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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 December 2012 (13.12.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/170431 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2012/040932

(22) International Filing Date:
5 June 2012 (05.06.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/493,918 6 June 2011 (06.06.2011) US

(71) Applicant *(for all designated States except US)*: **BLUEBIRD BIO, INC.** [US/US]; 840 Memorial Drive, Cambridge, Massachusetts 02139 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **KUTNER, Robert H.** [US/US]; 15 North Beacon Street #915, Allston, Massachusetts 02134 (US).

(74) Agents: **MCDONALD, Michael J.** et al.; Cooley LLP, 777 - 6th Street, NW, Suite 1100, Washington, District of Columbia 20001 (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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, 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, 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:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

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

(54) Title: IMPROVED GENESWITCH SYSTEMS

(57) Abstract: The present invention provides improved molecular switch gene expression systems and improved retroviral packaging cells and viral particle production. In various embodiments, the present invention also provides retroviral vectors that have increased control over gene expression. Further, the present invention provides vector systems that are useful in methods of gene therapy.

WO 2012/170431 A2

IMPROVED GENESWITCH SYSTEMS

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 61/493,918, filed June 6, 2011, which is incorporated by reference herein in its entirety.

STATEMENT REGARDING SEQUENCE LISTING

The Sequence Listing associated with this application is provided in text format in lieu of a paper copy, and is hereby incorporated by reference into the specification. The name of the text file containing the Sequence Listing is BLBD_002_01WO_ST25.txt. The text file is 17 KB, was created on June 5, 2011, and is being submitted electronically via EFS-Web, concurrent with the filing of the specification.

BACKGROUND

Technical Field

The present invention generally relates to improved molecular switch systems and improved retroviral packaging cells and viral particle production. In particular, the present invention comprises retroviral vectors that have increased control over gene expression.

Description of the Related Art

The Food and Drug Administration (FDA) has not yet approved any human gene therapy product for sale. Current gene therapy is experimental and has not proven very successful in clinical trials. Little progress has been made since the first gene therapy clinical trial began in 1990. In 1999, gene therapy suffered a major setback with the death of 18-year-old Jesse Gelsinger. Jesse was participating in a gene therapy trial for ornithine transcarboxylase deficiency (OTCD). He died from multiple

organ failures 4 days after starting the treatment. His death is believed to have been triggered by a severe immune response to the adenovirus carrier.

Another major blow came in January 2003, when the FDA placed a temporary halt on all gene therapy trials using retroviral vectors in blood stem cells. FDA took this action after it learned that a second child treated in a French gene therapy trial had developed a leukemia-like condition. Both this child and another who had developed a similar condition in August 2002 had been successfully treated by gene therapy for X-linked severe combined immunodeficiency disease (X-SCID), also known as "bubble baby syndrome." FDA's Biological Response Modifiers Advisory Committee (BRMAC) met at the end of February 2003 to discuss possible measures that could allow a number of retroviral gene therapy trials for treatment of life-threatening diseases to proceed with appropriate safeguards. In April of 2003, the FDA eased the ban on gene therapy trials using retroviral vectors in blood stem cells.

Recently, however, several groups have led moderately successful gene therapy trials in combating several diseases. In, 2008, UK researchers from the UCL Institute of Ophthalmology and Moorfields Eye Hospital NIHR Biomedical Research Centre announced a successful gene therapy clinical trial for treatment of Leber's congenital amaurosis, a type of inherited blindness. The results showed that the experimental treatment is safe and can improve sight (Maguire *et al.*, *N Engl J Med.* 358(21):2240 (2008)).

In 2011, Neurologix, Inc. announced positive results in a Phase 2 trial of its investigational gene therapy for advanced Parkinson's disease (PD), NLX-P101. Study participants who received NLX-P101 experienced statistically significant and clinically meaningful improvements in off-medication motor scores compared to control subjects who received sham surgery. In the trial, this benefit was seen at one month and continued virtually unchanged throughout the six month blinded study period. The results also demonstrated a positive safety profile for NLX-P101, with no serious adverse events related to the gene therapy or surgical procedure reported. Patients enrolled in the trial had moderate to advanced PD and were not adequately responsive to current therapies.

In 2009, a French groups of scientists reported using hematopoietic stem cell mediated gene therapy to successfully treat X-linked adrenoleukodystrophy (ALD). Autologous stem cells were removed from the patients, genetically corrected *ex vivo* and then re-infused into the patients after they had received myeloablative treatment. Over a span of 24 to 30 months of follow-up, polyclonal reconstitution, with 9 to 14% of granulocytes, monocytes, and T and B lymphocytes expressing the ALD protein was detected. These results strongly suggest that hematopoietic stem cells were transduced in the patients. Beginning 14 to 16 months after infusion of the genetically corrected cells, progressive cerebral demyelination in the two patients stopped.

Although before gene therapy can become a permanent cure for any condition, the therapeutic DNA introduced into target cells must remain functional and the cells containing the therapeutic DNA must be long-lived and stable. Problems with integrating therapeutic DNA into the genome and the rapidly dividing nature of many cells prevent gene therapy from many times achieving any long-term benefits. Patients will have to undergo multiple rounds of gene therapy. Anytime a foreign object is introduced into human tissues, the immune system is designed to attack the invader. The risk of stimulating the immune system in a way that reduces gene therapy effectiveness is always a potential risk. Furthermore, the immune system's enhanced response to invaders it has seen before makes it difficult for gene therapy to be repeated in patients.

Thus, while viruses are the carrier of choice, most gene therapy studies present a variety of potential problems to the patient--toxicity, immune and inflammatory responses, and gene control and targeting issues. In addition, there is always the fear that the viral vector, once inside the patient, may recover its ability to cause disease.

In addition, while production of retroviral particles has undergone iterative rounds of optimization in recent years, the packaging cell lines that presently exist in the art have essentially the same components as they did ten years ago. As a consequence, while marginal improvements in safety have been achieved, they have been achieved at the expense of efficient viral particle production of high titer viruses,

thus making retroviral-based gene therapy cost prohibitive in many contexts, *e.g.,* direct injection of viral particles to the brain.

Accordingly, the complete promise of gene therapy has yet to be realized due to issues with inconsistent therapeutic efficacy and inefficient production of viral particles, and thus, there exists a significant need in the art for improved viral vectors that have higher therapeutic efficacy in gene therapy methods and that can be used to efficiently produce viral particles. The present invention addresses these needs.

BRIEF SUMMARY

Thus, the appearances of the phrases "in one embodiment" or "in an embodiment" in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments.

In various embodiments, the present invention provides a mammalian packaging cell comprising a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to the reverse complement of one or more viral structural genes selected from the group consisting of gag, pol, and env; operably linked to an RNA dependent RNA polymerase (RDRP) promoter. In particular embodiments, the single stranded RNA virus is a minus-strand single stranded RNA virus. In certain embodiments, the first polynucleotide is stably integrated into the cell genome.

In particular other embodiments, the single stranded RNA virus is a plus-strand single stranded RNA virus.

In various embodiments, the present invention provides a mammalian packaging cell comprising a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to an RNA dependent RNA polymerase promoter of a single stranded RNA virus operably linked to one or more viral structural genes selected from the group consisting of gag, pol, and env; and an RNA polymerase recognition site (CSE). In various embodiments, the single stranded RNA virus is a plus-strand single stranded RNA virus. In certain embodiments, the

plus-strand single stranded RNA virus is a narnavirus or an alphavirus. In certain embodiments, the first polynucleotide is stably integrated into the cell genome.

In additional embodiments, the alphavirus is selected from the group consisting of: Aura virus, Bebaru virus, Cabassou virus, Chikungunya virus, Eastern equine encephalomyelitis virus, Fort Morgan virus, Getah virus, Kyzylagach virus, Mayaro virus, Middleburg virus, Mucambo virus, Ndumu virus, Pixuna virus, Ross River virus, Semliki Forest virus, Sindbis virus, Tonate virus, Triniti virus, Una virus, Venezuelan equine encephalomyelitis virus, Western equine encephalomyelitis virus, Whataroa virus, and Y-62-33 virus.

In one embodiment, the first polynucleotide comprises any one of a gag gene, a pol gene, or an env gene.

In another embodiment, the first polynucleotide comprises any two of a gag gene, a pol gene, or an env gene.

In a particular embodiment, the first polynucleotide comprises a gag gene, a pol gene, and an env gene.

In certain embodiments, the cell comprises a second polynucleotide comprising a second promoter operably linked to a viral RNA dependent RNA polymerase promoter operably linked to a viral structural gene not encoded by the first polynucleotide and a CSE.

In particular embodiments, the mammalian packaging cell comprises a second polynucleotide comprising a second promoter operably linked to the reverse complement of a viral structural gene not encoded by the first polynucleotide; operably linked to a viral RNA dependent RNA polymerase promoter.

In other particular embodiments, the mammalian packaging cell comprises a third polynucleotide comprising a third promoter operably linked to a viral RNA dependent RNA polymerase promoter operably linked to a viral structural gene not encoded by the first or second polynucleotides and a CSE.

In certain particular embodiments, the mammalian packaging cell comprises a third polynucleotide comprising a third promoter operably linked to the reverse complement of a viral structural gene not encoded by the first or second polynucleotides; operably linked to a viral RNA dependent RNA polymerase promoter.

In further embodiments, the first, second, third, and fourth promoters are independently selected from the group consisting of: an constitutive promoter and an inducible promoter.

In one embodiment, the first promoter is an inducible promoter and one or more of the second, third, and fourth promoters are the same inducible promoter or a different inducible promoter.

In a certain embodiment, the first promoter is an constitutive promoter and one or more of the second, third, and fourth promoters are the same constitutive promoter or a different constitutive promoter.

In further embodiments, the cell further comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more isolated non-structural protein genes of the single-stranded RNA virus.

In particular embodiments, the constitutive promoter is selected from the group consisting of: a cytomegalovirus immediate early gene promoter (CMV), an elongation factor 1 alpha promoter (EF1-α), a phosphoglycerate kinase-1 promoter (PGK), a ubiquitin-C promoter (UBQ-C), a cytomegalovirus enhancer/chicken beta-actin promoter (CAG), polyoma enhancer/herpes simplex thymidine kinase promoter (MC1), a beta actin promoter (β-ACT), and a simian virus 40 promoter (SV40).

In certain embodiments, the inducible promoter is selected from the group consisting of: a tetracycline responsive promoter, an ecdysone responsive promoter, a cumate responsive promoter, a glucocorticoid responsive promoter, and estrogen responsive promoter, and an RU-486 responsive promoter.

In certain particular embodiments, the cell is contacted with a composition comprising one or more isolated non-structural protein polypeptides of the single-stranded RNA virus.

In certain embodiments, the one or more non-structural protein polypeptides comprise a cell permeable peptide domain.

In additional embodiments, the cell permeable peptide domain is selected from the group consisting of: RKKRRQRRR, KKRRQRRR, RKKRRQRR, RRRRRRRRR, KKKKKKKKK, RQIKIWFQNRRMKWKK,

RQIKIWFQNRRMKSKK, RQIKIWFQNKRAKIKK, RQIKIWFQNRRMKWKK, and RVIRVWFQNKRCKDKK.

In various embodiments, the one or more viral structural genes are independently isolated from a virus selected from the group consisting of: human immunodeficiency virus type 1 (HIV-1), human immunodeficiency virus type 2 (HIV-2), visna virus, caprine arthritis-encephalitis virus (CAEV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), bovine immune deficiency virus (BIV), and simian immunodeficiency virus (SIV).

In further additional embodiments, the gag and pol viral structural genes are independently selected from the group consisting of: human immunodeficiency virus type 1 (HIV-1), human immunodeficiency virus type 2 (HIV-2), visna virus, caprine arthritis-encephalitis virus (CAEV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), bovine immune deficiency virus (BIV), and simian immunodeficiency virus (SIV).

In particular embodiments, the wherein the pol viral structural gene is isolated from the caprine arthritis-encephalitis virus (CAEV).

In further particular embodiments, the wherein the env viral structural gene encodes an envelope polypeptide selected from group consisting of FIV, HIV, VSV, MoMLV, GALV, JSRV, LCMV, Marburg virus, Mokola virus, Ebola virus, Rabies virus, Ross River Virus, Sendai virus, foul plague virus, influenza virus, and Lagos-bat virus.

In additional embodiments, the cell is selected from the group consisting of: CHO, BHK, MDCK, 10T1/2, WEHI cells, COS, BSC 1, BSC 40, BMT 10, VERO, W138, MRC5, A549, HT1080, 293, 293T, B-50, 3T3, NIH3T3, HepG2, Saos-2, Huh7, HeLa cells W163, HeLa, Vero, 211, and 211A.

In various embodiments, the present invention provides a mammalian packaging cell comprising:

a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to an alphavirus 26S subgenomic promoter operably linked to one or more viral structural genes selected from the group consisting of gag, pol, and env; and a CSE.

In various embodiments, the present invention provides a mammalian packaging cell comprising a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to the reverse complement one or more viral structural genes selected from the group consisting of gag, pol, and env; operably linked to an alphavirus 26S subgenomic promoter.

In particular embodiments, the cell further comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more alphavirus non-structural protein genes.

In more particular embodiments, the cell is contacted with a composition comprising one or more alphavirus non-structural protein polypeptides.

In certain embodiments, the one or more non-structural protein polypeptides comprise a cell permeable peptide domain.

In certain particular embodiments, the cell permeable peptide domain is selected from the group consisting of: RKKRRQRRR, KKRRQRRR, RKKRRQRR, RRRRRRRRR, KKKKKKKKK, RQIKIWFQNRRMKWKK, RQIKIWFQNRRMKSKK, RQIKIWFQNKRAKIKK, RQIKIWFQNRRMKWKK, and RVIRVWFQNKRCKDKK.

In additional embodiments, the one or more non-structural proteins is selected from the group consisting of: NSP1, NSP2, NSP3, and NSP4.

In particular additional embodiments, the one or more non-structural proteins are NSP2 and NSP4.

In certain additional embodiments, the non-structural protein is NSP4.

In various embodiments, the present invention provides a producer cell comprising a mammalian packaging cell according to any of the preceding packaging cells, wherein the cell is contacted with a composition comprising one or more non-structural proteins of the single-stranded RNA virus or, wherein the cell comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more non-structural genes isolated from the single-stranded RNA virus; and a transfer vector comprising a replication defective lentiviral genome, a packaging signal, and a mammalian promoter operably linked to a mammalian polypeptide; wherein the

mammalian cell is cultured for a time sufficient to package the lentiviral genome and produce lentiviral particles.

In various embodiments, the present invention provides a producer cell comprising the mammalian packaging cell according to any of the preceding packaging cells, wherein the cell is contacted with a composition comprising one or more non-structural proteins of the alphavirus or, wherein the cell comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more non-structural genes isolated from the alphavirus; and a transfer vector comprising a replication defective lentiviral genome, a packaging signal, and a mammalian promoter operably linked to a mammalian polypeptide; wherein the mammalian cell is cultured for a time sufficient to package the lentiviral genome and produce lentiviral particles.

In various embodiments, the present invention provides a method of producing a lentiviral particle comprising culturing a mammalian cell that expresses lentiviral gag and pol genes, a heterologous envelope gene, wherein expression of one or more of the gag, pol, and env genes is dependent upon expression of one or more non-structural proteins of a single stranded RNA virus in the cell; introducing a polynucleotide into the mammalian cell, the polynucleotide comprising a replication defective lentiviral genome; and culturing the mammalian cell for a time sufficient to package the lentiviral genome and produce lentiviral particles.

In various embodiments, the present invention provides a lentiviral vector comprising: a 5' LTR; a packaging signal, a promoter operative in mammalian cells; an RNA dependent RNA polymerase promoter of a single stranded RNA virus operably linked to a polynucleotide encoding a mammalian polypeptide; and a 3' LTR.

In various embodiments, the present invention provides a lentiviral vector comprising a 5' LTR; packaging signal; a mammalian promoter; the reverse complement sequence of a polynucleotide encoding a mammalian polypeptide operably linked to an RNA dependent RNA polymerase promoter of a single stranded RNA virus; and a 3' LTR. In particular embodiments, the single stranded RNA virus is a plus-strandor minus-strand single stranded RNA virus.

In various embodiments, the present invention provides a lentiviral vector comprising a 5' LTR; packaging signal; a mammalian promoter; an RNA

dependent RNA polymerase promoter of a single stranded RNA virus operably linked to a polynucleotide encoding a mammalian polypeptide; an RNA polymerase recognition site (CSE); and

In particular embodiments, the single stranded RNA virus is a plus-strand single stranded RNA virus.

In one embodiment, the RNA dependent RNA polymerase promoter comprises the alphavirus 26S subgenomic promoter.

In certain embodiments, the U3 region of the 5' LTR comprises a heterologous transcriptional regulatory element.

In additional embodiments, the heterologous transcriptional regulatory element is selected from the group consisting of: a cytomegalovirus enhancer, a cytomegalovirus promoter, a cytomegalovirus enhancer and promoter, a Rous sarcoma virus enhancer, a Rous sarcoma virus promoter, and a Rous sarcoma virus enhancer and promoter.

In various embodiments, the promoter operative in mammalian cells is selected from the group consisting of: a ubiquitous promoter, a tissue specific promoter, and an inducible promoter.

In further embodiments, the ubiquitous promoter is selected from the group consisting of: a cytomegalovirus immediate early gene promoter (CMV), an elongation factor 1 alpha promoter (EF1-α), a phosphoglycerate kinase-1 promoter (PGK), a ubiquitin-C promoter (UBQ-C), a cytomegalovirus enhancer/chicken beta-actin promoter (CAG), polyoma enhancer/herpes simplex thymidine kinase promoter (MC1), a beta actin promoter (β-ACT), and a simian virus 40 promoter (SV40).

In further particular embodiments, the tissue specific promoter expresses the polynucleotide encoding the mammalian polypeptide in ectodermal cells, endodermal cells, mesodermal cells, or mesenchymal cells.

In further additional embodiments, the inducible promoter is selected from the group consisting of: a tetracycline responsive promoter, an ecdysone responsive promoter, a cumate responsive promoter, a glucocorticoid responsive promoter, and estrogen responsive promoter, and an RU-486 promoter.

In further certain embodiments, the lentivirus is selected from the group consisting of: human immunodeficiency virus type 1 (HIV-1), human immunodeficiency virus type 2 (HIV-2), visna virus, caprine arthritis-encephalitis virus (CAEV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), bovine immune deficiency virus (BIV), and simian immunodeficiency virus (SIV).

In particular embodiments, the alphavirus is selected from the group consisting of: Aura virus, Bebaru virus, Cabassou virus, Chikungunya virus, Eastern equine encephalomyelitis virus, Fort Morgan virus, Getah virus, Kyzylagach virus, Mayaro virus, Middleburg virus, Mucambo virus, Ndumu virus, Pixuna virus, Ross River virus, Semliki Forest virus, Sindbis virus, Tonate virus, Triniti virus, Una virus, Venezuelan equine encephalomyelitis virus, Western equine encephalomyelitis virus, Whataroa virus, and Y-62-33 virus.

In more particular embodiments, the mammalian polypeptide is the mammalian polypeptide is adrenoleukodystrophy protein (ABCD1). In other particular embodiments, the mammalian polypeptide is β-globin.

In certain embodiments, one or more nucleotides of the U3 region of the 3' LTR are deleted.

In various embodiments, a vector comprises one or more vector sequences selected from the group consisting of: a central polypurine tract and the central termination sequence (cPPT/CTS), a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), hepatitis B virus posttranscriptional regulatory element (HPRE), and mammalian polyadenylation sequence 3' of the 3' R region.

In particular embodiments, a vector comprises one or more recombinase recognition sites selected from the group consisting of: LoxP, Lox511, Lox5171, Lox2272, m2, Lox71, Lox66, FRT, F1, F2, F3, F4, F5, FRT(LE), FRT(RE), attB, attP, attL, and attR.

DETAILED DESCRIPTION

A. Overview

The present invention contemplates, in part, inventive retroviral vector systems that are useful in developing improved gene switches and regulatable viral particle packaging cell lines that produce high viral titers. Retroviral transfer vectors and transfer vector systems that provide increased control and/or amplification of gene expression in transduced mammalian cells are also provided. Further, the present invention contemplates, in part, single transfer vector constructs, binary retroviral transfer vector systems, and polypeptide compositions used in combination with transfer vectors. Improved methods of gene therapy are also provided.

The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, *e.g.,* Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (1982); *DNA Cloning: A Practical Approach*, vol. I & II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Nucleic Acid Hybridization* (B. Hames & S. Higgins, eds., 1985); *Transcription and Translation* (B. Hames & S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); *A Practical Guide to Molecular Cloning* (B. Perbal, ed., 1984).

All publications, patents and patent applications cited herein are hereby incorporated by reference in their entirety.

B. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. For the purposes of the present invention, the following terms are defined below.

The articles "a," "an," and "the" are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

As used herein, the term "about" or "approximately" refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that varies by as much as 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% or 1% to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length. In one embodiment, the term "about" or "approximately" refers a range of quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length ± 15%, ± 10%, ± 9%, ± 8%, ± 7%, ± 6%, ± 5%, ± 4%, ± 3%, ± 2%, or ± 1% about a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

Throughout this specification, unless the context requires otherwise, the words "comprise", "comprises" and "comprising" will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of." Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory, and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that no other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements

Reference throughout this specification to "one embodiment" or "an embodiment" means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, the appearances of the phrases "in one embodiment" or "in an embodiment" in various places throughout this specification are not necessarily all

referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments.

