.

X. Methods of Use

[00410] A ceDNA vector for controlled transgene expressionas disclosed herein can also be used in a method for the delivery of a nucleotide sequence of interest (e.g., a transgene) to a target cell (e.g., a host cell). The method may in particular be a method for delivering a transgene to a cell of a subject in need thereof and treating a disease of interest. The invention allows for the *in vivo* expression of a transgene, e.g., a protein, antibody, nucleic acid such as miRNA etc. encoded in the ceDNA vector in a cell in a subject such that therapeutic effect of the expression of the transgene occurs. These results are seen with both *in vivo* and *in vitro* modes of ceDNA vector delivery.

[00411] In addition, the invention provides a method for the delivery of a transgene in a cell of a subject in need thereof, comprising multiple administrations of the ceDNA vector of the invention comprising said nucleic acid or transgene of interest to titrate the transgene expression to the desired level.

[00412] The ceDNA vector nucleic acid(s) are 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, intravenous (e.g., in a liposome formulation), direct delivery to the selected organ (*e.g.*, intraportal delivery to the liver), intramuscular, and other parental routes of administration. Routes of administration may be combined, if desired.

[00413] Closed-ended DNA vector (e.g. ceDNA vector) delivery is not limited to delivery gene replacements. For example, conventionally produced (e.g., using a cell-based production method or synthetically produced closed-ended DNA vectors) (e.g., 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 synthetically produced 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 transgene.

[00414] 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 for controlled transgene expressioncan be introduced in the presence of a carrier, such a carrier is not required. The ceDNA vector selected comprises a nucleotide sequence of interest useful for treating the 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.

[00415] The compositions and vectors provided herein can be used to deliver a transgene for various purposes. In some embodiments, the transgene encodes a protein or functional RNA 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 product. In another example, the transgene encodes a protein or functional RNA that is intended to be used to create an animal model of disease. In some embodiments, the transgene encodes one or more peptides, polypeptides, or proteins, which are useful for the treatment or prevention of disease states in a mammalian subject. The transgene 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.

[00416] In principle, the expression cassette can include a nucleic acid or any transgene that encodes a protein or polypeptide 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.

[00417] A ceDNA vector for controlled transgene expression is not limited to one species of ceDNA vector. As such, in another aspect, multiple ceDNA vectors comprising different transgenes or the same transgene 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 genes simultaneously. It is also possible to separate different portions of the transgene into separate ceDNA vectors (e.g., different domains and/or co-factors required for functionality of the transgene) which can be administered simultaneously or at different times, and can be separately regulatable, thereby adding an additional level of control of

expression of the transgene. 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.

[00418] 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 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 for controlled transgene expression can be administered via any suitable route as provided above, and elsewhere herein.

XI. Methods of Treatment

[00419] The technology described herein also demonstrates methods for making, as well as methods of using the disclosed ceDNA vectors in a variety of ways, including, for example, ex situ, *in vitro and in vivo* applications, methodologies, diagnostic procedures, and/or gene therapy regimens.

[00420] Provided herein is a method of treating a disease or disorder 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, 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 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 for controlled transgene expression can be administered via any suitable route as provided above, and elsewhere herein.

[00421] Disclosed herein are ceDNA vector compositions and formulations 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,

injury, disorder, trauma or dysfunction. In one aspect the disease, injury, disorder, trauma or dysfunction is a human disease, injury, disorder, trauma or dysfunction.

[00422] 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, 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 from the ceDNA vector thereby providing the subject with a diagnostically- or a therapeutically-effective amount of the protein, peptide, nucleic acid expressed by the ceDNA vector. In a further aspect, the subject is human.

[00423] 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 vectors, 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 a further aspect, the subject is human.

[00424] Another aspect is use of the ceDNA vector for controlled transgene expression 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 deficiency state diseases, ceDNA vectors can be used to deliver transgenes to bring a normal gene into affected tissues for replacement therapy, as well, in some embodiments, to create animal models for the disease using antisense mutations. For unbalanced disease states, ceDNA vectors 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 vectors and methods 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:

[00425] In some embodiments, the ceDNA vector for controlled transgene expression delivers the transgene into a subject host cell. 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.

[00426] The present disclosure also relates to recombinant host cells as mentioned above, including ceDNA vectors as described herein. Thus, one can use multiple host cells depending on the purpose as is obvious to the skilled artisan. A construct or ceDNA vector for controlled transgene expression 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. 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 the ceDNA vector for controlled transgene expression *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. Exemplary transgenes and diseases to be treated with a ceDNA vector

[00427] The ceDNA vectors are also useful for correcting a defective gene. As a non-limiting example, DMD gene of Duchene Muscular Dystrophy can be delivered using the ceDNA vectors as disclosed herein.

[00428] A ceDNA vector for controlled transgene expression or a composition thereof can be used in the treatment of any hereditary disease. As a non-limiting example, the ceDNA vector or a composition thereof e.g. can be used in the treatment of transthyretin amyloidosis (ATTR), an orphan disease where the mutant protein misfolds and aggregates in nerves, the heart, the gastrointestinal system etc. It is contemplated herein that the disease can be treated by deletion of the mutant disease gene (mutTTR) using the ceDNA vector systems described herein. Such treatments of hereditary diseases can halt disease progression and may enable regression of an established disease or reduction of at least one symptom of the disease by at least 10%.

[00429] In another embodiment, a ceDNA vector for controlled transgene expression can be used in the treatment of ornithine transcarbamylase deficiency (OTC deficiency), hyperammonaemia or other urea cycle disorders, which impair a neonate or infant's ability to detoxify ammonia. As with all diseases of inborn metabolism, it is contemplated herein that even a partial restoration of enzyme activity compared to wild-type controls (*e.g.,* at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or at least 99%) may be sufficient for reduction in at least one symptom OTC and/or an improvement in the quality of life for a subject having OTC deficiency. In one embodiment, a nucleic acid encoding OTC can be inserted behind the albumin endogenous promoter for *in vivo* protein replacement.

[00430] In another embodiment, a ceDNA vector for controlled transgene expression can be used in the treatment of phenylketonuria (PKU) by delivering a nucleic acid sequence encoding a phenylalanine hydroxylase enzyme to reduce buildup of dietary phenylalanine, which can be toxic to PKU sufferers. As with all diseases of inborn metabolism, it is contemplated herein that even a partial restoration of enzyme activity compared to wild-type controls (*e.g.,* at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or at least 99%) may be sufficient for reduction in at least one symptom of PKU and/or an improvement in the quality of life for a subject having PKU. In one embodiment, a nucleic acid encoding phenylalanine hydroxylase can be inserted behind the albumin endogenous promoter for *in vivo* protein replacement.

[00431] In another embodiment, a ceDNA vector for controlled transgene expression can be used in the treatment of glycogen storage disease (GSD) by delivering a nucleic acid sequence encoding an enzyme to correct aberrant glycogen synthesis or breakdown in subjects having GSD. Non-limiting examples of enzymes that can be delivered and expressed using the ceDNA vectors and methods as described herein include glycogen synthase, glucose-6-phosphatase, acid-alpha glucosidase, glycogen debranching enzyme, glycogen branching enzyme, muscle glycogen phosphorylase, liver glycogen phosphorylase, muscle phosphofructokinase, phosphorylase kinase, glucose transporter -2 (GLUT-2), aldolase A, beta-enolase, phosphoglucomutase-1 (PGM-1), and glycogenin-1. As with all diseases of inborn metabolism, it is contemplated herein that even a partial restoration of enzyme activity compared to wild-type controls (*e.g.,* at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or at least 99%) may be sufficient for reduction in at least one symptom of GSD and/or an improvement in the quality of life for a subject having GSD. In one embodiment, a nucleic acid encoding an enzyme to correct aberrant glycogen storage can be inserted behind the albumin endogenous promoter for *in vivo* protein replacement.

[00432] The ceDNA vectors described herein are also contemplated for use in the treatment of any of; of Leber congenital amaurosis (LCA), polyglutamine diseases, including polyQ repeats, and alpha-1

antitrypsin deficiency (A1AT). LCA is a rare congenital eye disease resulting in blindness, which can be caused by a mutation in any one of the following genes: GUCY2D, RPE65, SPATA7, AIPL1, LCA5, RPGRIP1, CRX, CRB1, NMNAT1, CEP290, IMPDH1, RD3, RDH12, LRAT, TULP1, KCNJ13, GDF6 and/or PRPH2. It is contemplated herein that the ceDNA vectors and compositions and methods as described herein can be adapted for delivery of one or more of the genes associated with LCA in order to correct an error in the gene(s) responsible for the symptoms of LCA. Polyglutamine diseases include, but are not limited to: dentatorubropallidoluysian atrophy, Huntington's disease, spinal and bulbar muscular atrophy, and spinocerebellar ataxia types 1, 2, 3 (also known as Machado-Joseph disease), 6, 7, and 17. A1AT deficiency is a genetic disorder that causes defective production of alpha-1 antitrypsin, leading to decreased activity of the enzyme in the blood and lungs, which in turn can lead to emphysema or chronic obstructive pulmonary disease in affected subjects. Treatment of a subject with an A1AT deficiency is specifically contemplated herein using the ceDNA vectors or compositions thereof as outlined herein. It is contemplated herein that a ceDNA vector for controlled transgene expression comprising a nucleic acid encoding a desired protein for the treatment of LCA, polyglutamine diseases or A1AT deficiency can be admininstered to a subject in need of treatment.

[00433] In further embodiments, the compositions comprising a ceDNA vector for controlled transgene expression as described herein can be used to deliver a viral sequence, a pathogen sequence, a chromosomal sequence, a translocation junction (*e.g.,* a translocation associated with cancer), a non-coding RNA gene or RNA sequence, a disease associated gene, among others.

[00434] Any nucleic acid or target gene of interest may be delivered or expressed by a ceDNA vector for controlled transgene expression as disclosed herein. Target nucleic acids and target genes include, but are not limited to nucleic acids encoding polypeptides, or non-coding nucleic acids (e.g., RNAi, miRs etc.) preferably therapeutic (e.g., for medical, diagnostic, or veterinary uses) or immunogenic (e.g., for vaccines) polypeptides. In certain embodiments, the target nucleic acids or target genes that are targeted by the ceDNA vectors as described herein encode one or more polypeptides, peptides, ribozymes, peptide nucleic acids, siRNAs, RNAis, antisense oligonucleotides, antisense polynucleotides, antibodies, antigen binding fragments, or any combination thereof.

[00435] In particular, a gene target or transgene for expression by the ceDNA vector for controlled transgene expression as disclosed herein can encode, for example, but is not limited to, protein(s), polypeptide(s), peptide(s), enzyme(s), antibodies, antigen binding fragments, as well as variants, and/or active fragments thereof, for use in the treatment, prophylaxis, and/or amelioration of one or more symptoms of a disease, dysfunction, injury, and/or disorder. In one aspect, the disease, dysfunction, trauma, injury and/or disorder is a human disease, dysfunction, trauma, injury, and/or disorder.

[00436] The expression cassette can also encode 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.

[00437] Sequences provided in the expression cassette, expression construct of a ceDNA vector for controlled transgene expression 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, 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.

[00438] Many organisms display a bias for use of particular codons to code for insertion of a particular amino acid in a growing peptide chain. Codon preference or codon bias, differences in codon usage between organisms, is afforded by degeneracy of the genetic code, and is well documented among many organisms. Codon bias often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, inter alia, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

[00439] Given the large number of gene sequences available for a wide variety of animal, plant and microbial species, it is possible to calculate the relative frequencies of codon usage (Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000)).

[00440] As noted herein, a ceDNA vector for controlled transgene expression as disclosed herein can encode a protein or peptide, or therapeutic nucleic acid sequence or therapeutic agent, including but not limited to one or more agonists, antagonists, anti-apoptosis factors, inhibitors, receptors, cytokines, cytotoxins, erythropoietic agents, glycoproteins, growth factors, growth factor receptors, hormones,

hormone receptors, interferons, interleukins, interleukin receptors, nerve growth factors, neuroactive peptides, neuroactive peptide receptors, proteases, protease inhibitors, protein decarboxylases, protein kinases, protein kinase inhibitors, enzymes, receptor binding proteins, transport proteins or one or more inhibitors thereof, serotonin receptors, or one or more uptake inhibitors thereof, serpins, serpin receptors, tumor suppressors, diagnostic molecules, chemotherapeutic agents, cytotoxins, or any combination thereof.

[00441] The ceDNA vectors are also useful for ablating gene expression. For example, in one embodiment a ceDNA vector for controlled transgene expression can be used to express an antisense nucleic acid or functional RNA to induce knockdown of a target gene. As a non-limiting example, expression of CXCR4 and CCR5, HIV receptors, have been successfully ablated in primary human T-cells, See Schumann *et al.* (2015), *PNAS* 112(33): 10437-10442, herein incorporated by reference in its entirety. Another gene for targeted inhibition is PD-1, where the ceDNA vector can express an inhibitory nucleic acid or RNAi or functional RNA to inhibit the expression of PD-1. PD-1 expresses an immune checkpoint cell surface receptor on chronically active T cells that happens in malignancy. See Schumann et al. *supra.*

[00442] In some embodiments, a ceDNA vectors is useful for correcting a defective gene by expressing a transgene that targets the diseased gene. Non-limiting examples of diseases or disorders amenable to treatment by a ceDNA vector as disclosed herein, are listed in Tables A-C along with their and their associated genes of US patent publication 2014/0170753, which is herein incorporated by reference in its entirety.

[00443] In alternative embodiments, the ceDNA vectors are used for insertion of an expression cassette for expression of a therapeutic protein or reporter protein in a safe harbor gene, *e.g.,* in an inactive intron. In certain embodiments, a promoter-less cassette is inserted into the safe harbor gene. In such embodiments, a promoter-less cassette can take advantage of the safe harbor gene regulatory elements (promoters, enhancers, and signaling peptides), a non-limiting example of insertion at the safe harbor locus is insertion into to the albumin locus that is described in *Blood* (2015) 126 (15): 1777-1784, which is incorporated herein by reference in its entirety. Insertion into Albumin has the benefit of enabling secretion of the transgene into the blood (See *e.g.,* Example 22). In addition, a genomic safe harbor site can be determined using techniques known in the art and described in, for example, Papapetrou, ER & Schambach, A. *Molecular Therapy* 24(4):678-684 (2016) or Sadelain et al. *Nature Reviews Cancer* 12:51-58 (2012), the contents of each of which are incorporated herein by reference in their entirety. It is specifically contemplated herein that safe harbor sites in an adeno associated virus (AAV) genome (*e.g.,* AAVS1 safe harbor site) can be used with the methods and compositions described herein (see *e.g.,* Oceguera-Yanez et al. *Methods* 101:43-55 (2016) or Tiyaboonchai, A et al. *Stem Cell Res* 12(3):630-7

(2014), the contents of each of which are incorporated by reference in their entirety). For example, the AAVS1 genomic safe harbor site can be used with the ceDNA vectors and compositions as described herein for the purposes of hematopoietic specific transgene expression and gene silencing in embryonic stem cells (*e.g.,* human embryonic stem cells) or induced pluripotent stem cells (iPS cells). In addition, it is contemplated herein that synthetic or commercially available homology-directed repair donor templates for insertion into an AASV1 safe harbor site on chromosome 19 can be used with the ceDNA vectors or compositions as described herein. For example, homology-directed repair templates, and guide RNA, can be purchased commercially, for example, from System Biosciences, Palo Alto, CA, and cloned into a ceDNA vector.

[00444] In some embodiments, the ceDNA vectors are used for expressing a transgene, or knocking out or decreasing expression of a target gene in a T cell, *e.g.,* to engineer the T cell for improved adoptive cell transfer and/or CAR-T therapies (see, *e.g.,* Example 24). In some embodiments, the ceDNA vector for controlled transgene expression as described herein can express transgenes that knock-out genes. Non-limiting examples of therapeutically relevant knock-outs of T cells are described in PNAS (2015) 112(33):10437-10442, which is incorporated herein by reference in its entirety.

C. Additional diseases for gene therapy:

[00445] In general, the ceDNA vector for controlled transgene expression as disclosed herein can be used to deliver any transgene in accordance with the description above to treat, prevent, or ameliorate the symptoms associated with any disorder related to gene expression. Illustrative disease states include, but are not-limited to: cystic fibrosis (and other diseases of the lung), hemophilia A, hemophilia B, thalassemia, anemia and other blood disorders, AIDS, Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, epilepsy, and other neurological disorders, cancer, diabetes mellitus, muscular dystrophies (e.g., Duchenne, Becker), Hurler's disease, adenosine deaminase deficiency, metabolic defects, retinal degenerative diseases (and other diseases of the eye), mitochondriopathies (e.g., Leber's hereditary optic neuropathy (LHON), Leigh syndrome, and subacute sclerosing encephalopathy), myopathies (e.g., facioscapulohumeral myopathy (FSHD) and cardiomyopathies), diseases of solid organs (e.g., brain, liver, kidney, heart), and the like. In some embodiments, the ceDNA vectors as disclosed herein can be advantageously used in the treatment of individuals with metabolic disorders (e.g., ornithine transcarbamylase deficiency).

[00446] In some embodiments, the ceDNA vector for controlled transgene expression described herein can be used to treat, ameliorate, and/or prevent a disease or disorder caused by mutation in a gene or gene product. Exemplary diseases or disorders that can be treated with a ceDNA vectors include, but are not limited to, metabolic diseases or disorders (e.g., Fabry disease, Gaucher disease, phenylketonuria

(PKU), glycogen storage disease); urea cycle diseases or disorders (e.g., ornithine transcarbamylase (OTC) deficiency); lysosomal storage diseases or disorders (e.g., metachromatic leukodystrophy (MLD), mucopolysaccharidosis Type II (MPSII; Hunter syndrome)); liver diseases or disorders (e.g., progressive familial intrahepatic cholestasis (PFIC); blood diseases or disorders (e.g., hemophilia (A and B), thalassemia, and anemia); cancers and tumors, and genetic diseases or disorders (e.g., cystic fibrosis).

[00447] As still a further aspect, a ceDNA vector for controlled transgene expression as disclosed herein may be employed to deliver a heterologous nucleotide sequence in situations in which it is desirable to regulate the level of transgene expression (e.g., transgenes encoding hormones or growth factors, as described herein).

[00448] Accordingly, in some embodiments, the ceDNA vector for controlled transgene expression described herein can be used to correct an abnormal level and/or function of a gene product (e.g., an absence of, or a defect in, a protein) that results in the disease or disorder. The ceDNA vector can produce a functional protein and/or modify levels of the protein to alleviate or reduce symptoms resulting from, or confer benefit to, a particular disease or disorder caused by the absence or a defect in the protein. For example, treatment of OTC deficiency can be achieved by producing functional OTC enzyme; treatment of hemophilia A and B can be achieved by modifying levels of Factor VIII, Factor IX, and Factor X; treatment of PKU can be achieved by modifying levels of phenylalanine hydroxylase enzyme; treatment of Fabry or Gaucher disease can be achieved by producing functional alpha galactosidase or beta glucocerebrosidase, respectively; treatment of MLD or MPSII can be achieved by producing functional arylsulfatase A or iduronate-2-sulfatase, respectively; treatment of cystic fibrosis can be achieved by producing functional cystic fibrosis transmembrane conductance regulator; treatment of glycogen storage disease can be achieved by restoring functional G6Pase enzyme function; and treatment of PFIC can be achieved by producing functional ATP8B1, ABCB11, ABCB4, or TJP2 genes.

[00449] In alternative embodiments, the ceDNA vectors as disclosed herein can be used to provide an antisense nucleic acid to a cell *in vitro* or *in vivo*. For example, where the transgene is a RNAi molecule, expression of the antisense nucleic acid or RNAi in the target cell diminishes expression of a particular protein by the cell. Accordingly, transgenes which are RNAi molecules or antisense nucleic acids may be administered to decrease expression of a particular protein in a subject in need thereof. Antisense nucleic acids may also be administered to cells *in vitro* to regulate cell physiology, e.g., to optimize cell or tissue culture systems.

[00450] In some embodiments, exemplary transgenes encoded by the ceDNA vector for controlled transgene expression include, but are not limited to: X, lysosomal enzymes (e.g., hexosaminidase A, associated with Tay-Sachs disease, or iduronate sulfatase, associated, with Hunter Syndrome/MPS II), erythropoietin, angiostatin, endostatin, superoxide dismutase, globin, leptin, catalase, tyrosine

hydroxylase, as well as cytokines (e.g., a interferon, β-interferon, interferon-γ, interleukin-2, interleukin-4, interleukin 12, granulocyte-macrophage colony stimulating factor, lymphotoxin, and the like), peptide growth factors and hormones (e.g., somatotropin, insulin, insulin-like growth factors 1 and 2, platelet derived growth factor (PDGF), epidermal growth factor (EGF), fibroblast growth factor (FGF), nerve growth factor (NGF), neurotrophic factor-3 and 4, brain-derived neurotrophic factor (BDNF), glial derived growth factor (GDNF), transforming growth factor-α and -β, and the like), receptors (e.g., tumor necrosis factor receptor).

[00451] In some exemplary embodiments, the transgene encodes a monoclonal antibody specific for one or more desired targets. Exemplary ceDNA vectors for controlled expression of antibodies and fusion proteins in the methods as disclosed herein are disclosed in International Application PCT/US19/18016, filed on February 14, 2019, which is incorporated herein in its entirety by reference.

[00452] In some exemplary embodiments, more than one transgene is encoded by the ceDNA vector. In some exemplary embodiments, the transgene encodes a fusion protein comprising two different polypeptides of interest. In some embodiments, the transgene encodes an antibody, including a full-length antibody or antibody fragment, as defined herein. In some embodiments, the antibody is an antigen-binding domain or an immunoglobulin variable domain sequence, as that is defined herein. Other illustrative transgene sequences encode suicide gene products (thymidine kinase, cytosine deaminase, diphtheria toxin, cytochrome P450, deoxycytidine kinase, and tumor necrosis factor), proteins conferring resistance to a drug used in cancer therapy, and tumor suppressor gene products.

[00453] In a representative embodiment, the transgene expressed by the ceDNA vector for controlled transgene expression can be used for the treatment of muscular dystrophy in a subject in need thereof, the method comprising: administering a treatment-, amelioration- or prevention-effective amount of ceDNA vector described herein, wherein the ceDNA vector comprises a heterologous nucleic acid encoding dystrophin, a mini-dystrophin, a micro-dystrophin, myostatin propeptide, follistatin, activin type II soluble receptor, IGF-1, anti-inflammatory polypeptides such as the Ikappa B dominant mutant, sarcospan, utrophin, a micro-dystrophin, laminin-α2, α-sarcoglycan, β-sarcoglycan, γ-sarcoglycan, δ-sarcoglycan, IGF-1, an antibody or antibody fragment against myostatin or myostatin propeptide, and/or RNAi against myostatin. In particular embodiments, the ceDNA vector can be administered to skeletal, diaphragm and/or cardiac muscle as described elsewhere herein.

[00454] In some embodiments, the ceDNA vector for controlled transgene expression can be used to deliver a transgene to skeletal, cardiac or diaphragm muscle, for production of a polypeptide (e.g., an enzyme) or functional RNA (e.g., RNAi, microRNA, antisense RNA) that normally circulates in the blood or for systemic delivery to other tissues to treat, ameliorate, and/or prevent a disorder (e.g., a metabolic disorder, such as diabetes (e.g., insulin), hemophilia (e.g., VIII), a mucopolysaccharide disorder

(e.g., Sly syndrome, Hurler Syndrome, Scheie Syndrome, Hurler-Scheie Syndrome, Hunter's Syndrome, Sanfilippo Syndrome A, B, C, D, Morquio Syndrome, Maroteaux-Lamy Syndrome, etc.) or a lysosomal storage disorder (such as Gaucher's disease [glucocerebrosidase], Pompe disease [lysosomal acid .alpha.-glucosidase] or Fabry disease [.alpha.-galactosidase A]) or a glycogen storage disorder (such as Pompe disease [lysosomal acid a glucosidase]). Other suitable proteins for treating, ameliorating, and/or preventing metabolic disorders are described above.

[00455] In other embodiments, the ceDNA vector for controlled transgene expression as disclosed herein can be used to deliver a transgene in a method of treating, ameliorating, and/or preventing a metabolic disorder in a subject in need thereof. Illustrative metabolic disorders and transgenes encoding polypeptides are described herein. Optionally, the polypeptide is secreted (e.g., a polypeptide that is a secreted polypeptide in its native state or that has been engineered to be secreted, for example, by operable association with a secretory signal sequence as is known in the art).

[00456] Another aspect of the invention relates to a method of treating, ameliorating, and/or preventing congenital heart failure or PAD in a subject in need thereof, the method comprising administering a ceDNA vector for controlled transgene expression as described herein to a mammalian subject, wherein the ceDNA vector comprises a transgene encoding, for example, a sarcoplasmic endoreticulum Ca^{2+}-ATPase (SERCA2a), an angiogenic factor, phosphatase inhibitor I (I-1), RNAi against phospholamban; a phospholamban inhibitory or dominant-negative molecule such as phospholamban S16E, a zinc finger protein that regulates the phospholamban gene, β2-adrenergic receptor, .beta.2-adrenergic receptor kinase (BARK), PI3 kinase, calsarcan, a .beta.-adrenergic receptor kinase inhibitor (βARKct), inhibitor 1 of protein phosphatase 1, S100A1, parvalbumin, adenylyl cyclase type 6, a molecule that effects G-protein coupled receptor kinase type 2 knockdown such as a truncated constitutively active βARKct, Pim-1, PGC-1α, SOD-1, SOD-2, EC-SOD, kallikrein, HIF, thymosin-β4, mir-1, mir-133, mir-206 and/or mir-208.

[00457] The ceDNA vectors 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.

[00458] In some embodiments, the ceDNA vectors can be administered to tissues of the CNS (e.g., brain, eye). In particular embodiments, the ceDNA vectors as disclosed herein may be administered to

treat, ameliorate, or prevent diseases of the CNS, including genetic disorders, neurodegenerative disorders, psychiatric disorders and tumors. Illustrative diseases of the CNS include, but are not limited to Alzheimer's disease, Parkinson's disease, Huntington's disease, Canavan disease, Leigh's disease, Refsum disease, Tourette syndrome, primary lateral sclerosis, amyotrophic lateral sclerosis, progressive muscular atrophy, Pick's disease, muscular dystrophy, multiple sclerosis, myasthenia gravis, Binswanger's disease, trauma due to spinal cord or head injury, Tay Sachs disease, Lesch-Nyan disease, epilepsy, cerebral infarcts, psychiatric disorders including mood disorders (e.g., depression, bipolar affective disorder, persistent affective disorder, secondary mood disorder), schizophrenia, drug dependency (e.g., alcoholism and other substance dependencies), neuroses (e.g., anxiety, obsessional disorder, somatoform disorder, dissociative disorder, grief, post-partum depression), psychosis (e.g., hallucinations and delusions), dementia, paranoia, attention deficit disorder, psychosexual disorders, sleeping disorders, pain disorders, eating or weight disorders (e.g., obesity, cachexia, anorexia nervosa, and bulemia) and cancers and tumors (e.g., pituitary tumors) of the CNS.

[00459] Ocular disorders that may be treated, ameliorated, or prevented with the ceDNA vectors of the invention 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 for controlled transgene expression 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 factors either intraocularly (e.g., in the vitreous) or periocularly (e.g., in the sub-Tenon's region). One or more neurotrophic factors may also be co-delivered, either intraocularly (e.g., intravitreally) or periocularly. 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, Stargardt disease, 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.