C. Polynucleotides

In various illustrative embodiments, the present invention contemplates, in part, polynucleotides sequences and compositions comprising the same. As used herein, the terms "polynucleotide" or "nucleic acid" refers to messenger RNA (mRNA), RNA, genomic RNA (gRNA), plus strand RNA (RNA(+)), minus strand RNA (RNA(-)), genomic DNA (gDNA), complementary DNA (cDNA) or DNA. Polynucleotides refer to polymeric form of nucleotides of at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 100, at least 200, at least 300, at least 400, at least 500, at least 1000, at least 5000, at least 10000, or at least 15000 or more nucleotides in length, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide. Polynucleotides include single and double stranded polynucleotides. Preferably, polynucleotides of the invention include polynucleotides or variants having at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% nucleotide identity thereto.

As used herein, the terms "polynucleotide variant" and "variant" and the like refer to polynucleotides displaying substantial sequence identity with a reference polynucleotide sequence or polynucleotides that hybridize with a reference sequence under stringent conditions that are defined hereinafter. These terms also encompass polynucleotides that are distinguished from a reference polynucleotide by the addition, deletion or substitution of at least one nucleotide. Accordingly, the terms "polynucleotide variant" and "variant" include polynucleotides in which one or more nucleotides have been added or deleted, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference polynucleotide whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide.

The recitations "sequence identity" or, for example, comprising a "sequence 50% identical to," as used herein, refer to the extent that sequences are

identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a "percentage of sequence identity" may be calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.,* A, T, C, G, I) or the identical amino acid residue (*e.g.,* Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. Included are nucleotides and polypeptides having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to any of the reference sequences described herein, typically where the polypeptide variant maintains at least one biological activity of the reference polypeptide.

Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include "reference sequence," "comparison window," "sequence identity," "percentage of sequence identity," and "substantial identity". A "reference sequence" is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (*i.e.*, only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal

alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (*i.e.*, resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, *Nucl. Acids Res.* 25:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons Inc, 1994-1998, Chapter 15.

As used herein, the term "isolated" means material that is substantially or essentially free from components that normally accompany it in its native state. For example, an "isolated polynucleotide," as used herein, refers to a polynucleotide that has been purified from the sequences which flank it in a naturally-occurring state, *e.g.,* a DNA fragment that has been removed from the sequences that are normally adjacent to the fragment.

Terms that describe the orientation of polynucleotides include: 5' (normally the end of the polynucleotide having a free phosphate group) and 3' (normally the end of the polynucleotide having a free hydroxyl (OH) group). Polynucleotide sequences can be annotated in the 5' to 3' orientation or the 3' to 5' orientation. For DNA and mRNA, the 5' to 3' strand is designated the "sense," "plus," or "coding" strand because its sequence is identical to the sequence of the premessenger (premRNA) [except for uracil (U) in RNA, instead of thymine (T) in DNA]. For DNA and mRNA, the complementary 3' to 5' strand which is the strand transcribed by the RNA polymerase is designated as "template," "antisense," "minus," or "non-coding" strand. As used herein, the term "reverse orientation" refers to a 5' to 3' sequence written in the 3' to 5' orientation or a 3' to 5' sequence written in the 5' to 3' orientation. In various embodiments, the reverse orientation of a minus-strand RDRP promoter is the 3' to 5' sequence of the promoter on the minus-strand written in the 5' to 3' orientation on the plus-strand.

Thus, in particular embodiments, a transcribed mRNA may comprise a non-functional or inoperable RDRP promoter or subgenomic promoter in the 5' to 3' orientation because it is in the reverse orientation; but the same mRNA can also be used as a minus-strand template because the RDRP or subgenomic promoter is operable or functional in the 3' to 5' orientation.

In virology, the genome of an RNA virus can be said to be either positive-sense, also known as a "plus-strand," or negative-sense, also known as a "minus-strand." In this context, the terms sense and strand can be used interchangeably, making such terms as positive-strand equivalent to positive-sense, and plus-strand equivalent to plus-sense. Positive-sense (5' to 3') viral RNA signifies that a particular viral RNA sequence may be directly translated into the desired viral proteins. Therefore, in positive-sense RNA viruses, the viral RNA genome can be considered viral mRNA, and can be immediately translated by the host cell. Unlike negative-sense RNA, positive-sense RNA is of the same sense as mRNA. Negative-sense (3' to 5') viral RNA is complementary to the viral mRNA and thus must be converted to positive-sense RNA by an RNA dependent RNA polymerase prior to translation. Negative-sense RNA (like DNA) has a polynucleotide sequence complementary to the mRNA that it encodes. Like DNA, this RNA is not translated into protein directly. Instead, it must first be transcribed into a positive-sense RNA that acts as an mRNA.

The terms "complementary" and "complementarity" refer to polynucleotides *(i.e.,* a sequence of nucleotides) related by the base-pairing rules. A strand is complementary to another in that the base pairs between them are non-covalently connected via two or three hydrogen bonds. For DNA, adenine (A) bases complement thymine (T) bases and vice versa; guanine (G) bases complement cytosine (C) bases and vice versa. With RNA, it is the same except that adenine (A) bases complement uracil (U) bases instead of thymine (T) bases. Complementarity can be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there can be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. For example, the complementary strand of the DNA sequence 5' A G T C A T

G 3' is 3' T C A G T A C 5'. The latter sequence is often written as the reverse complement with the 5' end on the left and the 3' end on the right, 5' C A T G A C T 3'. A sequence that is equal to its reverse complement is said to be a palindromic sequence.

In particular embodiments, polynucleotides are provided by this invention that encode at least about 5, 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 500, 1000 or more contiguous amino acid residues of a polypeptide of the invention, as well as all intermediate lengths. It will be readily understood that "intermediate lengths," in this context, means any length between the quoted values, such as 6, 7, 8, 9, *etc.*, 101, 102, 103, *etc.*; 151, 152, 153, *etc.*; 201, 202, 203, *etc.*

Polynucleotides include a polynucleotide-of-interest. As used herein, the term "polynucleotide-of-interest" refers to the polynucleotide, *e.g.,* a polynucleotide encoding a polypeptide (*i.e.,* a polypeptide-of-interest), inserted into an expression vector that is desired to be expressed. In certain embodiments, the polynucleotide-of-interest encodes a polypeptide that provides a therapeutic effect in the treatment or prevention of a disease or disorder, which may be referred to as a "therapeutic polypeptide." Polynucleotides-of-interest, and polypeptides encoded therefrom, include both polynucleotides that encode wild-type polypeptides, as well as functional variants and fragments thereof. In particular embodiments, a functional variant has at least 80%, at least 90%, at least 95%, or at least 99% identity to a corresponding wild-type reference polynucleotide or polypeptide sequence. In certain embodiments, a functional variant or fragment has at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of a biological activity of a corresponding wild-type polypeptide. In one embodiment, the polynucleotide-of-interest encodes a polypeptide which provides a therapeutic function for the treatment of a hemoglobinopathy.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide, or fragment of variant thereof, as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention, for example polynucleotides that are optimized for human and/or primate codon selection. Further, alleles of the genes

comprising the polynucleotide sequences provided herein may also be used. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides.

In various other embodiment, the polynucleotide-of-interest comprises an inhibitory nucleic acid sequence including, but not limited to: an siRNA, an miRNA, an shRNA, and a ribozyme.

As used herein, the terms "siRNA" or "short interfering RNA" refer to a short polynucleotide sequence that mediates a process of sequence-specific post-transcriptional gene silencing, translational inhibition, transcriptional inhibition, or epigenetic RNAi in animals (Zamore *et al.*, 2000, *Cell*, 101, 25-33; Fire *et al.*, 1998, *Nature*, 391, 806; Hamilton *et al.*, 1999, *Science*, 286, 950-951; Lin *et al.*, 1999, *Nature*, 402, 128-129; Sharp, 1999, *Genes & Dev.*, 13, 139-141; and Strauss, 1999, *Science*, 286, 886). In certain embodiments, a repressor, or RNAi oligonucleotide, is single stranded. In other embodiments, the repressor, or RNAi oligonucleotide, is double stranded. Certain embodiments may also employ short-interfering RNAs (siRNA). In certain embodiments, the first strand of the double-stranded oligonucleotide contains two more nucleoside residues than the second strand. In other embodiments, the first strand and the second strand have the same number of nucleosides; however, the first and second strands are offset such that the two terminal nucleosides on the first and second strands are not paired with a residue on the complimentary strand. In certain instances, the two nucleosides that are not paired are thymidine resides. The siRNA should include a region of sufficient homology to the target gene, and be of sufficient length in terms of nucleotides, such that the siRNA, or a fragment thereof, can mediate down regulation of the target gene. Thus, an siRNA includes a region which is at least partially complementary to the target RNA. It is not necessary that there be perfect complementarity between the siRNA and the target, but the correspondence must be sufficient to enable the siRNA, or a cleavage product thereof, to direct sequence specific silencing, such as by RNAi cleavage of the target RNA. Complementarity, or degree of homology with the target strand, is most critical in the antisense strand. While perfect complementarity, particularly in the antisense strand, is often desired, some embodiments include one or more, but preferably 10, 8, 6,

5, 4, 3, 2, or fewer mismatches with respect to the target RNA. The mismatches are most tolerated in the terminal regions, and if present are preferably in a terminal region or regions, *e.g.*, within 6, 5, 4, or 3 nucleotides of the 5' and/or 3' terminus. The sense strand need only be sufficiently complementary with the antisense strand to maintain the overall double-strand character of the molecule.

In addition, an siRNA may be modified or include nucleoside analogs. Single stranded regions of an siRNA may be modified or include nucleoside analogs, *e.g.*, the unpaired region or regions of a hairpin structure, *e.g.*, a region which links two complementary regions, can have modifications or nucleoside analogs. Modification to stabilize one or more 3'- or 5'-terminus of an siRNA, *e.g.*, against exonucleases, or to favor the antisense siRNA agent to enter into RISC are also useful. Modifications can include C3 (or C6, C7, C12) amino linkers, thiol linkers, carboxyl linkers, non-nucleotidic spacers (C3, C6, C9, C12, abasic, triethylene glycol, hexaethylene glycol), special biotin or fluorescein reagents that come as phosphoramidites and that have another DMT-protected hydroxyl group, allowing multiple couplings during RNA synthesis. Each strand of an siRNA can be equal to or less than 30, 25, 24, 23, 22, 21, or 20 nucleotides in length. The strand is preferably at least 19 nucleotides in length. For example, each strand can be between 21 and 25 nucleotides in length. Preferred siRNAs have a duplex region of 17, 18, 19, 29, 21, 22, 23, 24, or 25 nucleotide pairs, and one or more overhangs of 2-3 nucleotides, preferably one or two 3' overhangs, of 2-3 nucleotides.

As used herein, the terms "miRNA" or "microRNA" s refer to small non-coding RNAs of 20–22 nucleotides, typically excised from ~70 nucleotide foldback RNA precursor structures known as pre-miRNAs. miRNAs negatively regulate their targets in one of two ways depending on the degree of complementarity between the miRNA and the target. First, miRNAs that bind with perfect or nearly perfect complementarity to protein-coding mRNA sequences induce the RNA-mediated interference (RNAi) pathway. miRNAs that exert their regulatory effects by binding to imperfect complementary sites within the 3' untranslated regions (UTRs) of their mRNA targets, repress target-gene expression post-transcriptionally, apparently at the level of translation, through a RISC complex that is similar to, or possibly identical

with, the one that is used for the RNAi pathway. Consistent with translational control, miRNAs that use this mechanism reduce the protein levels of their target genes, but the mRNA levels of these genes are only minimally affected. miRNAs encompass both naturally occurring miRNAs as well as artificially designed miRNAs that can specifically target any mRNA sequence. For example, in one embodiment, the skilled artisan can design short hairpin RNA constructs expressed as human miRNA (*e.g.*, miR-30 or miR-21) primary transcripts. This design adds a Drosha processing site to the hairpin construct and has been shown to greatly increase knockdown efficiency (Pusch *et al.*, 2004). The hairpin stem consists of 22-nt of dsRNA (*e.g.*, antisense has perfect complementarity to desired target) and a 15-19-nt loop from a human miR. Adding the miR loop and miR30 flanking sequences on either or both sides of the hairpin results in greater than 10-fold increase in Drosha and Dicer processing of the expressed hairpins when compared with conventional shRNA designs without microRNA. Increased Drosha and Dicer processing translates into greater siRNA/miRNA production and greater potency for expressed hairpins.

As used herein, the terms "shRNA" or "short hairpin RNA" refer to double-stranded structure that is formed by a single self-complementary RNA strand. \shRNA constructs containing a nucleotide sequence identical to a portion, of either coding or non-coding sequence, of the target gene are preferred for inhibition. RNA sequences with insertions, deletions, and single point mutations relative to the target sequence have also been found to be effective for inhibition. Greater than 90% sequence identity, or even 100% sequence identity, between the inhibitory RNA and the portion of the target gene is preferred. In certain preferred embodiments, the length of the duplex-forming portion of an shRNA is at least 20, 21 or 22 nucleotides in length, *e.g.*, corresponding in size to RNA products produced by Dicer-dependent cleavage. In certain embodiments, the shRNA construct is at least 25, 50, 100, 200, 300 or 400 bases in length. In certain embodiments, the shRNA construct is 400-800 bases in length. shRNA constructs are highly tolerant of variation in loop sequence and loop size.

As used herein, the term "ribozyme" refers to a catalytically active RNA molecule capable of site-specific cleavage of target mRNA. Several subtypes have been described, *e.g.*, hammerhead and hairpin ribozymes. Ribozyme catalytic activity

and stability can be improved by substituting deoxyribonucleotides for ribonucleotides at noncatalytic bases. While ribozymes that cleave mRNA at site-specific recognition sequences can be used to destroy particular mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA has the following sequence of two bases: 5′-UG-3′. The construction and production of hammerhead ribozymes is well known in the art.

A preferred method of delivery of a polynucleotide-of-interest that comprises an siRNA, an miRNA, an shRNA, or a ribozyme comprises one or more regulatory sequences, such as, for example, a strong constitutive pol III, *e.g.*, human U6 snRNA promoter, the mouse U6 snRNA promoter, the human and mouse H1 RNA promoter and the human tRNA-val promoter, or a strong constitutive pol II promoter, as described elsewhere herein.

The polynucleotides of the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, untranslated regions (UTRs), Kozak sequences, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, internal ribosomal entry sites (IRES), recombinase recognition sites (*e.g.,* LoxP, FRT, and Att sites), termination codons, transcriptional termination signals, and polynucleotides encoding self-cleaving polypeptides, epitope tags, as disclosed elsewhere herein, such that their overall length may vary considerably. It is therefore contemplated that a polynucleotide fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

Polynucleotides can be prepared, manipulated and/or expressed using any of a variety of well established techniques known and available in the art. In order to express a desired polypeptide, a nucleotide sequence encoding the polypeptide, can be inserted into appropriate vector. The term vector includes viral vectors as disclosed elsewhere herein and "expression vectors" (*e.g.,* a plasmid, cosmid or phage chromosome) containing a gene construct in a form suitable for expression by a cell (*e.g.,* operably linked to a promoter). In particular embodiments, viral vectors are expression vectors.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector—origin of replication, selection cassettes, promoters, enhancers, translation initiation signals (Shine Dalgarno sequence or Kozak sequence) introns, a polyadenylation sequence, 5' and 3' untranslated regions—which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including ubiquitous promoters and inducible promoters may be used.

In particular embodiments, a vector for use in practicing the invention including, but not limited to expression vectors and viral vectors, will include exogenous, endogenous, or heterologous control sequences such as promoters and/or enhancers. The term "promoter" as used herein refers to a recognition site of a polynucleotide (DNA or RNA) to which an RNA polymerase binds. An RNA polymerase initiates and transcribes polynucleotides operably linked to the promoter. In particular embodiments, promoters operative in mammalian cells comprise an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated and/or another sequence found 70 to 80 bases upstream from the start of transcription, a CNCAAT region where N may be any nucleotide. The term "enhancer" refers to a segment of DNA which contains sequences capable of providing enhanced transcription and in some instances can function independent of their orientation relative to another control sequence. An enhancer can function cooperatively or additively with promoters and/or other enhancer elements. The term "promoter/enhancer" refers to a segment of DNA which contains sequences capable of providing both promoter and enhancer functions.

An "endogenous" control sequence is one which is naturally linked with a given gene in the genome. An "exogenous" control sequence is one which is placed in juxtaposition to a gene by means of genetic manipulation (*i.e.*, molecular biological techniques) such that transcription of that gene is directed by the linked enhancer/promoter. A "heterologous" control sequence is an exogenous sequence that is from a different species than the cell being genetically manipulated. The term

"operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. In one embodiment, the term refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, and/or enhancer) and a second polynucleotide sequence, *e.g.,* a polynucleotide-of-interest, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

In particular embodiments, an operable linkage may be operable on the sense or plus strand or on the non-sense or minus strand. For example, in one embodiment, a reverse complement of a polynucleotide encoding a polypeptide of interest is operably linked to an RDRP promoter or subgenomic promoter in the reverse orientation. This means that the linkage is operable in the 3′ to 5′ orientation but not in the 5′ to 3′ orientation. In certain embodiments, an operable linkage may be operable on both strands (independent of directionality).

In one embodiment, a polynucleotide of the invention comprises a constitutive promoter. As used herein, the term "constitutive promoter" refers to a promoter that continually or continuously allows for transcription of an operably linked sequence. Constitutive promoters may be a "ubiquitous promoter" that allows expression in a wide variety of cell and tissue types or a "tissue-specific promoter" that allows expression in a restricted variety of cell and tissue types.

Ubiquitous promoters can be selected based on strength of ubiquitous promoter activity, specificity for target tissue or other factors relating to desired control over expression, as is understood in the art. The sequences of these and numerous additional ubiquitous promoters are known in the art. The relevant sequences may be readily obtained from public databases and incorporated into vectors for use in practicing the present invention. Illustrative ubiquitous promoters include, but are not limited to, a cytomegalovirus (CMV) immediate early promoter, a viral simian virus 40 (SV40) (*e.g.*, early or late), a Moloney murine leukemia virus (MoMLV) LTR promoter, a Rous sarcoma virus (RSV) LTR, a herpes simplex virus (HSV) (thymidine kinase) promoter, H5, P7.5, and P11 promoters from vaccinia virus, an elongation factor 1-alpha (EF1a) promoter, early growth response 1 (EGR1), ferritin H (FerH), ferritin L (FerL), Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), eukaryotic

translation initiation factor 4A1 (EIF4A1), heat shock 70kDa protein 5 (HSPA5), heat shock protein 90kDa beta, member 1 (HSP90B1), heat shock protein 70kDa (HSP70), β-kinesin (β-KIN), the human ROSA 26 locus (Irions *et al.*, *Nature Biotechnology* 25, 1477 - 1482 (2007)), a Ubiquitin C promoter (UBC), a phosphoglycerate kinase-1 (PGK) promoter, a cytomegalovirus enhancer/chicken β-actin (CAG) promoter, and a β-actin promoter.

In a particular embodiment, it may be desirable to use a tissue-specific promoter to achieve cell type specific, lineage specific, or tissue-specific expression of a desired polynucleotide sequence (*e.g.*, to express a particular nucleic acid encoding a polypeptide in only a subset of cell types or tissues or during specific stages of development). In one embodiment, a vector of the present invention comprises a tissue specific promoter and/or enhancer that expresses a desired polypeptide. Illustrative examples of tissue specific promoters include, but are not limited to: an B29 promoter (B cell expression), a runt transcription factor (CBFa2) promoter (stem cell specific expression), an CD14 promoter (monocytic cell expression), an CD43 promoter (leukocyte and platelet expression), an CD45 promoter (hematopoietic cell expression), an CD68 promoter (macrophage expression), a CYP450 3A4 promoter (hepatocyte expression), an desmin promoter (muscle expression), an elastase 1 promoter (pancreatic acinar cell expression, an endoglin promoter (endothelial cell expression), a fibroblast specific protein 1 promoter (FSP1) promoter (fibroblast cell expression), a fibronectin promoter (fibroblast cell expression), a fms-related tyrosine kinase 1 (FLT1) promoter (endothelial cell expression), a glial fibrillary acidic protein (GFAP) promoter (astrocyte expression), an insulin promoter (pancreatic beta cell expression), an integrin, alpha 2b (ITGA2B) promoter (megakaryocytes), an intracellular adhesion molecule 2 (ICAM-2) promoter (endothelial cells), an interferon beta (IFN-β) promoter (hematopoietic cells), a keratin 5 promoter (keratinocyte expression), a myoglobin (MB) promoter (muscle expression), a myogenic differentiation 1 (MYOD1) promoter (muscle expression), a nephrin promoter (podocyte expression), a bone gamma-carboxyglutamate protein 2 (OG-2) promoter (osteoblast expression), an 3-oxoacid CoA transferase 2B (Oxct2B) promoter, (haploid-spermatid expression), a surfactant

protein B (SP-B) promoter (lung expression), a synapsin promoter (neuron expression), a Wiskott-Aldrich syndrome protein (WASP) promoter (hematopoietic cell expression).