[00460] In some embodiments, inflammatory ocular diseases or disorders (e.g., uveitis) can be treated, ameliorated, or prevented by the ceDNA vectors of the invention. One or more anti-inflammatory factors can be expressed by intraocular (e.g., vitreous or anterior chamber) administration of the ceDNA vector for controlled transgene expression as disclosed herein. In other embodiments, ocular diseases or

disorders characterized by retinal degeneration (e.g., retinitis pigmentosa) can be treated, ameliorated, or prevented by the ceDNA vectors of the invention. intraocular (e.g., vitreal administration) of the ceDNA vector as disclosed herein encoding one or more neurotrophic factors can be used to treat such retinal degeneration-based diseases. In some embodiments, diseases or disorders that involve both angiogenesis and retinal degeneration (e.g., age-related macular degeneration) can be treated with the ceDNA vectors of the invention. Age-related macular degeneration can be treated by administering the ceDNA vector as disclosed herein encoding one or more neurotrophic factors intraocularly (e.g., vitreous) and/or one or more anti-angiogenic factors intraocularly or periocularly (e.g., in the sub-Tenon's region). Glaucoma is characterized by increased ocular pressure and loss of retinal ganglion cells. Treatments for glaucoma include administration of one or more neuroprotective agents that protect cells from excitotoxic damage using the ceDNA vector as disclosed herein. Accordingly, such agents include N-methyl-D-aspartate (NMDA) antagonists, cytokines, and neurotrophic factors, can be delivered intraocularly, optionally intravitreally using the ceDNA vector as disclosed herein.

[00461] In other embodiments, the ceDNA vector for controlled transgene expression as disclosed herein may be used to treat seizures, e.g., to reduce the onset, incidence or severity of seizures. The efficacy of a therapeutic treatment for seizures can be assessed by behavioral (e.g., shaking, ticks of the eye or mouth) and/or electrographic means (most seizures have signature electrographic abnormalities). Thus, the ceDNA vector for controlled transgene expression as disclosed herein can also be used to treat epilepsy, which is marked by multiple seizures over time. In one representative embodiment, somatostatin (or an active fragment thereof) is administered to the brain using the ceDNA vector as disclosed herein to treat a pituitary tumor. According to this embodiment, the ceDNA vector as disclosed herein encoding somatostatin (or an active fragment thereof) is administered by microinfusion into the pituitary. Likewise, such treatment can be used to treat acromegaly (abnormal growth hormone secretion from the pituitary). The nucleic acid (e.g., GenBank Accession No. J00306) and amino acid (e.g., GenBank Accession No. P01166; contains processed active peptides somatostatin-28 and somatostatin-14) sequences of somatostatins as are known in the art. In particular embodiments, the ceDNA vector can encode a transgene that comprises a secretory signal as described in U.S. Pat. No. 7,071,172.

[00462] Another aspect of the invention relates to the use of a ceDNA vector for controlled transgene expression as described herein to produce antisense RNA, RNAi or other functional RNA (e.g., a ribozyme) for systemic delivery to a subject *in vivo*. Accordingly, in some embodiments, the ceDNA vector can comprise a transgene that encodes an antisense nucleic acid, a ribozyme (e.g., as described in U.S. Pat. No. 5,877,022), RNAs that affect spliceosome-mediated trans-splicing (see, Puttaraju et al., (1999) Nature Biotech. 17:246; U.S. Pat. No. 6,013,487; U.S. Pat. No. 6,083,702), interfering RNAs (RNAi) that mediate gene silencing (see, Sharp et al., (2000) Science 287:2431) or other non-translated

RNAs, such as "guide" RNAs (Gorman et al., (1998) Proc. Nat. Acad. Sci. USA 95:4929; U.S. Pat. No. 5,869,248 to Yuan et al.), and the like.

[00463] In some embodiments, the ceDNA vector for controlled transgene expression can further also comprise a transgene that encodes 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 comprising a transgene encoding a reporter polypeptide may be used for diagnostic purposes or as markers of the ceDNA vector's activity in the subject to which they are administered.

[00464] In some embodiments, the ceDNA vector for controlled transgene expression can comprise a transgene or a heterologous nucleotide sequence that shares homology with, and recombines with a locus on the host chromosome. This approach may be utilized to correct a genetic defect in the host cell.

[00465] In some embodiments, the ceDNA vector for controlled transgene expression can comprise a transgene that can be used to express an immunogenic polypeptide in a subject, e.g., for vaccination. The transgene may encode any immunogen of interest known in the art including, but not limited to, immunogens from human immunodeficiency virus, influenza virus, gag proteins, tumor antigens, cancer antigens, bacterial antigens, viral antigens, and the like.

[00466] D. Testing for successful gene expression using a ceDNA vector

[00467] Assays well known in the art can be used to test the efficiency of gene delivery by a ceDNA vector can be performed in both *in vitro* and *in vivo* models. Knock-in or knock-out of a desired transgene by ceDNA can be assessed by one skilled in the art by measuring mRNA and protein levels of the desired transgene (*e.g.,* reverse transcription PCR, western blot analysis, and enzyme-linked immunosorbent assay (ELISA)). Nucleic acid alterations by ceDNA (*e.g.,* point mutations, or deletion of DNA regions) can be assessed by deep sequencing of genomic target DNA. In one embodiment, ceDNA comprises a reporter protein that can be used to assess the expression of the desired transgene, 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 gene and/or gene product to determine if gene expression has successfully occurred. For example, it is envisioned that a point mutation in the cystic fibrosis transmembrane conductance regulator gene (CFTR) inhibits the capacity of CFTR to move anions (*e.g.,* Cl^-) through the anion channel, can be corrected by delivering a functional (i.e., non-mutated) CFTR gene to the subject with a ceDNA vector. Following administration of a ceDNA vector, one skilled in the art can assess the capacity for anions to move

through the anion channel to determine if the CFTR gene has been delivered and expressed. One skilled will be able to determine the best test for measuring functionality of a protein *in vitro* or *in vivo.*

[00468] It is contemplated herein that the effects of gene expression of the transgene 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.

[00469] In some embodiments, a transgene 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.

XII. Administration

[00470] Exemplary modes of administration of the ceDNA vector for controlled transgene expression 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), 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., to liver, eye, skeletal muscle, cardiac muscle, diaphragm muscle or brain).

[00471] Administration of the ceDNA vector for controlled transgene expression can be to any site in a subject, including, without limitation, a site selected from the group consisting of the brain, a skeletal muscle, a smooth muscle, the heart, the diaphragm, the airway epithelium, the liver, the kidney, the spleen, the pancreas, the skin, and the eye. Administration of the ceDNA vector for controlled transgene expression can also be to a tumor (e.g., in or near a tumor or a lymph node). 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 transgene in a single vector, or multiple ceDNA vectors (e.g. a ceDNA cocktail).

[00472] Administration of the ceDNA vector for controlled transgene expression disclosed herein to skeletal muscle according to the present invention includes but is not limited to administration to 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 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 certain embodiments, the ceDNA vector for controlled transgene expression as disclosed herein can be administered without employing "hydrodynamic" techniques.

[00473] Administration of the ceDNA vector for controlled transgene expression 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. Administration to diaphragm muscle can be by any suitable method including intravenous administration, intra-arterial administration, and/or intra-peritoneal administration. Administration 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.

[00474] In some embodiments, a ceDNA vector for controlled transgene expression according to the present invention is administered to skeletal muscle, diaphragm muscle and/or cardiac muscle (e.g., to treat, ameliorate and/or prevent muscular dystrophy or heart disease (e.g., PAD or congestive heart failure).

***A. Ex vivo* treatment**

[00475] In some embodiments, cells are removed from a subject, a ceDNA vector 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.

[00476] Cells transduced with a ceDNA vector 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.

[00477] In some embodiments, the ceDNA vector for controlled transgene expression can encode a transgene (sometimes called a heterologous nucleotide sequence) that is any polypeptide that is desirably produced in a cell *in vitro, ex vivo,* or *in vivo*. For example, in contrast to the use of the ceDNA vectors in a method of treatment as discussed herein, in some embodiments the ceDNA vectors may be introduced into cultured cells and the expressed gene product isolated therefrom, e.g., for the production of antigens or vaccines.

[00478] The ceDNA vectors 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.

[00479] One aspect of the technology described herein relates to a method of delivering a transgene to a cell. Typically, for *in vitro* methods, the ceDNA vector for controlled transgene expression may be introduced into the cell using the methods as disclosed herein, as well as other methods known in the art. ceDNA vectors 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 transgene in a target cell.

B. Unit dosage forms

[00480] In some embodiments, the pharmaceutical compositions 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 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 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.

XIII. Various applications

[00481] The compositions and ceDNA vectors provided herein can be used to deliver a transgene for various purposes as described above. In some embodiments, a transgene can encode a protein or be a functional RNA, and in some embodiments, can be a protein or functional RNA that is modified for research purposes, e.g., to create a somatic transgenic animal model harboring one or more mutations or a corrected gene sequence, e.g., to study the function of the target gene. In another example, the transgene encodes a protein or functional RNA to create an animal model of disease.

[00482] In some embodiments, the transgene encodes one or more peptides, polypeptides, or proteins, which are useful for the treatment, amelioration, or prevention of disease states in a mammalian subject. The transgene expressed by the ceDNA vector for controlled transgene expression 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: reduced expression, lack of expression or dysfunction of the target gene.

[00483] In some embodiments, the ceDNA vectors are envisioned for use in diagnostic and screening methods, whereby a transgene is transiently or stably expressed in a cell culture system, or alternatively, a transgenic animal model.

[00484] Another aspect of the technology described herein provides a method of transducing a population of mammalian cells. 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 disclosed herein.

[00485] Additionally, the present invention provides compositions, as well as therapeutic and/or diagnostic kits that include one or more of the disclosed ceDNA vectors or ceDNA compositions, formulated with one or more additional ingredients, or prepared with one or more instructions for their use.

[00486] A cell to be administered the ceDNA vector for controlled transgene expression 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.

[00487] *Some embodiments of the technology described herein can be defined according to any of the following numbered paragraphs:*

1. A composition comprising a ceDNA vector that can be re-dosed to increase the expression level of the transgene from a previous expression level, the ceDNA vector comprising asymmetric, or symmetric ITR sequences flanking a transgene polynucleotide sequence operatively linked to a promoter, wherein the at least one of the ITRs is a replication competent ITR.
2. The composition of paragraph 1, wherein the ceDNA vector expresses the transgene for a time period selected from at least 42 days, at least 84 days, and at least 132 days.
3. The composition of paragraph 1, wherein the transgene is a genetic medicine selected from any of: a nucleic acid, an inhibitor, peptide or polypeptide, antibody or antibody fragment, fusion protein, antigen, antagonist, agonist or RNAi molecule.
4. A method for increasing level of expression of a transgene in a cell, the method comprising:
    a. administering to a cell at a first time point, a priming dose of a composition to achieve expression of a heterologous nucleic acid sequence, and
    b. administering to the cell at a second time point, a dose of a composition to increase the expression level of the heterologous nucleic acid sequence as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the first time point, or to increase the expression level of the heterologous nucleic acid sequence to achieve a desired expression level,

    wherein the composition administered at the first and second time point comprises a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA), wherein the ceDNA vector comprises a heterologous nucleic acid sequence encoding a transgene operatively position between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

    wherein the ceDNA when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

5. The method of paragraph 4, wherein the two asymmetric or symmetric inverted terminal repeat sequences (ITRs) are AAV ITRs.
6. The method of paragraph 4-5, wherein the ITR comprising a comprising a functional terminal resolution site and a Rep binding site is a wild-type AAV ITR.
7. The method of any of paragraphs 4-6, wherein the AAV ITRs is are AAV-2 ITRs.
8. The method of any of paragraphs 4-7, wherein the two asymmetric or symmetric ITRS are a pair of ITRS selected from the group consisting of: (i) SEQ ID NO: 1 and SEQ ID NO: 4; and (ii) SEQ ID NO: 3 and SEQ ID NO: 2.
9. The method of any of paragraphs 4-8, wherein the ceDNA vector is administered in combination with a pharmaceutically acceptable carrier and/or excipient.
10. The method of any of paragraphs 4-9, wherein the second time point is at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months after the first time point.

11. The method of any of paragraphs 4-10, wherein the heterologous nucleic acid sequence encodes a therapeutic transgene and the desired level of expression of the transgene is a therapeutically effective amount.

12. The method of any of paragraphs 4-11, wherein the ceDNA vector is obtained from a vector polynucleotide, wherein the vector polynucleotide encodes a heterologous nucleic acid operatively positioned between two asymmetric or symmetric inverted terminal repeat sequences (ITRs), at least one of the ITRs comprising a functional terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution, relative to the other ITR; the presence of Rep protein inducing replication of the vector polynucleotide and production of the DNA vector in an insect cell, the DNA vector being obtainable from a process comprising the steps of:

a. incubating a population of insect cells harboring the vector polynucleotide, which is devoid of viral capsid coding sequences in the presence of Rep protein under conditions effective and for time sufficient to induce production of the capsid-free, non-viral DNA vector within the insect cells, wherein the insect cells do not comprise production capsid-free, non-viral DNA within the insect cells; and
b. harvesting and isolating the capsid-free, non-viral DNA from the insect cells;

wherein the presence of the capsid-free, non-viral DNA isolated from the insect cells can be confirmed by digesting DNA isolated from the insect cells 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.

13. The method of any of paragraphs 4-11, further comprising administering to the cell, at one or more time points after the second time point, a dose of the composition to increase the expression level of the heterologous nucleic acid sequence as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the second time point or previous time point, or to increase the expression level of the heterologous nucleic acid sequence to achieve a desired expression level,

wherein the composition administered at the one or more time points after the second time point comprises a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA), wherein the ceDNA vector comprises a heterologous nucleic acid sequence encoding a transgene operatively positioned between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

14. The method of any of paragraphs 4-13, wherein the ceDNA vector administered at any of: the first, second or any subsequent time point, is administered in combination with a pharmaceutically acceptable carrier and/or excipient.

15. The method of any of paragraphs 1-14, wherein the ceDNA vector administered at the first, second or any subsequent time point is the same ceDNA vector comprising the same transgene, or a modified transgene.

16. The method of any of paragraphs 1-15, wherein the ceDNA vector administered at the first, second or any subsequent time point is a different ceDNA vector comprising the same transgene, or a modified transgene.

17. The method of paragraph 16, wherein the different ceDNA vector has a different promoter operatively linked to the same transgene, or to a modified transgene.

18. The method of any of paragraphs 1-17, where the transgene is a genetic medicine.

19. A method for treating a disease in a subject, the method comprising:

    a. administering to the subject at a first time point, a priming dose of a composition comprising a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA) to achieve expression of a heterologous nucleic acid sequence, and

    b. administering to the subject at a second time point, a dose of a composition comprising a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA) to increase the expression level of the heterologous nucleic acid sequence as compared to the level of

expression of the heterologous nucleic acid achieved after administration of the composition at the first time point, or to increase the expression level of the heterologous nucleic acid sequence to achieve a desired expression level, thereby treating the disease in the subject,

wherein the ceDNA vector administered at the first and second time point comprises a heterologous nucleic acid sequence encoding a transgene operatively positioned between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA vector when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

20. The method of paragraph 19, wherein, the two asymmetric or symmetric inverted terminal repeat sequences (ITRs) are AAV ITRs.

21. The method of paragraph 19-20, wherein the ITR comprising a comprising a functional terminal resolution site and a Rep binding site is a wild-type AAV ITR.

22. The method of any of paragraphs 19-21, wherein the AAV ITRs is are AAV-2 ITRs.

23. The method of any of paragraphs 19-22, wherein the two asymmetric ITRs are a pair of ITRs selected from the group consisting of: (i) SEQ ID NO: 1 and SEQ ID NO: 4; and (ii) SEQ ID NO: 3 and SEQ ID NO: 2.

24. The method of any of paragraphs 19-23, wherein the ceDNA vector is administered in combination with a pharmaceutically acceptable carrier and/or excipient.

25. The method of any of paragraphs 19-24, wherein the second time point is at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months after the first time point.

26. The method of any of paragraphs 19-25, wherein the desired expression level of transgene achieved after the administration of the composition at the second time point is a therapeutically effective amount of the transgene.

27. The method of any of paragraphs 19-26, further comprising administering to the subject at one or more time points after the second time point, a dose of the composition comprising a ceDNA vector to increase the expression level of the heterologous nucleic acid sequence as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the second

time point, or previous time point, or to increase the expression level of the heterologous nucleic acid sequence to a desired expression level,

wherein the composition administered at the one or more time points after the second time point comprises a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA), wherein the ceDNA vector comprises a heterologous nucleic acid sequence encoding a transgene operatively positioned between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

28. The method of any of paragraphs 19-27, wherein the desired expression level of transgene achieved after the administration of the composition at one or more time points after the second time point is a therapeutically effective amount of the transgene.

29. The method of any of paragraphs 19-28, wherein the ceDNA vector is administered at the first, second or any subsequent time point is administered with a pharmaceutically acceptable carrier.

30. The method of any of paragraphs 19-29, wherein the one or more time points after the second time point is at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months after the proceeding time point.

31. The method of any of paragraphs 19 to 30, wherein the ceDNA vector administered at the first, second or any subsequent time point is the same ceDNA vector comprising the same transgene, or a modified transgene.

32. The method of any of paragraphs 19 to 30, wherein the ceDNA vector administered at the first, second or any subsequent time point is a different ceDNA vector comprising the same transgene, or a modified transgene.

33. The method of paragraph 31, wherein the different ceDNA vector has a different promoter operatively linked to the same transgene, or to a modified transgene.

34. The method of any of paragraphs 19 or 33, wherein the ceDNA vector is obtained from a vector polynucleotide, wherein the vector polynucleotide encodes a heterologous nucleic acid operatively positioned between two asymmetric or symmetric inverted terminal repeat sequences (ITRs), at least one of the ITRs comprising a functional terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution, relative to the other ITR; the presence of Rep protein inducing replication of the vector polynucleotide and production of the DNA vector in an insect cell, the DNA vector being obtainable from a process comprising the steps of:

a. incubating a population of insect cells harboring the vector polynucleotide, which is devoid of viral capsid coding sequences in the presence of Rep protein under conditions effective and for time sufficient to induce production of the capsid-free, non-viral DNA vector within the insect cells, wherein the insect cells do not comprise production capsid-free, non-viral DNA within the insect cells; and
b. harvesting and isolating the capsid-free, non-viral DNA from the insect cells;

wherein the presence of the capsid-free, non-viral DNA isolated from the insect cells can be confirmed by digesting DNA isolated from the insect cells 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.

35. The method of any of paragraphs 19 to 34, where the transgene is a genetic medicine.

36. A composition comprising a ceDNA vector that can be re-dosed to maintain a sustained expression level of the transgene, the ceDNA vector comprising asymmetric or symmetric ITR sequences flanking a transgene polynucleotide sequence operatively linked to a promoter, wherein the at least one of the ITRs is a replication competent ITR.

37. The composition of paragraph 36, wherein the ceDNA vector expresses the transgene for a time period selected from at least 42 days, at least 84 days, and at least 132 days.

38. The composition of paragraph 36, wherein the transgene is a genetic medicine selected from any of: a nucleic acid, an inhibitor, peptide or polypeptide, antibody or antibody fragment, antigen, antagonist, agonist or RNAi molecule.

39. A method for sustaining the level of expression of a transgene in a cell, the method comprising:

a. administering to a cell at a first time point, a priming dose of a composition to achieve expression of a heterologous nucleic acid sequence, and
b. administering to the cell at a second time point, a dose of a composition to compensate for any decrease in expression level of the heterologous nucleic acid sequence after administration of the composition at the first time point,

wherein the composition administered at the first and second time point comprises a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA), wherein the ceDNA vector comprises a heterologous nucleic acid sequence encoding a transgene operatively position between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a

functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

40. The method of paragraph 39, wherein the two asymmetric or symmetric inverted terminal repeat sequences (ITRs) are AAV ITRs.
41. The method of paragraph 39-40, wherein the ITR comprising a functional terminal resolution site and a Rep binding site is a wild-type AAV ITR.
42. The method of any of paragraphs 39-41, wherein the AAV ITRs are AAV-2 ITRs.
43. The method of any of paragraphs 39-42, wherein the two asymmetric ITRs are a pair of ITRs selected from the group consisting of: (i) SEQ ID NO: 1 and SEQ ID NO: 4; and (ii) SEQ ID NO: 3 and SEQ ID NO: 2.
44. The method of any of paragraphs 39-43, wherein the ceDNA vector is administered in combination with a pharmaceutically acceptable carrier and/or excipient.
45. The method of any of paragraphs 39-44, wherein the second time point is at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months after the first time point.
46. method of any of paragraphs 39-45, wherein the heterologous nucleic acid sequence encodes a therapeutic transgene and the sustained level of expression of the transgene is a therapeutically effective amount.
47. The method of any of paragraphs 39-46, wherein the ceDNA vector is obtained from a vector polynucleotide, wherein the vector polynucleotide encodes a heterologous nucleic acid operatively positioned between two asymmetric inverted terminal repeat sequences (ITRs), at least one of the ITRs comprising a functional terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution, relative to the other ITR; the presence of Rep protein inducing replication of the vector polynucleotide and production of the DNA vector in an insect cell, the DNA vector being obtainable from a process comprising the steps of:

    a. incubating a population of insect cells harboring the vector polynucleotide, which is devoid of viral capsid coding sequences in the presence of Rep protein under conditions effective and for time sufficient to induce production of the capsid-free, non-viral DNA vector within the insect cells, wherein the insect cells do not comprise production capsid-free, non-viral DNA within the insect cells; and

b. harvesting and isolating the capsid-free, non-viral DNA from the insect cells;

wherein the presence of the capsid-free, non-viral DNA isolated from the insect cells can be confirmed by digesting DNA isolated from the insect cells 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.

48. The method of any of paragraphs 39-47, further comprising administering to the cell, at one or more time points after the second time point, a further dose of the composition to increase the expression level of the heterologous nucleic acid sequence as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the second time point or previous time point, or to increase the expression level of the heterologous nucleic acid sequence to maintain a desired sustained expression level,

wherein the composition administered at the one or more time points after the second time point comprises a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA), wherein the ceDNA vector comprises a heterologous nucleic acid sequence encoding a transgene operatively positioned between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

49. The method of any of paragraphs 39-48, wherein the ceDNA vector administered at any of: the first, second or any subsequent time point, is administered in combination with a pharmaceutically acceptable carrier and/or excipient.

50. The method of any of paragraphs 39-49, wherein the ceDNA vector administered at the first, second or any subsequent time point is the same ceDNA vector comprising the same transgene, or a modified transgene.

51. The method of any of paragraphs 39-50, wherein the ceDNA vector administered at the first, second or any subsequent time point is a different ceDNA vector comprising the same transgene, or a modified transgene.

52. The method of paragraph 51, wherein the different ceDNA vector has a different promoter operatively linked to the same transgene, or to a modified transgene.

53. The method of any of paragraphs 1-52, where the transgene is a genetic medicine.

54. A method for treating a disease in a subject, the method comprising:

a. administering to the subject at a first time point, a priming dose of a composition comprising a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA) to achieve expression of a heterologous nucleic acid sequence, and

b. administering to the subject at a second time point, a dose of a composition comprising a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA) to maintain the expression level of the heterologous nucleic acid sequence at a desired sustained level as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the first time point, thereby treating the disease in the subject,

wherein the ceDNA vector administered at the first and second time point comprises a heterologous nucleic acid sequence encoding a transgene operatively positioned between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA vector when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

55. The method of paragraph 54, wherein, the two asymmetric or symmetric inverted terminal repeat sequences (ITRs) are AAV ITRs.

56. The method of paragraph 54 or 55, wherein the ITR comprising a comprising a functional terminal resolution site and a Rep binding site is a wild-type AAV ITR.

57. The method of any of paragraphs 54-56, wherein the AAV ITRs is are AAV-2 ITRs.

58. The method of any of paragraphs 54-57, wherein the two asymmetric ITRs are a pair of ITRs selected from the group consisting of: (i) SEQ ID NO: 1 and SEQ ID NO: 4; and (ii) SEQ ID NO: 3 and SEQ ID NO: 2.

59. The method of any of paragraphs 54-58, wherein the ceDNA vector is administered in combination with a pharmaceutically acceptable carrier and/or excipient.

60. The method of any of paragraphs 54-59, wherein the second time point is at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months after the first time point.

61. The method of any of paragraphs 54-60, wherein the desired expression level of transgene achieved after the administration of the composition at the second time point is a therapeutically effective amount of the transgene.

62. The method of any of paragraphs 54-61, further comprising administering to the subject at one or more time points after the second time point, a dose of the composition comprising a ceDNA vector to increase the expression level of the heterologous nucleic acid sequence as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the second time point such that the desired sustained expression level of the heterologous nucleic acid is maintained,

wherein the composition administered at the one or more time points after the second time point comprises a non-viral capsid-free DNA vector with covalently-closed ends (ceDNA), wherein the ceDNA vector comprises a heterologous nucleic acid sequence encoding a transgene operatively positioned between two asymmetric or symmetric AAV inverted terminal repeat sequences (ITRs), one of the ITRs comprising a functional AAV terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution relative to the other ITR,

wherein the ceDNA when digested with a restriction enzyme having a single recognition site on the ceDNA vector has the presence of characteristic bands of linear and continuous DNA as compared to linear and non-continuous DNA controls when analyzed on a non-denaturing gel.

63. The method of any of paragraphs 54-62, wherein the desired expression level of transgene achieved after the administration of the composition at one or more time points after the second time point is a therapeutically effective amount of the transgene.

64. The method of any of paragraphs 54-63, wherein the ceDNA vector is administered at the first, second or any subsequent time point is administered with a pharmaceutically acceptable carrier.

65. The method of any of paragraphs 54-64, wherein the one or more time points after the second time point is at least 30 days, or at least 60 days or between 60-90 days, or between 90-120 days, or between about 3-6 months after the proceeding time point.

66. The method of any of paragraphs 54 to 65, wherein the ceDNA vector administered at the first, second or any subsequent time point is the same ceDNA vector comprising the same transgene, or a modified transgene.

67. The method of any of paragraphs 54 to 65, wherein the ceDNA vector administered at the first, second or any subsequent time point is a different ceDNA vector comprising the same transgene, or a modified transgene.

68. The method of paragraph 67, wherein the different ceDNA vector has a different promoter operatively linked to the same transgene, or to a modified transgene.

69. The method of any of paragraphs 54-68, wherein the ceDNA vector is obtained from a vector polynucleotide, wherein the vector polynucleotide encodes a heterologous nucleic acid operatively positioned between two asymmetric or symmetric inverted terminal repeat sequences (ITRs), at least one of the ITRs comprising a functional terminal resolution site and a Rep binding site, and one of the ITRs comprising a deletion, insertion, or substitution, relative to the other ITR; the presence of Rep protein inducing replication of the vector polynucleotide and production of the DNA vector in an insect cell, the DNA vector being obtainable from a process comprising the steps of:

c. incubating a population of insect cells harboring the vector polynucleotide, which is devoid of viral capsid coding sequences in the presence of Rep protein under conditions effective and for time sufficient to induce production of the capsid-free, non-viral DNA vector within the insect cells, wherein the insect cells do not comprise production capsid-free, non-viral DNA within the insect cells; and
d. harvesting and isolating the capsid-free, non-viral DNA from the insect cells;

wherein the presence of the capsid-free, non-viral DNA isolated from the insect cells can be confirmed by digesting DNA isolated from the insect cells 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.

70. The method of any of paragraphs 54 to 69, where the transgene is a genetic medicine.

EXAMPLES

[00488] 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

[00489] 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.

[00490] 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.

[00491] Production of ceDNA-bacmids:

[00492] 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* (Φ80dlacZΔM15 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.

[00493] 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 μm filter, separating the infectious baculovirus particles from cells or cell debris.

[00494] 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 ~4.0E+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 μm filter.

[00495] 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 2.5E+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.

[00496] A "Rep-plasmid" was produced in a pFASTBAC™-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® 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* (Φ80dlacZΔM15 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.

[00497] 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 $2.5x10^6$ 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.

[00498] ceDNA vector generation and characterization

[00499] 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 (2.5E+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).

[00500] Yields of ceDNA vectors produced and purified from the Sf9 insect cells were initially determined based on UV absorbance at 260nm.

[00501] 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.

[00502] 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.

[00503] 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**).

[00504] 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).

[00505] 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.

[00506] For illustrative 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.

[00507] 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.

[00508] Production of ceDNA-bacmids:

[00509] 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* (Φ80dlacZΔM15 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.

[00510] 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 μm filter, separating the infectious baculovirus particles from cells or cell debris.

[00511] 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 ~4.0E+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 μm filter.

[00512] 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 2.5E+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.

[00513] A "Rep-plasmid" was produced in a pFASTBAC™-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® 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* (Φ80dlacZΔM15 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.

[00514] 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 $2.5x10^6$ 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.

[00515] ceDNA vector generation and characterization

[00516] 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 (2.5E+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).

[00517] 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

[00518] 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).

[00519] In some embodiments, a construct to make a ceDNA vector comprises a regulatory switch as described herein.

[00520] For illustrative purposes, Example 2 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.

[00521] 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.

[00522] 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. Upon ligation a closed-ended DNA vector is formed.

[00523] 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. 3B and 3D), and may have two or more hairpin loops or a single hairpin loop. The hairpin loop modified ITR can be generated by genetic modification of an existing oligo or by de novo biological and/or chemical synthesis.

[00524] 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

[00525] 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.

[00526] As disclosed herein, the ITR oligonucleotides can comprise WT-ITRs or modified ITRs (*e.g.,* see, FIGS. 6A, 6B, 7A and 7B of PCT/US19/14122, which is incorporated herein in its entirity). 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

[00527] 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.

[00528] 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.

[00529] 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.

[00530] 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.

[00531] 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

[00532] 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,

[00533] The following is an exemplary method for confirming the identity of ceDNA vectors.

[00534] ceDNA vectors can be assessed by identified by agarose gel electrophoresis under native or denaturing conditions as illustrated in **FIGS. 4C and 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.

[00535] 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.

[00536] 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 and 4E**).

[00537] 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.

[00538] 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.

[00539] 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) 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 paragams *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.

[00540] 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. The results are shown in **FIG. 6**. This figure is a graph showing luciferase transgene expression as measured by IVIS imaging for at least 132 days after 3 different administration protocols.

[00541] An extension study was performed to investigate the effect of a re-dose, e.g., a re-adminstration 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.

[00542] 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 administrationof 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 expresdsion 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).

[00543] 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. The results are shown in **FIG. 6,** which is a graph showing luciferase transgene expression 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).

[00544] 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. Unexpectedly, the second administration (i.e., re-dose) of the ceDNA vector increased expression by at least 7-fold, even up to 17-fold.

[00545] Unexpectedly, 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.

[00546] 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.

[00547] Taken together, these data in **FIG. 6** show that the expression level of a transgene 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

[00548] 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").

[00549] 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

[00550] 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.

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

[00551] 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*. 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.

[00552] 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/cm^2) 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.

[00553] 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.

[00554] 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. 10A)**.

The approximately 21 day time point of the observed decrease in the wild-type mice corresponds to the timeframe in which an adapative 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. 10B**). 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

[00555] 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.

[00556] 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.

[00557] 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.

[00558] 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. 11). 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.

REFERENCES

[00559] 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

1. A method of regulating expression of a transgene in a subject comprising:
   a. administering to a subject a sufficient amount of a capsid-free close-ended DNA (ceDNA) vector comprising a nucleic acid cassette containing at least one transgene operably linked to a promoter between flanking inverted terminal repeats (ITRs), to express a measurable level of the transgene, wherein the transgene encodes a desired protein to treat a disease; and
   b. titrating the ceDNA vector by administering to the subject at least a second dose of the ceDNA vector comprising the at least one transgene between flanking ITRs to obtain the transgene expression of the desired protein at a predetermined level for a predetermined time or to increase the transgene expression of the desired protein to a predetermined level.

2. The method of claim 1, wherein the subject is assessed after step (a) to determine the titrating dose.

3. The method of claim 2, wherein the assessment is determining the disease state after step (a) or the level of desired protein expressed in the subject.

4. A method of regulating expression of a transgene in a subject comprising:
   a. administering a sufficient amount of a capsid-free close-ended DNA (ceDNA) vector comprising a nucleic acid cassette containing at least one transgene operably linked to a promoter between flanking inverted terminal repeats (ITRs), to the subject to express a measurable level of the transgene, wherein the transgene encodes a desired protein; and
   b. administering to the subject at least a second dose of the ceDNA vector comprising the at least one transgene between flanking ITRs to (i) continue expression of the desired protein at a predetermined level for a predetermined time or (ii) modulate expression of the desired protein to a predetermined level.

5. The method of claim 4, wherein the second administration of the ceDNA vector does not generate an immune reaction sufficient to prevent obtaining the predetermined level of expression of the desired protein.

6. The method of any of claims 1-5, wherein the ceDNA vector is administered in combination with a pharmaceutically acceptable carrier.

7. The method of any of claims 1-6, wherein the second administration is at a time when the level of the expression of the transgene decreases from a desired level.

8. The method of any of claims 1-7, wherein the second administration is at least 90 days after the first administration.

9. The method of any of claims 1-8, wherein the transgene encodes a therapeutic protein and the desired level of expression of the transgene is a therapeutically effective amount of the therapeutic protein.

10. The method of any of claims 1-9, wherein there are at least three administrations of the ceDNA vector, and none of the administrations generate an immune response to the ceDNA vector that prevents the achieving the predetermined level of expression of the desired protein.

11. The method of any of claims 1-10wherein the administrations are on a periodic schedule.

12. The method of any of claims 1-11, wherein the second administration is to increase the level of expression of the desired protein.

13. The method of any of claims 1-12, wherein the second administration is to prolong the expression of the desired protein at a predetermined level of expression.

14. The method of any of claims 1-13, wherein the desired protein is an inhibitor protein.

15. The method of any of claims 1-14, wherein the inhibitor protein is an antibody or fusion protein.

16. The method of any of claims 1-15, wherein the desired protein replaces a defective protein or a protein that is not being expressed.

17. The method of any of claims 1-5, wherein the promoter is an inducible or repressible promoter.

18. The method of any of claims 1-17, wherein the transgene is under the control of a regulatory switch.

19. The method of any of claims 1-18, wherein the ceDNA vector administered at the first, second or any subsequent time point is the same type of ceDNA vector comprising the same transgene, or a modified transgene.

20. The method of any of claims 1-194, wherein the second ceDNA vector has a different promoter operatively linked to the same transgene, or a modified transgene.

21. The method of any of claims 1-20, wherein, the two inverted terminal repeat sequences (ITRs) are AAV ITRs.

22. The method of any of claims 1-21, wherein at least one ITR comprises a functional terminal resolution site and a Rep binding site.

23. The method of claim 22, wherein the AAV ITRs are AAV-2 ITRs.

24. The method of any of claims 1-23, wherein the flanking ITRs are symmetric or asymmetric.

25. The method of any of claims 1-23, wherein the flanking ITRs are symmetrical or substantially symmetrical.

26. The method of any of claims 1-23, wherein the flanking ITRs are asymmetric.

27. The method of any of claims 1-26, wherein one or both of the ITRs are wild type, or wherein both of the ITRs are wild-type.

28. The method of any of claims 1-27, wherein the flanking ITRs are from different viral serotypes.

29. The method of any of claims 1-285, wherein one or both of the ITRs comprises a sequence selected from the sequences in Tables 2, 4A, 4B or 5.

30. The method of any of claims 1-29, wherein at least one of the ITRs is altered from a wild-type AAV ITR sequence by a deletion, addition, or substitution that affects the overall three-dimensional conformation of the ITR.

31. The method of any of claims 1-30, wherein one or both of the ITRs are derived from an AAV serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, and AAV12.

32. The method of any of claims 1-31, wherein one or both of the ITRs are synthetic.

33. The method of any of claims 1-32, wherein one or both of the ITRs is not a wild type ITR, or wherein both of the ITRs are not wild-type.

34. The method of any of claims 1-33, wherein one or both of the ITRs is modified by a deletion, insertion, and/or substitution in at least one of the ITR regions selected from A, A', B, B', C, C', D, and D'.

35. The method of any of claims 1-34, wherein the deletion, insertion, and/or substitution results in the deletion of all or part of a stem-loop structure normally formed by the A, A', B, B' C, or C' regions.

36. The method of any of claims 1-34, wherein one or both of the ITRs are modified by a deletion, insertion, and/or substitution that results in the deletion of all or part of a stem-loop structure normally formed by the B and B' regions.

37. The method of any of claims 1-36, wherein one or both of the ITRs are modified by a deletion, insertion, and/or substitution that results in the deletion of all or part of a stem-loop structure normally formed by the C and C' regions.

38. The method of any of claims 1-37, wherein one or both of the ITRs are modified by a deletion, insertion, and/or substitution that results in the deletion of part of a stem-loop structure normally formed by the B and B' regions and/or part of a stem-loop structure normally formed by the C and C' regions.

39. The method of any of claims 1-38, wherein one or both of the ITRs comprise a single stem-loop structure in the region that normally comprises 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.

40. The method of any of claims 1-39, wherein one or both of the ITRs comprise a single stem and two loops in the region that normally comprises 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.

41. The method of any of claims 1-40, wherein both ITRs are altered in a manner that results in an overall three-dimensional symmetry when the ITRs are inverted relative to each other.

42. The method of any of claims 1-41, wherein at least one heterologous nucleotide sequence is under the control of at least one regulatory switch.

43. The method of claim 42, wherein at least one regulatory switch is selected from a binary regulatory switch, a small molecule regulatory switch, a passcode regulatory switch, a nucleic acid-based regulatory switch, a post-transcriptional regulatory switch, a radiation-controlled or ultrasound controlled regulatory switch, a hypoxia-mediated regulatory switch, an inflammatory response regulatory switch, a shear-activated regulatory switch, and a kill switch.

44. The method of any of claims 1-43, wherein the subject has a disease or disorder selected from cancer, autoimmune disease, a neurodegenerative disorder, hypercholesterolemia, acute organ rejection, multiple sclerosis, post-menopausal osteoporosis, skin conditions, asthma, or hemophilia.

45. The method of claim 44, wherein the cancer is selected from a solid tumor, soft tissue sarcoma, lymphoma, and leukemia.

46. The method of claim 44, wherein the autoimmune disease is selected from rheumatoid arthritis and Crohn's disease.

47. The method of claim 44, wherein the skin condition is selected from psoriasis and atopic dermatitis.

48. The method of claim 44, wherein the neurodegenerative disorder is Alzheimer's disease.

49. The method of claim 44, further comprising administering to the subject, at one or more time points after the second time point, a dose of the composition to increase the expression level of the heterologous nucleic acid sequence as compared to the level of expression of the heterologous nucleic acid achieved after administration of the composition at the second time point or previous time point, or to increase the expression level of the heterologous nucleic acid sequence to achieve a desired expression level.

50. The method of any of claims 1-49, wherein the transgene is a genetic medicine selected from any of: a nucleic acid, an inhibitor, peptide or polypeptide, antibody or antibody fragment, fusion protein, antigen, antagonist, agonist, RNAi molecule, etc.

51. The method of any of claims 1-50, wherein the predetermined dose of the composition administered at the second or any subsequent time point, is in an amount that is between 2-fold and 10-fold the dose of the composition administered at the first time point.

52. The method of any of claims 1-51, wherein the predetermined dose of the composition administered at the second or any subsequent time point, is in an amount that increases the expression of the transgene by at least 3-fold, or at least 5-fold, or least 10-fold, or between 2-15 fold or 2-20 fold as compared the expression of the transgene after administration of the composition at the first time point.

53. The method of any of claims 1-52, wherein the predetermined dose of the ceDNA vector administered at the second time point is determined using a dose-dependent relationship for the ceDNA vector to achieve the desired level of expression of the transgene in the cell.

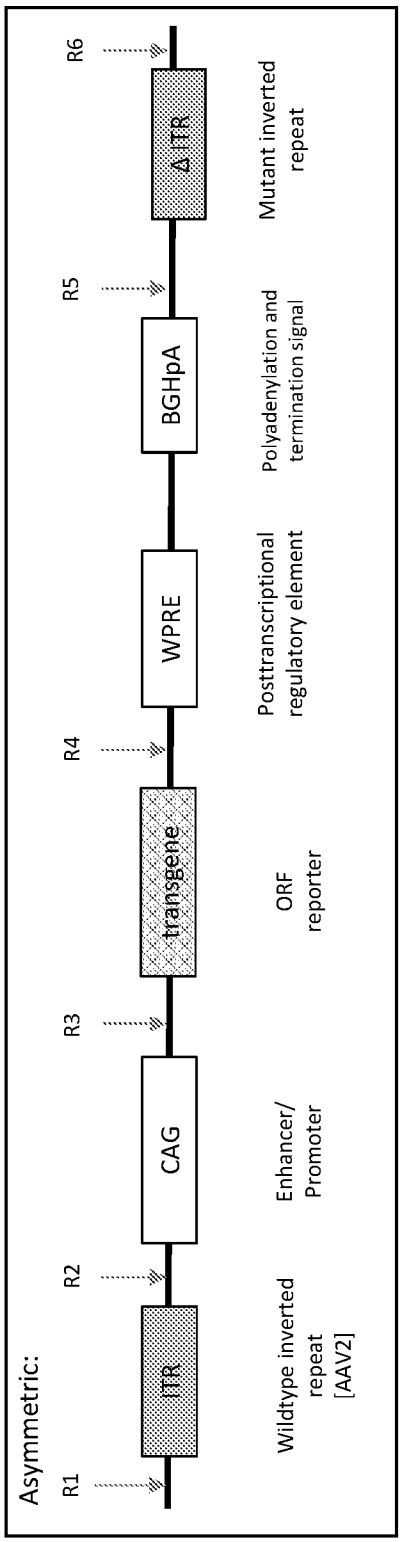


FIG. 1A

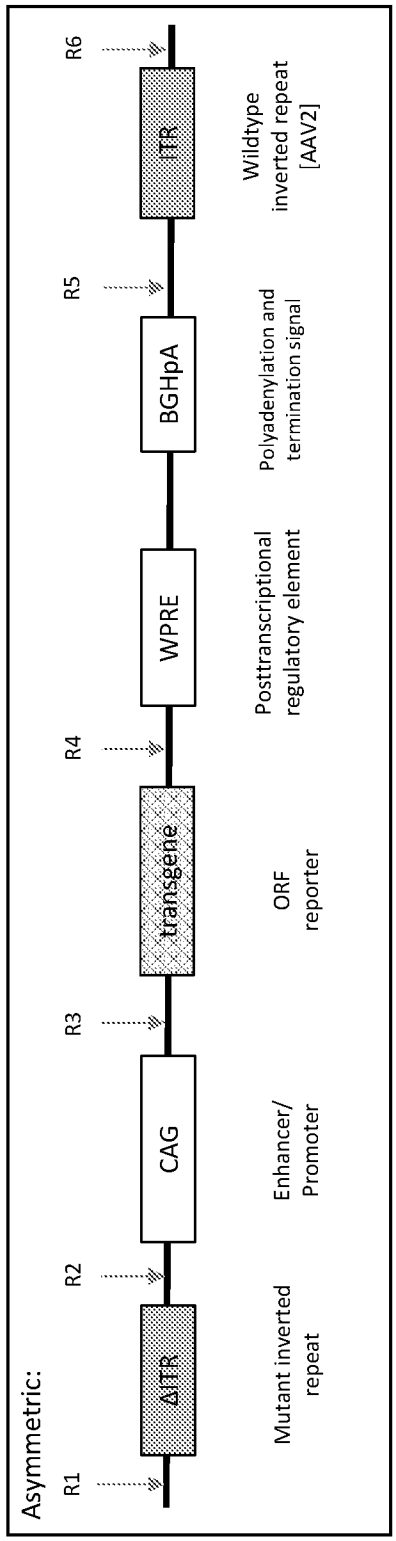


FIG. 1B

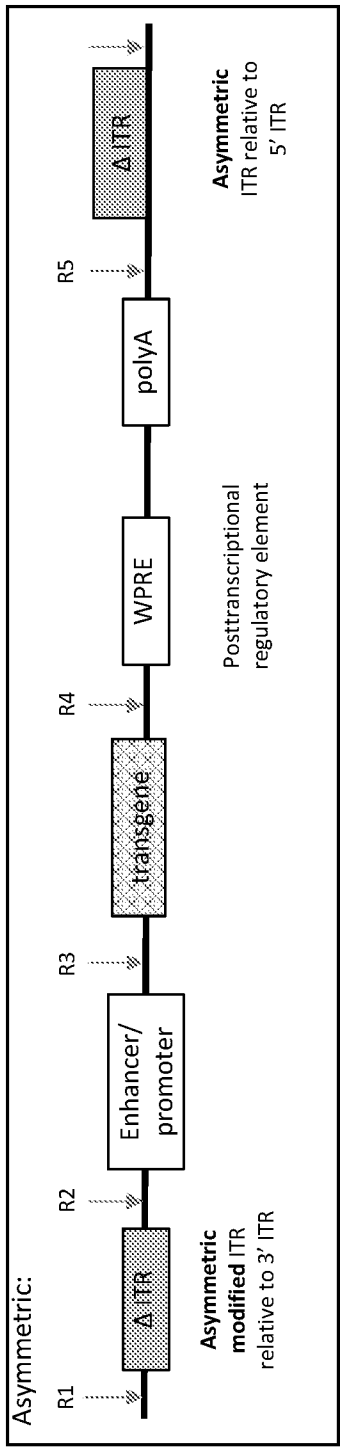


FIG. 1C

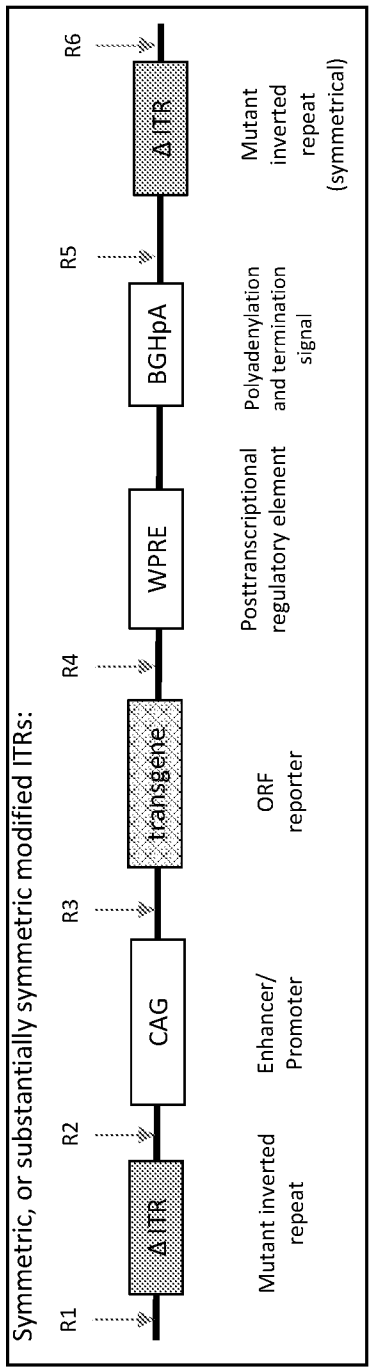

Symmetric, or substantially symmetric modified ITRs:
R1
R2
R3
R4
R5
R6
Δ ITR
CAG
transgene
WPRE
BGHpA
Δ ITR
Mutant inverted repeat
Enhancer/ Promoter
ORF reporter
Posttranscriptional regulatory element
Polyadenylation and termination signal
Mutant inverted repeat (symmetrical)

FIG. 1D

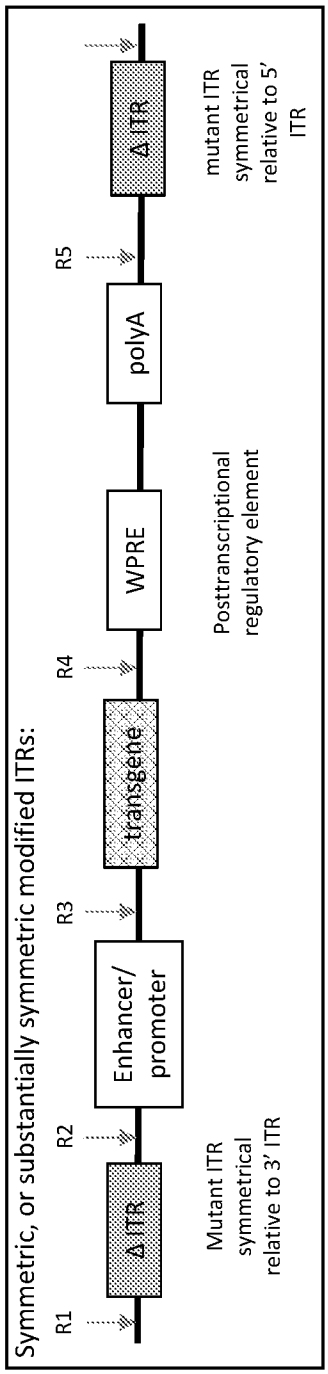

Symmetric, or substantially symmetric modified ITRs:
R1
R2
R3
R4
R5
Δ ITR
Enhancer/ promoter
transgene
WPRE
polyA
Δ ITR
Mutant ITR symmetrical relative to 3' ITR
Posttranscriptional regulatory element
mutant ITR symmetrical relative to 5' ITR

FIG. 1E

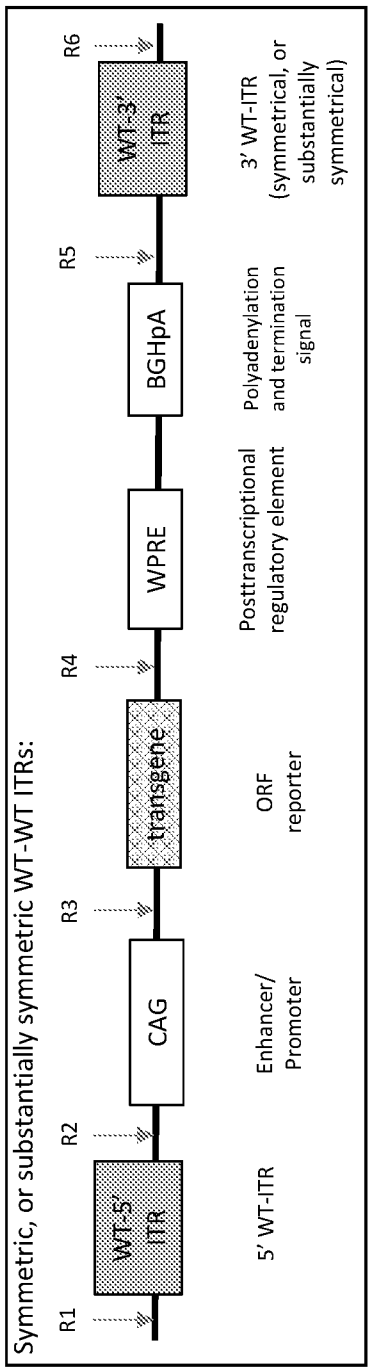

Symmetric, or substantially symmetric WT-WT ITRs:
R1
R2
R3
R4
R5
R6
WT-5' ITR
CAG
transgene
WPRE
BGHpA
WT-3' ITR
5' WT-ITR
Enhancer/ Promoter
ORF reporter
Posttranscriptional regulatory element
Polyadenylation and termination signal
3' WT-ITR (symmetrical, or substantially symmetrical)

FIG. 1F

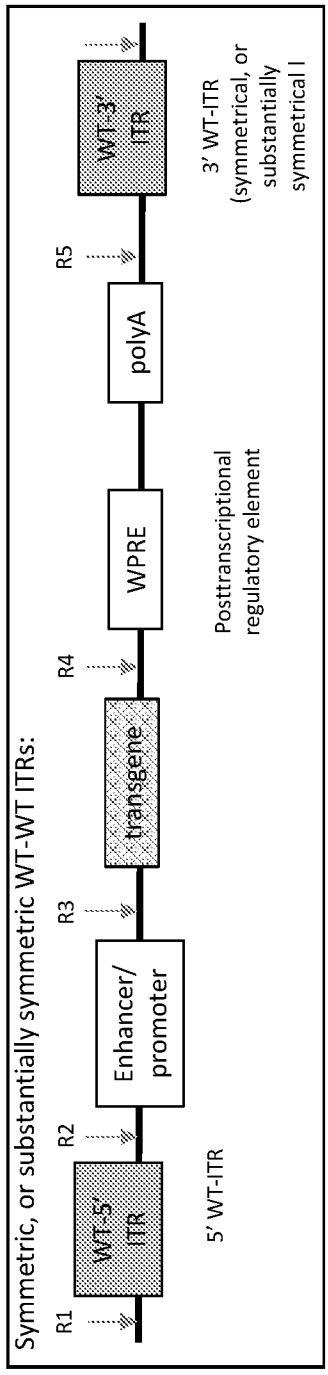

Symmetric, or substantially symmetric WT-WT ITRs:
R1
R2
R3
R4
R5
WT-5' ITR
Enhancer/ promoter
transgene
WPRE
polyA
WT-3' ITR
5' WT-ITR
Posttranscriptional regulatory element
3' WT-ITR (symmetrical, or substantially symmetrical I

FIG. 1G

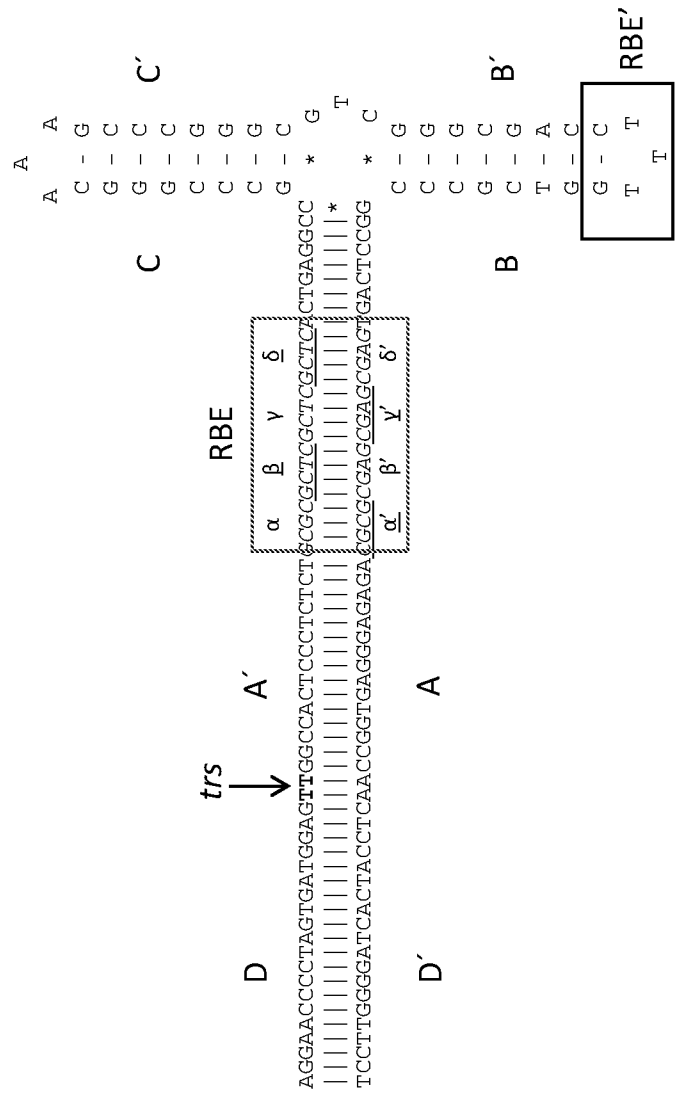


FIG. 2A

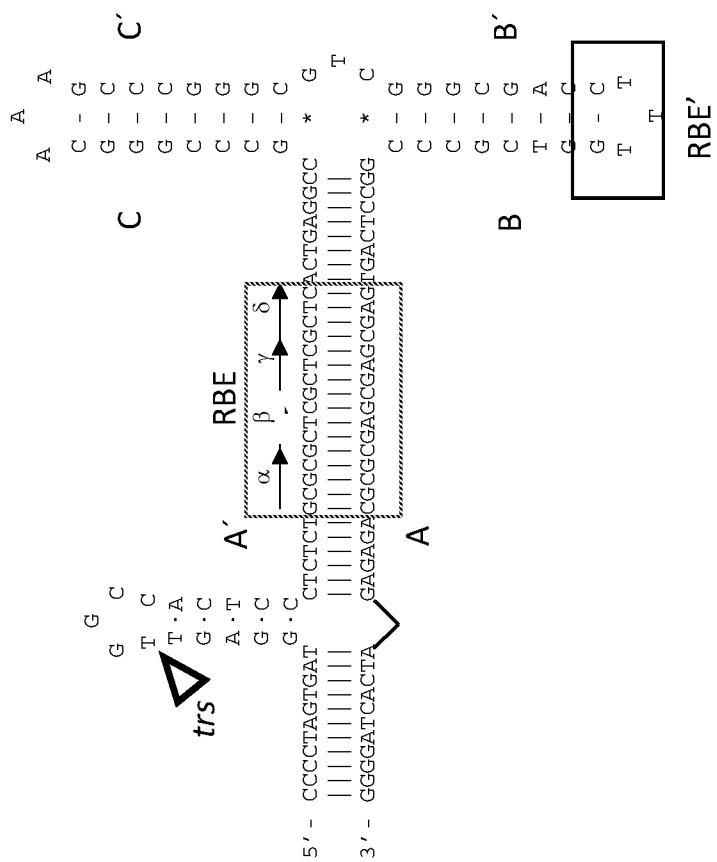

trs
RBE
A
A'
B
B'
C
C'
RBE'
α
β
γ
δ
5' - CCCCTAGTGAT
3' - GGGGATCACTA
CTCTCTGCGCGCTCGCTCGCTCACTGAGGCC
GAGAGACGCGCGAGCGAGCGAGTGACTCCGG