According to certain embodiments of the invention a cell type specific promoter is specific for cell types found in the brain (*e.g.,* neurons, glial cells), liver (*e.g.,* hepatocytes), pancreas, skeletal muscle (*e.g.,* myocytes), immune system (*e.g.,* T cells, B cells, macrophages), heart (*e.g.,* cardiac myocytes), retina, skin (*e.g.,* keratinocytes), bone (*e.g.,* osteoblasts or osteoclasts), *etc.*

Numerous cell type or tissue-specific promoters are known, and one of ordinary skill in the art will readily be able to identify tissue specific promoters (or to determine whether any particular promoter is a tissue specific promoter) from the literature or by performing experiments such as Northern blots, immunoblots, *etc.* in which expression of either an endogenous gene or a reporter gene operably linked to the promoter is compared in different cell or tissue types). A lineage specific promoter directs transcription in cells of a particular lineage and not in fully differentiated cells of most or all other lineages. For example, the promoter may direct transcription in cells types of the B cell lineage, T cell lineage, macrophage lineage, *etc.* Erythrocyte-specific expression can be achieved by using the human β-globin promoter region and locus control region (LCR).

The invention therefore provides polynucleotides as described elsewhere herein, *supra* and *infra*, including but not limited to expression and viral vectors, comprising a cell type or tissue-specific promoter and methods of using the vectors and/or viral particles derived therefrom to achieve cell type or tissue specific expression of genes, *e.g.,* a polynucleotide-of-interest.

Certain embodiments of the invention provide conditional expression of a polynucleotide-of-interest, *e.g.,* expression is controlled by subjecting a cell, tissue, organism, *etc.*, to a treatment or condition that causes the polynucleotide to be expressed or that causes an increase or decrease in expression of the polynucleotide encoded by the polynucleotide-of-interest. As used herein, "conditional expression" may refer to any type of conditional expression including, but not limited to, inducible expression; repressible expression; expression in cells or tissues having a particular

physiological, biological, or disease state, *etc.* This definition is not intended to exclude cell type or tissue-specific expression.

One approach to achieving conditional expression involves the use of inducible promoters. As used herein, the term "inducible promoter" refers to a control element (*e.g.,* a promoter, enhancer, promoter/enhancer, or portion thereof) whose transcriptional activity may be regulated by exposing a cell or tissue comprising a nucleic acid sequence operably linked to the promoter to a treatment or condition that alters the transcriptional activity of the promoter, resulting in increased transcription of the nucleic acid sequence. For convenience, as used herein, the term "inducible promoter" also includes repressible promoters, *i.e.*, promoters whose transcriptional activity may be regulated by exposing a cell or tissue comprising a nucleic acid sequence operably linked to the promoter to a treatment or condition that alters the transcriptional activity of the promoter, resulting in decreased transcription of the nucleic acid sequence. Typical inducible promoters are active in mammalian cells. Inducible promoters/systems include, but are not limited to, steroid-inducible promoters such as promoters for genes encoding glucocorticoid or estrogen receptors (inducible by treatment with the corresponding hormone), metallothionine promoter (inducible by treatment with various heavy metals), MX-1 promoter (inducible by interferon), the "GeneSwitch" mifepristone-regulatable system (Sirin *et al.*, 2003, *Gene*, 323:67), the cumate inducible gene switch (WO 2002/088346), tetracycline-dependent regulatory systems, *etc.*

For example, in tetracycline-dependent regulatory systems (Gossen & Bujard. 1992. *Proc. Natl Acad. Sci. USA* 89, 5547-5551), the effector is a fusion of sequences that encode the VP16 transactivation domain and the *Escherichia coli* tetracycline repressor (TetR) protein, which specifically binds both tetracycline and the 19-bp operator sequences (tetO) of the tet operon in the target transgene, resulting in its transcription. In the original system, the tetracycline-controlled transactivator (tTA) cannot bind DNA when the inducer is present, while in a modified version, the "reverse tTA" (rtTA) binds DNA only when the inducer is present ("tet-on"; Gossen *et al.*, 1995. *Science*. 268:766). The current inducer of choice is doxycycline (Dox). The invention therefore provides polynucleotides comprising a tetracycline-controlled transactivator

or reverse tetracycline-controlled transactivator, vectors comprising operator sequences of the tet operon to which the tetracycline-controlled transactivator or reverse tetracycline-controlled transactivator specifically bind, and methods of using the polynucleotides and/or viral particles derived therefrom to achieve conditional expression.

Conditional expression can also be achieved by using a site-specific DNA recombinase. According to certain embodiments of the invention the vector comprises at least one (typically two) site(s) for recombination mediated by a site-specific recombinase. As used herein, the terms "recombinase" or "site-specific recombinase" include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites (*e.g.,* two, three, four, five, seven, ten, twelve, fifteen, twenty, thirty, fifty, *etc.*), which may be wild-type proteins (see Landy, Current Opinion in Biotechnology 3:699-707 (1993)), or mutants, derivatives (*e.g.,* fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof. Illustrative examples of recombinases suitable for use in particular embodiments of the present invention include, but are not limited to: Cre, Int, IHF, Xis, Flp, Fis, Hin, Gin, ΦC31, Cin, Tn3 resolvase, TndX, XerC, XerD, TnpX, Hjc, Gin, SpCCE1, and ParA.

The vectors may comprise one or more recombination sites for any of a wide variety of site-specific recombinases. It is to be understood that the target site for a site-specific recombinase is in addition to any site(s) required for integration of a vector, *e.g.,* a retroviral vector or lentiviral vector. As used herein, the terms "recombination sequence," "recombination site," or "site-specific recombination site" refer to a particular nucleic acid sequence to which a recombinase recognizes and binds.

For example, one recombination site for Cre recombinase is loxP which is a 34 base pair sequence comprising two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence (see FIG. 1 of Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994)). Other exemplary loxP sites include, but are not limited to: lox511 (Hoess *et al.*, 1996; Bethke and Sauer, 1997),

lox5171 (Lee and Saito, 1998), lox2272 (Lee and Saito, 1998), m2 (Langer *et al.*, 2002), lox71 (Albert *et al.*, 1995), and lox66 (Albert *et al.*, 1995).

Suitable recognition sites for the FLP recombinase include, but are not limited to: FRT (McLeod, *et al.*, 1996), F_1, F_2, F_3 (Schlake and Bode, 1994), F_4, F_5 (Schlake and Bode, 1994), FRT(LE) (Senecoff *et al.*, 1988), FRT(RE) (Senecoff *et al.*, 1988).

Other examples of recognition sequences are the attB, attP, attL, and attR sequences, which are recognized by the recombinase enzyme λ Integrase, *e.g,* phi-c31. The φC31 SSR mediates recombination only between the heterotypic sites *attB* (34 bp in length) and *attP* (39 bp in length) (Groth *et al.*, 2000). *attB* and *attP*, named for the attachment sites for the phage integrase on the bacterial and phage genomes, respectively, both contain imperfect inverted repeats that are likely bound by φC31 homodimers (Groth *et al.*, 2000). The product sites, *attL* and *attR*, are effectively inert to further φC31-mediated recombination (Belteki *et al.*, 2003), making the reaction irreversible. For catalyzing insertions, it has been found that *attB*-bearing DNA inserts into a genomic *attP* site more readily than an *attP* site into a genomic *attB* site (Thyagarajan *et al.*, 2001; Belteki *et al.*, 2003). Thus, typical strategies position by homologous recombination an *attP*-bearing "docking site" into a defined locus, which is then partnered with an *attB*-bearing incoming sequence for insertion.

Numerous recombination systems from various organisms have been described. See, *e.g.,* Hoess, *et al.*, *Nucleic Acids Research* 14(6):2287 (1986); Sauer *et al.*, *Nucleic Acids Res.*, 17:147 (1989); Abremski, *et al.*, *J. Biol. Chem.* 261(1):391 (1986); Gorman *et al.*, *Curr. Op. Biotechnol.*, 11:455 (2000); Campbell, *J. Bacteriol.* 174(23):7495 (1992); O'Gorman *et al.*, *Science*, 251:1351 (1991); Qian, *et al.*, *J. Biol. Chem.* 267(11):7794 (1992); Kolb, *Cloning Stem Cells*, 4:65 (2002); Araki, *et al.*, J. Mol. Biol. 225(1):25 (1992); Kuhn *et al.*, *Methods Mol. Biol.,* 180:175 (2002); Maeser and Kahnmann, *Mol. Gen. Genet.* 230:170-176 (1991); Esposito, *et al.*, *Nucl. Acids Res.* 25(18):3605 (1997); and Branda and Dymecki. *Dev Cell.* 6(1):7-28 (2004).

On one embodiment, conditional expression of a polynucleotide sequence, *e.g.,* a polynucleotide-of-interest operably linked to a promoter element, a promoter element or both, can be achieved by positioning the polynucleotide sequence

between recombinase recognition sites ("floxed" in the case of LoxP sites), *e.g.,* Ventura *et al.*, 2004. *Proc Natl Acad Sci U S A.* 101(28):10380-5). This arrangement allows for controlled expression of the polynucleotide sequence following transfer into a cell. By inducing expression of the recombinase in the cell, the flanked polynucleotide sequence is excised, thus preventing further transcription and/or expression of a polynucleotide-of-interest and effectively eliminating expression of the sequence. This system has a number of applications including recombinase-mediated gene activation and inactivation.

Recombinase expression may be induced in any of a variety of ways. For example, the recombinase may be present in the cells under control of an inducible promoter, and recombinase expression may be induced by activating the promoter. Alternatively or additionally, recombinase expression may be induced by introducing an expression vector that directs expression of recombinase into the cell. Any suitable expression vector can be used, including, but not limited to, viral vectors such as lentiviral or retroviral vectors.

Vectors comprising two recombinase recognition sites are useful in any applications for which standard vectors comprising two recombinase recognition sites can be used. For example, selectable markers may be placed between the recombinase recognition sites. This allows for sequential and repeated targeting of multiple genes to a single cell (or its progeny). For example, after introduction of a vector comprising a floxed polynucleotide into a cell, stable transfectants may be selected. After isolation of a stable transfectant, the polynucleotide can be excised by induction of Cre or replaced by another polynucleotide sequence harboring a different selectable marker. The orientation of the recombinase recognition sites dictates whether excision or replacement occurs.

In particular embodiments, the vectors contemplated by the invention, include one or more polynucleotides-of-interest that encode one or more polypeptides. To achieve efficient translation of each of the plurality of polypeptides, the polynucleotide sequences can be separated by one or more IRES sequences or polynucleotide sequences encoding self-cleaving polypeptides. As used herein, an "internal ribosome entry site" or "IRES" refers to an element that promotes direct

internal ribosome entry to the initiation codon, such as ATG, of a cistron (a protein encoding region), thereby leading to the cap-independent translation of the gene. *See, e.g.,* Jackson *et al.*, 1990. *Trends Biochem Sci* 15(12):477-83) and Jackson and Kaminski. 1995. *RNA* 1(10):985-1000. Examples of IRES generally employed by those of skill in the art include those described in U.S. Pat. No. 6,692,736. Examples of "IRES" known in the art include, but are not limited to IRES obtainable from picornavirus (Jackson *et al.*, 1990) and IRES obtainable from viral or cellular mRNA sources, such as for example, immunoglobulin heavy-chain binding protein (BiP), the vascular endothelial growth factor (VEGF) (Huez *et al.* 1998. *Mol. Cell. Biol.* 18(11):6178-6190), the fibroblast growth factor 2 (FGF-2), and insulin-like growth factor (IGFII), the translational initiation factor eIF4G and yeast transcription factors TFIID and HAP4, the encephelomycarditis virus (EMCV) which is commercially available from Novagen (Duke *et al.*, 1992. *J. Virol* 66(3):1602-9) and the VEGF IRES (Huez *et al.*, 1998. *Mol Cell Biol* 18(11):6178-90). IRES have also been reported in different viruses such as cardiovirus, rhinovirus, aphthovirus, HCV, Friend murine leukemia virus (FrMLV) and Moloney murine leukemia virus (MoMLV). As used herein, "IRES" encompasses functional variations of IRES sequences as long as the variation is able to promote direct internal ribosome entry to the initiation codon of a cistron. An IRES may be mammalian, viral or protozoan. "Under translational control of an IRES" as used herein means that translation is associated with the IRES and proceeds in a cap-independent manner.

Efficient expression of polynucleotides can be increased in some embodiments, by using sequences that increase translational efficiency, *e.g.,* through an increase in mRNA ribosomal binding or an increase in mRNA stability. In certain embodiments, polynucleotides encoding a polypeptide of the invention comprise a short recognition sequence, *i.e.,* a Kozak sequence, that greatly facilitates the initial binding of mRNA to the small subunit of the ribosome and increases translation. The consensus Kozak sequence is (GCC)RCCATGG, where R is a purine (A or G) (Kozak, 1986. *Cell.* 44(2):283-92, and Kozak, 1987. *Nucleic Acids Res.* 15(20):8125-48).

In particular embodiments, vectors comprise a polyadenylation sequence 3′ of a polynucleotide encoding a polypeptide to be expressed. Polyadenylation

sequences can promote mRNA stability by addition of a polyA tail to the 3′ end of the coding sequence and thus, contribute to increased translational efficiency. Recognized polyadenylation sites include an ideal polyA sequence (*e.g.,* AATAAA, ATTAAA AGTAAA), an SV40 polyA sequence, a bovine growth hormone polyA sequence (BGHpA), a rabbit β-globin polyA sequence (rβgpA), or another suitable heterologous or endogenous polyA sequence known in the art.

In certain embodiments, vectors comprise a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.,* ampicillin, neomycin, hygromycin, methotrexate, Zeocin, Blastocidin, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.,* the gene encoding D-alanine racemase for Bacilli. Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, 1977. *Cell* 11:223-232) and adenine phosphoribosyltransferase (Lowy *et al.*, 1990. *Cell* 22:817-823) genes which can be employed in tk- or aprt- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler *et al.*, 1980. *Proc. Natl. Acad. Sci. U.S.A.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, 1981. *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, 1988. *Proc. Natl. Acad. Sci. U.S.A.* 85:8047-51). Trp1 and/or Leu2 deficient yeast strains provide a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan (*e.g.,* ATCC No. 44076 or PEP4-1) or leucine (*e.g.,* ATCC 20,622 or 38,626).

A variety of expression prokaryotic and eukaryotic vector/host systems are known and may be utilized to contain and express polynucleotide sequences. The term "host cell" includes an isolated primary cell or a cell in culture which can be or has been a recipient of any recombinant vector(s) or isolated polynucleotide of the

invention. Host cells include eukaryotic cells and prokaryotic cells. In particular exemplary embodiments, eukaryotic cells include, plant cells (*e.g., Arabidopsis*), yeast cells (*e.g., Pichia Saccharomyces*, and *Schizosaccharomyces*), insect cells (*e.g., Drosophila*), and mammalian cells (*e.g.,* murine and human). Exemplary prokaryotic cells include *Escherichia coli*, among others. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation and/or genetic change. A host cell includes cells transfected or infected *in vivo* or *in vitro* with a recombinant vector or a polynucleotide of the invention. The term "transfection" refers to the introduction of foreign DNA into eukaryotic cells. Transfection may be accomplished by a variety of means known in the art including but not limited to calcium phosphate-DNA co-precipitation, DEAE-dextran-mediated transfection, polybrene-mediated transfection, electroporation, microinjection, liposome fusion, lipofection, protoplast fusion, retroviral infection, and biolistics. A host cell which comprises a recombinant vector of the invention is a recombinant host cell. A host cell also includes cells transduced *in vivo* or *in vitro* with a recombinant polypeptide or polypeptide fusion protein of the invention, as described elsewhere herein.

D. Polypeptides

The present invention contemplates, in part, compositions comprising polypeptides and vectors that express polypeptides. "Polypeptide," "polypeptide fragment," "peptide" and "protein" are used interchangeably, unless specified to the contrary, and according to conventional meaning, *i.e.,* as a sequence of amino acids. Polypeptides are not limited to a specific length, *e.g.,* they may comprise a full length protein sequence or a fragment of a full length protein, and may include post-translational modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. Polypeptides can be prepared using any of a variety of well known recombinant and/or synthetic techniques. An "isolated peptide" or an "isolated polypeptide" and the like, as used herein, refer to *in vitro*

isolation and/or purification of a peptide or polypeptide molecule from a cellular environment, and from association with other components of the cell, *i.e.*, it is not significantly associated with *in vivo* substances. Similarly, an "isolated cell" refers to a cell that has been obtained from an *in vivo* tissue or organ and is substantially free of extracellular matrix.

Polypeptides include polypeptide variants. Polypeptide variants may differ from a naturally occurring polypeptide in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences used in the methods of the invention and evaluating their effects using any of a number of techniques well known in the art. Preferably, polypeptides of the invention include polypeptides having at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% amino acid identity thereto.

In certain embodiments, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Modifications may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, variant polypeptide of the invention, one skilled in the art, for example, can change one or more of the codons of the encoding DNA sequence, *e.g.,* according to Table 1.

TABLE 1- Amino Acid Codons

Amino Acids	**One letter code**	**Three letter code**	**Codons**					
Alanine	A	Ala	GCA	GCC	GCG	GCU		
Cysteine	C	Cys	UGC	UGU				
Aspartic acid	D	Asp	GAC	GAU				
Glutamic acid	E	Glu	GAA	GAG				
Phenylalanine	F	Phe	UUC	UUU				
Glycine	G	Gly	GGA	GGC	GGG	GGU		
Histidine	H	His	CAC	CAU				
Isoleucine	I	Iso	AUA	AUC	AUU			
Lysine	K	Lys	AAA	AAG				
Leucine	L	Leu	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	M	Met	AUG					
Asparagine	N	Asn	AAC	AAU				
Proline	P	Pro	CCA	CCC	CCG	CCU		
Glutamine	Q	Gln	CAA	CAG				
Arginine	R	Arg	AGA	AGG	CGA	CGC	CGG	CGU
Serine	S	Ser	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	T	Thr	ACA	ACC	ACG	ACU		
Valine	V	Val	GUA	GUC	GUG	GUU		
Tryptophan	W	Trp	UGG					
Tyrosine	Y	Tyr	UAC	UAU				

Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological activity can be found using computer programs well known in the art, such as DNASTAR™ software. Preferably, amino acid changes in the protein variants disclosed herein are conservative amino acid changes, *i.e.,* substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-polar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cystine,

serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and generally can be made without altering a biological activity of a resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.,* Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p.224). Exemplary conservative substitutions are described in U.S. Provisional Patent Application No. 61/241,647 , the disclosure of which is herein incorporated by reference.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.,* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity.

As detailed in U.S. Patent No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine

(+0.2); glycine (0); threonine (–0.4); proline (–0.5 ± 1); alanine (–0.5); histidine (–0.5); cysteine (–1.0); methionine (–1.3); valine (–1.5); leucine (–1.8); isoleucine (–1.8); tyrosine (–2.3); phenylalanine (–2.5); tryptophan (–3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions may be based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like.

Polypeptide variants further include glycosylated forms, aggregative conjugates with other molecules, and covalent conjugates with unrelated chemical moieties (*e.g.,* pegylated molecules). Covalent variants can be prepared by linking functionalities to groups which are found in the amino acid chain or at the N- or C-terminal residue, as is known in the art. Variants also include allelic variants, species variants, and muteins. Truncations or deletions of regions which do not affect functional activity of the proteins are also variants.

In one embodiment, where expression of two or more polypeptides is desired, the polynucleotide sequences encoding them can be separated by and IRES sequence as discussed elsewhere herein. In another embodiment, two or more polypeptides can be expressed as a fusion protein that comprises one or more self-cleaving polypeptide sequences.

Polypeptides of the present invention include fusion polypeptides. In preferred embodiments, fusion polypeptides and polynucleotides encoding fusion polypeptides are provided. Fusion polypeptides and fusion proteins refer to a polypeptide having at least two, three, four, five, six, seven, eight, nine, or ten polypeptide segments.

Fusion polypeptides can comprise one or more polypeptide domains or segments including, but are not limited to cell permeable peptide domains (CPP), epitope tags (*e.g.,* maltose binding protein ("MBP"), glutathione S transferase (GST),

HIS6, MYC, FLAG, V5, VSV-G, and HA), polypeptide linkers, and polypeptide cleavage signals. Fusion polypeptides are typically linked C-terminus to N-terminus, although they can also be linked C-terminus to C-terminus, N-terminus to N-terminus, or N-terminus to C-terminus. The polypeptides of the fusion protein can be in any order. Fusion polypeptides or fusion proteins can also include conservatively modified variants, polymorphic variants, alleles, mutants, subsequences, and interspecies homologs, so long as the desired transcriptional activity of the fusion polypeptide is preserved. Fusion polypeptides may be produced by chemical synthetic methods or by chemical linkage between the two moieties or may generally be prepared using other standard techniques. Ligated DNA sequences comprising the fusion polypeptide are operably linked to suitable transcriptional or translational control elements as discussed elsewhere herein.

In one embodiment, a fusion partner comprises a sequence that assists in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments or to facilitate transport of the fusion protein through the cell membrane.

In various embodiments, fusion polypeptides comprise one or more CPPs. An important factor in the administration of polypeptide compounds is ensuring that the polypeptide has the ability to traverse the plasma membrane of a cell, or the membrane of an intra-cellular compartment such as the nucleus. Cellular membranes are composed of lipid-protein bilayers that are freely permeable to small, nonionic lipophilic compounds and are inherently impermeable to polar compounds, macromolecules, and therapeutic or diagnostic agents. However, proteins, lipids and other compounds, which have the ability to translocate polypeptides across a cell membrane, have been described.

Examples of peptide sequences which can facilitate protein uptake into cells include, but are not limited to: HIV TAT polypeptides; a 20 residue peptide sequence which corresponds to amino acids 84-103 of the p16 protein (see Fahraeus *et al.*, 1996. *Curr. Biol.* 6:84); the third helix of the 60-amino acid long homeodomain of *Antennapedia* (Derossi *et al.*, 1994. *J. Biol. Chem.* 269:10444); the h region of a signal

peptide, such as the Kaposi fibroblast growth factor (K-FGF) h region; and the VP22 translocation domain from HSV (Elliot *et al.*, 1997. *Cell* 88:223-233). In addition, Several bacterial toxins, including *Clostridium perfringens* iota toxin, diphtheria toxin (DT), *Pseudomonas* exotoxin A (PE), *Bordetella pertussis* toxin (PT), *Bacillus anthracis* toxin, and *Bordetella pertussis* adenylate cyclase (CYA), have been used to deliver peptides to the cell cytosol as internal or amino-terminal fusions. Arora *et al.,* 1993. *J. Biol. Chem.* 268:3334-3341; Perelle *et al.*, 1993. *Infect. Immun.* 61:5147-5156; Stenmark *et al.*, 1991. *J. Cell Biol.* 113:1025-1032; Donnelly *et al.*, 1993. *Proc. Natl. Acad. Sci. USA* 90:3530-3534; Carbonetti *et al.*, 1995. *Abstr. Annu. Meet. Am. Soc. Microbiol.* 95:295; Sebo *et al.*, 1995. *Infect. Immun.* 63:3851-3857; Klimpel *et al.*, 1992. *Proc. Natl. Acad. Sci. USA*. 89:10277-10281; and Novak *et al.*, 1992. *J. Biol. Chem.* 267:17186-17193.

Other exemplary CPP amino acid sequences include, but are not limited to: RKKRRQRRR (SEQ ID NO: 1), KKRRQRRR (SEQ ID NO: 2), and RKKRRQRR (SEQ ID NO: 3) (derived from HIV TAT protein); RRRRRRRRR (SEQ ID NO: 4); KKKKKKKKK (SEQ ID NO: 5); RQIKIWFQNRRMKWKK (SEQ ID NO: 6) (from *Drosophila Antp* protein); RQIKIWFQNRRMKSKK (SEQ ID NO: 7) (from *Drosophila Ftz* protein); RQIKIWFQNKRAKIKK (SEQ ID NO: 8) (from *Drosophila Engrailed* protein); RQIKIWFQNRRMKWKK (SEQ ID NO: 9) (from human *Hox-A5* protein); and RVIRVWFQNKRCKDKK (SEQ ID NO: 10) (from human *Isl-1* protein). Such subsequences can be used to facilitate polypeptide translocation, including the fusion polypeptides contemplated herein, across a cell membrane.

Fusion polypeptides may optionally comprises a linker that can be used to link the one or more polypeptides. A peptide linker sequence may be employed to separate any two or more polypeptide components by a distance sufficient to ensure that each polypeptide folds into its appropriate secondary and tertiary structures so as to allow the polypeptide domains to exert their desired functions. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second

polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*, *Gene* 40:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. Linker sequences are not required when a particular fusion polypeptide segment contains non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference. Preferred linkers are typically flexible amino acid subsequences which are synthesized as part of a recombinant fusion protein. Linker polypeptides can be between 1 and 200 amino acids in length, between 1 and 100 amino acids in length, or between 1 and 50 amino acids in length, including all integer values in between.

Exemplary linkers include, but are not limited to the following amino acid sequences: DGGGS (SEQ ID NO: 11); TGEKP (SEQ ID NO: 12) (see, *e.g.,* Liu *et al.*, *PNAS* 5525-5530 (1997)); GGRR (SEQ ID NO: 13) (Pomerantz *et al.* 1995, supra); $(GGGGS)_n$ (SEQ ID NO: 14) (Kim *et al.*, *PNAS* 93, 1156-1160 (1996.); EGKSSGSGSESKVD (SEQ ID NO: 15) (Chaudhary *et al.*, 1990, *Proc. Natl. Acad. Sci. U.S.A.* 87:1066-1070); KESGSVSSEQLAQFRSLD (SEQ ID NO: 16) (Bird *et al.*, 1988, *Science* 242:423-426), GGRRGGGS (SEQ ID NO: 17); LRQRDGERP (SEQ ID NO: 18); LRQKDGGGSERP (SEQ ID NO: 19); $LRQKD(GGGS)_2ERP$ (SEQ ID NO: 20). Alternatively, flexible linkers can be rationally designed using a computer program capable of modeling both DNA-binding sites and the peptides themselves (Desjarlais & Berg, *PNAS* 90:2256-2260 (1993), *PNAS* 91:11099-11103 (1994) or by phage display methods.

Fusion polypeptides may further comprise a polypeptide cleavage signal between each of the polypeptide domains described herein. In addition, polypeptide site can be put into any linker peptide sequence. Exemplary polypeptide cleavage signals include polypeptide cleavage recognition sites such as protease cleavage sites, nuclease cleavage sites (*e.g.,* rare restriction enzyme recognition sites, self-cleaving

ribozyme recognition sites), and self-cleaving viral oligopeptides (see deFelipe and Ryan, 2004. *Traffic*, 5(8); 616-26).

Suitable protease cleavages sites and self-cleaving peptides are known to the skilled person (*see, e.g.,* in Ryan *et al.*, 1997. *J. Gener. Virol.* 78, 699-722; Scymczak *et al.* (2004) Nature Biotech. 5, 589-594). Exemplary protease cleavage sites include, but are not limited to the cleavage sites of potyvirus NIa proteases (*e.g.,* tobacco etch virus protease), potyvirus HC proteases, potyvirus P1 (P35) proteases, byovirus NIa proteases, byovirus RNA-2-encoded proteases, aphthovirus L proteases, enterovirus 2A proteases, rhinovirus 2A proteases, picorna 3C proteases, comovirus 24K proteases, nepovirus 24K proteases, RTSV (rice tungro spherical virus) 3C-like protease, PYVF (parsnip yellow fleck virus) 3C-like protease, heparin, thrombin, factor Xa and enterokinase. Due to its high cleavage stringency, TEV (tobacco etch virus) protease cleavage sites are preferred in one embodiment, *e.g.,* EXXYXQ(G/S) (SEQ ID NO: 21), for example, ENLYFQG (SEQ ID NO: 22) and ENLYFQS (SEQ ID NO: 23), wherein X represents any amino acid (cleavage by TEV occurs between Q and G or Q and S).

In a particular embodiment, self-cleaving peptides include those polypeptide sequences obtained from potyvirus and cardiovirus 2A peptides, FMDV (foot-and-mouth disease virus), equine rhinitis A virus, Thosea asigna virus and porcine teschovirus.

In certain embodiments, the self-cleaving polypeptide site comprises a 2A or 2A-like site, sequence or domain (Donnelly *et al.*, 2001. *J. Gen. Virol.* 82:1027-1041).

Table 2: Exemplary 2A sites include the following sequences:

SEQ ID NO: 24	LLNFDLLKLAGDVESNPGP
SEQ ID NO: 25	TLNFDLLKLAGDVESNPGP
SEQ ID NO: 26	LLKLAGDVESNPGP
SEQ ID NO: 27	NFDLLKLAGDVESNPGP
SEQ ID NO: 28	QLLNFDLLKLAGDVESNPGP
SEQ ID NO: 29	APVKQTLNFDLLKLAGDVESNPGP

SEQ ID NO: 30	VTELLYRMKRAETYCPRPLLAIHPTEARHKQKIVAPVKQT
SEQ ID NO: 31	LNFDLLKLAGDVESNPGP
SEQ ID NO: 32	LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP
SEQ ID NO: 33	EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP

E. Viral Vector Elements

In various embodiments, the present invention contemplates, in part, engineering effective therapeutic viral vector systems by taking advantage of the desirable features of various types of viruses. In one non-limiting example, retroviral, or optionally lentiviral genomes, are used as the transfer vector backbone and heterologous RNA dependent RNA polymerase systems, *e.g.,* of alphavirus or narnavirus, are used to control and amplify transcription of one or more polynucleotides-of-interest.

In another non-limiting example, a therapeutic vector system comprises at least two vectors. In this example, at least one of the vectors encodes the RNA dependent RNA polymerase, *e.g.,* from alphavirus or narnavirus, either from an expression vector or a retroviral, *e.g.,* lentiviral vector, and at least one of the vectors is a retroviral or lentiviral vector comprising a RNA dependent RNA polymerase promoter operably linked to one or more polynucleotides-of-interest.

1. Retrovirus/Lentivirus Vector Elements

Retroviral vectors are a common tool for gene delivery (Miller, 2000, *Nature*. 357: 455-460). As used herein, the term "retrovirus" refers an RNA virus that utilizes reverse transcriptase during its replication cycle. Illustrative retroviruses include, but are not limited to: type c retroviruses, such as Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), feline leukemia virus (FLV), spumavirus, Friend, Murine Stem Cell Virus (MSCV) and Rous Sarcoma Virus (RSV)). Retroviruses of the invention also include human T cell leukemia viruses, HTLV-1 and HTLV-2, and the lentiviral family of retroviruses. The core sequence of the retroviral vectors of the present invention may be readily derived from a wide variety of

retroviruses, including for example, B, C, and D type retroviruses as well as spumaviruses and lentiviruses (RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985). An example of a retrovirus suitable for use in the compositions and methods of the present invention includes, but is not limited to, a lentivirus. Other retroviruses suitable for use in the compositions and methods of the present invention include, but are not limited to, Avian Leukosis Virus, Bovine Leukemia Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis virus and Rous Sarcoma Virus. Preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe, 1976. *J. Virol.* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998), and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be readily obtained from depositories or collections such as the American Type Culture Collection ("ATCC"; Rockville, Md.), or isolated from known sources using commonly available techniques.

Retroviral vectors and more particularly lentiviral vectors may be used in practicing the present invention. Accordingly, the term "retrovirus" or "retroviral vector", as used herein is meant to include "lentivirus" and "lentiviral vectors" respectively. As used herein, the term "lentivirus" refers to a group (or genus) of retroviruses that give rise to slowly developing disease. Viruses included within this group include HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2), the etiologic agent of the human acquired immunodeficiency syndrome (AIDS); visna-maedi, which causes encephalitis (visna) or pneumonia (maedi) in sheep (VMV) virus; the caprine arthritis-encephalitis virus (CAEV), which causes immune deficiency, arthritis, and encephalopathy in goats; equine infectious anemia virus (EIAV), which causes autoimmune hemolytic anemia, and encephalopathy in horses; feline immunodeficiency virus (FIV), which causes immune deficiency in cats; bovine immune deficiency virus (BIV), which causes lymphadenopathy, lymphocytosis, and possibly central nervous system infection in cattle; and simian immunodeficiency virus (SIV), which cause immune deficiency and encephalopathy in sub-human primates. It will be appreciated that each of these viruses exists in multiple variants or strains.

Diseases caused by these viruses are characterized by a long incubation period and protracted course. Usually, the viruses latently infect monocytes and macrophages, from which they spread to other cells. HIV, FIV, and SIV also readily infect T lymphocytes (*i.e.*, T-cells). In one embodiment, HIV based vector backbones (*i.e.*, HIV cis-acting sequence elements) are preferred.

Retroviral vectors have been tested and found to be suitable delivery vehicles for the stable introduction of genes of interest into the genome of a broad range of target cells.

The term "vector" is used herein to refer to a nucleic acid molecule capable transferring or transporting another nucleic acid molecule. The transferred nucleic acid is generally linked to, *e.g.*, inserted into, the vector nucleic acid molecule. A vector may include sequences that direct autonomous replication in a cell, or may include sequences sufficient to allow integration into host cell DNA. Useful vectors include, for example, plasmids (*e.g.*, DNA plasmids or RNA plasmids), cosmids, bacterial artificial chromosomes, and viral vectors. Useful viral vectors include, *e.g.*, replication defective retroviruses and lentiviruses.

As will be evident to one of skill in the art, the term "viral vector" is widely used refer either to a nucleic acid molecule (*e.g.*, a transfer plasmid) that includes virus-derived nucleic acid elements that typically facilitate transfer of the nucleic acid molecule or integration into the genome of a cell or to a viral particle that mediates nucleic acid transfer. Viral particles will typically include various viral components and sometimes also host cell components in addition to nucleic acid(s).

The term viral vector may refer either to a virus or viral particle capable of transferring a nucleic acid into a cell or to the transferred nucleic acid itself. Viral vectors and transfer plasmids contain structural and/or functional genetic elements that are primarily derived from a virus. The term "retroviral vector" refers to a viral vector or plasmid containing structural and functional genetic elements that are primarily derived from a retrovirus. The term "lentiviral vector" refers to a viral vector or plasmid containing structural and functional genetic elements, including LTRs that are primarily derived from a lentivirus. The term "hybrid" refers to a vector, LTR or other nucleic acid containing both retroviral, *e.g.*, lentiviral, sequences and non-lentiviral

viral sequences. In one embodiment, a hybrid vector refers to a vector or transfers plasmid comprising retroviral *e.g.*, lentiviral, sequences for reverse transcription, replication, integration and/or packaging and alphavirus subgenomic promoter sequences, non-structural proteins, and/or polymerase recognition sites.

In particular embodiments, the terms "lentiviral vector," "lentiviral expression vector" may be used to refer to lentiviral transfer plasmids and/or infectious lentiviral particles. Where reference is made herein to elements such as cloning sites, promoters, regulatory elements, heterologous nucleic acids, *etc.*, it is to be understood that the sequences of these elements are present in RNA form in the lentiviral particles of the invention and are present in DNA form in the DNA plasmids of the invention.

The ability of retroviral vectors to deliver unrearranged, single copy transgenes into cells makes retroviral vectors well suited for transferring genes into cells. The retroviral genomic RNA is converted into double-stranded DNA by reverse transcriptase. This double-stranded DNA form of the virus is capable of being integrated into the chromosome of the infected cell. The delivery of a gene(s) or other polynucleotide sequence using a viral or retroviral vector by means of viral infection rather than by transfection is referred to as "transduction." In one embodiment, retroviral vectors are transduced into a cell through infection and provirus integration. In certain embodiments, a cell is "transduced" if it comprises a gene or other polynucleotide sequence delivered to the cell by infection using a viral or retroviral vector. In particular embodiments, a transduced cell comprises the gene or other polynucleotide sequence delivered to by a viral or retroviral vector in its cellular genome.

Once the virus is integrated into the host genome, it is referred to as a "provirus." The provirus serves as a template for RNA polymerase II and directs the expression of RNA molecules which encode the structural proteins and enzymes needed to produce new viral particles.

At each end of the provirus are structures called "long terminal repeats" or "LTRs." The term "long terminal repeat (LTR)" refers to domains of base pairs located at the ends of retroviral DNAs which, in their natural sequence context, are direct repeats and contain U3, R and U5 regions. LTRs generally provide functions

fundamental to the expression of retroviral genes (*e.g.,* promotion, initiation and polyadenylation of gene transcripts) and to viral replication. The LTR contains numerous regulatory signals including transcriptional control elements, polyadenylation signals and sequences needed for replication and integration of the viral genome. The viral LTR is divided into three regions called U3, R and U5. The U3 region contains the enhancer and promoter elements. The U5 region is the sequence between the primer binding site and the R region and contains the polyadenylation sequence. The R (repeat) region is flanked by the U3 and U5 regions. The LTR composed of U3, R and U5 regions, appears at both the both the 5' and 3' ends of the viral genome. Adjacent to the 5' LTR are sequences necessary for reverse transcription of the genome (the tRNA primer binding site) and for efficient encapsidation of viral RNA into particles (the Psi site).

In various embodiments, the 5' LTR and/or 3' LTR is modified. Modifications of the 3' LTR are often made to improve the safety of lentiviral or retroviral systems by rendering viruses replication-defective. As used herein, the term "replication-defective" refers to virus that is not capable of complete, effective replication such that infective virions are not produced (*e.g.,* replication-defective lentiviral progeny). The term "replication-competent" refers to wild-type virus or mutant virus that is capable of replication, such that viral replication of the virus is capable of producing infective virions (*e.g.,* replication-competent lentiviral progeny). "Self-inactivating" (SIN) vectors refers to replication-defective vectors, *e.g.,* retroviral or lentiviral vectors, in which the right (3') LTR enhancer-promoter region, known as the U3 region, has been modified (*e.g.,* by deletion or substitution) to prevent viral transcription beyond the first round of viral replication. Consequently, the vectors are capable of infecting and then integrating into the host genome only once, and cannot be passed further. This is because the right (3') LTR U3 region is used as a template for the left (5') LTR U3 region during viral replication and, thus, the viral transcript cannot be made without the U3 enhancer-promoter. If the viral transcript is not made, it cannot be processed or packaged into virions, hence the life cycle of the virus ends. Accordingly, SIN vectors greatly reduce risk of creating unwanted replication-competent virus since the right (3') LTR U3 region has been modified to prevent viral

transcription beyond the first round of replication, hence eliminating the ability of the virus to be passed. In a further embodiment of the invention, the 3' LTR is modified such that the U5 region is replaced, for example, with an ideal poly(A) sequence. It should be noted that modifications to the LTRs such as modifications to the 3' LTR, the 5' LTR, or both 3' and 5' LTRs, are also included in the invention.

An additional safety enhancement is provided by replacing the U3 region of the 5' LTR with a heterologous promoter to drive transcription of the viral genome during production of viral particles. Examples of heterologous promoters which can be used include, for example, viral simian virus 40 (SV40) (*e.g.,* early or late), cytomegalovirus (CMV) (*e.g.,* immediate early), Moloney murine leukemia virus (MoMLV) or Rous sarcoma virus (RSV), and herpes simplex virus (HSV) (thymidine kinase) promoters. Typical promoters are able to drive high levels of transcription in a Tat-independent manner. This replacement reduces the possibility of recombination to generate replication-competent virus because there is no complete U3 sequence in the virus production system.

In some embodiments, viral vectors comprise a TAR element. The term "TAR" refers to the "trans-activation response" genetic element located in the R region of lentiviral (*e.g.,* HIV) LTRs. This element interacts with the lentiviral trans-activator (tat) genetic element to enhance viral replication. However, this element is not required in embodiments wherein the U3 region of the 5' LTR is replaced by a heterologous promoter.

The "R region" refers to the region within retroviral LTRs beginning at the start of the capping group (*i.e.*, the start of transcription) and ending immediately prior to the start of the poly A tract. The R region is also defined as being flanked by the U3 and U5 regions. The R region plays a role during reverse transcription in permitting the transfer of nascent DNA from one end of the genome to the other.

In particular embodiments, the retroviral or lentiviral vector backbones comprise one or more FLAP elements upstream or downstream of the heterologous genes of interest in the vectors. For example, in particular embodiments a transfer plasmid includes a FLAP element. As used herein, the term "FLAP element" refers to a nucleic acid whose sequence includes the central polypurine tract and central

termination sequences (cPPT and CTS) of a retrovirus. Typically the retrovirus is a lentivirus, *e.g.,* HIV-1 or HIV-2. Suitable FLAP elements are described in U.S. Pat. No. 6,682,907 and in Zennou, *et al.*, (2000, Cell, 101:173). During HIV-1 reverse transcription, central initiation of the plus-strand DNA at the central polypurine tract (cPPT) and central termination at the central termination sequence (CTS) lead to the formation of a three-stranded DNA structure: the HIV-1 central DNA flap. While not wishing to be bound by any theory, the DNA flap may act as a cis-active determinant of lentiviral genome nuclear import and/or may increase the titer of the virus.

In one embodiment, retroviral or lentiviral transfer vectors comprise one or more export elements. The term "export element" refers to a cis-acting post-transcriptional regulatory element which regulates the transport of an RNA transcript from the nucleus to the cytoplasm of a cell. Examples of RNA export elements include, but are not limited to, the human immunodeficiency virus (HIV) rev response element (RRE) (*see e.g.,* Cullen *et al.*, 1991. *J. Virol.* 65: 1053; and Cullen *et al.*, 1991. *Cell* 58: 423), and the hepatitis B virus post-transcriptional regulatory element (HPRE). Generally, the RNA export element is placed within the 3' UTR of a gene, and can be inserted as one or multiple copies.

In particular embodiments, expression of heterologous sequences in viral vectors is increased by incorporating posttranscriptional regulatory elements, and efficient polyadenylation sites and optionally, transcription termination signals into the vectors. A variety of posttranscriptional regulatory elements can increase expression of a heterologous nucleic acid at the protein level. One example is the posttranscriptional regulatory element (PRE) is the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE) as described (Zufferey *et al.*, 1999, *J. Virol.*, 73:2886). Other posttranscriptional regulatory elements that may be used include the posttranscriptional processing element present within the genome of various viruses such as that present within the thymidine kinase gene of herpes simplex virus (Liu *et al.,* 1995, *Genes Dev.*, 9:1766), and the posttranscriptional regulatory element present in hepatitis B virus (HPRE) (Huang *et al., Mol. Cell. Biol.*, 5:3864). The posttranscriptional regulatory element is positioned at the 3' end the heterologous nucleic acid sequence. This configuration results in synthesis of a transcript whose 5'

portion comprises the heterologous nucleic acid coding sequences and whose 3' portion comprises the posttranscriptional regulatory element sequence.