FIG. 2B

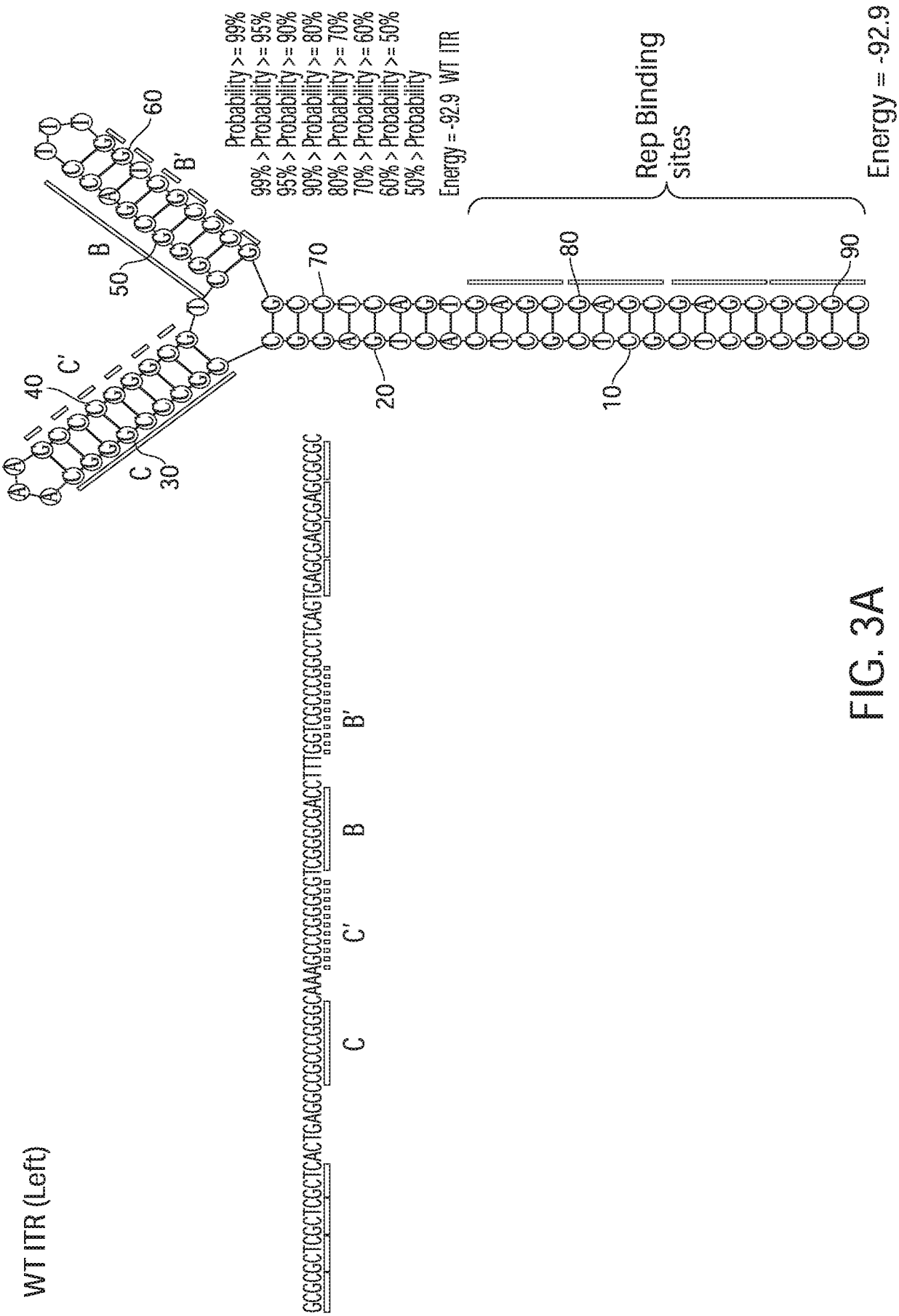

WT ITR (Left)
GCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC
C
C'
B
B'
Probability >= 99%
99% > Probability >= 95%
95% > Probability >= 90%
90% > Probability >= 80%
80% > Probability >= 70%
70% > Probability >= 60%
60% > Probability >= 50%
50% > Probability
Energy = -92.9 WT ITR
Rep Binding sites
Energy = -92.9
10
20
30
40
50
60
70
80
90
A
B
B'
C
C'

FIG. 3A

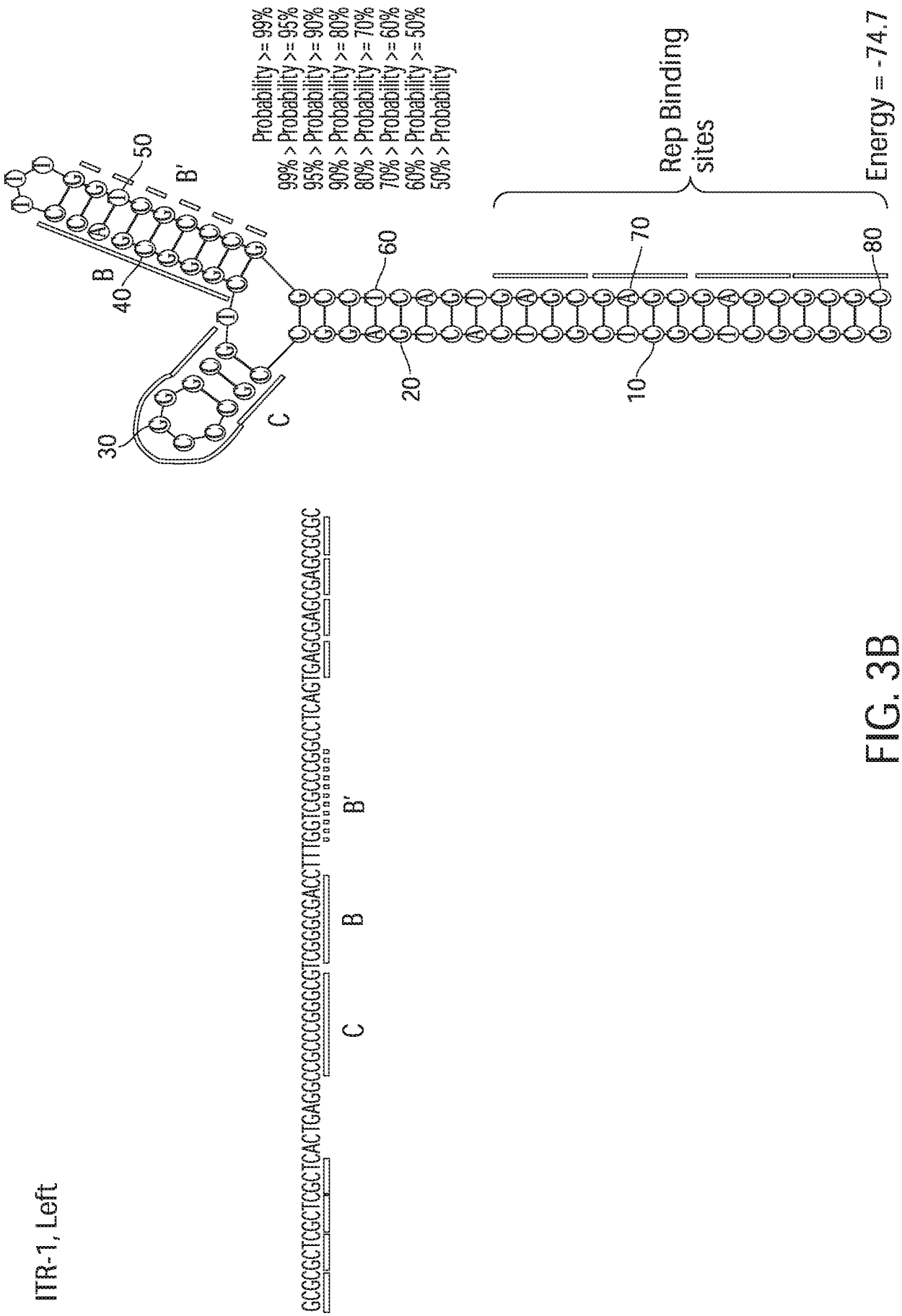

ITR-1, Left
GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC
C
B
B'
Probability >= 99%
99% > Probability >= 95%
95% > Probability >= 90%
90% > Probability >= 80%
80% > Probability >= 70%
70% > Probability >= 60%
60% > Probability >= 50%
50% > Probability
Rep Binding sites
Energy = -74.7
10
20
30
40
50
60
70
80
B
B'
C

FIG. 3B

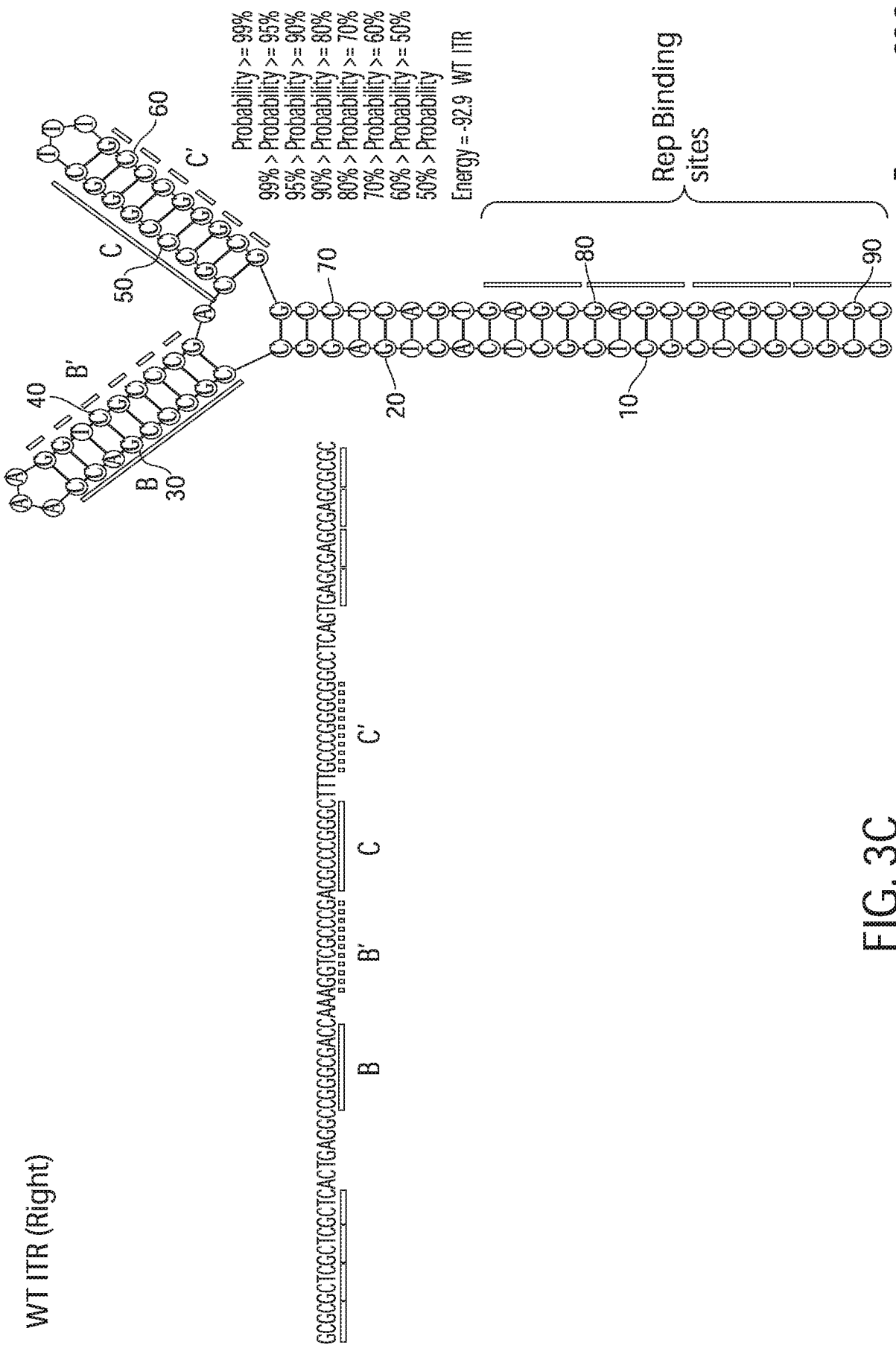

WT ITR (Right)
GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC
B
B'
C
C'
Probability >= 99%
99% > Probability >= 95%
95% > Probability >= 90%
90% > Probability >= 80%
80% > Probability >= 70%
70% > Probability >= 60%
60% > Probability >= 50%
50% > Probability
Energy = -92.9 WT ITR
Rep Binding sites
Energy = -92.9
10
20
30
40
50
60
70
80
90

FIG. 3C

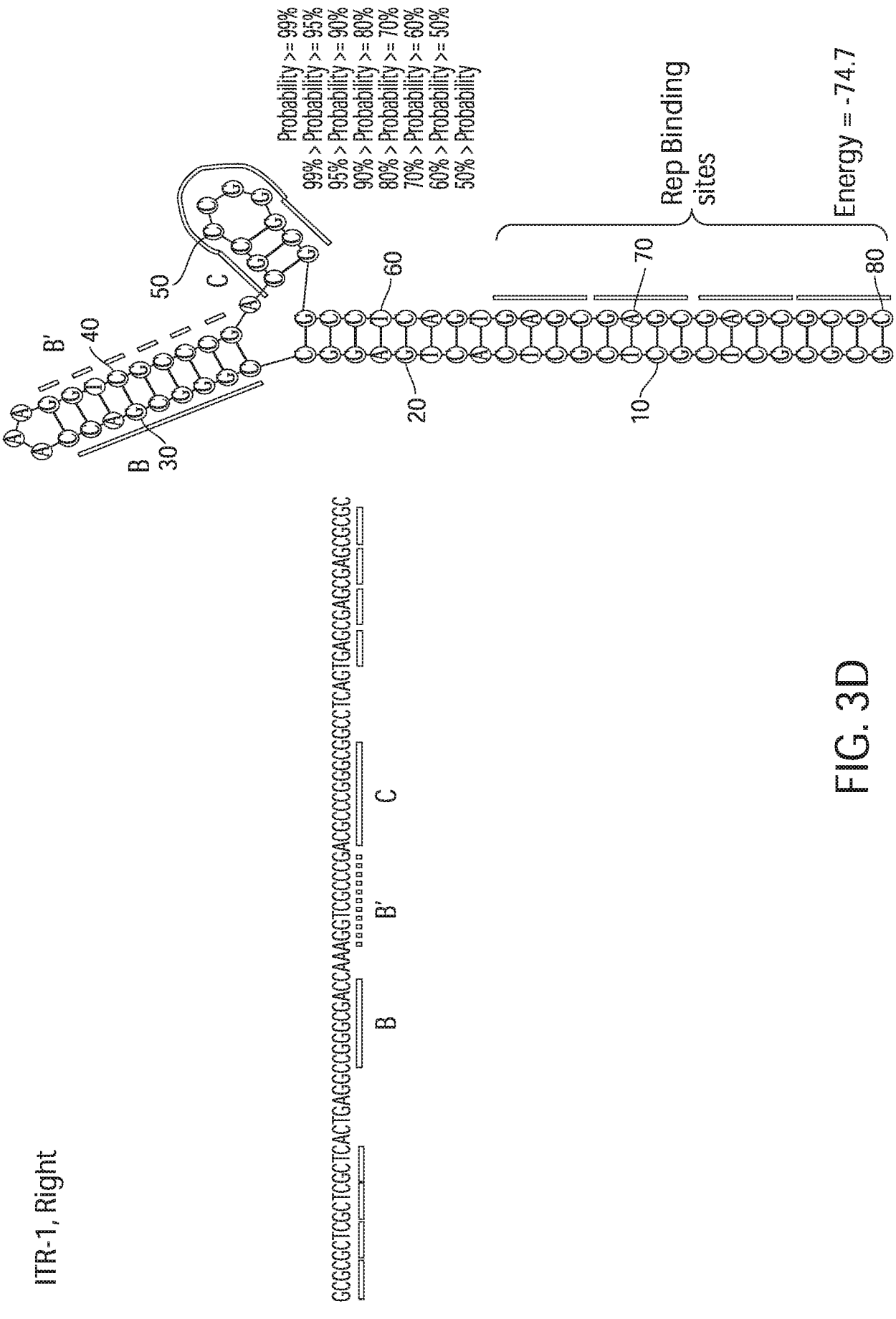

ITR-1, Right
GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC
B
B'
C
Probability >= 99%
99% > Probability >= 95%
95% > Probability >= 90%
90% > Probability >= 80%
80% > Probability >= 70%
70% > Probability >= 60%
60% > Probability >= 50%
50% > Probability
Rep Binding sites
Energy = -74.7
10
20
30
40
50
60
70
80
B
B'
C

FIG. 3D

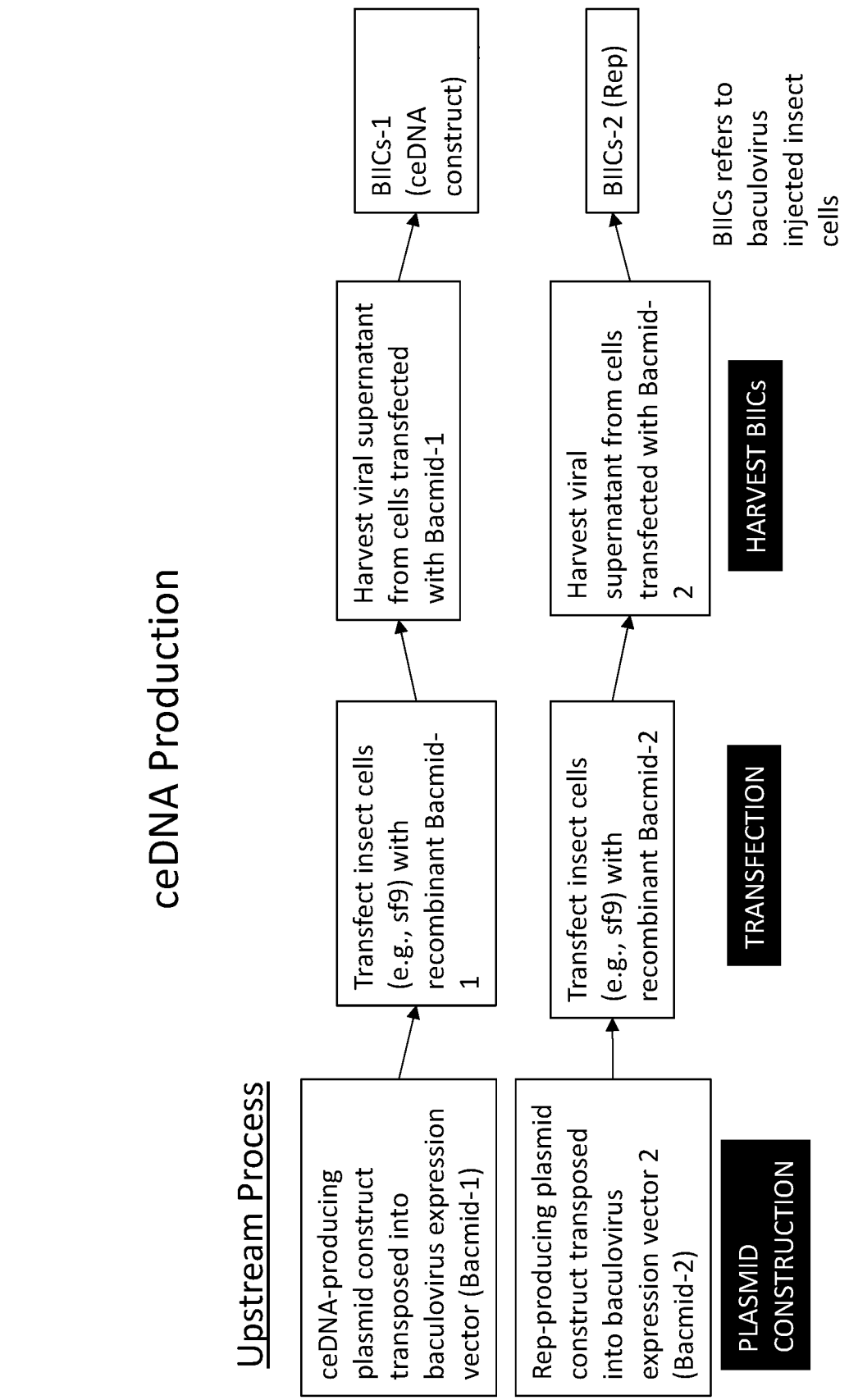

ceDNA Production
Upstream Process
ceDNA-producing plasmid construct transposed into baculovirus expression vector (Bacmid-1)
Rep-producing plasmid construct transposed into baculovirus expression vector 2 (Bacmid-2)
PLASMID CONSTRUCTION
Transfect insect cells (e.g., sf9) with recombinant Bacmid-1
Transfect insect cells (e.g., sf9) with recombinant Bacmid-2
TRANSFECTION
Harvest viral supernatant from cells transfected with Bacmid-1
Harvest viral supernatant from cells transfected with Bacmid-2
HARVEST BIICs
BIICs-1 (ceDNA construct)
BIICs-2 (Rep)
BIICs refers to baculovirus injected insect cells

FIG. 4A

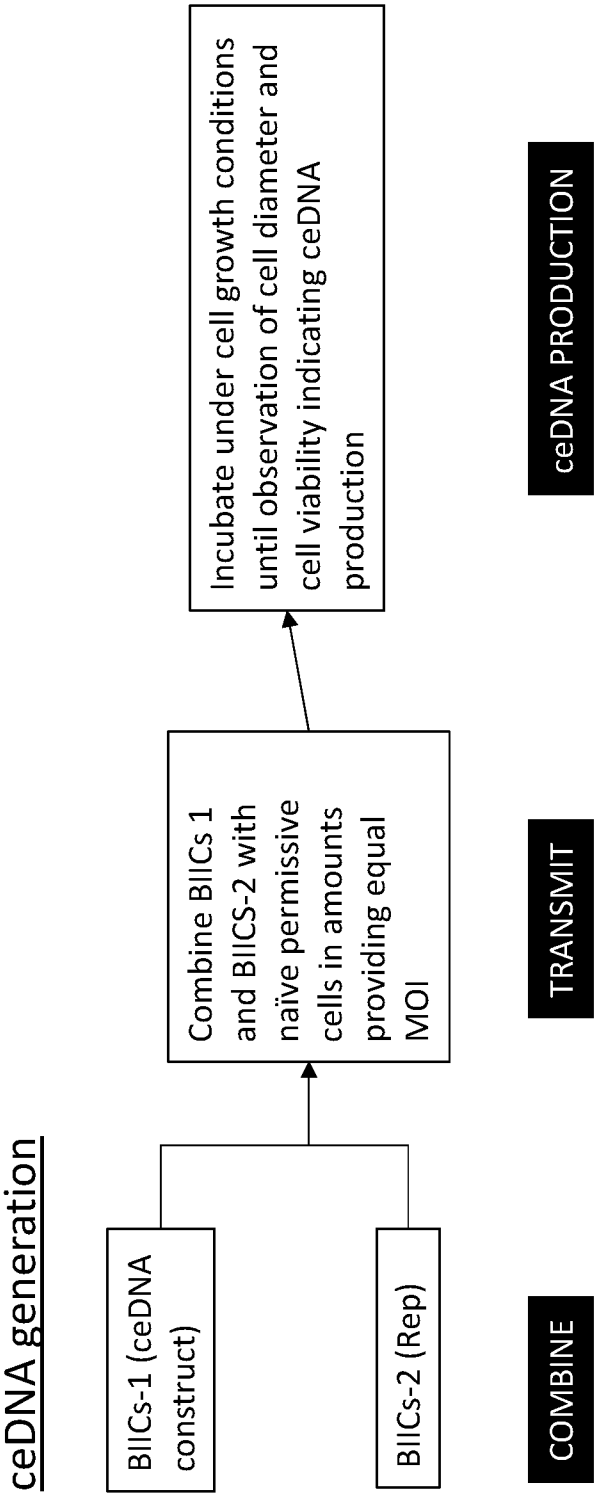

ceDNA Production
ceDNA generation
BIICs-1 (ceDNA construct)
BIICs-2 (Rep)
COMBINE
Combine BIICs 1 and BIICS-2 with naïve permissive cells in amounts providing equal MOI
TRANSMIT
Incubate under cell growth conditions until observation of cell diameter and cell viability indicating ceDNA production
ceDNA PRODUCTION

FIG. 4B

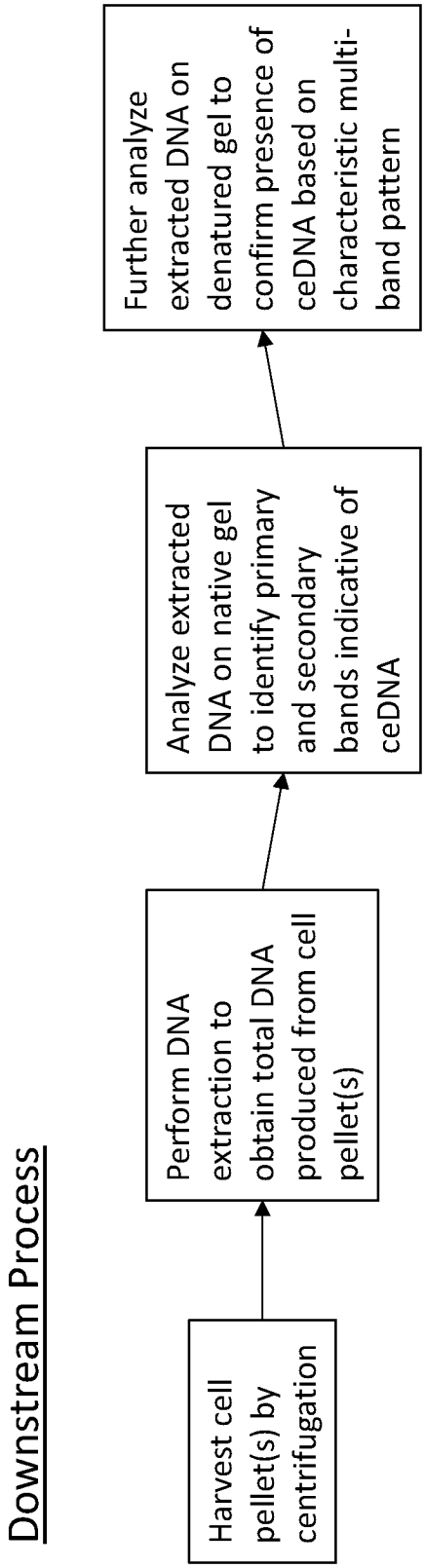

ceDNA Production
Downstream Process
Harvest cell pellet(s) by centrifugation
Perform DNA extraction to obtain total DNA produced from cell pellet(s)
Analyze extracted DNA on native gel to identify primary and secondary bands indicative of ceDNA
Further analyze extracted DNA on denatured gel to confirm presence of ceDNA based on characteristic multi-band pattern
HARVEST
EXTRACT
Confirm ceDNA Production

FIG. 4C

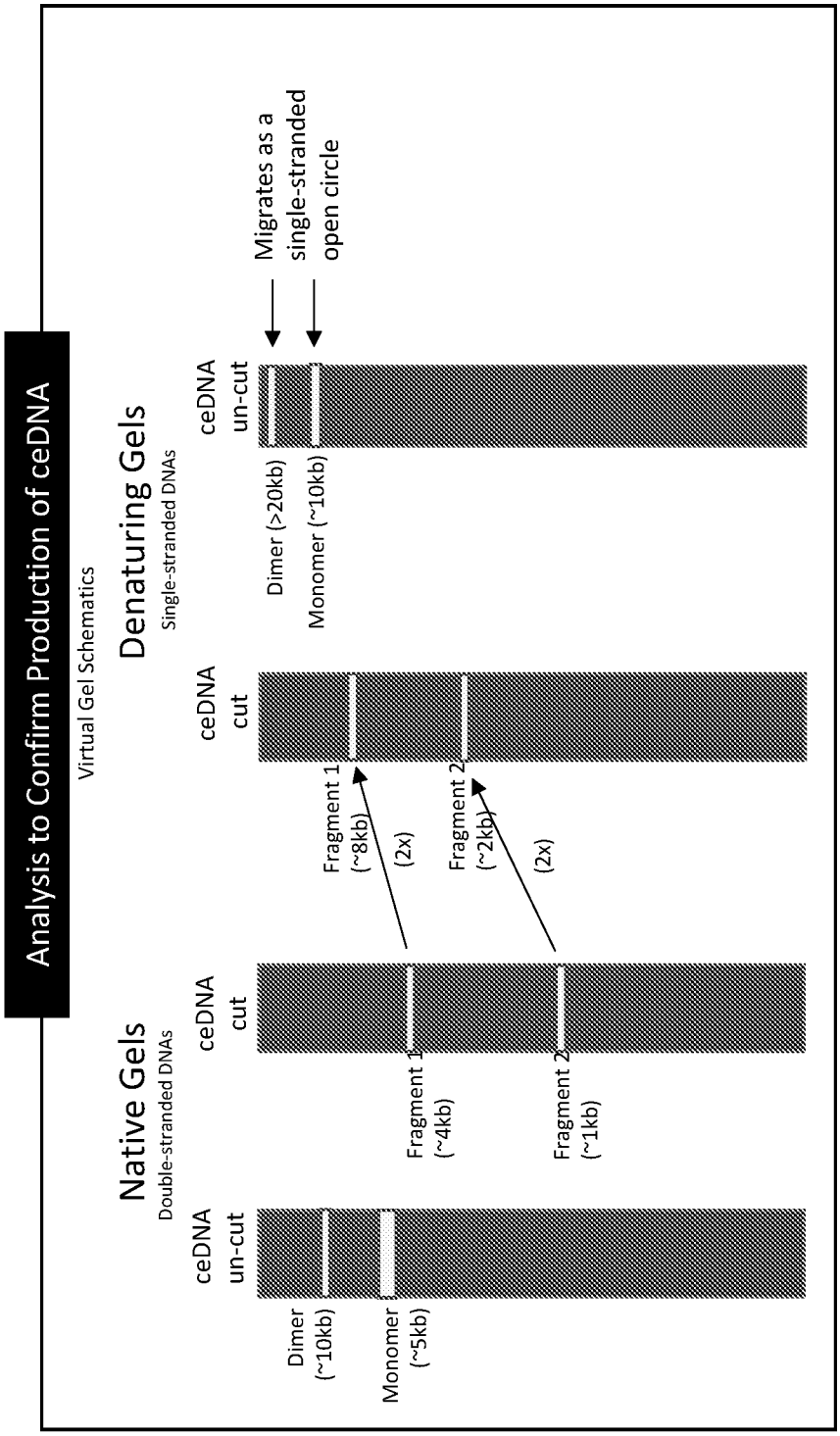
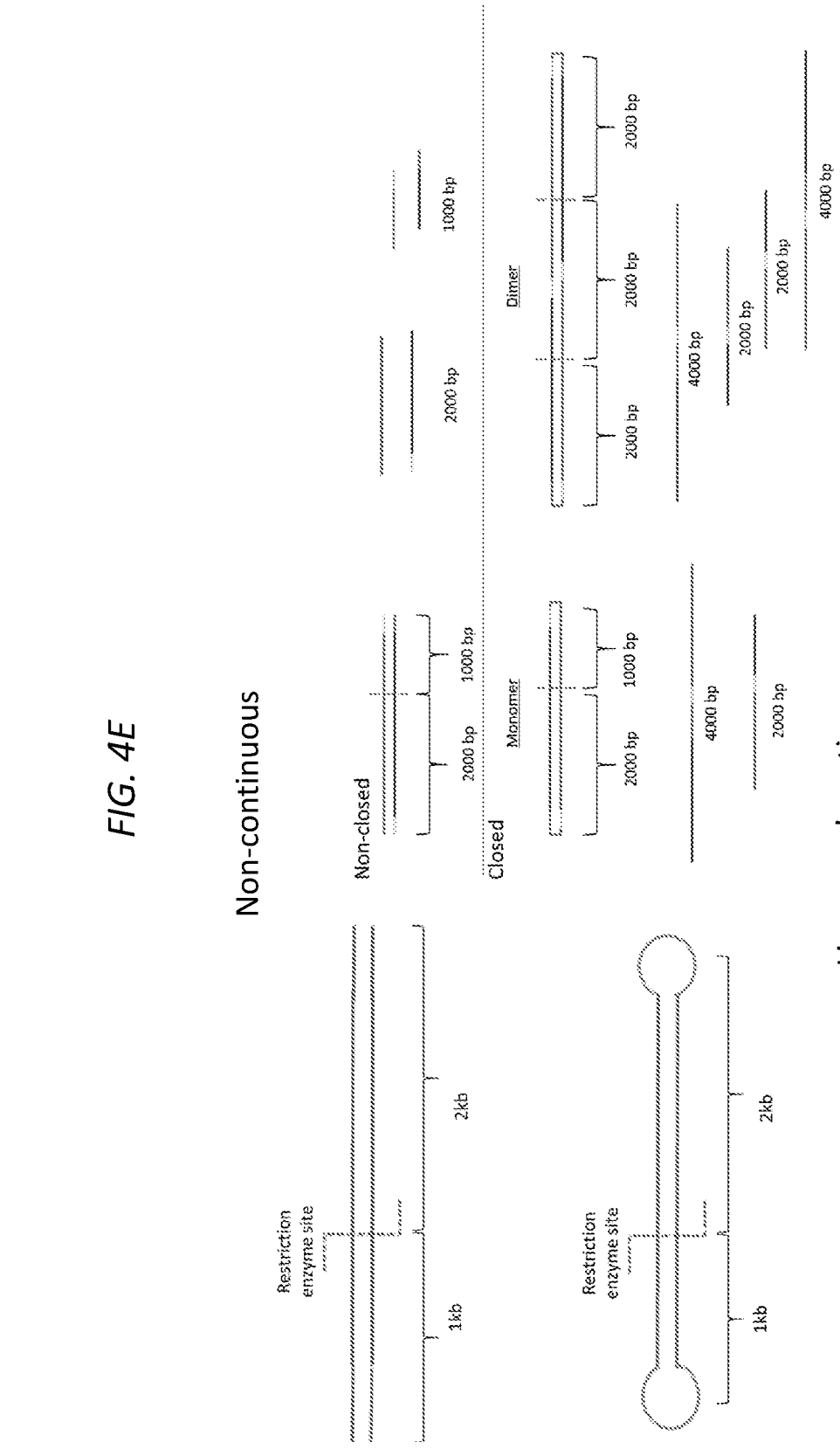

Analysis to Confirm Production of ceDNA
Virtual Gel Schematics
Native Gels
Double-stranded DNAs
ceDNA un-cut
Dimer (~10kb)
Monomer (~5kb)
ceDNA cut
Fragment 1 (~4kb)
Fragment 2 (~1kb)
Denaturing Gels
Single-stranded DNAs
ceDNA un-cut
Dimer (>20kb)
Monomer (~10kb)
Migrates as a single-stranded open circle
ceDNA cut
Fragment 1 (~8kb)
(2x)
Fragment 2 (~2kb)
(2x)

FIG. 4D

FIG. 4E


Non-continuous
Non-closed
Closed
Monomer
Dimer
Restriction enzyme site
1kb
2kb
2000 bp
1000 bp
4000 bp
Linear and continuous

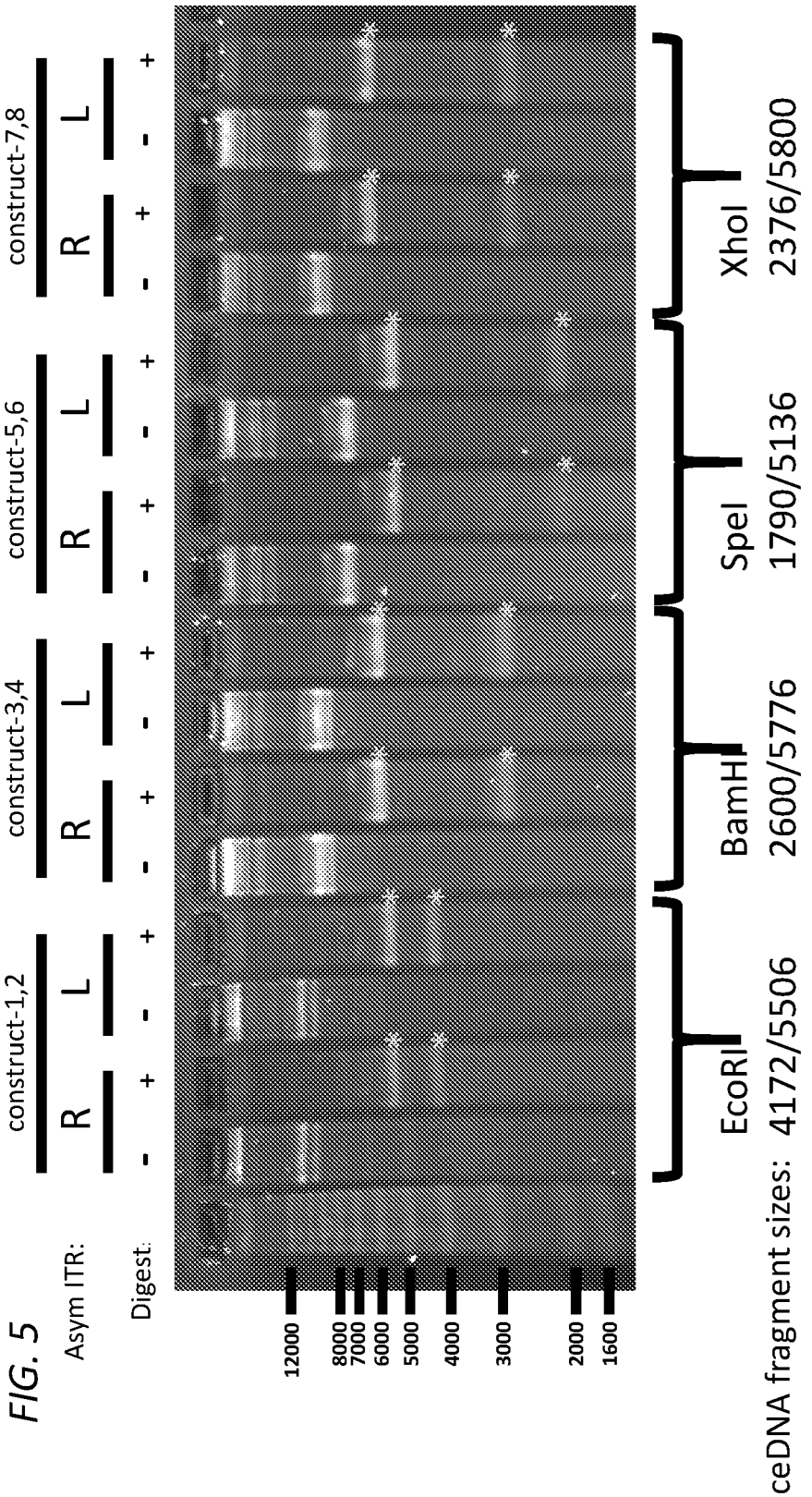
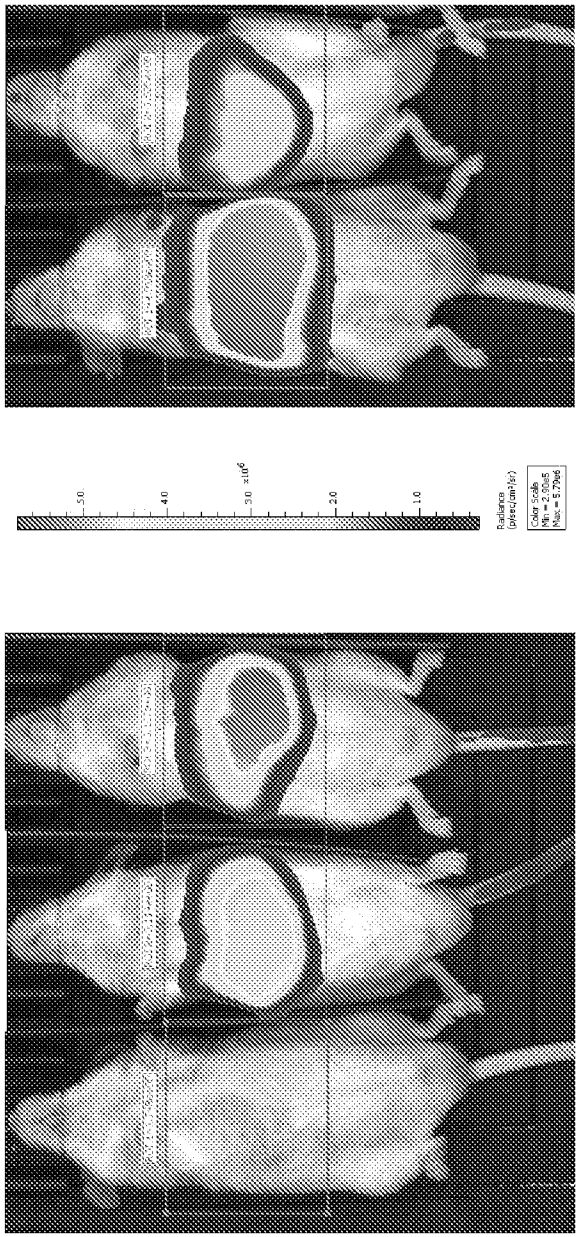


FIG. 5
Asym ITR:
Digest:
construct-1,2
construct-3,4
construct-5,6
construct-7,8
R
L
-
+
12000
8000
7000
6000
5000
4000
3000
2000
1600
EcoRI
BamHI
SpeI
XhoI
ceDNA fragment sizes: 4172/5506 2600/5776 1790/5136 2376/5800

FIG. 6


Repeat Dosing
re-dose 10mg/kg
re-dose 3mg/kg
LNPceDNA 1mg/kg
LNPceDNA 0.5mg/kg
LNPceDNA 0.1mg/kg
LNP Only
Mean Radiance Flux (ph/s) + SEM
DAYS POST DOSE
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140

FIG. 7



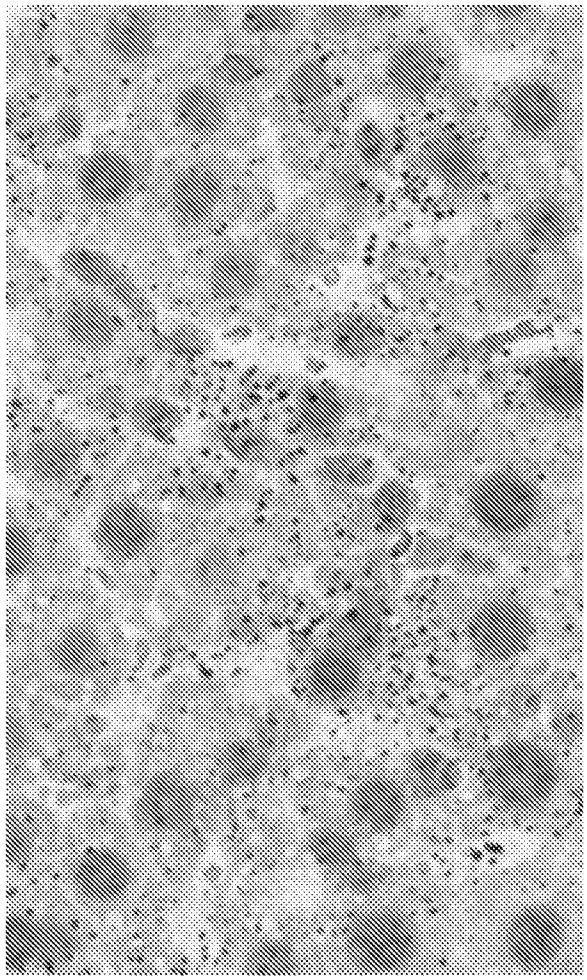


FIG. 8

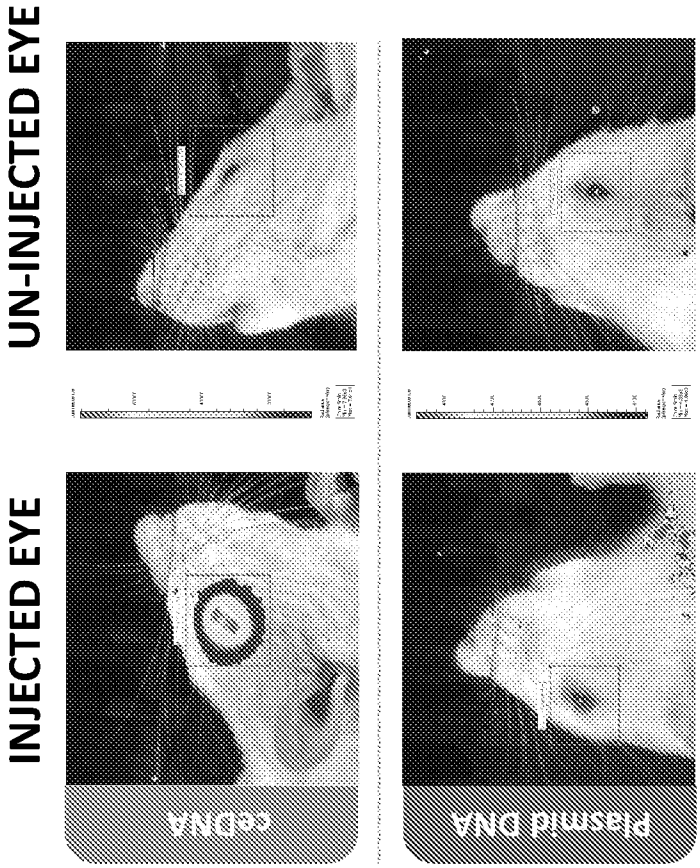


FIG. 9A

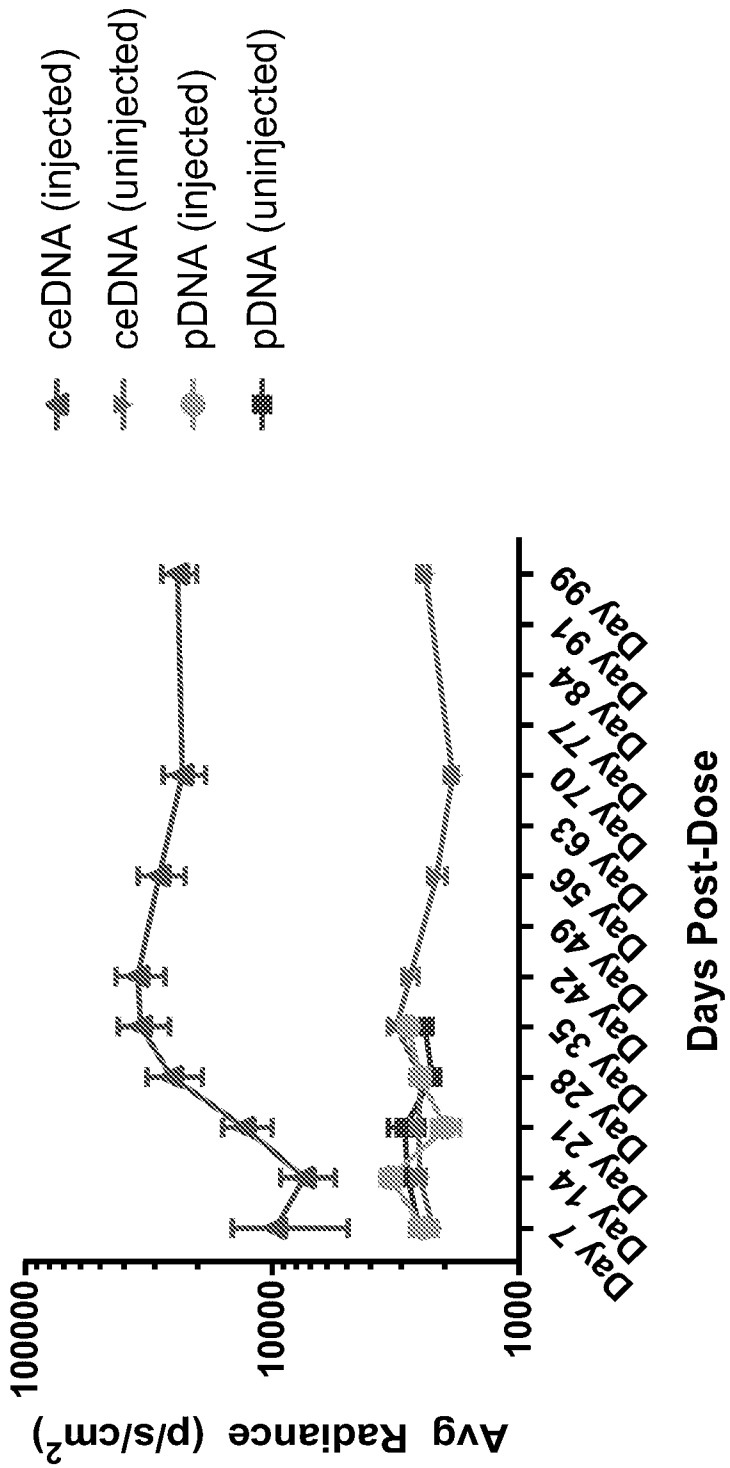

ceDNA (injected)
ceDNA (uninjected)
pDNA (injected)
pDNA (uninjected)
Avg Radiance (p/s/cm²)
100000
10000
1000
Day 7
Day 14
Day 21
Day 28
Day 35
Day 42
Day 49
Day 56
Day 63
Day 70
Day 77
Day 84
Day 91
Day 99
Days Post-Dose

FIG. 9B

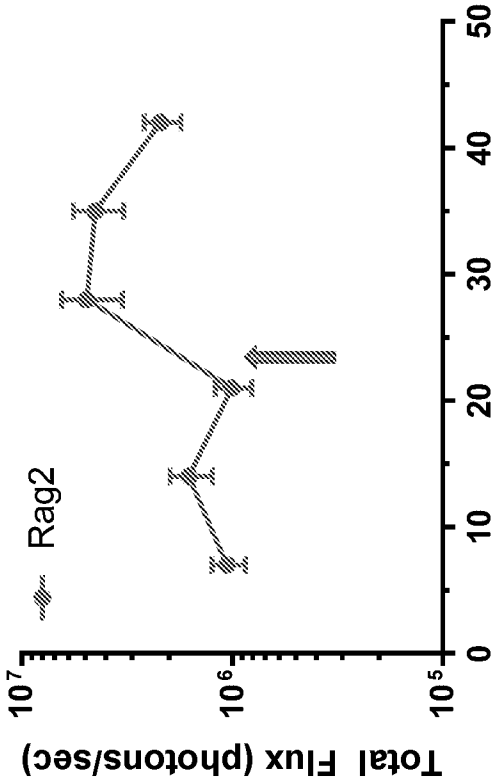


FIG. 10B

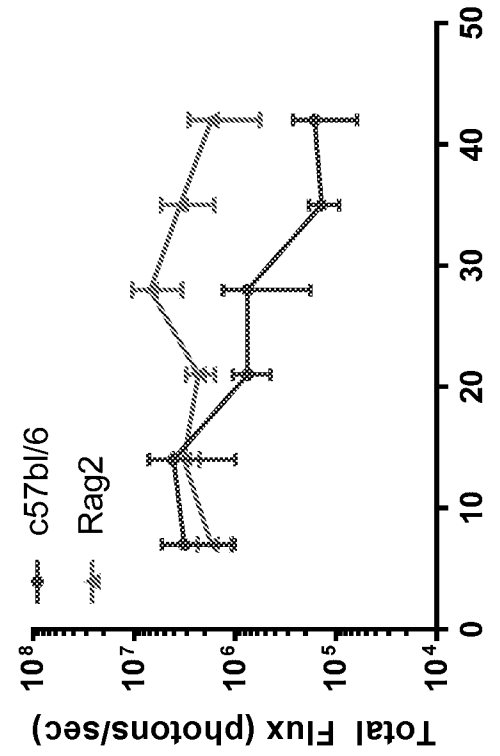


FIG. 10A

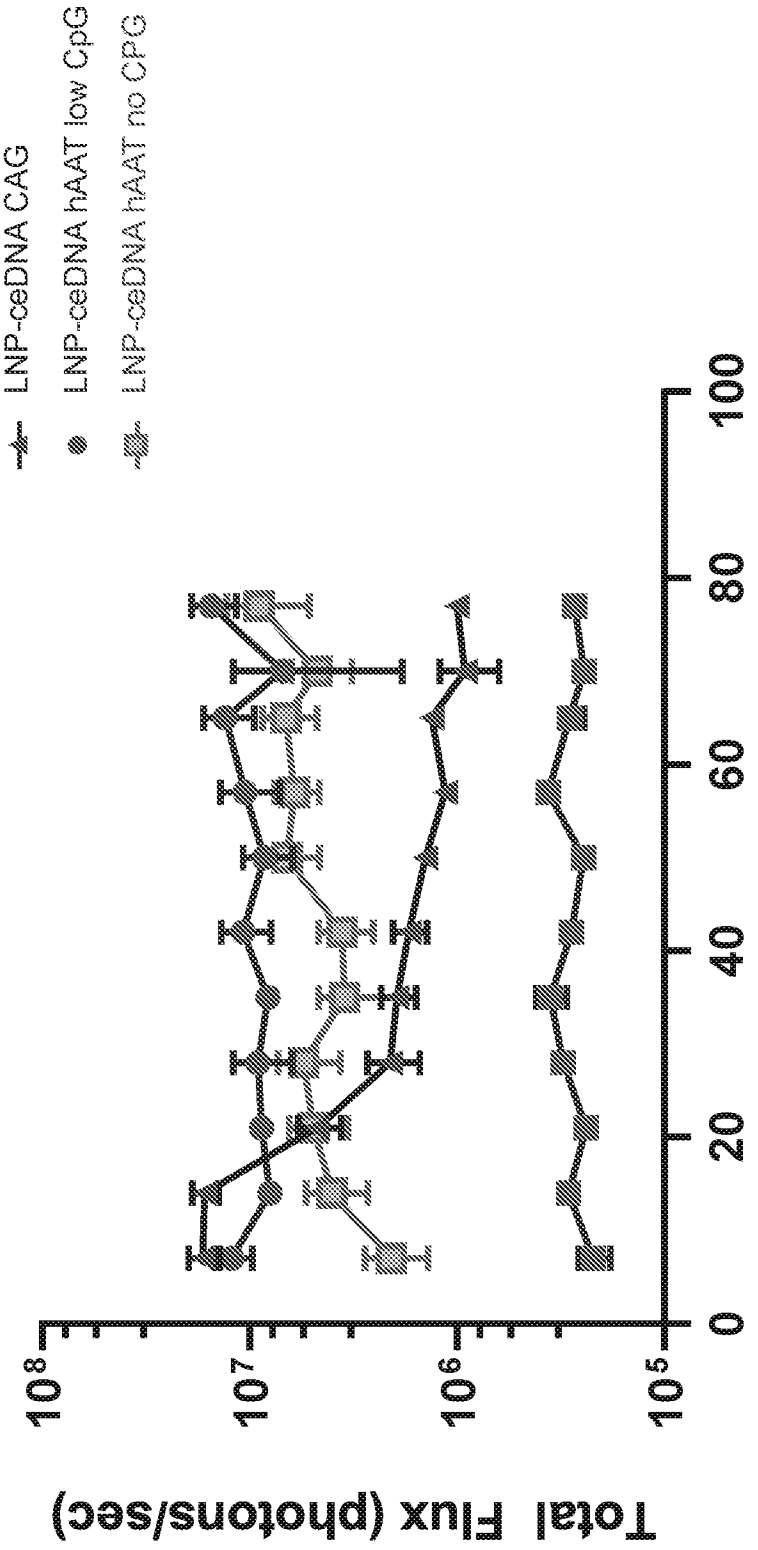

LNP-PolyC
LNP-ceDNA CAG
LNP-ceDNA hAAT low CpG
LNP-ceDNA hAAT no CPG
Total Flux (photons/sec)
10^8
10^7
10^6
10^5
0
20
40
60
80
100
Study Day

FIG. 11

SEQUENCE LISTING

<110> GENERATION BIO INC.