Elements directing the efficient termination and polyadenylation of the heterologous nucleic acid transcripts increases heterologous gene expression. Transcription termination signals are generally found downstream of the polyadenylation signal. The term "polyA site" or "polyA sequence" as used herein denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript by RNA polymerase II. Efficient polyadenylation of the recombinant transcript is desirable as transcripts lacking a poly A tail are unstable and are rapidly degraded. The polyA signal utilized in an expression vector may be heterologous or endogenous polyA signal.

Retroviruses have a different distribution of favored integration sites in the human genome than is seen with lentiviruses. Lentiviruses favor integration within active transcription units, while retroviruses favor integration near transcription start sites and associated features such as CpG islands and DNAseI cleavage sites. The association of retroviral integration with gene 5′-ends may be part of the reason insertional activation is seen with these vectors. However, both retroviruses and lentiviruses may be subject to position effects of surrounding gene expression elements, depending on where the proviral integration site is in the genome. Accordingly, in certain embodiments, a retroviral or lentiviral vector further comprises an insulator element. Insulator elements act to preserve the independent function of genes or transcription units embedded in a genome or genetic context in which their expression may otherwise be influenced by regulatory signals within the genome or genetic context (*i.e.,* position effect; *see, e.g.,* Burgess-Beusse *et al.*, 2002, *Proc. Natl. Acad. Sci., USA*, 99:16433; and Zhan *et al.*, 2001, *Hum. Genet.*, 109:471). Thus, insulators may contribute to protecting lentivirus-expressed sequences from integration site effects, which may be mediated by cis-acting elements present in genomic DNA and lead to deregulated expression of transferred sequences. In some embodiments, transfer vectors comprise an insulator element in one or both LTRs or elsewhere in the region of the vector that integrates into the cellular genome. Suitable insulators for use in the invention include, but are not limited to, the chicken β-Globin insulator (see Chung *et*

al., 1993. *Cell* 74:505; Chung *et al.,* 1997. *PNAS* 94:575; and Bell *et al.*, 1999. *Cell* 98:387, incorporated by reference herein). Examples of insulator elements include, but are not limited to, an insulator from an β-globin locus, such as chicken HS4.

According to certain specific embodiments of the invention, most or all of the viral vector backbone sequences are derived from a lentivirus, *e.g.,* HIV-1. However, it is to be understood that many different sources of lentiviral sequences can be used, and numerous substitutions and alterations in certain of the lentiviral sequences may be accommodated without impairing the ability of a transfer plasmid to perform the functions described herein. Moreover, a variety of lentiviral vectors are known in the art, see Naldini *et al.*, (1996a, 1996b, and 1998); Zufferey *et al.*, (1997); Dull *et al.*, 1998, U.S. Pat. Nos. 6,013,516; and 5,994,136, any of which may be adapted to produce a transfer vector of the present invention.

2. RNA dependent RNA polymerase Systems

In various embodiments, the present invention provides improved gene switches that solve the problems associated with poorly controlled or regulatable gene expression systems. In preferred embodiments, the invention provides systems that are not leaky, in contrast to other inducible gene expression systems such a cumate or tetracycline inducible gene expression systems, which are very leaky. Accordingly, the gene switches of the present invention are superior to and provide numerous advantages over existing gene expression systems in the art.

Temporal control of gene expression, as well as control of the amplitude of gene expression, are both important considerations in many viral-based methods of gene therapy. For example, tight regulation of cell-toxic viral packaging genes is important to preserve the safety, longevity, and fitness of packaging and producer cell lines. Illustrative embodiments of the present invention, provide state-of-the-art packaging and producer cell lines that provide for maximum safety, longevity, and health of cell lines. This results in maximum scalable production of safe virus at increased viral titers compared to existing methods in the art.

Additionally, embodiments of the present invention provide robust systems for controlling the temporal expression and amplitude of gene expression *in*

vivo, in methods of gene therapy. Tight control of therapeutic gene expression *in vivo* is extremely important due to the fact that insertion of viral vectors comprising the therapeutic genes in the genome is semi-random. Thus, in situations where insertion results in abnormal or dangerous levels of therapeutic gene expression, embodiments of the present invention advantageously provide mechanisms for damping gene expression to safe therapeutic levels and/or eliminating gene expression all together. Accordingly, illustrative embodiments of the invention provide for the maximum therapeutic efficacy while preserving maximum safety for the patient.

The present invention contemplates, in part, using viral RNA dependent RNA polymerase (RDRP) systems. As used herein, the term "RNA dependent RNA polymerase" refers to any RNA-based virus protein that uses RNA as a template for replication of RNA sequences. Illustrative examples of suitable RNA viruses for use in RDRP systems of the invention include, but are not limited to plus-strand single stranded RNA viruses (*i.e.,* plus single stranded RNA viruses, plus or positive sense ssRNA viruses, or (+) ssRNA viruses) and minus-strand single stranded RNA viruses (*i.e.,* minus single stranded RNA viruses, minus or negative sense ssRNA viruses, or (-) ssRNA viruses).

Plus single stranded RNA viruses have their genome directly utilized as if it were mRNA, producing one or more proteins or polyprotein which are modified by host and viral proteins to form the various proteins needed for replication. One of these proteins, RDRP catalyzes the synthesis of large numbers of "antisense" replicative intermediates. These intermediates serve as templates for the synthesis of large numbers of mRNA molecules that are translated by the host cell machinery into the proteins needed to make fresh virions. In various embodiments, the templates for amplification are minus-strand RNAs comprising minus-strand RDRP promoters, *e.g.,* junctional or subgenomic promoters. The minus-strand subgenomic promoters are operable or functional on the minus-strand, but not on the plus-strand.

In other embodiments, the plus-strand mRNAs comprise the minus-RDRP promoters in the reverse orientation, which allows the RDRP to use the plus-strand like a minus-strand template (*i.e.,* transcribe RNAs in the 3′ to 5′ direction), thus, amplifying the pool of translatable RNAs.

Minus single stranded RNA viruses must have their genome copied by an RNA polymerase to form positive-sense RNA. This means that the virus comprises an RNA-dependent RNA polymerase enzyme along with its genomic RNA. The positive-sense RNA molecule then acts as viral mRNA, which is translated into proteins by the host ribosomes. The resultant protein goes on to direct the synthesis of new virions, such as capsid proteins and RNA replicase, which is used to produce new negative-sense RNA molecules.

Exemplary plus single stranded RNA viruses suitable for use in the RDRP systems of the present invention are present in Table 3.

Table 3: Plus Single Stranded RNA Viruses

Family	**Exemplary Virus**	**Genome Accession**
Alphaflexiviridae	Allexivirus (Shallot virus X (ShVX))	NC_003795.1
	Botrexvirus (Botrytis virus X (BotV-X))	NC_005132.1
	Lolavirus (Lolium latent virus (LoLV))	NC_010434.1
	Mandarivirus (Indian citrus ringspot virus (ICRSV))	NC_003093.1
	Potexvirus (Potato virus X (PVX))	NC_011620.1
	Sclerodarnavirus (Sclerotinia sclerotiorum debilitation-associated RNA virus (SsDRV))	NC_007415.1
Betaflexiviridae	Capillovirus (Apple stem grooving virus (ASGV))	NC_001749.2
	Carlavirus (Potato virus M (PVM))	NC_001361.2
	Citrivirus (Citrus leaf blotch virus (CLBV))	NC_003877.1
	Foveavirus (Apple stem pitting virus (ASPV))	NC_003462.1
	Trichovirus (Apple chlorotic leaf spot virus (ACLSV))	NC_001409.1
	Vitivirus (Grapevine virus A (GVA))	NC_003604.2
Gammaflexiviridae	Mycoflexivirus (Botrytis virus F (BotV-F))	NC_002604.1
Tymoviridae	Maculavirus (Grapevine fleck virus (GFkV))	NC_003347.1
	Marafivirus (Maize rayado fino virus (MRFV))	NC_002786.1
	Tymovirus (Turnip yellow mosaic virus (TYMV))	NC_004063.1
Arteriviridae	Arterivirus (Equine arteritis virus (EAV))	NC_002532.2
Coronavirinae	Alphacoronavirus (Human coronavirus 229E (HCoV-	NC_002645.1

	229E))	
	Betacoronavirus (Human SARS coronavirus)	NC_004718.3
	Gammacoronavirus (Avian Infectious bronchitis virus (strain Beaudette))	NC_001451.1
Torovirinae	Bafinivirus (White bream virus)	NC_008516.1
	Torovirus (Breda virus)	NC_007447.1
Roniviridae	Okavirus (Gill-associated virus)	NC_010306.1
Dicistroviridae	Cripavirus (Cricket paralysis virus (CrPV))	NC_003924.1
Flaviridae	Flavirus (Infectious flacherie virus (IFV))	NC_003781.1
Marnaviridae	Marnavirus (Heterosigma akashiwo RNA virus (HaRNAV))	NC_005281.1
Picornaviridae	Aphthovirus (Foot-and-mouth disease virus O (FMDV))	NC_004004.1
	Avihepatovirus (Duck hepatitis A virus (DHAV))	NC_008250.1
	Cardiovirus (Encephalomyocarditis virus (EMCV))	NC_001479.1
	Enterovirus (Poliovirus (Human enterovirus C serotype PV-1))	NC_002058.3
	Erbovirus (Equine rhinitis B virus 1 (ERBV-1))	NC_003983.1
	Hepatovirus (Hepatitis A virus (HAV))	NC_001489.1
	Kobuvirus (Aichi virus (AiV))	NC_001918.1
	Parechovirus (Human parechovirus 2 (HPeV-2))	NC_001897.1
	Sapelovirus (Porcine sapelovirus)	NC_003987.1
	Senecavirus (Seneca Valley virus (SVV))	NC_011349.1
	Teschovirus (Porcine teschovirus 1 (PTV-1))	NC_003985.1
	Tremovirus (Avian encephalomyelitis virus (AEV))	NC_003990.1
Secoviridae	Cheravirus (Cherry rasp leaf virus (CRLV))	NC_006271.1
	Sadwavirus (Satsuma dwarf virus (SDV))	NC_003785.1
	Sequivirus (Parsnip yellow fleck virus (PYFV))	NC_003628.1
	Torradovirus (Tomato torrado virus (ToTV))	NC_009013.1
	Waikivirus (Rice tungro spherical virus (RTSV))	NC_001632.1

Comovirinae	Comovirus (Cowpea mosaic virus (CPMV))	NC_003549.1
	Fabavirus (Broad bean wilt virus 1 (BBWV-1))	NC_005289.1
	Nepovirus (Tomatop ringspot virus)	NC_003840.1
Astroviridae	Avastrovirus (Turkey astrovirus 1 (TAstV-1))	Y15936.2
	Mamastrovirus (Human astrovirus (HAstV))	Z25771.1
Barnaviridae	Barnavirus (Mushroom bacilliform virus (MBV))	U07551.1
Bromoviridae	Alfamovirus (Alfalfa mosaic virus (AMV))	NC_001495.1
	Anulavirus (Pelargonium zonate spot virus (PZSV))	NC_003649.1
	Bromovirus (Brome mosaic virus (BMV))	NC_002026.1
	Cucumovirus (Cucumber mosaic virus (strain FNY))	NC_002034.1
	Ilarvirus (Tobacco streak virus (TSV))	NC_003844.1
	Oleavirus (Olive latent virus 2 (OLV-2))	NC_003673.1
Caliciviridae	Lagovirus (Rabbit haemorrhagic disease virus (RHDV))	NC_001543.1
	Nebovirus (Newbury-1 virus)	NC_007916.1
	Norovirus (Norwalk virus (NV))	NC_001959.2
	Sapovirus (Sapporo virus (SV))	NC_006269.1
	Vesivirus (Vesicular exanthema of swine virus (VESV))	NC_002551.1
Closteroviridae	Ampelovirus (Grapevine leafroll-associated virus 3 (GLRaV-3))	NC_004667.1
	Closterovirus (Beet yellows virus (BYV))	NC_001598.1
	Crinivirus (Lettuce infectious yellows virus (LIYV))	NC_003617.1
Flaviviridae	Flavivirus(Yellow fever virus (strain 17D vaccine))	NC_002031.1
	Hepacivirus (Hepatitis C virus (HCV))	NC_004102.1
	Pestivirus (Bovine diarrhea virus 1 (BVDV-1))	NC_001461.1
Hepeviridae	Hepevirus (Hepatitis E virus (HEV))	NC_001434.1
Leviviridae	Allolevivirus (Enterobacteria phage Qbeta (Qbeta))	NC_001890.1
	Levivirus (Enterobacteria phage MS2 (MS2))	NC_001417.2
Luteoviridae	Enamovirus (Pea enation mosaic virus-1 (PEMV-1))	NC_003629.1
	Luteovirus Barley yellow dwarf virus-PAV (BYDV-	NC_004750.1

	PAV))	
	Polerovirus (Potato leafroll virus (PLRV))	NC_001747.1
Narnaviridae	Narnavirus (Saccharomyces cerevisiae narnavirus 20S (ScNV-20S)	NC_004051.1
	Mitovirus (Cryphonectria parasitica mitovirus 1-NB631)	NC_004046.1
Nodaviridae	Alphanodavirus (Nodamura virus (NoV))	NC_002690.1
	Betanodavirus (Striped jack nervous necrosis virus (SJNNV))	NC_003448.1
Potyviridae	Brambyvirus (Blackberry virus Y (BVY))	NC_008558.1
	Bymovirus (Barley yellow mosaic virus (BaYMV))	NC_002990.1
	Ipomovirus (Sweet potato mild mottle virus (SPMMV))	NC_003797.1
	Macluravirus (Maclura mosaic virus (MacMV))	Not yet
	Potyvirus (Potato virus Y (PVY))	NC_001616.1
	Rymovirus (Ryegrass mosaic virus (RGMV))	NC_001814.1
	Tritimovirus (Brome streak virus)	NC_003501.1
Tetraviridae	Betatetravirus (Nudaurelia capensis beta virus (NbetaV))	NC_001990.1
	Omegatetravirus (Helicoverpa armigera stunt virus)	NC_001981.1
Togaviridae	Alphavirus (Sindbis virus (SINV)), see also Table X	NC_001547.1
	Rubivirus (Rubella virus (RUBV))	NC_001545.1
Tombusviridae	Aureusvirus (Pothos latent virus)	NC_000939.1
	Avenavirus (Oat chlorotic stunt virus)	NC_003633.1
	Carmovirus (Carnation mottle virus)	NC_001265.1
	Dianthovirus (Red clover necrotic mosaic virus)	NC_003756.1
	Machlomovirus (Maize chlorotic mottle virus)	NC_003627.1
	Necrovirus (Tobacco necrosis virus A)	NC_001777.1
	Panicovirus (Panicum mosaic virus)	NC_002598.1
	Tombusvirus (Tomato bushy stunt virus)	NC_001554.1
Virgaviridae	Furovirus (Soil-borne wheat mosaic virus (SBWMV))	NC_002041.1

	Hordeivirus (Barley stripe mosaic virus (BSMV))	NC_003469.1
	Pecluvirus (Peanut clump virus (PCV))	NC_003672.1
	Pomovirus (Potato mop-top virus (PMTV))	NC_003723.1
	Tobamovirus Tobacco mosaic virus	NC_001367.1
	Tobravirus (Tobacco rattle virus (TRV))	NC_003805.1
None	Benyvirus (Beet necrotic yellow vein virus)	NC_003514.1
	Cilevirus (Citrus leprosis virus C (CiLV-C))	NC_008169.1
	Idaeovirus (Raspberry bushy dwarf virus (RBDV))	NC_003739.1
	Ourmiavirus (Ourmia melon virus (OuMV))	NC_011068.1
	Polemovirus (Poinsettia latent virus (PnLV))	NC_011543.1
	Sobemovirus (Southern bean mosaic virus (SBMV))	NC_004060.1
	Umbravirus (Groundnut rosette virus (GRV))	NC_003603.1

Exemplary minus single stranded RNA viruses suitable for use in the RDRP systems of the present invention are present in Table 4.

Table 4: Minus Single Stranded RNA Viruses

Family	**Exemplary Virus**	**Genome Accession**
Bornaviridae	Bornavirus (Borna disease virus (BDV))	NC_001607.1
Rhabdoviridae	Cytorhabdovirus (Lettuce necrotic yellows virus (LNYV))	NC_007642.1
	Ephemerovirus (Bovine ephemeral fever virus (BEFV))	NC_002526.1
	Lyssavirus (Rabies virus (RABV))	NC_001542.1
	Novirhabdovirus (Infectious hematopoietic necrosis virus (IHNV))	NC_001652.1
	Nucleorhabdovirus (Maize mosaic virus (MMV))	NC_005975.1
	Vesiculovirus (Vesicular stomatitis Indiana virus (VSIV))	NC_001560.1
Filoviridae	Ebolavirus (Zaire ebolavirus (ZEBOV))	NC_002549.1
	Marburgvirus (Lake Victoria marburgvirus (MARV))	NC_001608.3
Paramyxovirinae	Avulavirus (Newcastle disease virus (NDV))	NC_002617.1
	Henipavirus (Hendra virus (HeV))	NC_001906.2

	Morbillivirus (Measles virus (MeV))	NC_001498.1
	Respirovirus (Sendai virus (SeV))	NC_001552.1
	Rubulavirus (Mumps virus (MuV))	NC_002200.1
Pneumovirinae	Metapneumovirus (Human metapneumovirus (HMPV))	NC_004148.2
	Pneumovirus (Human respiratory syncytial virus (HRSV))	NC_001781.1
Arenaviridae	Arenavirus (Lymphocytic choriomeningitis virus (LCMV))	NC_004291.1
Bunyaviridae	Hantavirus (Hantaan virus (HTNV))	NC_005222.1
	Nairovirus (Dugbe virus (DUGV))	NC_004159.1
	Orthobunyavirus (Bunyamwera virus (BUNV))	NC_001925.1
	Phlebovirus (Rift valley fever virus (RVFV))	NC_014397.1
	Tospovirus (Tomato spotted wilt virus (TSWV))	NC_002052.1
Ophioviridae	Ophiovirus (Citrus psorosis virus (CPsV))	NC_006314.1
Orthomyxoviridae	H1N1 (Influenza A/Puerto Rico/8/34 H1N1 (FLUAV))	NC_002023.1
	H5N1 (Influenza A virus A/Goose/Guangdong/1/96 H5N1 (FLUAV))	NC_007357.1
	Influenzavirus B (Influenza B/Lee/1940 (FLUBV))	NC_002204.1
	Influenzavirus C (Influenza C virus (C/Ann Arbor/1/50))	NC_006307.1
	Thogotovirus (Thogoto virus (THOV))	NC_006508.1
	Isavirus (Infectious salmon anemia virus (ISAV))	NC_006505.1
None	Deltavirus (Hepatitis D virus (HDV))	NC_001653.2
	Emaravirus (European mountain ash ringspot-associated virus)	NC_013106.1
	Tenuivirus (Rice stripe virus (RSV))	NC_003755.1
	Varicosavirus (Lettuce big-vein virus (LBVaV))	NC_011558.1

Although particular illustrative embodiments include more detailed description of alphavirus and narnavirus RNA dependent RNA polymerase systems, the invention should not be considered to be limited by this disclosure. One having skill in

the art would readily appreciate that the principles of viral RNA dependent RNA polymerase systems of the present invention disclosed herein that are illustrated with alphavirus and narnavirus can be applied to the aforementioned single stranded viruses listed in Tables 3 and 4 for use in particular embodiments of RNA dependent RNA polymerase systems of the present invention.

a. Alphavirus Vector Elements

Alphaviruses comprise a set of genetically, structurally, and serologically related arthropod-borne viruses of the *Togaviridae* family. These viruses are distributed worldwide, and persist in nature through a mosquito to vertebrate cycle. Birds, rodents, horses, primates, and humans are among the defined alphavirus vertebrate reservoir/hosts. Twenty-six known viruses and virus subtypes have been classified within the alphavirus genus utilizing the hemagglutination inhibition (HI) assay. Sequences encoding wild-type alphaviruses suitable for use in preparing the vectors disclosed herein can be readily obtained from naturally-occurring sources, or from depositories (*e.g.,* the American Type Culture Collection, Rockville, Md.).

Representative examples of suitable alphaviruses include Aura virus (ATCC VR-368), Bebaru virus (ATCC VR-600, ATCC VR-1240), Cabassou virus (ATCC VR-922), Chikungunya virus (ATCC VR-64, ATCC VR-1241), Eastern equine encephalomyelitis virus (ATCC VR-65, ATCC VR-1242), Fort Morgan virus (ATCC VR-924), Getah virus (ATCC VR-369, ATCC VR-1243), Kyzylagach virus (ATCC VR-927), Mayaro virus (ATCC VR-66, ATCC VR-1277), Middleburg virus (ATCC VR-370), Mucambo virus (ATCC VR-580, ATCC VR-1244), Ndumu virus (ATCC VR-371), Pixuna virus (ATCC VR-372, ATCC VR-1245), Ross River virus (ATCC VR-373, ATCC VR-1246), Semliki Forest virus (ATCC VR-67, ATCC VR-1247), Sindbis virus (ATCC VR-68, ATCC VR-1248; see also CMCC #464), Tonate virus (ATCC VR-925), Triniti virus (ATCC VR-469). Una virus (ATCC VR-374), Venezuelan equine encephalomyelitis virus (ATCC VR-69, ATCC VR-923, ATCC VR-1250 ATCC VR-1249, ATCC VR-532), Western equine encephalomyelitis virus (ATCC VR-70, ATCC VR-1251, ATCC VR-622, ATCC VR-1252), Whataroa virus (ATCC VR-926), and Y-62-33 virus (ATCC VR-375).

The genome of an alphavirus is an approximately 12 kb single-stranded positive-sense RNA molecule which is capped and polyadenylated, and contained within a virus-encoded capsid protein shell. After release of viral genomic RNA into the cytoplasm, the replicative process is initiated by translation of four nonstructural proteins (NSP1-NSP4) from the 5' two-thirds of the viral genome that participate in the RNA replicative process, which includes binding to the conserved nucleotide sequence elements (CSEs) present at the 5' and 3' ends, and a junction region subgenomic promoter (*i.e.,* a 26S subgenomic promoter) located internally in the genome.

The positive strand genomic RNA, *i.e.,* the sense strand, serves as template for the NSP-catalyzed synthesis of a full-length complementary negative strand, *i.e.,* the non-sense or antisense strand. Synthesis of a complementary negative strand is catalyzed by NSPs binding to the 3' terminal CSE of the positive strand genomic RNA. The negative strand, in turn, serves as template for the synthesis of additional positive strand genomic RNA and an abundantly expressed 26S subgenomic RNA, initiated internally at the junction region promoter. Thus, the 5' end CSE and junction region promoter are functional only after they are transcribed into the negative strand, *i.e.,* the complement of the positive strand genomic RNA (*i.e.,* the 5' end CSE is functional when it is the 3' end of the genomic negative stranded complement). Additional positive strand genomic RNA synthesis from the 5 'CSE and abundant expression of 26S subgenomic RNA from the 26S promoter results in a spectacular increase in expression of the RNA downstream of the 26S promoter, normally structural proteins (SPs) in the context of a complete alphavirus genome.

A wide variety of sequences which encode alphavirus nonstructural proteins, in addition to those explicitly provided herein, may be utilized in the present invention, and are therefore deemed to fall within the scope of the phrase "alphavirus nonstructural proteins." For example, due to the degeneracy of the genetic code, more than one codon may code for a given amino acid. Therefore, a wide variety of nucleic acid sequences which encode alphavirus nonstructural proteins may be generated. Furthermore, amino acid substitutions, additions, or deletions at any of numerous positions may still provide functional or biologically active nonstructural proteins.

Methods for making such derivatives are readily accomplished by one of ordinary skill in the art given the disclosure provided herein.

Alphaviruses express four nonstructural proteins, designated NSP1, NSP2, NSP3, and NSP4. Experiments examining growth characteristics of alphavirus nonstructural protein cleavage mutants have indicated that the polyproteins are involved in the synthesis of the genomic negative stranded RNA, while the individual monomeric proteins catalyze the synthesis of the genomic and subgenomic positive stranded RNA species (Shirako and Strauss, 1994. *J. Virol.* 68:1874-1885). In particular embodiments, retroviral or lentiviral transfer vectors may comprise polynucleotide sequences encoding one, two, three, or all four of the nonstructural proteins of an alphavirus. In preferred embodiments, retroviral or lentiviral transfer vectors comprise a polynucleotide-of-interest operably linked to a 26S promoter.

NSP1 can play a role in the initiation of (or continuation of) minus-strand RNA synthesis and capping of the 5' terminus of genomic and subgenomic alphavirus RNAs during transcription. NSP1 has both methyltransferase (Mi and Stollar, 1991. *Vir.* 184:423-427) and guanyltransferase activity (Strauss and Strauss, 1994. *Microbiol. Rev.* 58(3):491-562). NSP1 also modulates the proteinase activity of NSP2, as polyproteins containing NSP1 inefficiently cleave between NSP2 and NSP3 (de Groot *et al.*, 1990. *EMBO J.* 9:2631-2638).

NSP2 is a multifunctional protein, involved in the replication of the viral RNA and processing of the nonstructural polyprotein. The N-terminal domain of the protein (spanning about the first 460 amino acids) can act as a helicase, which is active in duplex unwinding during RNA replication and transcription. Synthesis of RNA operably linked to the 26S subgenomic promoter and encoding the gene(s) of interest, requires functional NSP2. A number of mutations in the proteinase domain of NSP2 have been found to render the synthesis of 26S RNA temperature sensitive (Hahn *et al.* 1989. *J. Virol.* 63:3142-3150; Sawicki and Sawicki. 1985. *Virology* 144:20-34; Sawicki and Sawicki. 1993. *J. Virol.* 67:3605-3610).

NSP3 contains two distinct domains. The N-terminal domain ranges in length from 322 to 329 residues in different alphaviruses and exhibits a minimum of 51% amino acid sequence identity among any two alphaviruses. The C-terminal

domain, however, is not conserved among known alphaviruses in length or in sequence, and multiple changes are tolerated (Li *et al.*, 1990. *Virology*, 179:416-427). NSP3 play a role in minus-strand viral RNA synthesis.

NSP4 is the virus-encoded RNA polymerase and contains the GDD motif characteristic of such enzymes (Karner and Argos, 1984. *Nucleic Acids Res.* 12:7269-7282). Thus, NSP4 plays a role in alphavirus RNA replication, positive strand genomic RNA synthesis from the 5 'CSE, and abundant expression of 26S subgenomic RNA from the 26S promoter. Additionally, NSP4 can be metabolically unstable, through degradation by the N-end rule pathway (Gonda *et al.*, 1989. *J. Biol. Chem.* 264:16700-16712). However, NSP4 is stabilized through its association with replication complexes which conceal degradation signals. Further stabilization of the enzyme may be achieved by altering the amino terminal residue of NSP4. Stabilizing amino terminal residues include methionine, alanine, and tyrosine.

In particular embodiments, retroviral or lentiviral transfer vectors may comprise polynucleotide sequences encoding one, two, three, or more junction region promoters (26S promoters). In preferred embodiments, retroviral or lentiviral transfer vectors comprise a 26S subgenomic promoter operably linked to one or more heterologous polynucleotide(s)-of-interest (*e.g.,* gene encoding a therapeutic polypeptide).

The alphavirus viral junction region normally controls transcription initiation of the subgenomic mRNA; thus, this element is also referred to as the subgenomic mRNA promoter. The following table contains an exemplary list of junctional region promoter (26S promoter) sequences for the following alphaviruses: Sindbis virus (SIN), Aura virus (AURA), Eastern equine encephalomyelitis virus (EEE), Middleburg virus (MBV), Ross River virus (RRV), Semliki Forest virus (SFV, SFV1), O'Nyong-Nyong virus (ONN, ONN*), Whataroa virus (WHA), Western equine encephalomyelitis virus (WEE), Venezuelan equine encephalomyelitis virus (VEE, VEE*) and Mayaro virus (MAY); and the Rubella virus (RUB), which is not an alphavirus but has a sequence that resembles the minimal junction region.

Table 5: Junction region promoter sequences.

SEQ ID	Virus	Strand	Sequence
34	SIN	PLUS STRAND	5' AUCUCUACGGUGGUCCUAAAUAGU 3'
35		MINUS STRAND	3' UAGAGAUGCCACCAGGAUUUAUCA 5'
36	SIN/RRV	PLUS STRAND	5' AUCUCUACGGUGGUCCUAAAUAGA 3'
37		MINUS STRAND	3' UAGAGAUGCCACCAGGAUUUAUCU 5'
38	MBV	PLUS STRAND	5' ACCUCUACGGCGGUCCUAAAUAGU 3'
39		MINUS STRAND	3' UGGAGAUGCCGCCAGGAUUUAUCA 5'
40	AURA	PLUS STRAND	5' ACCUCUACGGUGGUCCUAAAUAGA 3'
41		MINUS STRAND	3' UGGAGAUGCCACCAGGAUUUAUCU 5'
42	WHA	PLUS STRAND	5' AGCUCUACGGCGGUCCUAAAUAGU 3'
43		MINUS STRAND	3' UCGAGAUGCCGCCAGGAUUUAUCA 5'
44	WEE	PLUS STRAND	5' CCCUCUACGGCUGACCUAAAUAGG 3'
45		MINUS STRAND	3' GGGAGAUGCCGACUGGAUUUAUCC 5'
46	EEE	PLUS STRAND	5' CCCUCUACGGCUGACCUAAAUAGG 3'
47		MINUS STRAND	3' GGGAGAUGCCGACUGGAUUUAUCC 5'
48	VEE	PLUS STRAND	5' CUCUCUACGGCUAACCUAAAUGGA 3'
49		MINUS STRAND	3' GAGAGAUGCCGAUUGGAUUUACCU 5'
50	VEE*	PLUS STRAND	5' CCCUCUACGGCUAACCUAAAUGGU 3'
51		MINUS STRAND	3' GGGAGAUGCCGAUUGGAUUUACCA 5'
52	ONN	PLUS STRAND	5' CCUUGUACGGCGGACCUAAAUAGG 3'
53		MINUS STRAND	3' GGAACAUGCCGCCUGGAUUUAUCC 5'
54	ONN*	PLUS STRAND	5' UCUUGUACGGCGGACCUAAAUAGG 3'
55		MINUS STRAND	3' AGAACAUGCCGCCUGGAUUUAUCC 5'
56	SFV	PLUS STRAND	5' ACCUCUACGGCGGUCCUAAAUUGG 3'
57		MINUS STRAND	3' UGGAGAUGCCGCCAGGAUUUAACC 5'
58	SFV1	PLUS STRAND	5' ACCUCUACGGCGGUCCUAGAUUGG 3'
59		MINUS STRAND	3' UGGAGAUGCCGCCAGGAUCUAACC 5'
60	RRV	PLUS STRAND	5' ACCUCUACGGCGGUCCUAAAUAGA 3'
61		MINUS STRAND	3' UGGAGAUGCCGCCAGGAUUUAUCU 5'

62	MAY	PLUS STRAND	5' AUCUGUACGGCGGUCCUAAAUAGG 3'
63		MINUS STRAND	3' UAGACAUGCCGCCAGGAUUUAUCC 5'

It would be appreciated that any of the foregoing 26S promoter sequences is suitable for use in particular embodiments of the invention. Further, it would be within the skill of one in the art to modify the foregoing 26S promoter sequences or determine optimal 26S promoter sequences based on the consensus of the promoter sequences shown in Table 5. Accordingly, in certain embodiments, vectors of the invention comprise variants or optimal 26S promoter sequences that increase RNA levels compared to unmodified 26S promoter sequences.

In particular embodiments, vectors of the invention that comprise a 26S promoter sequence also include an alphavirus RNA polymerase recognition sequence (also termed "alphavirus replicase recognition sequence", "3' terminal CSE," or "3' cis replication sequence"). Briefly, the alphavirus RNA polymerase recognition sequence, which is located at the 3' end region of positive stranded genomic RNA, provides a recognition site at which the virus begins replication by synthesis of the negative strand. A wide variety of sequences may be utilized as an alphavirus RNA polymerase recognition sequence. For example, within one embodiment, virus vector constructs in which the polymerase recognition is truncated to the smallest region that can still function as a recognition sequence can be utilized.

b. Narnavirus Elements

20S and 23S RNAs are positive strand RNA viruses found in *Saccharomyces cerevisiae* and belong to the genus Narnavirus. The 20S and 23S RNA genomes are small (2514 and 2891 nucleotides, respectively), and each RNA encodes only a single protein: a 91-kDa protein (p91) and a 104-kDa (p104), respectively. Both proteins contain four amino acid motifs well conserved among RNA-dependent RNA polymerases. The cis-acting signals for replication in 23S RNA and 20S have been determined (Fujimura and Esteban. *J Biol Chem*. 282(26):19011-9 (2007), Esteban *et al.*, *J Biol Chem*. 280(40):33725-34 (2005), and Esteban and Fujimura. *PNAS*. 100(5):2568-73 (2003)).

The positive strand genomic RNA, *i.e.,* the sense strand, serves as template for the RDRP (*i.e.,* p91 or p104)-catalyzed synthesis of a full-length complementary negative strand, *i.e.,* the non-sense or antisense strand. Synthesis of a complementary negative strand is catalyzed by RDRP binding to the 3' terminal cis-acting sequence of the positive strand genomic RNA. The negative strand, in turn, serves as template for the synthesis of additional positive strand genomic RNA catalyzed by RDRP binding to the 5' terminal cis-acting sequence. Thus, the 5' end cis-acting sequence is functional only after it transcribed into the negative strand, *i.e.,* the complement of the positive strand genomic RNA (*i.e.,* the 5' end cis-acting sequence is functional when it is the 3' end of the genomic negative stranded complement).

20S and 23S RNA genomes share five-nucleotide inverted repeats at the 5′ and 3′ termini (5′-GGGGC... GCCCC-OH). *Id.* The RNA genomes contain a bipartite cis-signal in the 3′-region that consists of the cluster of the terminal four Cs and a mismatched pair of purines present in a stem structure adjacent to the 3′-end. Although the 3′-terminal and penultimate Cs are dispensable for launching 20S or 23S RNA virus, the generated viruses recover these Cs. This indicates that the virus has an efficient 3′-terminal repair mechanism(s). In addition, the bipartite 3′-cis-signal for replication is also essential for formation of ribonucleoprotein complexes *in vivo* with its RNA-dependent RNA polymerase, *i.e.,* p91 or p104. Thus, transcription of both the 20S and 23S negative genomic RNAs are catalyzed by p91 or p104, respectively, binding to the 3′ cis-acting sequences.

Briefly, the narnavirus RNA 3′ cis-acting sequence that binds the RDRP is located at the 3' end region of positive stranded genomic RNA and provides a recognition site at which the virus begins replication by synthesis of the negative strand. Once the negative genomic RNA has been transcribed, RDRP binds to the 5′ cis-acting sequence (5′-GGGGC... GCCCC-OH; at the 3′ end of the negative genomic RNA) and transcribes the positive strand RNA encoding the RDRP. In particular embodiments, retroviral or lentiviral transfer vectors may comprise polynucleotide sequences encoding the RDRP of the 23S RNA narnavirus, p104 or an active fragment or variant thereof. In preferred embodiments, retroviral or lentiviral transfer vectors comprise a polynucleotide-of-interest operably linked to the 5′ cis-acting sequence that binds p104.

In other preferred embodiments, retroviral or lentiviral transfer vectors comprise a 5′ cis-acting sequence that binds p104 operably linked to one or more heterologous gene(s)-of-interest. In certain embodiments, vectors of the invention that comprise a 5′ cis-acting sequence that binds p104 also include a 3′ cis-acting sequence that binds p104.

In particular embodiments, retroviral or lentiviral transfer vectors may comprise polynucleotide sequences encoding the RDRP of the 20S RNA narnavirus, p91 or an active fragment or variant thereof. In preferred embodiments, retroviral or lentiviral transfer vectors comprise a polynucleotide-of-interest operably linked to the 5′ cis-acting sequence that binds p91. In other preferred embodiments, retroviral or lentiviral transfer vectors comprise a 5′ cis-acting sequence that binds p91 operably linked to one or more heterologous gene(s)-of-interest. In certain embodiments, vectors of the invention that comprise a 5′ cis-acting sequence that binds p91 also include a 3′ cis-acting sequence that binds p91.

F. Vector Systems for Use in Practicing the Invention

In various embodiments, vectors of the invention may be used to achieve ubiquitous, conditional, reversible, or tissue-specific expression or one or more polypeptides, *e.g.,* therapeutic polypeptide or polypeptide-of-interest, in cells, tissues, or organisms. Vectors, including lentiviral vectors and transfer plasmids of the invention may comprise a polynucleotide-of-interest, including, for example, polynucleotide sequences that express a polypeptide of interest or a therapeutic polypeptide. In particular embodiments, a therapeutic polypeptide is provided to a patient in whom said polypeptide is expressed at a reduced level or in a mutated, less functional form. In one non-limiting example, a polynucleotide-of-interest and its expressed polypeptide includes, but is not limited to, any gene, for which a deficiency results in, or contributes to a disease, disorder, or condition or another disease, *e.g.,* ATP-binding cassette, sub-family D (ALD), member 1 (ABCD1) in adrenoleukodystrophy and adrenomyeloneuropathy; or β-globin in hemoglobinopathies such as sickle cell disease and β-thalassemia.

Transfer vectors, including lentiviral vectors, optionally in combination with polypeptide compositions of the invention can be used in gene therapy, including for the treatment of various diseases, disorders, and conditions. The invention also includes host cells comprising, *e.g.,* transfected with, the vectors of the invention. In one embodiment, the host cell is an embryonic stem cell, a somatic stem cell, or a progenitor cell.

1. Single Vector System

In various embodiments, the present invention contemplates, in part, engineering effective therapeutic viral vector systems by using lentiviral genomes as the transfer vector backbone and heterologous RNA dependent RNA polymerase (RDRP) systems, *e.g.,* of plus-strand or minus-strand ssRNA viruses, to control and amplify transcription of one or more polynucleotides-of-interest.

In various embodiments, the present invention contemplates several single vector strategies to amplify and control gene expression. In one embodiment, the DNA sense strand of an integrated transfer vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of a plus-strand ssRNA virus RDRP promoter, *e.g.,* 5′ cis-acting sequence and/or subgenomic promoter, operably linked to the sense strand of a polynucleotide-of-interest and a plus strand sequence of a 3′ cis-acting sequence. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide-of-interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, this arrangement allows for controlled amplification of gene expression. When, present, an RDRP binds to the 3′ cis-acting sequence at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, the RDRP binds to the subgenomic promoter and/or 5′ cis-acting sequence and amplifies sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are then translated by the ribosome. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible

promoter and the subgenomic promoter, albeit using different RNA templates. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In various related embodiments, the present invention contemplates several single vector strategies to amplify and control gene expression. In one embodiment, the DNA sense strand of an integrated transfer vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of a plus-strand ssRNA virus RDRP promoter, *e.g.,* a subgenomic promoter, operably linked to the sense strand of a polynucleotide-of-interest and a plus strand sequence of a 3′ cis-acting sequence. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide-of-interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, this arrangement allows for controlled amplification of gene expression. When, present, an RDRP binds to the 3′ cis-acting sequence at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, the RDRP binds to the subgenomic promoter and amplifies sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are then translated by the ribosome. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the subgenomic promoter, albeit using different RNA templates. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In various other embodiments, the present invention contemplates several single vector strategies to amplify and control gene expression. In one embodiment, the DNA sense strand of an integrated transfer vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of an alphavirus subgenomic promoter, *e.g.,* 26S subgenomic promoter, operably linked to the sense strand of a polynucleotide-of-

interest and a plus strand sequence of a CSE. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide-of-interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, this arrangement allows for controlled amplification of gene expression. When, present, one or more NSPs bind to the CSE at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, one or more NSPs bind to the alphavirus subgenomic promoter and amplify sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are then translated by the ribosome. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the subgenomic promoter, albeit using different RNA templates. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation, *e.g.,* 5′ cis-acting sequence and/or subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the polypeptide of interest and an RDRP promoter in the reverse orientation: the reverse complement sequence encoding a polypeptide of interest cannot be directly translated into the polypeptide of interest. However, the transcribed mRNA can be used as a minus strand RNA template because it comprises an RDRP promoter in an operable orientation. When present, the RDRP binds to the RDRP promoter sequence (now in the proper orientation) on the transcribed strand and amplifies translatable RNA copies using the transcribed strand as a template in the 3′ to 5′ orientation. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are now in the proper orientation and are then translated by the ribosome. Accordingly, the 3′ cis-acting sequence (CSE) is not needed to

generate a minus strand because the RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In an additional embodiment, the DNA sense strand of an integrated vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation, *e.g.,* a subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the polypeptide of interest and the subgenomic promoter in the reverse orientation: the reverse complement sequence encoding the polypeptide of interest cannot be directly translated into the polypeptide of interest. However, the transcribed mRNA can be used as a minus strand RNA template. When present, RDRP binds to the subgenomic promoter sequence, which is operable in the 3′ to 5′ orientation on the transcribed mRNA, and amplifies translatable RNA copies using the transcribed strand as a template in the 3′ to 5′ orientation. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are now in the proper orientation and are then translated by the ribosome. Accordingly, the CSE is not needed to generate a minus strand because the RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to an alphavirus subgenomic promoter in the reverse orientation, *e.g.,* 26S subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the polypeptide of interest and an alphavirus subgenomic promoter in reverse orienation: the reverse complement encoding the polypeptide of interest cannot be

directly translated into the polypeptide of interest. However, the transcribed mRNA can be used as a minus strand RNA template in the 3′ to 5′ orientation. When present, one or more NSPs bind to the alphavirus subgenomic promoter sequence on the transcribed strand and amplify translatable RNA copies using the transcribed strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are now in the proper orientation and are then translated by the ribosome. Accordingly, the CSE is not needed to generate a minus strand because the RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template.. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated vector (an integrated provirus) comprises: 1) a ubiquitous, tissue specific, or inducible promoter operably linked to the sense strand of a polynucleotide of interest and 2) the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand single strand ssRNA virus RDRP promoter in the reverse orientation, *e.g.,* 5′ cis-acting sequence or subgenomic promoter (*e.g.,* 26S subgenomic promoter). In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide of interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, in this arrangement, the mRNA comprises a reverse complement of the mRNA that encodes the polypeptide of interest and an RDRP promoter in the reverse orientation. Thus, in the presence of the RDRPs, this arrangement allows for controlled amplification of gene expression using the transcribed mRNA sequence as a minus strand RNA template in the 3′ to 5′ orientation. To avoid viral recombination at formation of RNA secondary stem loop structures between the sense and reverse complement sequences, either one or both of the sequences can be codon optimized so that the identity between the sequences is neglible. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the RDRP promoter. In this manner, gene expression of the one or more polypeptides of interest

can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In one embodiment, a vector, *e.g.,* viral vector or transfer plasmid, of the invention comprises a left (5') retroviral LTR; a promoter operably linked to one or more polynucleotides encoding an RNA dependent RNA polymerase (RDRP), an RDRP promoter (*e.g.* 5′ cis-acting sequence or subgenomic promoter) operably linked to at least one polynucleotide encoding a mammalian polypeptide, and a CSE for production of a minus strand RNA; and a right (3') retroviral LTR. In some embodiments, the RNA dependent RDRP and RDRP promoter can be selected from any plus-strand ssRNA virus. In one embodiment, the plus-strand ssRNA virus is an alphavirus or narnavirus.