<120> CONTROLLED EXPRESSION OF TRANSGENES USING CLOSE-ENDED DNA (CEDNA) VECTORS

<130> 080170-091190-WOPT

<140>
<141>

<150> 62/746,762
<151> 2018-10-17

<150> 62/633,882
<151> 2018-02-22

<150> 62/633,795
<151> 2018-02-22

<150> 62/633,757
<151> 2018-02-22

<160> 190

<170> PatentIn version 3.5

<210> 1
<211> 141
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 1
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgcag ctgcctgcag g 141

<210> 2
<211> 141
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 2
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc t 141

<210> 3
<211> 130
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 3
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gcggcctcag tgagcgagcg agcgcgcagc 120

tgcctgcagg 130

<210> 4
<211> 130
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 4
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgaccttt 60

ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120

aggggttcct 130

<210> 5
<211> 143
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 5
ttgcccactc cctctctgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc 60

agacggcaga ggtctcctct gccggcccca ccgagcgagc gacgcgcgca gagagggagt 120

gggcaactcc atcactaggg taa 143

<210> 6
<211> 144
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 6
ttggccactc cctctatgcg cactcgctcg ctcggtgggg cctggcgacc aaaggtcgcc 60

agacggacgt gggtttccac gtccggcccc accgagcgag cgagtgcgca tagagggagt 120

ggccaactcc atcactagag gtat 144

<210> 7
<211> 127
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 7
ttggccactc cctctatgcg cgctcgctca ctcactcggc cctggagacc aaaggtctcc 60

agactgccgg cctctggccg gcagggccga gtgagtgagc gagcgcgcat agagggagtg 120

gccaact 127

<210> 8
<211> 166
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 8
tccccctgt cgcgttcgct cgctcgctgg ctcgtttggg ggggcgacgg ccagagggcc 60

gtcgtctggc agctctttga gctgccaccc ccccaaacga gccagcgagc gagcgaacgc 120

gacaggggg agagtgccac actctcaagc aaggggttt tgtaag 166

<210> 9
<211> 144
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 9
ttgcccactc cctctaatgc gcgctcgctc gctcggtggg gcctgcggac caaaggtccg 60

cagacggcag aggtctcctc tgccggcccc accgagcgag cgagcgcgca tagagggagt 120

gggcaactcc atcactaggg gtat 144

<210> 10
<211> 143
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 10
ttaccctagt gatggagttg cccactccct ctctgcgcgc gtcgctcgct cggtggggcc 60

ggcagaggag acctctgccg tctgcggacc tttggtccgc aggccccacc gagcgagcga 120

gcgcgcagag agggagtggg caa 143

<210> 11
<211> 144
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 11
atacctctag tgatggagtt ggccactccc tctatgcgca ctcgctcgct cggtggggcc 60

ggacgtggaa acccacgtcc gtctggcgac ctttggtcgc caggccccac cgagcgagcg 120

agtgcgcata gagggagtgg ccaa 144

<210> 12
<211> 127
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 12
agttggccac attagctatg cgcgctcgct cactcactcg gccctggaga ccaaaggtct 60

ccagactgcc ggcctctggc cggcagggcc gagtgagtga gcgagcgcgc atagagggag 120

tggccaa 127

<210> 13
<211> 166
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 13
cttacaaaac ccccttgctt gagagtgtgg cactctcccc cctgtcgcgt tcgctcgctc 60

gctggctcgt ttgggggggt ggcagctcaa agagctgcca gacgacggcc ctctggccgt 120

cgccccccca aacgagccag cgagcgagcg aacgcgacag ggggga 166

<210> 14
<211> 144
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 14
ataccctag tgatggagtt gcccactccc tctatgcgcg ctcgctcgct cggtggggcc 60

ggcagaggag acctctgccg tctgcggacc tttggtccgc aggccccacc gagcgagcga 120

gcgcgcatta gagggagtgg gcaa 144

<210> 15
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 15
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

cgcacgcccg ggtttcccgg gcggcctcag tgagcgagcg agcgcgcagc tgcctgcagg 120

<210> 16
<211> 122
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 16
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgacgcccg ggctttgccc gggcggcctc agtgagcgag cgagcgcgca gctgcctgca 120

gg 122

<210> 17
<211> 129
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 17
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gcgcctcagt gagcgagcga gcgcgcagct 120

gcctgcagg 129

<210> 18
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 18
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ctttgcctca gtgagcgagc gagcgcgcag ctgcctgcag g 101

<210> 19
<211> 139
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 19
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgaca aagtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga 120

gcgcgcagct gcctgcagg 139

<210> 20
<211> 137
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 20
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgaaa atcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc 120

gcgcagctgc ctgcagg 137

<210> 21
<211> 135
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 21
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgaaa cgcccgacgc ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc 120

gcagctgcct gcagg 135

<210> 22
<211> 133
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 22
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcaaag cccgacgccc gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc 120

agctgcctgc agg 133

<210> 23
<211> 139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 23
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gtttcccggg cggcctcagt gagcgagcga 120

gcgcgcagct gcctgcagg 139

<210> 24
<211> 137
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 24
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg tttccgggcg gcctcagtga gcgagcgagc 120

gcgcagctgc ctgcagg 137

<210> 25
<211> 135
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 25
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgt ttcgggcggc ctcagtgagc gagcgagcgc 120

gcagctgcct gcagg 135

<210> 26
<211> 133
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 26
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgccctt tgggcggcct cagtgagcga gcgagcgcgc 120

agctgcctgc agg 133

<210> 27
<211> 131
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 27
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgccttt ggcggcctca gtgagcgagc gagcgcgcag 120

ctgcctgcag g 131

<210> 28
<211> 129
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 28
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgctttg cggcctcagt gagcgagcga gcgcgcagct 120

gcctgcagg 129

<210> 29
<211> 127
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 29
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgtttcg gcctcagtga gcgagcgagc gcgcagctgc 120

ctgcagg 127

<210> 30
<211> 122
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 30
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacggcctc agtgagcgag cgagcgcgca gctgcctgca 120

gg 122

<210> 31
<211> 130
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 31
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gcggcctcag tgagcgagcg agcgcgcagc 120

tgcctgcagg 130

<210> 32
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 32
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtgcg 60

cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact aggggttcct 120

<210> 33
<211> 122
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 33

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgt cgggcgacct ttggtcgccc 60

ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca ctaggggttc 120

ct 122

<210> 34
<211> 122
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 34
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca ctaggggttc 120

ct 122

<210> 35
<211> 129
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 35
cctgcaggca gctgcgcgct cgctcgctca ctgaggcgcc cgggcgtcgg gcgacctttg 60

gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120

ggggttcct 129

<210> 36
<211> 101
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 36
cctgcaggca gctgcgcgct cgctcgctca ctgaggcaaa gcctcagtga gcgagcgagc 60

gcgcagagag ggagtggcca actccatcac taggggttcc t 101

<210> 37
<211> 139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 37
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacttt gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac 120

tccatcacta ggggttcct 139

<210> 38
<211> 137
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 38
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgatttt cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc 120

catcactagg ggttcct 137

<210> 39
<211> 135
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 39
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgtttcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg gccaactcca 120

tcactagggg ttcct 135

<210> 40
<211> 133
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 40
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggctttgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120

actaggggtt cct 133

<210> 41
<211> 139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 41
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtcgg 60

gcgacctttg gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac 120

tccatcacta ggggttcct 139

<210> 42
<211> 137
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 42
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccggaaaccg ggcgtcgggc 60

gacctttggt cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc 120

catcactagg ggttcct 137

<210> 43
<211> 135
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 43
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgaaacggg cgtcgggcga 60

cctttggtcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg gccaactcca 120

tcactagggg ttcct 135

<210> 44
<211> 133
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 44
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccaaagggcg tcgggcgacc 60

tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120

actaggggtt cct 133

<210> 45
<211> 131
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 45
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc caaaggcgtc gggcgacctt 60

tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac 120

taggggttcc t 131

<210> 46
<211> 129
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 46
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc aaagcgtcgg gcgacctttg 60

gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120

ggggttcct 129

<210> 47
<211> 127
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 47
cctgcaggca gctgcgcgct cgctcgctca ctgaggccga aacgtcgggc gacctttggt 60

cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc catcactagg 120

ggttcct 127

<210> 48
<211> 122
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 48
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacggcctc agtgagcgag cgagcgcgca gctgcctgca 120

gg 122

<210> 49
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 49
cgatcgttcg at 12

<210> 50
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 50
atcgaaccat cg 12

<210> 51
<211> 12

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 51
atcgaacgat cg 12

<210> 52
<211> 165
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 52
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgcccgggc aaagcccggg cgtcgggcga cctttggtcg cccggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaactcca tcactagggg ttcct 165

<210> 53
<211> 140
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 53
cccctagtga tggagttggc cactccctct ctgcgcgctc gctcgctcac tgaggccgcc 60

cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg 120

cgcagagaga tcactagggg 140

<210> 54
<211> 91
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 54
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg 60

tcgcccggcc tcagtgagcg agcgagcgcg c 91

<210> 55
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 55
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggctttgc 60

ccgggcggcc tcagtgagcg agcgagcgcg c 91

<210> 56
<211> 1662
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 56
gccgccacca tggaagacgc caaaaacata aagaaaggcc cggcgccatt ctatccgctg 60

gaagatggaa ccgctggaga gcaactgcat aaggctatga agagatacgc cctggttcct 120

ggaacaattg cttttacaga tgcacatatc gaggtggaca tcacttacgc tgagtacttc 180

gaaatgtccg ttcggttggc agaagctatg aaacgatatg ggctgaatac aaatcacaga 240

atcgtcgtat gcagtgaaaa ctctcttcaa ttctttatgc cggtgttggg cgcgttattt 300

atcggagttg cagttgcgcc cgcgaacgac atttataatg aacgtgaatt gctcaacagt 360

atgggcattt cgcagcctac cgtggtgttc gtttccaaaa aggggttgca aaaaattttg 420

aacgtgcaaa aaaagctccc aatcatccaa aaaattatta tcatggattc taaaacggat 480

taccagggat ttcagtcgat gtacacgttc gtcacatctc atctacctcc cggttttaat 540

gaatacgatt ttgtgccaga gtccttcgat agggacaaga caattgcact gatcatgaac 600

tcctctggat ctactggtct gcctaaaggt gtcgctctgc ctcatagaac tgcctgcgtg 660

agattctcgc atgccagaga tcctattttt ggcaatcaaa tcattccgga tactgcgatt 720

ttaagtgttg ttccattcca tcacggtttt ggaatgttta ctacactcgg atatttgata 780

tgtggatttc gagtcgtctt aatgtataga tttgaagaag agctgtttct gaggagcctt 840

caggattaca agattcaaag tgcgctgctg gtgccaaccc tattctcctt cttcgccaaa 900

agcactctga ttgacaaata cgatttatct aatttacacg aaattgcttc tggtggcgct 960

cccctctcta aggaagtcgg ggaagcggtt gccaagaggt tccatctgcc aggtatcagg 1020

caaggatatg ggctcactga gactacatca gctattctga ttacacccga gggggatgat 1080

aaaccgggcg cggtcggtaa agttgttcca ttttttgaag cgaaggttgt ggatctggat 1140

accgggaaaa cgctgggcgt taatcaaaga ggcgaactgt gtgtgagagg tcctatgatt 1200

atgtccggtt atgtaaacaa tccggaagcg accaacgcct tgattgacaa ggatggatgg 1260

ctacattctg gagacatagc ttactgggac gaagacgaac acttcttcat cgttgaccgc 1320

ctgaagtctc tgattaagta caaaggctat caggtggctc ccgctgaatt ggaatccatc 1380

ttgctccaac accccaacat cttcgacgca ggtgtcgcag gtcttcccga cgatgacgcc 1440

ggtgaacttc ccgccgccgt tgttgttttg gagcacggaa agacgatgac ggaaaaagag 1500

atcgtggatt acgtcgccag tcaagtaaca accgcgaaaa agttgcgcgg aggagttgtg 1560

tttgtggacg aagtaccgaa aggtcttacc ggaaaactcg acgcaagaaa aatcagagag 1620

atcctcataa aggccaagaa gggcggaaag atcgccgtgt aa 1662

<210> 57
<211> 453

<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polypeptide

<220>
<221> MOD_RES
<222> (1)..(1)
<223> Any amino acid

<400> 57

```
Xaa Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
1               5                  10                  15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Ala Phe Ser Ser Tyr
            20                  25                  30

Gly Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
        35                  40                  45

Ala Val Ile Trp Phe Asp Gly Thr Lys Lys Tyr Tyr Thr Asp Ser Val
    50                  55                  60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65                  70                  75                  80

Leu Gln Met Asn Thr Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
                85                  90                  95

Ala Arg Asp Arg Gly Ile Gly Ala Arg Arg Gly Pro Tyr Tyr Met Asp
            100                 105                 110

Val Trp Gly Lys Gly Thr Thr Val Thr Val Ser Ser Ala Ser Thr Lys
        115                 120                 125

Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly
    130                 135                 140
```

```
Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro
145                 150                 155                 160

Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr
                165                 170                 175

Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val
            180                 185                 190

Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn
        195                 200                 205

Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Arg Val Glu Pro
    210                 215                 220

Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
225                 230                 235                 240

Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
                245                 250                 255

Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
            260                 265                 270

Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly
        275                 280                 285

Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn
    290                 295                 300

Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
305                 310                 315                 320

Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro
                325                 330                 335
```

```
Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
            340                 345                 350

Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
        355                 360                 365

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
    370                 375                 380

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
385                 390                 395                 400

Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
                405                 410                 415

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
            420                 425                 430

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
        435                 440                 445

Ser Leu Ser Pro Gly
    450


<210> 58
<211> 214
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      polypeptide

<400> 58
Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1               5                   10                  15
```

```
Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr
            20                  25                  30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
        35                  40                  45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65                  70                  75                  80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro Leu
                85                  90                  95

Thr Phe Gly Gly Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
            100                 105                 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
        115                 120                 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
    130                 135                 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145                 150                 155                 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
                165                 170                 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
            180                 185                 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
        195                 200                 205
```

Phe Asn Arg Gly Glu Cys
210

<210> 59
<211> 1310
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 59
ggagccgaga gtaattcata caaaaggagg gatcgccttc gcaaggggag agcccaggga 60

ccgtccctaa attctcacag acccaaatcc ctgtagccgc cccacgacag cgcgaggagc 120

atgcgctcag ggctgagcgc ggggagagca gagcacacaa gctcatagac cctggtcgtg 180

gggggggagga ccggggagct ggcgcggggc aaactgggaa agcggtgtcg tgtgctggct 240

ccgccctctt cccgagggtg ggggagaacg gtatataagt gcggcagtcg ccttggacgt 300

tctttttcgc aacgggtttg ccgtcagaac gcaggtgagg ggcgggtgtg gcttccgcgg 360

gccgccgagc tggaggtcct gctccgagcg ggccgggccc cgctgtcgtc ggcggggatt 420

agctgcgagc attcccgctt cgagttgcgg gcggcgcggg aggcagagtg cgaggcctag 480

cggcaacccc gtagcctcgc ctcgtgtccg gcttgaggcc tagcgtggtg tccgcgccgc 540

cgccgcgtgc tactccggcc gcactctggt ctttttttttt tttgttgttg ttgccctgct 600

gccttcgatt gccgttcagc aataggggct aacaaaggga gggtgcgggg cttgctcgcc 660

cggagcccgg agaggtcatg gttggggagg aatggaggga caggagtggc ggctggggcc 720

cgcccgcctt cggagcacat gtccgacgcc acctggatgg ggcgaggcct ggggtttttc 780

ccgaagcaac caggctgggg ttagcgtgcc gaggccatgt ggccccagca cccggcacga 840

tctggcttgg cggcgccgcg ttgccctgcc tccctaacta gggtgaggcc atcccgtccg 900

gcaccagttg cgtgcgtgga aagatggccg ctcccgggcc ctgttgcaag gagctcaaaa 960

tggaggacgc ggcagcccgg tggagcgggc gggtgagtca cccacacaaa ggaagagggc 1020

ctggtccctc accggctgct gcttcctgtg accccgtggt cctatcggcc gcaatagtca 1080

cctcgggctt ttgagcacgg ctagtcgcgg cgggggggagg ggatgtaatg gcgttggagt 1140

ttgttcacat ttggtgggtg gagactagtc aggccagcct ggcgctggaa gtcatttttg 1200

gaatttgtcc ccttgagttt tgagcggagc taattctcgg gcttcttagc ggttcaaagg 1260

tatcttttaa acccttttt aggtgttgtg aaaaccaccg ctaattcaaa 1310

<210> 60
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 60
gcgcgctcgc tcgctc 16

<210> 61
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 61
ggttga 6

<210> 62
<211> 4
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 62
agtt 4

<210> 63
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 63
ggttgg 6

<210> 64
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 64
agttgg 6

<210> 65
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 65
agttga 6

<210> 66
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 66
rrttrr 6

<210> 67
<211> 581
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 67
gagcatctta ccgccattta ttcccatatt tgttctgttt ttcttgattt gggtatacat 60

ttaaatgtta ataaaacaaa atggtggggc aatcatttac atttttaggg atatgtaatt 120

actagttcag gtgtattgcc acaagacaaa catgttaaga aactttcccg ttatttacgc 180

tctgttcctg ttaatcaacc tctggattac aaaatttgtg aaagattgac tgatattctt 240

aactatgttg ctccttttac gctgtgtgga tatgctgctt tatagcctct gtatctagct 300

attgcttccc gtacggcttt cgttttctcc tccttgtata aatcctggtt gctgtctctt 360

ttagaggagt tgtggcccgt tgtccgtcaa cgtggcgtgg tgtgctctgt gtttgctgac 420

gcaaccccca ctggctgggg cattgccacc acctgtcaac tcctttctgg gactttcgct 480

ttccccctcc cgatcgccac ggcagaactc atcgccgcct gccttgcccg ctgctggaca 540

ggggctaggt tgctgggcac tgataattcc gtggtgttgt c 581

<210> 68
<211> 225
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 68
tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct 60

ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct 120

gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg 180

ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggc 225

<210> 69
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 69
actgaggc 8

<210> 70
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 70
gcctcagt 8

<210> 71
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 71
gagcgagcga gcgcgc 16

<210> 72

<211> 1923
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 72
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60

ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120

aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180

gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240

gcctggctga ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat 300

agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360

ccacttggca gtacatcaag tgtatcatat gccaagtccg ccccctattg acgtcaatga 420

cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg 480

gcagtacatc tacgtattag tcatcgctat taccatggtc gaggtgagcc ccacgttctg 540

cttcactctc cccatctccc ccccctcccc acccccaatt ttgtatttat ttatttttta 600

attattttgt gcagcgatgg gggcgggggg gggggggggg cgcgcgccag gcggggcggg 660

gcggggcgag gggcggggcg gggcgaggcg gagaggtgcg gcggcagcca atcagagcgg 720

cgcgctccga aagtttcctt ttatggcgag gcggcggcgg cggcggccct ataaaaagcg 780

aagcgcgcgg cgggcgggag tcgctgcgac gctgccttcg ccccgtgccc cgctccgccg 840

ccgcctcgcg ccgcccgccc cggctctgac tgaccgcgtt actcccacag gtgagcgggc 900

gggacggccc ttctcctccg ggctgtaatt agcgcttggt ttaatgacgg cttgtttctt 960

ttctgtggct gcgtgaaagc cttgaggggc tccgggaggg ccctttgtgc gggggggagc 1020

ggctcggggg gtgcgtgcgt gtgtgtgtgc gtggggagcg ccgcgtgcgg cccgcgctgc 1080

ccggcggctg tgagcgctgc gggcgcggcg cggggctttg tgcgctccgc agtgtgcgcg 1140

aggggagcgc ggccgggggc ggtgccccgc ggtgcggggg gggctgcgag gggaacaaag 1200

gctgcgtgcg gggtgtgtgc gtgggggggt gagcaggggg tgtgggcgcg gcggtcgggc 1260

tgtaaccccc ccctgcaccc ccctccccga gttgctgagc acggcccggc ttcgggtgcg 1320

gggctccgta cggggcgtgg cgcggggctc gccgtgccgg gcgggggggtg gcggcaggtg 1380

ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcgggggа ggggcgcggc 1440

ggcccccgga gcgccggcgg ctgtcgaggc gcggcgagcc gcagccattg ccttttatgg 1500

taatcgtgcg agagggcgca gggacttcct ttgtcccaaa tctgtgcgga gccgaaatct 1560

gggaggcgcc gccgcacccc ctctagcggg cgcggggcga agcggtgcgg cgccggcagg 1620

aaggaaatgg gcggggaggg ccttcgtgcg tcgccgcgcc gccgtcccct tctccctctc 1680

cagcctcggg gctgtccgcg gggggacggc tgccttcggg ggggacgggg cagggcgggg 1740

ttcggcttct ggcgtgtgac cggcggctct agagcctctg ctaaccatgt tttagccttc 1800

ttctttttcc tacagctcct gggcaacgtg ctggttattg tgctgtctca tcatttgtcg 1860

acagaattcc tcgaagatcc gaaggggttc aagcttggca ttccggtact gttggtaaag 1920

cca 1923

<210> 73
<211> 1272
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 73
aggctcagag gcacacagga gtttctgggc tcaccctgcc cccttccaac ccctcagttc 60

ccatcctcca gcagctgttt gtgtgctgcc tctgaagtcc acactgaaca aacttcagcc 120

tactcatgtc cctaaaatgg gcaaacattg caagcagcaa acagcaaaca cacagccctc 180

cctgcctgct gaccttggag ctggggcaga ggtcagagac ctctctgggc ccatgccacc 240

tccaacatcc actcgacccc ttggaatttc ggtggagagg agcagaggtt gtcctggcgt 300

ggtttaggta gtgtgagagg gtccgggttc aaaaccactt gctgggtggg gagtcgtcag 360

taagtggcta tgccccgacc ccgaagcctg tttccccatc tgtacaatgg aaatgataaa 420

gacgcccatc tgatagggtt tttgtggcaa ataaacattt ggtttttttg ttttgttttg 480

ttttgttttt tgagatggag gtttgctctg tcgcccaggc tggagtgcag tgacacaatc 540

tcatctcacc acaaccttcc cctgcctcag cctcccaagt agctgggatt acaagcatgt 600

gccaccacac ctggctaatt ttctattttt agtagagacg ggtttctcca tgttggtcag 660

cctcagcctc ccaagtaact gggattacag gcctgtgcca ccacacccgg ctaatttttt 720

ctatttttga cagggacggg gtttcaccat gttggtcagg ctggtctaga ggtaccggat 780

cttgctacca gtggaacagc cactaaggat tctgcagtga gagcagaggg ccagctaagt 840

ggtactctcc cagagactgt ctgactcacg ccacccccctc caccttggac acaggacgct 900

gtggtttctg agccaggtac aatgactcct ttcggtaagt gcagtggaag ctgtacactg 960

cccaggcaaa gcgtccgggc agcgtaggcg ggcgactcag atcccagcca gtggacttag 1020

cccctgtttg ctcctccgat aactggggtg accttggtta atattcacca gcagcctccc 1080

ccgttgcccc tctggatcca ctgcttaaat acggacgagg acagggccct gtctcctcag 1140

cttcaggcac caccactgac ctgggacagt gaatccggac tctaaggtaa atataaaatt 1200

tttaagtgta taatgtgtta aactactgat tctaattgtt tctctctttt agattccaac 1260

ctttggaact ga 1272

<210> 74
<211> 1177
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 74
ggctcagagg ctcagaggca cacaggagtt tctgggctca ccctgccccc ttccaacccc 60

tcagttccca tcctccagca gctgtttgtg tgctgcctct gaagtccaca ctgaacaaac 120

ttcagcctac tcatgtccct aaaatgggca aacattgcaa gcagcaaaca gcaaacacac 180

agccctccct gcctgctgac cttggagctg gggcagaggt cagagacctc tctgggccca 240

tgccacctcc aacatccact cgacccctg gaatttcggt ggagaggagc agaggttgtc 300

ctggcgtggt ttaggtagtg tgagagggtc cgggttcaaa accacttgct gggtggggag 360

tcgtcagtaa gtggctatgc cccgaccccg aagcctgttt ccccatctgt acaatggaaa 420

tgataaagac gcccatctga tagggttttt gtggcaaata aacatttggt ttttttgttt 480

tgttttgttt tgttttttga gatggaggtt tgctctgtcg cccaggctgg agtgcagtga 540

cacaatctca tctcaccaca accttcccct gcctcagcct cccaagtagc tgggattaca 600

agcatgtgcc accacacctg gctaattttc tatttttagt agagacgggt ttctccatgt 660

tggtcagcct cagcctccca agtaactggg attacaggcc tgtgccacca cacccggcta 720

atttttttcta tttttgacag ggacggggtt tcaccatgtt ggtcaggctg gtctagaggt 780

accggatctt gctaccagtg gaacagccac taaggattct gcagtgagag cagagggcca 840

gctaagtggt actctcccag agactgtctg actcacgcca ccccctccac cttggacaca 900

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 960

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 1020

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 1080

gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 1140

tcctcagctt caggcaccac cactgacctg ggacagt 1177

<210> 75
<211> 547
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 75

ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc 60

tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc 120

cactcgaccc cttggaattt ttcggtggag aggagcagag gttgtcctgg cgtggtttag 180

gtagtgtgag aggggaatga ctcctttcgg taagtgcagt ggaagctgta cactgcccag 240

gcaaagcgtc cgggcagcgt aggcgggcga ctcagatccc agccagtgga cttagcccct 300

gtttgctcct ccgataactg gggtgacctt ggttaatatt caccagcagc ctcccccgtt 360

gcccctctgg atccactgct taaatacgga cgaggacagg gccctgtctc ctcagcttca 420

ggcaccacca ctgacctggg acagtgaatc cggactctaa ggtaaatata aaatttttaa 480

gtgtataatg tgttaaacta ctgattctaa ttgtttctct cttttagatt ccaacctttg 540

gaactga 547

<210> 76
<211> 556
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 76
ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc 60

tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc 120

cactcgaccc cttggaattt cggtggagag gagcagaggt tgtcctggcg tggtttaggt 180

agtgtgagag gggaatgact cctttcggta agtgcagtgg aagctgtaca ctgcccaggc 240

aaagcgtccg ggcagcgtag gcgggcgact cagatcccag ccagtggact tagcccctgt 300

ttgctcctcc gataactggg gtgaccttgg ttaatattca ccagcagcct cccccgttgc 360

ccctctggat ccactgctta aatacggacg aggacactcg agggccctgt ctcctcagct 420

tcaggcacca ccactgacct gggacagtga atccggacat cgattctaag gtaaatataa 480

aatttttaag tgtataattt gttaaactac tgattctaat tgtttctctc ttttagattc 540

caacctttgg aactga 556

<210> 77
<211> 1179
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 77
ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cctttttccc gagggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtaagtgcc 240

gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgcgt gccttgaatt 300

acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg aagtgggtgg 360

gagagttcga ggccttgcgc ttaaggagcc ccttcgcctc gtgcttgagt tgaggcctgg 420

cctgggcgct ggggccgccg cgtgcgaatc tggtggcacc ttcgcgcctg tctcgctgct 480

ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct ttttttctgg 540

caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt ttttggggcc 600

gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg gggcctgcga 660

gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc tctggtgcct 720

ggtctcgcgc cgccgtgtat cgccccgccc tgggcggcaa ggctggcccg gtcggcacca 780

gttgcgtgag cggaaagatg gccgcttccc ggccctgctg cagggagctc aaaatggagg 840

acgcggcgct cgggagagcg ggcgggtgag tcacccacac aaaggaaaag ggcctttccg 900

tcctcagccg tcgcttcatg tgactccacg gagtaccggg cgccgtccag gcacctcgat 960

tagttctcga gcttttggag tacgtcgtct ttaggttggg gggaggggtt ttatgcgatg 1020

gagtttcccc acactgagtg ggtggagact gaagttaggc cagcttggca cttgatgtaa 1080

ttctccttgg aatttgccct ttttgagttt ggatcttggt tcattctcaa gcctcagaca 1140

gtggttcaaa gttttttttct tccatttcag gtgtcgtga 1179

<210> 78
<211> 141
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 78
aataaacgat aacgccgttg gtggcgtgag gcatgtaaaa ggttacatca ttatcttgtt 60

cgccatccgg ttggtataaa tagacgttca tgttggtttt tgtttcagtt gcaagttggc 120

tgcggcgcgc gcagcacctt t 141

<210> 79
<211> 317
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 79
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 120

tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa 180

ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag 240

aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag 300

gcctaggctt ttgcaaa 317

<210> 80

<211> 241
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 80
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60

ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120

aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180

atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

c 241

<210> 81
<211> 215
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 81
gaacgctgac gtcatcaacc cgctccaagg aatcgcgggc ccagtgtcac taggcgggaa 60

cacccagcgc gcgtgcgccc tggcaggaag atggctgtga gggacagggg agtggcgccc 120

tgcaatattt gcatgtcgct atgtgttctg ggaaatcacc ataaacgtga aatgtctttg 180

gatttgggaa tcgtataaga actgtatgag accac 215

<210> 82
<211> 546
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 82
ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc 60

tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc 120

cactcgaccc cttggaattt ttcggtggag aggagcagag gttgtcctgg cgtggtttag 180

gtagtgtgag aggggaatga ctcctttcgg taagtgcagt ggaagctgta cactgcccag 240

gcaaagcgtc cgggcagcgt aggcgggcga ctcagatccc agccagtgga cttagcccct 300

gtttgctcct ccgataactg gggtgacctt ggttaatatt caccagcagc ctcccccgtt 360

gcccctctgg atccactgct taaatacgga cgaggacagg gccctgtctc ctcagcttca 420

ggcaccacca ctgacctggg acagtgaatc cggactctaa ggtaaatata aaatttttaa 480

gtgtataatg tgttaaacta ctgattctaa ttgtttctct cttttagatt ccaacctttg 540

gaactg 546

<210> 83
<211> 576
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 83
tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 60

cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 120

atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag 180

tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc 240

cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 300

tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catggtgatg 360

cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 420

ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 480

aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt acggtgggag 540

gtctatataa gcagagctgg tttagtgaac cgtcag 576

<210> 84
<211> 150
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 84
ataaacgata acgccgttgg tggcgtgagg catgtaaaag gttacatcat tatcttgttc 60

gccatccggt tggtataaat agacgttcat gttggttttt gtttcagttg caagttggct 120

gcggcgcgcg cagcaccttt gcggccatct 150

<210> 85
<211> 1313
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 85
ggagccgaga gtaattcata caaaaggagg gatcgccttc gcaaggggag agcccaggga 60

ccgtccctaa attctcacag acccaaatcc ctgtagccgc cccacgacag cgcgaggagc 120

atgcgcccag ggctgagcgc gggtagatca gagcacacaa gctcacagtc cccggcggtg 180

gggggagggg cgcgctgagc gggggccagg gagctggcgc ggggcaaact gggaaagtgg 240

tgtcgtgtgc tggctccgcc ctcttcccga gggtgggggа gaacggtata taagtgcggt 300

agtcgccttg gacgttcttt ttcgcaacgg gtttgccgtc agaacgcagg tgagtggcgg 360

gtgtggcttc cgcgggcccc ggagctggag ccctgctctg agcgggccgg gctgatatgc 420

gagtgtcgtc cgcagggttt agctgtgagc attcccactt cgagtggcgg gcggtgcggg 480

ggtgagagtg cgaggcctag cggcaacccc gtagcctcgc ctcgtgtccg gcttgaggcc 540

tagcgtggtg tccgccgccg cgtgccactc cggccgcact atgcgttttt tgtccttgct 600

gccctcgatt gccttccagc agcatgggct aacaaaggga gggtgtgggg ctcactctta 660

aggagcccat gaagcttacg ttggatagga atggaagggc aggaggggcg actggggccc 720

gcccgccttc ggagcacatg tccgacgcca cctggatggg gcgaggcctg tggctttccg 780

aagcaatcgg gcgtgagttt agcctacctg ggccatgtgg ccctagcact gggcacggtc 840

tggcctggcg gtgccgcgtt cccttgcctc ccaacaaggg tgaggccgtc ccgcccggca 900

ccagttgctt gcgcggaaag atggccgctc ccggggccct gttgcaagga gctcaaaatg 960

gaggacgcgg cagcccggtg gagcgggcgg gtgagtcacc cacacaaagg aagagggcct 1020

tgcccctcgc cggccgctgc ttcctgtgac cccgtggtct atcggccgca tagtcacctc 1080

gggcttctct tgagcaccgc tcgtcgcggc ggggggaggg gatctaatgg cgttggagtt 1140

tgttcacatt tggtgggtgg agactagtca ggccagcctg gcgctggaag tcattcttgg 1200

aatttgcccc tttgagtttg gagcgaggct aattctcaag cctcttagcg gttcaaaggt 1260

attttctaaa cccgtttcca ggtgttgtga aagccaccgc taattcaaag caa 1313

<210> 86
<211> 213
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 86
taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa aaatgcttta 60

tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc aataaacaag 120

ttaacaacaa caattgcatt cattttatgt ttcaggttca gggggaggtg tgggaggttt 180

tttaaagcaa gtaaaacctc tacaaatgtg gta 213

<210> 87
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 87
Pro Lys Lys Lys Arg Lys Val
1 5

<210> 88
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 88
Met Asp Trp Thr Trp Arg Ile Leu Phe Leu Val Ala Ala Ala Thr Gly
1 5 10 15

Ala His Ser

<210> 89
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 89
Met Leu Pro Ser Gln Leu Ile Gly Phe Leu Leu Leu Trp Val Pro Ala
1 5 10 15

Ser Arg Gly

<210> 90
<211> 7
<212> PRT
<213> Simian virus 40

<400> 90
Pro Lys Lys Lys Arg Lys Val
1 5

<210> 91
<211> 21
<212> DNA
<213> Simian virus 40

<400> 91
cccaagaaga agaggaaggt g 21

<210> 92
<211> 16
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
Nucleoplasmin bipartite NLS sequence

<400> 92
Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 93
<211> 9
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
C-myc NLS sequence

<400> 93
Pro Ala Ala Lys Arg Val Lys Leu Asp
1 5

<210> 94
<211> 11
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
C-myc NLS sequence

<400> 94
Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro
1 5 10

<210> 95
<211> 38
<212> PRT
<213> Homo sapiens

<400> 95
Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly
1 5 10 15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro
20 25 30

Arg Asn Gln Gly Gly Tyr
35

<210> 96
<211> 42
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
IBB domain from importin-alpha sequence

<400> 96
Arg Met Arg Ile Glx Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu
1 5 10 15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys

20 25 30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val
35 40

<210> 97
<211> 8
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
Myoma T protein sequence

<400> 97
Val Ser Arg Lys Arg Pro Arg Pro
1 5

<210> 98
<211> 8
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown:
Myoma T protein sequence

<400> 98
Pro Pro Lys Lys Ala Arg Glu Asp
1 5

<210> 99
<211> 8
<212> PRT
<213> Homo sapiens

<400> 99
Pro Gln Pro Lys Lys Lys Pro Leu
1 5

<210> 100
<211> 12
<212> PRT

<213> Mus musculus

<400> 100
Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro
1 5 10

<210> 101
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 101
gcgcgctcgc tcgctcactg aggccgcccg ggaaacccgg gcgtgcgcct cagtgagcga 60

gcgagcgcgc 70

<210> 102
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 102
gcgcgctcgc tcgctcactg aggcgcacgc ccgggtttcc cgggcggcct cagtgagcga 60

gcgagcgcgc 70

<210> 103
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 103
gcgcgctcgc tcgctcactg aggccgtcgg gcgacctttg gtcgcccggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 104
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 104
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 105
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 105
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 106
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 106
gcgcgctcgc tcgctcactg aggccgacgc ccgggctttg cccgggcggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 107
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 107
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg ctttgcccgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 108
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 108
gcgcgctcgc tcgctcactg aggccgggca aagcccgacg cccgggcttt gcccgggcgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 109
<211> 77
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 109
gcgcgctcgc tcgctcactg aggccgaaac gtcgggcgac ctttggtcgc ccggcctcag 60

tgagcgagcg agcgcgc 77

<210> 110
<211> 77

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 110
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgtt tcggcctcag 60

tgagcgagcg agcgcgc 77

<210> 111
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 111
gcgcgctcgc tcgctcactg aggcaaagcc tcagtgagcg agcgagcgcg c 51

<210> 112
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 112
gcgcgctcgc tcgctcactg aggctttgcc tcagtgagcg agcgagcgcg c 51

<210> 113
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 113
gcgcgctcgc tcgctcactg aggccgcccg ggcgtcgggc gacctttggt cgcccggcct 60

cagtgagcga gcgagcgcgc 80

<210> 114
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 114
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 115
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 115
gcgcgctcgc tcgctcactg aggcgcccgg gcgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgc 79

<210> 116
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 116
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcgcctc 60

agtgagcgag cgagcgcgc 79

<210> 117
<211> 5
<212> PRT
<213> Influenza virus

<400> 117
Asp Arg Leu Arg Arg
1 5

<210> 118
<211> 7
<212> PRT
<213> Influenza virus

<400> 118
Pro Lys Gln Lys Lys Arg Lys
1 5

<210> 119
<211> 10
<212> PRT
<213> Hepatitis delta virus

<400> 119
Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu
1 5 10

<210> 120
<211> 10
<212> PRT
<213> Mus musculus

<400> 120
Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg
1 5 10

<210> 121
<211> 20
<212> PRT
<213> Homo sapiens

<400> 121
Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys
1   5   10   15

Lys Ser Lys Lys
            20

<210> 122
<211> 17
<212> PRT
<213> Homo sapiens

<400> 122
Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys
1   5   10   15

Lys

<210> 123
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      oligonucleotide

<400> 123
gtttaaac   8

<210> 124
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      oligonucleotide

<400> 124
ttaattaa   8

<210> 125
<211> 141
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 125
aataaacgat aacgccgttg gtggcgtgag gcatgtaaaa ggttacatca ttatcttgtt 60

cgccatccgg ttggtataaa tagacgttca tgttggtttt tgtttcagtt gcaagttggc 120

tgcggcgcgc gcagcacctt t 141

<210> 126
<211> 317
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 126
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 120

tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa 180

ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag 240

aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag 300

gcctaggctt ttgcaaa 317

<210> 127
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 127
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 128
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 128
gagacagaca cactcctgct atgggtactg ctgctctggg ttccaggttc cactggtgac 60

<210> 129
<211> 1260
<212> DNA
<213> Adeno-associated virus - 2

<400> 129
atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag 60

gaggaccagg cctcatacat ctccttcaat gcggcctcca actcgcggtc ccaaatcaag 120

gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg 180

gtgggccagc agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta 240

aacgggtacg atccccaata tgcggcttcc gtctttctgg gatgggccac gaaaaagttc 300

ggcaagagga acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg 360

gaggccatag cccacactgt gcccttctac gggtgcgtaa actggaccaa tgagaacttt 420

cccttcaacg actgtgtcga caagatggtg atctggtggg aggaggggaa gatgaccgcc 480

aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa 540

tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctccaa caccaacatg 600

tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg 660

atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag 720

gaagtcaaag acttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc 780

tacgtcaaaa agggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag 840

cccaaacggg tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgatc 900

aactacgcag acaggtacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg 960

tttccctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga 1020

cagaaagact gtttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa 1080

aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc 1140

actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataaatgatt 1200

taaatcaggt atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga 1260

<210> 130
<211> 1932
<212> DNA
<213> Adeno-associated virus - 2

<400> 130
atgccggggt tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc 60

ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat 120

tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag 180

cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg 240

caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac caccggggtg 300

aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt 360

taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac cagaaatggc 420

gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctccccaaa 480

acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag cgcctgtttg 540

aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag 600

gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag atcaaaaact 660

tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac ctcggagaag 720

cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcgg 780

tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc 840

cccgactacc tggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa 900

attttggaac taaacgggta cgatccccaa tatgcggctt ccgtctttct gggatgggcc 960

acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac taccgggaag 1020

accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc 1080

aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg 1140

aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc 1200

gtggaccaga aatgcaagtc ctcggcccag atagacccga ctcccgtgat cgtcacctcc 1260

aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca ccagcagccg 1320

ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga ctttgggaag 1380

gtcaccaagc aggaagtcaa agacttttc cggtgggcaa aggatcacgt ggttgaggtg 1440

gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa gacccgcccc cagtgacgca 1500

gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac gtcagacgcg 1560

gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg 1620

aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc 1680

ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc tcaacccgtt 1740

tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaaggtg 1800

ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa 1860

caataaatga tttaaatcag gtatggctgc cgatggttat cttccagatt ggctcgagga 1920

cactctctct ga 1932

<210> 131

<211> 1876
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 131
cgcagccacc atggcggggt tttacgagat tgtgattaag gtccccagcg accttgacgg 60

gcatctgccc ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt 120

gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga 180

gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct 240

tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac 300

caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat 360

tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac 420

cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt 480

gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag 540

cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc 600

gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag 660

atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac 720

ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc 780

caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac 840

taaaaccgcc cccgactacc tggtgggcca gcagcccgtg gaggacattt ccagcaatcg 900

gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt ccgtctttct 960

gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac 1020

taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt 1080

aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg 1140

ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag 1200

caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga ctcccgtgat 1260

cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca 1320

ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga 1380

ctttgggaag gtcaccaagc aggaagtcaa agacttttc cggtgggcaa aggatcacgt 1440

ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa gacccgcccc 1500

cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac 1560

gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca 1620

cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc 1680

aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc 1740

tcaacccgtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat 1800

gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg 1860

catctttgaa caataa 1876

<210> 132
<211> 1194
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 132
atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag 60

gaggaccagg cctcatacat ctccttcaat gcggcctcca actcgcggtc ccaaatcaag 120

gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg 180

gtgggccagc agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta 240

aacgggtacg atccccaata tgcggcttcc gtctttctgg gatgggccac gaaaaagttc 300

ggcaagagga acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg 360

gaggccatag cccacactgt gcccttctac gggtgcgtaa actggaccaa tgagaacttt 420

cccttcaacg actgtgtcga caagatggtg atctggtggg aggaggggaa gatgaccgcc 480

aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa 540

tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctccaa caccaacatg 600

tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg 660

atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag 720

gaagtcaaag actttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc 780

tacgtcaaaa agggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag 840

cccaaacggg tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgatc 900

aactacgcag accgctacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg 960

tttccctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga 1020

cagaaagact gtttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa 1080

aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc 1140

actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataa 1194

<210> 133
<211> 1876
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 133
cgcagccacc atggcggggt tttacgagat tgtgattaag gtccccagcg accttgacgg 60

gcatctgccc ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt 120

gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga 180

gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct 240

tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac 300

caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat 360

tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac 420

cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt 480

gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag 540

cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc 600

gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag 660

atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac 720

ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc 780

caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac 840

taaaaccgcc cccgactacc tggtgggcca gcagcccgtg gaggacattt ccagcaatcg 900

gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt ccgtctttct 960

gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac 1020

taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt 1080

aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg 1140

ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag 1200

caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga ctcccgtgat 1260

cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca 1320

ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga 1380

ctttgggaag gtcaccaagc aggaagtcaa agacttttc cggtgggcaa aggatcacgt 1440

ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa gacccgcccc 1500

cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac 1560

gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca 1620

cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc 1680

aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc 1740

tcaacccgtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat 1800

gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg 1860

catctttgaa caataa 1876

<210> 134
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 134
ctaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg c 51

<210> 135
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 135
ctaggactga ggccgcccgg gcaaagcccg ggcgtcgggc gacctttggt cgcccggcct 60

cagtc 65

<210> 136
<211> 67
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 136
ggactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 60

tcctgca 67

<210> 137
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 137
gtgcgggcga ccaaaggtcg cccgacgccc gggcgcactc a 41

<210> 138
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 138
ggactgaggc cgggcgacca aaggtcgccc gacgcccggg cggcctcagt cctgca 56

<210> 139
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 139
ctaggactga ggccgcccgg gcgtcgggcg acctttggtc gcccggcctc agtc 54

<210> 140
<211> 48
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 140
ggactgaggc cgggcgacca aaggtcgccc gacggcctca gtcctgca 48

<210> 141
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 141
ctaggactga ggccgtcggg cgacctttgg tcgcccggcc tcagtc 46

<210> 142
<211> 67
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 142
ggactgaggc ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcgcctcag 60

tcctgca 67

<210> 143
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 143
atacctaggc acgcgtgtta ctagttatta atagtaatca attacgg 47

<210> 144
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 144
atacctaggg gccgcacgcg tgttactag 29

<210> 145
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 145
atacactcag tgcctgcagg cacgtggtcc ggagatccag ac 42

<210> 146
<211> 3754
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 146
cctaggtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact aggggttcct 60

tgtagttaat gattaacccg ccatgctact tatcgcggcc gctcaatatt ggccattagc 120

catattattc attggttata tagcataaat caatattggc tattggccat tgcatacgtt 180

gtatctatat cataatatgt acatttatat tggctcatgt ccaatatgac cgccatgttg 240

gcattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 300

atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccaa 360

cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 420

tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtacatca 480

agtgtatcat atgccaagtc cgccccctat tgacgtcaat gacggtaaat ggcccgcctg 540

gcattatgcc cagtacatga ccttacggga ctttcctact tggcagtaca tctacgtatt 600

agtcatcgct attaccatgg tcgaggtgag ccccacgttc tgcttcactc tccccatctc 660

ccccccctcc ccacccccaa ttttgtattt atttattttt taattatttt gtgcagcgat 720

gggggcgggg gggggggggg ggcgcgcgcc aggcggggcg gggcggggcg aggggcgggg 780

cggggcgagg cggagaggtg cggcggcagc caatcagagc ggcgcgctcc gaaagtttcc 840

ttttatggcg aggcggcggc ggcggcggcc ctataaaaag cgaagcgcgc ggcgggcggg 900

agtcgctgcg acgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc 960

cccggctctg actgaccgcg ttactcccac aggtgagcgg gcgggacggc ccttctcctc 1020

cgggctgtaa ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa 1080

gccttgaggg gctccgggag ggccctttgt gcgggggggа gcggctcggg gggtgcgtgc 1140

gtgtgtgtgt gcgtggggag cgccgcgtgc ggcccgcgct gcccggcggc tgtgagcgct 1200

gcgggcgcgg cgcggggctt tgtgcgctcc gcagtgtgcg cgaggggagc gcggccgggg 1260

gcggtgcccc gcggtgcggg gggggctgcg aggggaacaa aggctgcgtg cggggtgtgt 1320

gcgtgggggg gtgagcaggg ggtgtgggcg cggcggtcgg gctgtaaccc ccccctgcac 1380

ccccctcccc gagttgctga gcacggcccg gcttcgggtg cggggctccg tacggggcgt 1440

ggcgcggggc tcgccgtgcc gggcgggggg tggcggcagg tgggggtgcc gggcggggcg 1500

gggccgcctc gggccgggga gggctcgggg gaggggcgcg gcggcccccg gagcgccggc 1560

ggctgtcgag gcgcggcgag ccgcagccat tgccttttat ggtaatcgtg cgagagggcg 1620

cagggacttc ctttgtccca aatctgtgcg gagccgaaat ctgggaggcg ccgccgcacc 1680

ccctctagcg ggcgcggggc gaagcggtgc ggcgccggca ggaaggaaat gggcggggag 1740

ggccttcgtg cgtcgccgcg ccgccgtccc cttctccctc tccagcctcg gggctgtccg 1800

cgggggggacg gctgccttcg ggggggacgg ggcagggcgg ggttcggctt ctggcgtgtg 1860

accggcggct ctagagcctc tgctaaccat gttttagcct tcttcttttt cctacagctc 1920

ctgggcaacg tgctggttat tgtgctgtct catcatttgt cgacagaatt cctcgaagat 1980

ccgaaggggt tcaagcttgg cattccggta ctgttggtaa agccagttta aacgccgcca 2040

ccatggtgag caagggcgag gagctgttca ccggggtggt gcccatcctg gtcgagctgg 2100

acggcgacgt aaacggccac aagttcagcg tgtccggcga gggcgagggc gatgccacct 2160

acggcaagct gaccctgaag ttcatctgca ccaccggcaa gctgcccgtg ccctggccca 2220

ccctcgtgac caccctgacc tacggcgtgc agtgcttcag ccgctacccc gaccacatga 2280

agcagcacga cttcttcaag tccgccatgc ccgaaggcta cgtccaggag cgcaccatct 2340

tcttcaagga cgacggcaac tacaagaccc gcgccgaggt gaagttcgag ggcgacaccc 2400

tggtgaaccg catcgagctg aagggcatcg acttcaagga ggacggcaac atcctggggc 2460

acaagctgga gtacaactac aacagccaca acgtctatat catggccgac aagcagaaga 2520

acggcatcaa ggtgaacttc aagatccgcc acaacatcga ggacggcagc gtgcagctcg 2580

ccgaccacta ccagcagaac acccccatcg gcgacggccc cgtgctgctg cccgacaacc 2640

actacctgag cacccagtcc gccctgagca aagaccccaa cgagaagcgc gatcacatgg 2700

tcctgctgga gttcgtgacc gccgccggga tcactctcgg catggacgag ctgtacaagt 2760

aattaattaa gagcatctta ccgccattta ttcccatatt tgttctgttt ttcttgattt 2820

gggtatacat ttaaatgtta ataaaacaaa atggtggggc aatcatttac atttttaggg 2880

atatgtaatt actagttcag gtgtattgcc acaagacaaa catgttaaga aactttcccg 2940

ttatttacgc tctgttcctg ttaatcaacc tctggattac aaaatttgtg aaagattgac 3000

tgatattctt aactatgttg ctccttttac gctgtgtgga tatgctgctt tatagcctct 3060

gtatctagct attgcttccc gtacggcttt cgttttctcc tccttgtata aatcctggtt 3120

gctgtctctt ttagaggagt tgtggcccgt tgtccgtcaa cgtggcgtgg tgtgctctgt 3180

gtttgctgac gcaaccccca ctggctgggg cattgccacc acctgtcaac tcctttctgg 3240

gactttcgct ttccccctcc cgatcgccac ggcagaactc atcgccgcct gccttgcccg 3300

ctgctggaca ggggctaggt tgctgggcac tgataattcc gtggtgttgt ctgtgccttc 3360

tagttgccag ccatctgttg tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc 3420

cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg 3480

tcattctatt ctgggggggtg gggtggggca ggacagcaag ggggaggatt gggaagacaa 3540

tagcaggcat gctggggatg cggtgggctc tatggctcta gagcatggct acgtagataa 3600

gtagcatggc gggttaatca ttaactacac ctgcagcagg aacccctagt gatggagttg 3660

gccactccct ctctgcgcgc tcgctcgctc cctgcaggac tgaggccggg cgaccaaagg 3720

tcgcccgacg cccgggcggc ctcagtcctg cagg 3754

<210> 147
<211> 8418
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 147
ggcagctgcg cgctcgctcg ctcacctagg ccgcccgggc aaagcccggg cgtcgggcga 60