In another embodiment, a vector, *e.g.,* viral vector or transfer plasmid, of the invention comprises a left (5') retroviral LTR; a promoter operably linked to one or more polynucleotides encoding an RDRP, the reverse complement sequence of a polynucleotide encoding a mammalian polypeptide operably linked to an RDRP promoter in the reverse orientation (*e.g.* 5′ cis-acting sequence or subgenomic promoter); and a right (3') retroviral LTR. In some embodiments, the RDRP and RDRP promoter can be selected from any minus-strand or plus-strand ssRNA virus. In one embodiment, the plus-strand ssRNA virus is an alphavirus or narnavirus.

In yet another embodiment, a vector, *e.g.,* viral vector or transfer plasmid, of the invention comprises a left (5') retroviral LTR; a promoter operably linked to one or more polynucleotides encoding an RDRP, a polynucleotide encoding a mammalian polypeptide, the reverse complement of a polynucleotide encoding the mammalian polypeptide operably linked to the an RDRP promoter in the reverse orientation; and a right (3') retroviral LTR, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized. In some embodiments, the RDRP and RDRP promoter can be selected from any minus-strand or plus-strand ssRNA virus. In one embodiment, the plus-strand ssRNA virus is an alphavirus or narnavirus.

In a particular embodiment, a vector of the invention comprises: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous

promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding one or more alphavirus non-structural proteins (NSP1-NSP4), an alphavirus subgenomic promoter, *e.g.,* a 26S subgenomic promoter, at least one polynucleotide encoding a mammalian therapeutic polypeptide, and a CSE for production of a minus strand RNA. In some embodiments, one, two, three, or all four NSPs are operably linked to a ubiquitous, tissue specific, or inducible promoter. The polynucleotide sequences encoding NSPs can be contiguous or they can be separated by one or more IRES sequences and or polynucleotide sequences that include self-cleaving polypeptide sites, *e.g.,* a 2A or 2A polypeptides as disclosed elsewhere herein. In some embodiments, NSPs can be expressed as a polyprotein that undergoes self-cleaving or, a polyprotein comprising individual NSP/self-cleaving 2A peptide fusion polypeptides. In a preferred embodiment, a ubiquitous, tissue specific, or inducible promoter is operably linked to an NSP2-2A-NSP4 fusion polypeptide. In another preferred embodiment, a ubiquitous, tissue specific, or inducible promoter is operably linked to NSP4.

In another particular embodiment, a vector of the invention comprises: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; and a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding one or more alphavirus non-structural proteins (NSP1-NSP4), the reverse complement of a polynucleotide encoding a mammalian therapeutic polypeptide operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation.

In yet another particular embodiment, a vector of the invention comprises: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; and a ubiquitous, tissue specific, or inducible promoter operably

linked to polynucleotides encoding one or more alphavirus non-structural proteins (NSP1-NSP4), a polynucleotide encoding a mammalian therapeutic polypeptide, the reverse complement of a polynucleotide encoding the mammalian therapeutic polypeptide operably linked ton alphavirus 26S subgenomic promoter in the reverse orientation, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized.

In a certain embodiment, a vector of the invention comprises: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding NSP4 or NSP2 and NSP4, optionally encoded as a fusion polypeptide, *e.g.,*NSP2-2A-NSP4; and another ubiquitous, tissue specific, or inducible promoter operably linked to an alphavirus 26S subgenomic promoter and at least one polynucleotide encoding a mammalian therapeutic polypeptide, *e.g.,* a globin polypeptide or variant thereof, and a CSE for production of a minus strand RNA.

In another certain embodiment, a vector of the invention comprises: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding NSP4 or NSP2 and NSP4, optionally encoded as a fusion polypeptide, *e.g.,*NSP2-2A-NSP4; and another ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding a mammalian therapeutic polypeptide, *e.g.,* a globin polypeptide or variant thereof operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation.

In yet another certain embodiment, a vector of the invention comprises: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to

polynucleotides encoding NSP4 or NSP2 and NSP4, optionally encoded as a fusion polypeptide, *e.g.,*NSP2-2A-NSP4; and another ubiquitous, tissue specific, or inducible promoter operably linked a polynucleotide encoding a mammalian therapeutic polypeptide, *e.g.,* a globin polypeptide or variant thereof, the reverse complement of a polynucleotide encoding a mammalian therapeutic polypeptide operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized.

The foregoing vectors, *e.g.,* viral vectors or transfer plasmids, may further comprise one or more of each of the following: packaging sequences, posttranscriptional recognition elements (*e.g.,* WPRE, HPRE), insulator elements, cPPT/FLAP sequences, selectable markers, a retroviral export element, *e.g.,* a lentiviral REV response element (RRE), and other lentiviral vector features as disclosed elsewhere herein. In addition, although the preceding single transfer vector systems recite an ordered list of features, one having ordinary skill in the art would appreciate that other orders of the features are possible and in some cases preferred.

2. Binary Vector System

In various embodiments, the present invention contemplates, in part, engineering effective therapeutic vector systems by using at least two vectors: at least one of the vectors encoding the RDRP either from an expression vector or a retroviral vector, *e.g.,* lentiviral vector, and at least one retroviral or lentiviral vector comprising a RDRP promoter operably linked to one or more polynucleotides-of-interest.

In various embodiments, the present invention contemplates, in part, engineering effective therapeutic vector systems by using a vector that encodes a heterologous RDRP and a lentiviral vector comprising a RDRP promoter sequence, *e.g.,* 5′ cis-acting sequence or subgenomic promoter (*e.g.*, 26S subgenomic promoter) operably linked to one or more polynucleotides-of-interest. When the second vector has integrated as provirus into a cell and the first vector is also present in the same cell, either as an expression plasmid or integrated viral vector, the RDRP controls and/or amplifies transcription of the one or more polynucleotides-of-interest. In this way

expression is tightly controlled by using the RDRP expression as a gene switch. In the absence of the RDRP, the gene expression system is "off" and in the presence of RDRP, the gene expression system is switched "on."

In various embodiments, the present invention contemplates several binary vector strategies to amplify and control gene expression. In one embodiment, a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding an RDRP or RDRP complex. In a certain embodiment, the first vector comprises a ubiquitous, tissue specific, or inducible promoter operably linked to one, two, three, or four members of an RDRP complex. The polynucleotide sequences encoding members of the RDRP complex can be contiguous or they can be separated by one or more IRES sequences and or polynucleotide sequences that include self-cleaving polypeptide sites, *e.g.,* a 2A or 2A polypeptides as disclosed elsewhere herein

In various embodiments, the present invention contemplates several binary vector strategies to amplify and control gene expression. In one embodiment, a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding an RNA dependent RNA polymerase or polymerase complex. In particular embodiments, the first vector comprises a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding one or more alphavirus nonstructural proteins (NSP1-NSP4). In a certain embodiment, the first vector comprises a ubiquitous, tissue specific, or inducible promoter operably linked to one, two, three, or four NSPs. In a certain particular embodiment, the first vector comprises a ubiquitous, tissue specific, or inducible promoter operably linked to NSP4 or NSP2 and NSP4. The polynucleotide sequences encoding NSPs can be contiguous or they can be separated by one or more IRES sequences and or polynucleotide sequences that include self-cleaving polypeptide sites, *e.g.,* a 2A or 2A polypeptides as disclosed elsewhere herein. In some embodiments, NSPs can be expressed as a polyprotein that undergoes self-cleaving or, a polyprotein comprising individual NSP/self-cleaving 2A peptide fusion polypeptides. In a preferred

embodiment, a ubiquitous, tissue specific, or inducible promoter is operably linked to NSP4 or an NSP2-2A-NSP4 fusion polypeptide.

In various embodiments, the second vector of the binary vector strategy is a lentiviral or retroviral vector comprising a RDRP promoter sequence operably linked to one or more polynucleotides-of-interest.

In one embodiment, the DNA sense strand of an integrated second vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of a plus-strand ssRNA virus RDRP promoter, *e.g.,* 5′ cis-acting sequence and/or subgenomic promoter, operably linked to the sense strand of a polynucleotide-of-interest and a plus strand sequence of a 3′ cis-acting sequence. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide-of-interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, this arrangement allows for controlled amplification of gene expression. When, present, an RDRP binds to the 3′ cis-acting sequence at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, the RDRP binds to the subgenomic promoter and/or 5′ cis-acting sequence and amplifies sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are then translated by the ribosome. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the subgenomic promoter, albeit using different RNA templates. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated second vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of a plus-strand ssRNA virus RDRP promoter, *e.g.,* a subgenomic promoter, operably linked to the sense strand of a polynucleotide-of-interest and a plus strand sequence of a 3′ cis-acting sequence. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a

sense (plus-strand) mRNA that encodes the polypeptide-of-interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, this arrangement allows for controlled amplification of gene expression. When, present, an RDRP binds to the 3′ cis-acting sequence at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, the RDRP binds to the subgenomic promoter and amplifies sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are then translated by the ribosome. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the subgenomic promoter, albeit using different RNA templates. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In various other embodiments, the DNA sense strand of an integrated second vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of an alphavirus subgenomic promoter, *e.g.,* 26S subgenomic promoter, operably linked to the sense strand of a polynucleotide-of-interest and a plus strand sequence of a CSE. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide-of-interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, this arrangement allows for controlled amplification of gene expression. When, present, one or more NSPs bind to the CSE at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, one or more NSPs bind to the alphavirus subgenomic promoter and amplify sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are then translated by the ribosome. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the subgenomic promoter, albeit using different RNA templates. In this manner, gene

expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated second vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to the a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation, *e.g.,* 5′ cis-acting sequence and/or subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the polypeptide of interest and an RDRP promoter in the reverse orientation: the reverse complement encoding the polypeptide of interest cannot be directly translated into the polypeptide of interest. However, the transcribed mRNA can be used as a minus strand RNA template in the 3′ to 5′ orientation. When present, the RDRP binds to the RDRP promoter sequence on the transcribed strand and amplifies translatableRNA copies using the transcribed strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are now in the proper orientation and are then translated by the ribosome. Accordingly, the 3′ cis-acting sequence (CSE) is not needed to generate a minus strand because the RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In an additional embodiment, the DNA sense strand of an integrated second vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation, *e.g.,* a subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the polypeptide of interest and a subgenomic promoter in the reverse orientation: the reverse complement encoding the polypeptide of interest cannot be directly translated into the polypeptide of interest.

However, the transcribed mRNA can be used as a minus strand RNA template in the 3′ to 5′ orientation. When present, RDRP binds to the subgenomic promoter sequence on the transcribed strand and amplifies translatable RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are now in the proper orientation and are then translated by the ribosome. Accordingly, the CSE is not needed to generate a minus strand because RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated second vector (an integrated provirus) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to the an alphavirus subgenomic promoter in the reverse orientation, *e.g.,* 26S subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the polypeptide of interest and an alphavirus subgenomic promoter in the reverse orientation: the reverse complement encoding the polypeptide of interest cannot be directly translated into the polypeptide of interest. However, the transcribed mRNA can be used as a minus strand RNA template in the 3′ to 5′ orientation. When present, one or more NSPs bind to the alphavirus subgenomic promoter sequence on the transcribed strand and amplify translatable RNA copies using the transcribed strand as a template. These amplified sense (plus-strand) RNA copies of the polypeptide of interest are now in the proper orientation and are then translated by the ribosome. Accordingly, the CSE is not needed to generate a minus strand because RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In another embodiment, the DNA sense strand of an integrated second vector (an integrated provirus) comprises: 1) a ubiquitous, tissue specific, or inducible

promoter operably linked to the sense strand of a polynucleotide of interest and 2) the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand single strand ssRNA virus RDRP promoter in the reverse orientation, *e.g.,* 5′ cis-acting sequence or subgenomic promoter (*e.g.,* 26S subgenomic promoter). In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA that encodes the polypeptide of interest and that can be translated by the ribosome to produce the polypeptide of interest. In addition, in this arrangement, the mRNA comprises a reverse complement of the mRNA that encodes the polypeptide of interest and an RDRP promoter in the reverse orientation. Thus, in the presence of the RDRPs, this arrangement allows for controlled amplification of gene expression using the transcribed mRNA sequence as a minus strand RNA template in the 3′ to 5′ orientation. To avoid viral recombination at formation of RNA secondary stem loop structures between the sense and reverse complement sequences, either one or both of the sequences can be codon optimized so that the identity between the sequences is neglible. Accordingly, this design allows for expression of a translatable RNA from both the ubiquitous, tissue specific, or inducible promoter and the RDRP promoter. In this manner, gene expression of the one or more polypeptides of interest can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In one embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding an RDRP or RDRP complex; and a second vector comprising a left (5') retroviral LTR; a promoter operably linked to an RDRP promoter operably linked to at least one polynucleotide encoding a mammalian polypeptide, and a CSE for production of a minus strand RNA; and a right (3') retroviral LTR. In some embodiments, the RNA dependent RDRP and RDRP promoter can be selected from any plus-strand ssRNA virus. In one embodiment, the plus-strand ssRNA virus is an alphavirus or narnavirus.

In another embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector)

comprising a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding an RDRP or RDRP complex; and a second vector comprising a left (5') retroviral LTR; a promoter operably linked to the reverse complement sequence of a polynucleotide encoding a mammalian polypeptide operably linked to an RDRP promoter in the reverse orientation; and a right (3') retroviral LTR. In some embodiments, the RDRP and RDRP promoter can be selected from any minus-strand or plus-strand ssRNA virus. In one embodiment, the plus-strand ssRNA virus is an alphavirus or narnavirus.

In yet another embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding an RDRP or RDRP complex; and a second vector comprising a left (5') retroviral LTR; a promoter operably linked to a polynucleotide encoding a mammalian polypeptide, the reverse complement of a polynucleotide encoding the mammalian polypeptide operably linked to an RDRP promoter in the reverse orientation; and a right (3') retroviral LTR, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized. In some embodiments, the RDRP and RDRP promoter can be selected from any minus-strand or plus-strand ssRNA virus. In one embodiment, the plus-strand ssRNA virus is an alphavirus or narnavirus.

In a particular embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding one or more alphavirus non-structural proteins (NSP1-NSP4); and a second vector comprising: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to an alphavirus 26S subgenomic promoter, at least one polynucleotide encoding a mammalian therapeutic polypeptide, and a CSE for production of a minus strand RNA.

In another particular embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding one or more alphavirus non-structural proteins (NSP1-NSP4); and a second vector comprising: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; and a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding a mammalian therapeutic polypeptide operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation.

In yet another particular embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding one or more alphavirus non-structural proteins (NSP1-NSP4); and a second vector comprising: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; and a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding a mammalian therapeutic polypeptide, the reverse complement of a polynucleotide encoding the mammalian therapeutic polypeptide operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized.

In a certain embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding NSP4 or NSP2 and NSP4, optionally encoded as a fusion polypeptide, *e.g.,*NSP2-2A-NSP4; and a second vector comprising: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion

and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to an alphavirus 26S subgenomic promoter and at least one polynucleotide encoding a mammalian therapeutic polypeptide, *e.g.,* a globin polypeptide or variant thereof, and a CSE for production of a minus strand RNA.

In another certain embodiment a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding NSP4 or NSP2 and NSP4, optionally encoded as a fusion polypeptide, *e.g.,*NSP2-2A-NSP4; and a second vector comprising: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding a mammalian therapeutic polypeptide, *e.g.,* a globin polypeptide or variant thereof operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation.

In yet another certain embodiment, a binary vector system comprises a first vector, either an expression vector or viral vector (*e.g.,* retroviral or lentiviral vector) comprising a ubiquitous, tissue specific, or inducible promoter operably linked to polynucleotides encoding NSP4 or NSP2 and NSP4, optionally encoded as a fusion polypeptide, *e.g.,*NSP2-2A-NSP4; and a second vector comprising: a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence; a ubiquitous, tissue specific, or inducible promoter operably linked to a polynucleotide encoding a mammalian therapeutic polypeptide, *e.g.,* a globin polypeptide or variant thereof, the reverse complement of a polynucleotide encoding a mammalian therapeutic polypeptide operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized.

In particular embodiments, wherein the first and/or second vectors are viral vectors, *e.g.,* retroviral or lentiviral transfer vectors, the vectors may further comprise one or more of each of the following: packaging sequences, posttranscriptional recognition elements (*e.g.,* WPRE, HPRE), insulator elements, cPPT/FLAP sequences, selectable markers, a retroviral export element, *e.g.,* a lentiviral REV response element (RRE), and other lentiviral vector features as disclosed elsewhere herein. In addition, although the preceding transfer vectors recite an ordered list of features, one having ordinary skill in the art would appreciate that other orders of the features are possible and in some cases preferred.

3. Polypeptide Composition and Vector System

In various embodiments, the present invention contemplates, in part, engineering effective therapeutic vector systems that are used in combination with a polypeptide composition.

In various embodiments, the present invention contemplates, in part, engineering effective therapeutic vector systems comprising: a composition comprising an RDRP or RDRP complex and a lentiviral vector comprising a RDRP promoter sequence operably linked to one or more polynucleotides-of-interest. When the composition and integrated viral vector are present in the same cell, the RDRP or RDRP complex controls and/or amplifies transcription of the one or more polynucleotides-of-interest. In this way expression is tightly controlled by using the RDRP polypeptides as a gene switch. In the absence of the RDRP, the gene expression system is "off" and in the presence of RDRP, the gene expression system is switched "on."

Accordingly, in this aspect of the technology the vector component can be identical to the second vector disclosed in the binary vector system, see *supra*. Once the vector has integrated into a cell as a provirus, the cell can be contacted with a composition comprising an RDRP or RDRP complex. In various embodiments, the RDRP, polypeptides of the RDRP complex, and/or one or more RDRP polypeptides are fusion polypeptides comprising one or more cell permeable peptides (CPP), as

discussed elsewhere herein, to facilitate entry of the polypeptides into the transduced cell.

In various embodiments, a composition comprising an RDRP or RDRP complex contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of a plus-strand ssRNA virus RDRP promoter operably linked to the sense strand of a polynucleotide of interest and the plus strand of a 3′ cis-acting element. In various embodiments, a composition comprising an RDRP or RDRP complex contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of an alphavirus 26S subgenomic promoter operably linked to the sense strand of a polynucleotide of interest and a CSE. In particular embodiments, a composition comprising one or more alphavirus nonstructural proteins (NSP1-NSP4) contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of an alphavirus 26S subgenomic promoter operably linked to the sense strand of a polynucleotide of interest and a CSE. In certain embodiments, a composition comprising NSP4 or NSP2 and NSP4 contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the plus strand sequence of an alphavirus 26S subgenomic promoter operably linked to the sense strand of a polynucleotide of interest and a CSE.

In various embodiments, a composition comprising an RNA dependent RNA polymerase or polymerase complex vector strategies contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation. In various embodiments, a composition comprising an RDRP or RDRP complex vector strategies contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation. In particular embodiments, a composition comprising one or more alphavirus nonstructural proteins

(NSP1-NSP4) contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation. In certain embodiments, a composition comprising NSP4 or NSP2 and NSP4 contacts a cell transduced with a vector comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide of interest, which is operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation.

In various embodiments, a composition comprising an RDRP or RDRP complex vector strategies contacts a cell transduced with a vector comprising: 1) a ubiquitous, tissue specific, or inducible promoter operably linked to the sense strand of a polynucleotide of interest and 2) the reverse complement sequence of a polynucleotide of interest, which is operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation. In various embodiments, a composition comprising an RDRP or RDRP complex contacts a cell transduced with a vector comprising: 1) a ubiquitous, tissue specific, or inducible promoter operably linked to the sense strand of a polynucleotide of interest and 2) the reverse complement sequence of a polynucleotide of interest, which is operably linked to the an alphavirus 26S subgenomic promoter in the reverse orientation. In particular embodiments, a composition comprising one or more alphavirus nonstructural proteins (NSP1-NSP4) contacts a cell transduced with a vector comprising: 1) a ubiquitous, tissue specific, or inducible promoter operably linked to the sense strand of a polynucleotide of interest and 2) the reverse complement sequence of a polynucleotide of interest, which is operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation. In certain embodiments, a composition comprising NSP4 or NSP2 and NSP4 contacts a cell transduced with a vector comprising: 1) a ubiquitous, tissue specific, or inducible promoter operably linked to the sense strand of a polynucleotide of interest and 2) the reverse complement sequence of a polynucleotide of interest, which is operably linked to an alphavirus 26S subgenomic promoter in the reverse orientation.