cctttggtcg cccggcctag gtgagcgagc gagcgcgcag agagggagtg gccaactcca 120

tcactagggg ttccttgtag ttaatgatta acccgccatg ctacttatcg cggccgctca 180

atattggcca ttagccatat tattcattgg ttatatagca taaatcaata ttggctattg 240

gccattgcat acgttgtatc tatatcataa tatgtacatt tatattggct catgtccaat 300

atgaccgcca tgttggcatt gattattgac tagttattaa tagtaatcaa ttacggggtc 360

attagttcat agcccatata tggagttccg cgttacataa cttacggtaa atggcccgcc 420

tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg ttcccatagt 480

aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt aaactgccca 540

cttggcagta catcaagtgt atcatatgcc aagtccgccc cctattgacg tcaatgacgg 600

taaatggccc gcctggcatt atgcccagta catgacctta cgggactttc ctacttggca 660

gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca cgttctgctt 720

cactctcccc atctcccccc cctccccacc cccaattttg tatttattta ttttttaatt 780

attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg gggcggggcg 840

gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc agagcggcgc 900

gctccgaaag tttcctttta tggcgaggcg gcggcggcgg cggccctata aaaagcgaag 960

cgcgcggcgg gcgggagtcg ctgcgacgct gccttcgccc cgtgccccgc tccgccgccg 1020

cctcgcgccg cccgccccgg ctctgactga ccgcgttact cccacaggtg agcgggcggg 1080

acggcccttc tcctccgggc tgtaattagc gcttggttta atgacggctt gtttcttttc 1140

tgtggctgcg tgaaagcctt gaggggctcc gggagggccc tttgtgcggg ggggagcggc 1200

tcgggggggtg cgtgcgtgtg tgtgtgcgtg gggagcgccg cgtgcggccc gcgctgcccg 1260

gcggctgtga gcgctgcggg cgcggcgcgg ggctttgtgc gctccgcagt gtgcgcgagg 1320

ggagcgcggc cgggggcggt gccccgcggt gcgggggggg ctgcgagggg aacaaaggct 1380

gcgtgcgggg tgtgtgcgtg gggggggtgag caggggggtgt gggcgcggcg gtcgggctgt 1440

aacccccccc tgcacccccc tccccgagtt gctgagcacg gcccggcttc gggtgcgggg 1500

ctccgtacgg ggcgtggcgc ggggctcgcc gtgccgggcg gggggtggcg gcaggtgggg 1560

gtgccgggcg gggcggggcc gcctcgggcc ggggagggct cgggggaggg gcgcggcggc 1620

ccccggagcg ccggcggctg tcgaggcgcg gcgagccgca gccattgcct tttatggtaa 1680

tcgtgcgaga gggcgcaggg acttcctttg tcccaaatct gtgcggagcc gaaatctggg 1740

aggcgccgcc gcacccccctc tagcgggcgc ggggcgaagc ggtgcggcgc cggcaggaag 1800

gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtccccttct ccctctccag 1860

cctcggggct gtccgcgggg ggacggctgc cttcgggggg gacggggcag ggcggggttc 1920

ggcttctggc gtgtgaccgg cggctctaga gcctctgcta accatgtttt agccttcttc 1980

tttttcctac agctcctggg caacgtgctg gttattgtgc tgtctcatca tttgtcgaca 2040

gaattcctcg aagatccgaa ggggttcaag cttggcattc cggtactgtt ggtaaagcca 2100

gtttaaacgc cgccaccatg gtgagcaagg gcgaggagct gttcaccggg gtggtgccca 2160

tcctggtcga gctggacggc gacgtaaacg gccacaagtt cagcgtgtcc ggcgagggcg 2220

agggcgatgc cacctacggc aagctgaccc tgaagttcat ctgcaccacc ggcaagctgc 2280

ccgtgccctg gcccaccctc gtgaccaccc tgacctacgg cgtgcagtgc ttcagccgct 2340

accccgacca catgaagcag cacgacttct tcaagtccgc catgcccgaa ggctacgtcc 2400

aggagcgcac catcttcttc aaggacgacg gcaactacaa gacccgcgcc gaggtgaagt 2460

tcgagggcga caccctggtg aaccgcatcg agctgaaggg catcgacttc aaggaggacg 2520

gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc tatatcatgg 2580

ccgacaagca gaagaacggc atcaaggtga acttcaagat ccgccacaac atcgaggacg 2640

gcagcgtgca gctcgccgac cactaccagc agaacacccc catcggcgac ggccccgtgc 2700

tgctgcccga caaccactac ctgagcaccc agtccgccct gagcaaagac cccaacgaga 2760

agcgcgatca catggtcctg ctggagttcg tgaccgccgc cgggatcact ctcggcatgg 2820

acgagctgta caagtaatta attaagagca tcttaccgcc atttattccc atatttgttc 2880

tgtttttctt gatttgggta tacatttaaa tgttaataaa acaaaatggt ggggcaatca 2940

tttacatttt tagggatatg taattactag ttcaggtgta ttgccacaag acaaacatgt 3000

taagaaactt tcccgttatt tacgctctgt tcctgttaat caacctctgg attacaaaat 3060

ttgtgaaaga ttgactgata ttcttaacta tgttgctcct tttacgctgt gtggatatgc 3120

tgctttatag cctctgtatc tagctattgc ttcccgtacg gctttcgttt tctcctcctt 3180

gtataaatcc tggttgctgt ctcttttaga ggagttgtgg cccgttgtcc gtcaacgtgg 3240

cgtggtgtgc tctgtgtttg ctgacgcaac ccccactggc tggggcattg ccaccacctg 3300

tcaactcctt tctgggactt tcgctttccc cctcccgatc gccacggcag aactcatcgc 3360

cgcctgcctt gcccgctgct ggacaggggc taggttgctg ggcactgata attccgtggt 3420

gttgtctgtg ccttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt 3480

gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca 3540

ttgtctgagt aggtgtcatt ctattctggg gggtggggtg gggcaggaca gcaaggggga 3600

ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg ctctagagca 3660

tggctacgta gataagtagc atggcgggtt aatcattaac tacacctgca gcaggaaccc 3720

ctagtgatgg agttggccac tccctctctg cgcgctcgct cgctccctgc aggactgagg 3780

ccgggcgacc aaaggtcgcc cgacgcccgg gcggcctcag tcctgcaggg agcgagcgag 3840

cgcgcagctg cctgcacggg cgcgccggta ccgggagatg ggggaggcta actgaaacac 3900

ggaaggagac aataccggaa ggaacccgcg ctatgacggc aataaaaaga cagaataaaa 3960

cgcacgggtg ttgggtcgtt tgttcataaa cgcggggttc ggtcccaggg ctggcactct 4020

gtcgataccc caccgagacc ccattgggac caatacgccc gcgtttcttc cttttcccca 4080

ccccaacccc caagttcggg tgaaggccca gggctcgcag ccaacgtcgg ggcggcaagc 4140

cctgccatag ccactacggg tacgtaggcc aaccactaga actatagcta gagtcctggg 4200

cgaacaaacg atgctcgcct tccagaaaac cgaggatgcg aaccacttca tccggggtca 4260

gcaccaccgg caagcgccgc gacggccgag gtctaccgat ctcctgaagc cagggcagat 4320

ccgtgcacag caccttgccg tagaagaaca gcaaggccgc caatgcctga cgatgcgtgg 4380

agaccgaaac cttgcgctcg ttcgccagcc aggacagaaa tgcctcgact tcgctgctgc 4440

ccaaggttgc cgggtgacgc acaccgtgga aacggatgaa ggcacgaacc cagttgacat 4500

aagcctgttc ggttcgtaaa ctgtaatgca agtagcgtat gcgctcacgc aactggtcca 4560

gaaccttgac cgaacgcagc ggtggtaacg gcgcagtggc ggttttcatg gcttgttatg 4620

actgtttttt tgtacagtct atgcctcggg catccaagca gcaagcgcgt tacgccgtgg 4680

gtcgatgttt gatgttatgg agcagcaacg atgttacgca gcagcaacga tgttacgcag 4740

cagggcagtc gccctaaaac aaagttaggt ggctcaagta tgggcatcat tcgcacatgt 4800

aggctcggcc ctgaccaagt caaatccatg cgggctgctc ttgatctttt cggtcgtgag 4860

ttcggagacg tagccaccta ctcccaacat cagccggact ccgattacct cgggaacttg 4920

ctccgtagta agacattcat cgcgcttgct gccttcgacc aagaagcggt tgttggcgct 4980

ctcgcggctt acgttctgcc caggtttgag cagccgcgta gtgagatcta tatctatgat 5040

ctcgcagtct ccggcgagca ccggaggcag ggcattgcca ccgcgctcat caatctcctc 5100

aagcatgagg ccaacgcgct tggtgcttat gtgatctacg tgcaagcaga ttacggtgac 5160

gatcccgcag tggctctcta tacaaagttg ggcatacggg aagaagtgat gcactttgat 5220

atcgacccaa gtaccgccac ctaacaattc gttcaagccg agatcggctt cccggccgcg 5280

gagttgttcg gtaaattgtc acaacgccgc gaatatagtc tttaccatgc ccttggccac 5340

gcccctcttt aatacgacgg gcaatttgca cttcagaaaa tgaagagttt gctttagcca 5400

taacaaaagt ccagtatgct ttttcacagc ataactggac tgatttcagt ttacaactat 5460

tctgtctagt ttaagacttt attgtcatag tttagatcta ttttgttcag tttaagactt 5520

tattgtccgc ccacacccgc ttacgcaggg catccattta ttactcaacc gtaaccgatt 5580

ttgccaggtt acgcggctgg tctgcggtgt gaaataccgc acagatgcgt aaggagaaaa 5640

taccgcatca ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 5700

ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 5760

gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 5820

gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca caaaaatcga 5880

cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttcccccc 5940

ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 6000

tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg 6060

gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 6120

tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 6180

ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 6240

ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 6300

ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc 6360

accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 6420

tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca 6480

cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat 6540

taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 6600

caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt 6660

gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt 6720

gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag 6780

ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct 6840

attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt 6900

gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc 6960

tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt 7020

agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg 7080

gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg 7140

actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct 7200

tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc 7260

attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt 7320

tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt 7380

tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg 7440

aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta tcagggttat 7500

tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 7560

cgcacatttc cccgaaaagt gccacctgaa attgtaaacg ttaatatttt gttaaaattc 7620

gcgttaaatt tttgttaaat cagctcattt tttaaccaat aggccgaaat cggcaaaatc 7680

ccttataaat caaaagaata gaccgagata gggttgagtg ttgttccagt ttggaacaag 7740

agtccactat taaagaacgt ggactccaac gtcaaagggc gaaaaaccgt ctatcagggc 7800

gatggcccac tacgtgaacc atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa 7860

gcactaaatc ggaaccctaa agggagcccc cgatttagag cttgacgggg aaagccggcg 7920

aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc gctggcaagt 7980

gtagcggtca cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgcc gctacagggc 8040

gcgtcccatt cgccattcag gctgcaaata agcgttgata ttcagtcaat tacaaacatt 8100

aataacgaag agatgacaga aaaattttca ttctgtgaca gagaaaaagt agccgaagat 8160

gacggtttgt cacatggagt tggcaggatg tttgattaaa aacataacag gaagaaaaat 8220

gccccgctgt gggcggacaa aatagttggg aactgggagg ggtggaaatg gagtttttaa 8280

ggattattta gggaagagtg acaaaataga tgggaactgg gtgtagcgtc gtaagctaat 8340

acgaaaatta aaaatgacaa aatagtttgg aactagattt cacttatctg gttcggatct 8400

cctagtgagc tccctgca 8418

<210> 148
<211> 225
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 148
tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct 60

ggaaggtgcc actcccactg tcctttcctа ataaaatgag gaaattgcat cgcattgtct 120

gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg 180

ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggc 225

<210> 149
<211> 1177
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 149
ggctcagagg ctcagaggca cacaggagtt tctgggctca ccctgccccc ttccaacccc 60

tcagttccca tcctccagca gctgtttgtg tgctgcctct gaagtccaca ctgaacaaac 120

ttcagcctac tcatgtccct aaaatgggca aacattgcaa gcagcaaaca gcaaacacac 180

agccctccct gcctgctgac cttggagctg gggcagaggt cagagacctc tctgggccca 240

tgccacctcc aacatccact cgacccctg gaatttcggt ggagaggagc agaggttgtc 300

ctggcgtggt ttaggtagtg tgagagggtc cgggttcaaa accacttgct gggtggggag 360

tcgtcagtaa gtggctatgc cccgaccccg aagcctgttt ccccatctgt acaatggaaa 420

tgataaagac gcccatctga tagggttttt gtggcaaata aacatttggt ttttttgttt 480

tgttttgttt tgtttttga gatggaggtt tgctctgtcg cccaggctgg agtgcagtga 540

cacaatctca tctcaccaca accttcccct gcctcagcct cccaagtagc tgggattaca 600

agcatgtgcc accacacctg gctaattttc tatttttagt agagacgggt ttctccatgt 660

tggtcagcct cagcctccca agtaactggg attacaggcc tgtgccacca cacccggcta 720

atttttcta tttttgacag ggacggggtt tcaccatgtt ggtcaggctg gtctagaggt 780

accggatctt gctaccagtg gaacagccac taaggattct gcagtgagag cagagggcca 840

gctaagtggt actctcccag agactgtctg actcacgcca ccccctccac cttggacaca 900

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 960

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 1020

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 1080

gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 1140

tcctcagctt caggcaccac cactgacctg ggacagt 1177

<210> 150
<211> 1326

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 150
ctgcagggcc cactagtgga gccgagagta attcatacaa aaggagggat cgccttcgca 60

aggggagagc ccagggaccg tccctaaatt ctcacagacc caaatccctg tagccgcccc 120

acgacagcgc gaggagcatg cgcccagggc tgagcgcggg tagatcagag cacacaagct 180

cacagtcccc ggcggtgggg ggaggggcgc gctgagcggg ggccagggag ctggcgcggg 240

gcaaactggg aaagtggtgt cgtgtgctgg ctccgccctc ttcccgaggg tgggggagaa 300

cggtatataa gtgcggtagt cgccttggac gttctttttc gcaacgggtt tgccgtcaga 360

acgcaggtga gtggcgggtg tggcttccgc gggccccgga gctggagccc tgctctgagc 420

gggccgggct gatatgcgag tgtcgtccgc agggtttagc tgtgagcatt cccacttcga 480

gtggcgggcg gtgcgggggt gagagtgcga ggcctagcgg caaccccgta gcctcgcctc 540

gtgtccggct tgaggcctag cgtggtgtcc gccgccgcgt gccactccgg ccgcactatg 600

cgtttttgt ccttgctgcc ctcgattgcc ttccagcagc atgggctaac aaagggaggg 660

tgtggggctc actcttaagg agcccatgaa gcttacgttg gataggaatg gaagggcagg 720

aggggcgact ggggcccgcc cgccttcgga gcacatgtcc gacgccacct ggatggggcg 780

aggcctgtgg ctttccgaag caatcgggcg tgagtttagc ctacctgggc catgtggccc 840

tagcactggg cacggtctgg cctggcggtg ccgcgttccc ttgcctccca acaagggtga 900

ggccgtcccg cccggcacca gttgcttgcg cggaaagatg gccgctcccg gggccctgtt 960

gcaaggagct caaaatggag gacgcggcag cccggtggag cgggcgggtg agtcacccac 1020

acaaaggaag agggccttgc ccctcgccgg ccgctgcttc ctgtgacccc gtggtctatc 1080

ggccgcatag tcacctcggg cttctcttga gcaccgctcg tcgcggcggg gggaggggat 1140

ctaatggcgt tggagtttgt tcacatttgg tgggtggaga ctagtcaggc cagcctggcg 1200

ctggaagtca ttcttggaat ttgccccttt gagtttggag cgaggctaat tctcaagcct 1260

cttagcggtt caaaggtatt ttctaaaccc gtttccaggt gttgtgaaag ccaccgctaa 1320

ttcaaa 1326

<210> 151
<211> 573
<212> DNA
<213> Mus musculus

<400> 151
gtaagagttt tatgtttttt catctctgct tgtatttttc tagtaatgga agcctggtat 60

tttaaaatag ttaaattttc ctttagtgct gatttctaga ttattattac tgttgttgtt 120

gttattattg tcattatttg catctgagaa cccttaggtg gttatattat tgatatattt 180

ttggtatctt tgatgacaat aatgggggat tttgaaagct tagctttaaa tttcttttaa 240

ttaaaaaaaa atgctaggca gaatgactca aattacgttg gatacagttg aatttattac 300

ggtctcatag ggcctgcctg ctcgaccatg ctatactaaa aattaaaagt gtgtgttact 360

aattttataa atggagtttc catttatatt tacctttatt tcttatttac cattgtctta 420

gtagatattt acaaacatga cagaaacact aaatcttgag tttgaatgca cagatataaa 480

cacttaacgg gttttaaaaa taataatgtt ggtgaaaaaa tataactttg agtgtagcag 540

agaggaacca ttgccacctt cagattttcc tgt 573

<210> 152
<211> 1993
<212> DNA
<213> Mus musculus

<400> 152
acgatcggga actggcatct tcagggagta gcttaggtca gtgaagagaa gaacaaaaag 60

cagcatatta cagttagttg tcttcatcaa tctttaaata tgttgtgtgg tttttctctc 120

cctgtttcca cagacaagag tgagatcgcc catcggtata atgatttggg agaacaacat 180

ttcaaaggcc tgtaagttat aatgctgaaa gcccacttaa tatttctggt agtattagtt 240

aaagttttaa aacacctttt tccaccttga gtgtgagaat tgtagagcag tgctgtccag 300

tagaaatgtg tgcattgaca gaaagactgt ggatctgtgc tgagcaatgt ggcagccaga 360

gatcacaagg ctatcaagca ctttgcacat ggcaagtgta actgagaagc acacattcaa 420

ataatagtta attttaattg aatgtatcta gccatgtgtg gctagtagct cctttcctgg 480

agagagaatc tggagcccac atctaacttg ttaagtctgg aatcttattt tttatttctg 540

gaaaggtcta tgaactatag ttttgggggc agctcactta ctaactttta atgcaataag 600

atctcatggt atcttgagaa cattattttg tctctttgta gtactgaaac cttatacatg 660

tgaagtaagg ggtctatact taagtcacat ctccaacctt agtaatgttt taatgtagta 720

aaaaaatgag taattaattt atttttagaa ggtcaatagt atcatgtatt ccaaataaca 780

gaggtatatg gttagaaaag aaacaattca aaggacttat ataatatcta gccttgacaa 840

tgaataaatt tagagagtag tttgcctgtt tgcctcatgt tcataaatct attgacacat 900

atgtgcatct gcacttcagc atggtagaag tccatattcc tttgcttgga aaggcaggtg 960

ttcccattac gcctcagaga atagctgacg ggaagaggct ttctagatag ttgtatgaaa 1020

gatatacaaa atctcgcagg tatacacagg catgatttgc tggttgggag agccacttgc 1080

ctcatactga ggtttttgtg tctgcttttc agagtcctga ttgccttttc ccagtatctc 1140

cagaaatgct catacgatga gcatgccaaa ttagtgcagg aagtaacaga ctttgcaaag 1200

acgtgtgttg ccgatgagtc tgccgccaac tgtgacaaat cccttgtgag taccttctga 1260

ttttgtggat ctactttcct gctttctgga actctgtttc aaagccaatc atgactccat 1320

cacttaaggc cccgggaaca ctgtggcaga gggcagcaga gagattgata aagccagggt 1380

gatgggaatt ttctgtggga ctccatttca tagtaattgc agaagctaca atacactcaa 1440

aaagtctcac cacatgactg cccaaatggg agcttgacag tgacagtgac agtagatatg 1500

ccaaagtgga tgagggaaag accacaagag ctaaaccctg taaaaagaac tgtaggcaac 1560

taaggaatgc agagagaaga agttgccttg gaagagcata ccaactgcct ctccaatacc 1620

aatggtcatc cctaaaacat acgtatgaat aacatgcaga ctaagcaggc tacatttagg 1680

aatatacatg tatttacata aatgtatatg catgtaacaa caatgaatga aaactgaggt 1740

catggatctg aaagagagca agggggctta catgagaggg tttggaggga ggggttggag 1800

ggagggaggt attattcttt agttttacag ggaacgtagt aaaaacatag gcttctccca 1860

aaggagcaga gcccatgagg agctgtgcaa ggttccccag cttgatttta cctgctcctc 1920

aaattccctt gatttgtttt tattataatg actttactcc tagcttttag tgtcagatag 1980

aaaacatgga agg 1993

<210> 153
<211> 1350
<212> DNA
<213> Homo sapiens

<400> 153
taggaggctg aggcaggagg atcgcttgag cccaggagtt cgagaccagc ctgggcaaca 60

tagtgtgatc ttgtatctat aaaaataaac aaaattagct tggtgtggtg gcgcctgtag 120

tccccagcca cttggagggg tgaggtgaga ggattgcttg agcccgggat ggtccaggct 180

gcagtgagcc atgatcgtgc cactgcactc cagcctgggc gacagagtga gaccctgtct 240

cacaacaaca acaacaacaa caaaaaggct gagctgcacc atgcttgacc cagtttctta 300

aaattgttgt caaagcttca ttcactccat ggtgctatag agcacaagat tttatttggt 360

gagatggtgc tttcatgaat tcccccaaca gagccaagct ctccatctag tggacaggga 420

agctagcagc aaaccttccc ttcactacaa aacttcattg cttggccaaa aagagagtta 480

attcaatgta gacatctatg taggcaatta aaaacctatt gatgtataaa acagtttgca 540

ttcatggagg gcaactaaat acattctagg actttataaa agatcacttt ttatttatgc 600

acagggtgga acaagatgga ttatcaagtg tcaagtccaa tctatgacat caattattat 660

acatcggagc cctgccaaaa aatcaatgtg aagcaaatcg cagcccgcct cctgcctccg 720

ctctactcac tggtgttcat ctttggtttt gtgggcaaca tgctggtcat cctcatcctg 780

ataaactgca aaaggctgaa gagcatgact gacatctacc tgctcaacct ggccatctct 840

gacctgtttt tccttcttac tgtccccttc tgggctcact atgctgccgc ccagtgggac 900

tttggaaata caatgtgtca actcttgaca gggctctatt ttataggctt cttctctgga 960

atcttcttca tcatcctcct gacaatcgat aggtacctgg ctgtcgtcca tgctgtgttt 1020

gctttaaaag ccaggacggt cacctttggg gtggtgacaa gtgtgatcac ttgggtggtg 1080

gctgtgtttg cgtctctccc aggaatcatc tttaccagat ctcaaaaaga aggtcttcat 1140

tacacctgca gctctcattt tccatacagt cagtatcaat tctggaagaa tttccagaca 1200

ttaaagatag tcatcttggg gctggtcctg ccgctgcttg tcatggtcat ctgctactcg 1260

ggaatcctaa aaactctgct tcggtgtcga aatgagaaga agaggcacag ggctgtgagg 1320

cttatcttca ccatcatgat tgtttatttt 1350

<210> 154
<211> 1223
<212> DNA
<213> Homo sapiens

<400> 154
tgacagagac tcttgggatg acgcactgct gcatcaaccc catcatctat gcctttgtcg 60

gggagaagtt cagaaactac ctcttagtct tcttccaaaa gcacattgcc aaacgcttct 120

gcaaatgctg ttctattttc cagcaagagg ctcccgagcg agcaagctca gtttacaccc 180

gatccactgg ggagcaggaa atatctgtgg gcttgtgaca cggactcaag tgggctggtg 240

acccagtcag agttgtgcac atggcttagt tttcatacac agcctgggct gggggtgggg 300

tgggagaggt ctttttaaa aggaagttac tgttatagag ggtctaagat tcatccattt 360

atttggcatc tgtttaaagt agattagatc ttttaagccc atcaattata gaaagccaaa 420

tcaaaatatg ttgatgaaaa atagcaacct ttttatctcc ccttcacatg catcaagtta 480

ttgacaaact ctcccttcac tccgaaagtt ccttatgtat atttaaaaga aagcctcaga 540

gaattgctga ttcttgagtt tagtgatctg aacagaaata ccaaaattat ttcagaaatg 600

tacaactttt tacctagtac aaggcaacat ataggttgta aatgtgttta aaacaggtct 660

ttgtcttgct atggggagaa aagacatgaa tatgattagt aaagaaatga cacttttcat 720

gtgtgatttc ccctccaagg tatggttaat aagtttcact gacttagaac caggcgagag 780

acttgtggcc tgggagagct ggggaagctt cttaaatgag aaggaatttg agttggatca 840

tctattgctg gcaaagacag aagcctcact gcaagcactg catgggcaag cttggctgta 900

gaaggagaca gagctggttg ggaagacatg gggaggaagg acaaggctag atcatgaaga 960

accttgacgg cattgctccg tctaagtcat gagctgagca gggagatcct ggttggtgtt 1020

gcagaaggtt tactctgtgg ccaaaggagg gtcaggaagg atgagcattt agggcaagga 1080

gaccaccaac agccctcagg tcagggtgag gatggcctct gctaagctca aggcgtgagg 1140

atgggaagga gggaggtatt cgtaaggatg ggaaggaggg aggtattcgt gcagcatatg 1200

aggatgcaga gtcagcagaa ctg 1223

<210> 155
<211> 215
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 155
gaacgctgac gtcatcaacc cgctccaagg aatcgcgggc ccagtgtcac taggcgggaa 60

cacccagcgc gcgtgcgccc tggcaggaag atggctgtga gggacagggg agtggcgccc 120

tgcaatattt gcatgtcgct atgtgttctg ggaaatcacc ataaacgtga aatgtctttg 180

gatttgggaa tcttataagt tctgtatgag accac 215

<210> 156
<211> 141
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 156

cctgcaggca gctgcgcgct cgctcgctca cctaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctaggtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc t 141

<210> 157
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 157
gcgcgctcgc tcgctcacc 19

<210> 158
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 158
ctaggtgagc gagcgagcgc gc 22

<210> 159
<211> 75
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 159
cctgcaggac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg 60

cctcagtcct gcagg 75

<210> 160
<211> 130
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 160
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagt 60

gcgggcgacc aaaggtcgcc cgacgcccgg gcgcactcag tgagcgagcg agcgcgcagc 120

tgcctgcagg 130

<210> 161
<211> 142
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 161
cctgcaggca gctgcgcgct cgctcgctcc ctaggactga ggccgcccgg gcgtcgggcg 60

acctttggtc gcccggcctc agtcctaggg agcgagcgag cgcgcagaga gggagtggcc 120

aactccatca ctaggggttc ct 142

<210> 162
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 162
gcgcgctcgc tcgctcactg agtgcgggcg accaaaggtc gcccgacgcc cgggcgcact 60

cagtgagcga gcgagcgcgc 80

<210> 163
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 163
gcgcgctcgc tcgctcactg a 21

<210> 164
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 164
gtgagcgagc gagcgcgc 18

<210> 165
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 165
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgactttgtc 60

gcccggcctc agtgagcgag cgagcgcgc 89

<210> 166
<211> 89
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 166
gcgcgctcgc tcgctcactg aggccgggcg acaaagtcgc ccgacgcccg ggctttgccc 60

gggcggcctc agtgagcgag cgagcgcgc 89

<210> 167
<211> 87
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 167
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgattttcgc 60

ccggcctcag tgagcgagcg agcgcgc 87

<210> 168
<211> 87
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 168
gcgcgctcgc tcgctcactg aggccgggcg aaaatcgccc gacgcccggg ctttgcccgg 60

gcggcctcag tgagcgagcg agcgcgc 87

<210> 169
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 169
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgtttcgccc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 170
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 170
gcgcgctcgc tcgctcactg aggccgggcg aaacgcccga cgcccgggct ttgcccgggc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 171
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 171
gcgcgctcgc tcgctcactg aggccgcccg ggaaacccgg gcgtcgggcg acctttggtc 60

gcccggcctc agtgagcgag cgagcgcgc 89

<210> 172
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 172
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggtttccc 60

gggcggcctc agtgagcgag cgagcgcgc 89

<210> 173
<211> 87
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 173
gcgcgctcgc tcgctcactg aggccgcccg gaaaccgggc gtcgggcgac ctttggtcgc 60

ccggcctcag tgagcgagcg agcgcgc 87

<210> 174
<211> 87
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 174
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cggtttccgg 60

gcggcctcag tgagcgagcg agcgcgc 87

<210> 175
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 175
gcgcgctcgc tcgctcactg aggccgcccg aaacgggcgt cgggcgacct ttggtcgccc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 176
<211> 85
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 176
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgtttcgggc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 177
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 177
gcgcgctcgc tcgctcactg aggccgccca aagggcgtcg ggcgaccttt ggtcgcccgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 178
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 178
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc ctttgggcgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 179
<211> 81
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 179
gcgcgctcgc tcgctcactg aggccgccaa aggcgtcggg cgacctttgg tcgcccggcc 60

tcagtgagcg agcgagcgcg c 81

<210> 180
<211> 81
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 180
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc tttggcggcc 60

tcagtgagcg agcgagcgcg c 81

<210> 181
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 181
gcgcgctcgc tcgctcactg aggccgcaaa gcgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgc 79

<210> 182
<211> 79
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 182
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgct ttgcggcctc 60

agtgagcgag cgagcgcgc 79

<210> 183
<211> 81
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 183
ctgcgcgctc gctcgctcac tgaggccgaa acgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgca g 81

<210> 184
<211> 81
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 184
ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg tttcggcctc 60

agtgagcgag cgagcgcgca g 81

<210> 185
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 185
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 186
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 186
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 187
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 187
gcgcgctcgc tcgctcactg aggcgcccgg gcgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgc 79

<210> 188
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 188
ggagtcaaag ttctgtttgc cctgatctgc atcgctgtgg ccgaggcc 48

<210> 189
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 189
attcatacca acttgaagaa aaagttcagc ctcttcatcc tggtctttct cctgttcgca 60

gtcatctgtg tttggaagaa agggagcgac tatgaggcc 99

<210> 190
<211> 588
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 190
gcccctctcc ctccccccc cctaacgtta ctggccgaag ccgcttggaa taaggccggt 60

gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggcaat gtgagggccc 120

ggaaacctgg ccctgtcttc ttgacgagca ttcctagggg tctttccct ctcgccaaag 180

gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct tcttgaagac 240

aaacaacgtc tgtagcgacc ctttgcaggc agcggaaccc cccacctggc gacaggtgcc 300

tctgcggcca aaagccacgt gtataagata cacctgcaaa ggcggcacaa ccccagtgcc 360

acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctcctcaagc gtattcaaca 420

aggggctgaa ggatgcccag aaggtacccc attgtatggg atctgatctg gggcctcggt 480

gcacatgctt tacatgtgtt tagtcgaggt taaaaaaacg tctaggcccc ccgaaccacg 540

gggacgtggt tttcctttga aaaacacgat gataatatgg ccacaacc 588