In one embodiment, a vector comprises a left (5') retroviral LTR; a promoter operably linked to an RDRP promoter operably linked to at least one

polynucleotide encoding a mammalian polypeptide, and a CSE for production of a minus strand RNA; and a right (3') retroviral LTR. In another embodiment, a vector comprises a left (5') retroviral LTR; a promoter operably linked to the reverse complement sequence of a polynucleotide encoding a mammalian polypeptide operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation; and a right (3') retroviral LTR. In yet another embodiment, a vector comprises a left (5') retroviral LTR; a promoter operably linked to a polynucleotide encoding a mammalian polypeptide, the reverse complement of a polynucleotide encoding the mammalian polypeptide operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation; and a right (3') retroviral LTR, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized.

In one embodiment, a vector comprises a left (5') retroviral LTR; a promoter operably linked to an RDRP promoter, *e.g.,* an alphavirus 26S subgenomic promoter, operably linked to at least one polynucleotide encoding a mammalian polypeptide, and a CSE for production of a minus strand RNA; and a right (3') retroviral LTR. In another embodiment, a vector comprises a left (5') retroviral LTR; a promoter operably linked to the reverse complement sequence of a polynucleotide encoding a mammalian polypeptide operably linked to an RDRP promoter in the reverse orientation, *e.g.,* an alphavirus 26S subgenomic promoter; and a right (3') retroviral LTR. In yet another embodiment, a vector comprises a left (5') retroviral LTR; a promoter operably linked to a polynucleotide encoding a mammalian polypeptide, the reverse complement of a polynucleotide encoding the mammalian polypeptide operably linked to an RDRP in the reverse orientation, *e.g.,* an alphavirus 26S subgenomic promoter; and a right (3') retroviral LTR, wherein one or both polynucleotide sequences of the mammalian therapeutic polypeptide are codon optimized.

In some embodiments, the vectors comprise a modified left (5') retroviral LTR comprising a deletion in U3 and a heterologous promoter, *e.g.,* CMV, RSV, MoMLV and/or a modified right (3') retroviral LTR comprising a U3 deletion and/or U5 deletion and a heterologous polyA tail addition sequence. In certain embodiments,

the transfer vectors comprise a mammalian therapeutic polypeptide, such as, for example, a globin polypeptide or variant thereof.

In particular embodiments, wherein the vectors, *e.g.,* retroviral or lentiviral transfer vectors may further comprise one or more of each of the following: packaging sequences, posttranscriptional recognition elements (*e.g.,* WPRE, HPRE), insulator elements, cPPT/FLAP sequences, selectable markers, and other lentiviral vector features as disclosed elsewhere herein. In addition, although the preceding transfer vectors recite an ordered list of features, one having ordinary skill in the art would appreciate that other orders of the features are possible and in some cases preferred.

G. Compositions and Formulations

The compositions of the invention may comprise one or more polypeptides, polynucleotides, vectors comprising same, *etc.*, as described herein, formulated in pharmaceutically-acceptable or physiologically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy. It will also be understood that, if desired, the compositions of the invention may be administered in combination with other agents as well, such as, *e.g.,* other proteins, polypeptides, small molecules or various pharmaceutically-active agents. There is virtually no limit to other components that may also be included in the compositions, provided that the additional agents do not adversely affect the ability of the composition to deliver the intended gene therapy.

In the pharmaceutical compositions of the invention, 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, including *e.g.,* oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

In certain applications, the compositions disclosed herein may be delivered via oral administration to a subject. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be

enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

In certain circumstances it will be desirable to deliver the compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described, for example, in U.S. Pat. No. 5,543,158; U.S. Pat. No. 5,641,515 and U.S. Pat. No. 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U.S. Pat. No. 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form should be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and should be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.,* glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion (see, *e.g.,* Remington's Pharmaceutical Sciences, 15th Edition, pp. 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

Sterile injectable solutions can be prepared by incorporating the active compounds in the required amount in the appropriate solvent with the various other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine,

trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

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 pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

In certain embodiments, the compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, polynucleotides, and peptide compositions directly to the lungs via nasal aerosol sprays has been described *e.g.,* in U.S. Pat. No. 5,756,353 and U.S. Pat. No. 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-glycerol compounds (U.S. Pat. No. 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroetheylene support matrix is described in U.S. Pat. No. 5,780,045 (specifically incorporated herein by reference in its entirety).

In certain embodiments, the delivery may occur by use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, optionally mixing

with CPP polypeptides, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, a nanoparticle or the like. The formulation and use of such delivery vehicles can be carried out using known and conventional techniques. The formulations and compositions of the invention may comprise one or more repressors and/or activators comprised of a combination of any number of polypeptides, polynucleotides, and small molecules, as described herein, formulated in pharmaceutically-acceptable or physiologically-acceptable solutions (*e.g.,* culture medium) for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy. It will also be understood that, if desired, the compositions of the invention may be administered in combination with other agents as well, such as, *e.g.,* cells, other proteins or polypeptides or various pharmaceutically-active agents.

In a particular embodiment, a formulation or composition according to the present invention comprises a cell contacted with a combination of any number of polypeptides, polynucleotides, and small molecules, as described herein.

In certain aspects, the present invention provides formulations or compositions suitable for the delivery of viral vector systems (*i.e.*, viral-mediated transduction) including, but not limited to, retroviral (*e.g.,* lentiviral) vectors.

Exemplary formulations for *ex vivo* delivery may also include the use of various transfection agents known in the art, such as calcium phosphate, electoporation, heat shock and various liposome formulations (*i.e.,* lipid-mediated transfection). Liposomes, as described in greater detail below, are lipid bilayers entrapping a fraction of aqueous fluid. DNA spontaneously associates to the external surface of cationic liposomes (by virtue of its charge) and these liposomes will interact with the cell membrane.

In certain aspects, the present invention provides pharmaceutically acceptable compositions which comprise a therapeutically-effective amount of one or more polynucleotides or polypeptides, as described herein, formulated together with

one or more pharmaceutically acceptable carriers (additives) and/or diluents (*e.g.,* pharmaceutically acceptable cell culture medium).

Particular embodiments of the invention may comprise other formulations, such as those that are well known in the pharmaceutical art, and are described, for example, in Remington: The Science and Practice of Pharmacy, 20th Edition. Baltimore, MD: Lippincott Williams & Wilkins, 2000.

H. Methods of Delivery

In one embodiment, cells are contacted with a composition comprising an RDRP polymerase, polypeptides of an RDRP complex, or one or more NSP polypeptides (NSP1-NSP4), or polynucleotides encoding the same. In another embodiment, cells are contacted with a fusion polypeptide comprising one or more cell permeable peptides (CPP) fused to an RDRP, polypeptides of an RDRP complex, or one or more NSP polypeptides. It is contemplated that the cells of the invention may be contacted *in vitro*, *ex vivo*, or *in vivo* with compositions of the present invention.

Once formulated, the compositions of the invention can be administered (as proteins/polypeptides, or in the context of expression vectors, including viral vectors, for gene therapy) directly to the subject or delivered *ex vivo*, to cells derived from the subject (*e.g.,* as in *ex vivo* gene therapy). Direct *in vivo* delivery of the compositions will generally be accomplished by parenteral injection, *e.g.,* subcutaneously, intraperitoneally, intravenously or intramuscularly, myocardial, intratumoral, peritumoral, or to the interstitial space of a tissue. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment can be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in, for example, International PCT Publication No. WO 93/14778. Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, transduction, protoplast fusion, electroporation, encapsulation of the polynucleotide(s)

in liposomes, direct microinjection of the DNA into nuclei, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres.

Illustrative, but non-limiting methods of nucleic acid and polypeptide delivery are further discussed below.

In certain embodiments, it will be preferred to deliver one or more polynucleotides of the invention to a cell using a viral vector, *e.g.,* a retroviral vector or an expression vector, or other *in vivo* polynucleotide delivery technique. In a preferred embodiment, the viral vector is a lentiviral vector. This may be achieved using any of a variety or well-known approaches.

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA that stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. Most retrovirus will only infect actively dividing cells; thus, limiting their utility. However, another class of retrovirus, the lentivirus, is able to infect dividing as well as non-dividing cells; thus making lentivirus the vector of choice in virally mediated transgenesis.

Exemplary lentiviral vectors include, but are not limited to HIV-1, HIV-2, VMV, CAEV, EIAV, FIV, BIV, and SIV. HIV-1 has long been known to form pseudotypes by the incorporation of heterologous glycoproteins (GPs) through phenotypic mixing, and thus, broadening the tropism of the virus. Exemplary pseudotyping glycoproteins include, but are not limited to glycoproteins derived from the following viruses: Ebola, GALV, JSRV, LCMV, Marburg, Mokola, Rabies, RD114, RRV, SeV F, and VSV (Cronin *et al.*, 2006). Recent advances have provided episomal forms of retroviral vectors based on lentiviruses. The nonintegrating lentiviral vectors retain the high transduction efficiency and broad tropism of conventional lentiviruses but avoid the potential problems associated with the nonspecific integration of a transgene. In this respect they are particularly useful from a safety standpoint, and in certain embodiments, are preferred.

In another embodiment, cells are contacted with a fusion polypeptide comprising an RDRP, polypeptides of an RDRP complex, or one or more NSP

polypeptides and one or more cell permeable peptides (CPP) are incubated in growth medium with cells that will be used to provide a cell-based gene therapy. The incubation step can be repeated one, two, three, four, or five or more times in order to provide a continuous supply of the polypeptides to the cell.

In one embodiment, a polynucleotide may be administered directly to a cell via microinjection. Dubensky *et al.*, (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty & Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that polynucleotides encoding an RDRP, polypeptides of an RDRP complex, or one or more NSP polypeptides may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

In another embodiment, polynucleotides are administered to cells via electroporation.

In related embodiments, liposomes act as gene and or polypeptide delivery vehicles and are described in U.S. Patent No. 5,422,120; WO 95/13796; WO 94/23697; WO 91/14445; and EP 0524968. Additional approaches are described in Philip, *Mol. Cell Biol.* 14:2411 (1994), and in Woffendin, *Proc. Natl. Acad. Sci.* (1994) 91:11581-11585. The liposome fuses with the plasma membrane, thereby releasing the compound into the cytosol. Alternatively, the liposome is phagocytosed or taken up by the cell in a transport vesicle. Once in the endosome or phagosome, the liposome is

either degraded or it fuses with the membrane of the transport vesicle and releases its contents.

For use with the methods and compositions disclosed herein, liposomes typically comprise a polypeptide or fusion polypeptide as disclosed herein, a lipid component, *e.g.,* a neutral and/or cationic lipid, and optionally include a receptor-recognition molecule such as an antibody that binds to a predetermined cell surface receptor or ligand (*e.g.,* an antigen). A variety of methods are available for preparing liposomes as described in, *e.g.*; U.S. Pat. Nos.4,186,183; 4,217,344; 4,235,871; 4,261,975; 4,485,054; 4,501,728; 4,774,085; 4,837,028; 4,235,871; 4,261,975; 4,485,054; 4,501,728; 4,774,085; 4,837,028; 4,946,787; PCT Publication No. WO 91/17424; Szoka *et al.* (1980) *Ann. Rev. Biophys. Bioeng.* 9:467; Deamer *et al.* (1976) *Biochim. Biophys. Acta* 443:629-634; Fraley, *et al.* (1979) *Proc. Natl. Acad. Sci. USA* 76:3348-3352; Hope *et al.* (1985) *Biochim. Biophys. Acta* 812:55-65; Mayer *et al.* (1986) *Biochim. Biophys. Acta* 858:161-168; Williams *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:242-246; *Liposomes*, Ostro (ed.), 1983, Chapter 1); Hope *et al.* (1986) *Chem. Phys. Lip.* 40:89; Gregoriadis, *Liposome Technology* (1984) and Lasic, *Liposomes: from Physics to Applications* (1993). Suitable methods include, for example, sonication, extrusion, high pressure/homogenization, microfluidization, detergent dialysis, calcium-induced fusion of small liposome vesicles and ether-fusion methods, all of which are well known in the art.

In certain embodiments, it may be desirable to target a liposome using targeting moieties that are specific to a particular cell type, tissue, and the like. Targeting of liposomes using a variety of targeting moieties (*e.g.,* ligands, receptors, and monoclonal antibodies) has been previously described. *See*, *e.g.,* U.S. Pat. Nos. 4,957,773 and 4,603,044. Standard methods for coupling targeting agents to liposomes are used. These methods generally involve the incorporation into liposomes of lipid components, *e.g.,* phosphatidylethanolamine, which can be activated for attachment of targeting agents, or incorporation of derivatized lipophilic compounds, such as lipid derivatized bleomycin. Antibody targeted liposomes can be constructed using, for instance, liposomes which incorporate protein A. *See* Renneisen *et al.* (1990) *J. Biol.*

Chem. 265:16337-16342 and Leonetti *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:2448-2451.

I. Packaging Cell Lines of the Invention

Retroviral vector safety has been the impetus underlying the progress in packaging cell line development. The prospects of generating replication-competent retroviruses and the potential for vector mobilization continue to be the driving force for the advancement of packaging cell lines. The advent of stable packaging cell lines has provided an avenue by which to generate vector in a safe and reproducible manner for gene delivery in animal models, as well as clinical applications.

However, existing packaging cell lines suffer from the cytotoxic expression of viral packaging genes, such as gag, pol, and env. The art has evolved to include inducible expression systems to increase packaging cell line viability, but inducible cell lines using tetracycline, hormone (dexamethasone), or cumate gene switches are leaky and do not provide the requisite control of gene expression for optimum viral packaging. The end result is poor efficiency in generating high titer virus.

In contrast, the present invention provides improved gene switches that solve the problems associated with poorly controlled or regulatable gene expression systems. In preferred embodiments, the invention provides systems that are not leaky compared to existing inducible gene expression systems. Accordingly, the gene switches of the present invention are superior to and provide numerous advantages over existing gene expression systems in the art.

Temporal control of gene expression, as well as control of the amplitude of gene expression, are both important considerations in packaging viral particles. For example, tight regulation of cell-toxic viral packaging genes is important to preserve the safety, longevity, and fitness of packaging and producer cell lines. Illustrative embodiments of the present invention, provide state-of-the-art packaging and producer cell lines that provide for maximum safety, longevity, and health of cell lines. This results in maximum scalable production of safe virus at increased viral titers compared to existing methods in the art.

First generation lentiviral vector packaging systems provide separate packaging constructs for gag/pol and env, and typically employ a heterologous or functionally modified envelope protein for safety reasons. *See e.g.,* Miller and Buttimore, 1986. *Molec. Cell. Biol.* 6(8): 2895-2902. In second generation lentiviral vector systems, the accessory genes, vif, vpr, vpu and nef, are deleted or inactivated and the packaging functions are divided into two genomes: one genome expresses the gag and pol gene products, and the other genome expresses the env gene product (*see e.g.,* Bosselman *et al.*, 1987. *Molec. Cell. Biol.* 7(5):1797-1806; Markowitz *et al.*, 1988. *J. Virol.* 62(4):1120-1124; Danos and Mulligan, 1988. *Proc. Nat'l. Acad. Sci. (USA)* 85:6460-6464). In addition, the deletion of the 3' LTR on both packaging function constructs further reduces the ability to form functional recombinants. Third generation lentiviral vector systems include those from which the tat gene has been deleted or otherwise inactivated (*e.g.,* via mutation). Compensation for the regulation of transcription normally provided by tat can be provided by the use of a strong ubiquitous promoter, such as the human cytomegalovirus immediate early (HCMV-IE) enhancer/promoter. The gene encoding rev is preferably provided on a separate expression construct, such that a typical third generation lentiviral vector system will involve four plasmids: one each for gagpol, rev, envelope and the transfer vector.

However, large scale viral particle production is often necessary to achieve a reasonable viral titer because existing retroviral packaging cell lines are inefficient and/or expression of the viral proteins required for packaging retrovirus are toxic to the packaging cells. Another contributing factor to inefficient viral particle production or low viral titers is the absence of full length retroviral genomic transcripts produced by transfer plasmids (vectors). The present invention, contemplates, in part, safer and more efficient retroviral packaging and producer cell lines that provide solutions to these and other problems that exist in the art with regard to retroviral packaging and producer cells.

As used herein, the term "packaging" refers to the process of sequestering (or packaging) a viral genome inside a protein capsid, whereby a virion or viral particle is formed. This process is also known as encapsidation. As used herein, the term "packaging signal" or "packaging sequence" refers to sequences located within

the retroviral genome which are required for insertion of the viral RNA into the viral capsid or particle, see *e.g.,* Clever *et al.,* 1995. *J. of Virology*, Vol. 69, No. 4; pp. 2101–2109. Several retroviral vectors use the minimal packaging signal (also referred to as the psi [Ψ] sequence) needed for encapsidation of the viral genome. Thus, as used herein, the terms "packaging sequence," "packaging signal," "psi" and the symbol "Ψ," are used in reference to the non-coding sequence required for encapsidation of retroviral RNA strands during viral particle formation. As used herein, the term "packaging vector" refers to an expression vector or viral vector that lacks a packaging signal and comprises a polynucleotide encoding one, two, three, four or more gag, pol, env, tat , rev, vif, vpr, vpu, vpx, or nef genes or other retroviral genes.

As used herein, the term "packaging cell lines" is used in reference to cell lines that do not contain a packaging signal, but do stably or transiently express viral structural proteins and replication enzymes (*e.g.,* gag, pol and env) which are necessary for the correct packaging of viral particles.

The retroviral gag gene encodes structural proteins CA (p24), MA (p17) and NC (p7-11; and the pol gene encodes reverse transcriptase (RT), protease (PR) and integrase (IN). HIV-1 and HIV-2 contain accessory and other proteins involved in regulation of synthesis and processing virus RNA and other replicative functions. In addition to encoding the necessary retroviral proteins for production and assembly of core virions (*e.g.,* gag and pol proteins), packaging cell lines of the invention also encode viral envelope proteins (env) which determine the range of host cells which can ultimately be infected and transformed by recombinant retroviruses generated from the cell lines. In the case of lentiviruses, such as HIV-1, HIV-2, SIV, FIV and EIV, the env proteins include gp41 and gp120. Preferably, the viral env proteins expressed by packaging cells of the invention are encoded on a separate vector from the viral gag and pol genes, as has been previously described.

Integration-incompetent lentiviral vectors are obtained by modifying the pol gene encoding the Integrase, resulting in a mutated pol gene encoding an integrative deficient integrase, said modified pol gene being contained in the encapsidation plasmid. Such integration-incompetent lentiviral vectors have been described in patent application WO 2006/010834. Accordingly, the integrase capacity of the protein is

altered whereas the correct expression from the encapsidation plasmid of the GAG, PRO and POL proteins and/or the formation of the capsid and hence of the vector particles, as well as other steps of the viral cycle, preceding or subsequent to the integration step, such as the reverse transcription, the nuclear import, stay intact. An integrase is said defective when the integration that it should enable is altered in a way that an integration step takes place less than 1 over 1000, preferably less than 1 over 10000, when compared to a lentiviral vector containing a corresponding wild-type integrase.

In a particular embodiment of the invention, a defective integrase results from a mutation of class 1 , preferably amino acid substitutions (one- amino acid substitution) or short deletions fulfilling the requirements of the expression of a defective integrase. The mutation is carried out within the pol gene. These vectors may carry a defective integrase with the mutation D64V in the catalytic domain of the enzyme, which specifically blocks the DNA cleaving and joining reactions of the integration step. The D64V mutation decreases integration of pseudotyped HIV-1 up to 1/10,000 of wild type, but keep their ability to transduce non dividing cells, allowing efficient transgene expression.

Other mutations in the pol gene which are suitable to affect the integrase capacity of the integrase of HIV-1 are the following: H12N, H12C, H16C, H16V, S81 R, D41A, K42A, H51A, Q53C, D55V, D64E, D64V, E69A, K71A, E85A, E87A, D116N, D116I, D116A, N120G, N120I, N120E, E152G, E152A, D-35-E, K156E, K156A, E157A, K159E, K159A, K160A, R166A, D167A, E170A, H171A, K173A, K186Q, K186T, K188T, E198A, R199C, R199T, R199A, D202A, K211A, Q214L, Q216L, Q221 L, W235F, W235E, K236S, K236A, K246A, G247W, D253A, R262A, R263A and K264H.

In a particular embodiment, mutation in the pol gene is performed at either of the following positions D64, D1 16 or E152, or at several of these positions which are in the catalytic site of the protein. Any substitution at these positions is suitable, including those described above. Another proposed substitution is the replacement of the amino acids residues RRK (positions 262 to 264) by the amino acids residues AAH.

In a certain embodiment, the lentiviral vector is integration-incompetent. The lentiviral genome further comprises an origin of replication (ori), whose sequence is dependent on the nature of cells where the lentiviral genome has to be expressed. Said origin of replication may be from eukaryotic origin, preferably of mammalian origin, most preferably of human origin. It may alternatively be of viral origin, especially coming from DNA circular episomic viruses, such as SV40 or RPS. It is an advantageous embodiment of the invention to have an origin or replication inserted in the lentiviral genome of the lentiviral vector of the invention. Indeed, since the lentiviral genome does not integrate into the cell host genome (because of the defective integrase), the lentiviral genome is lost in cells undergoing frequent cell divisions. The presence of an origin of replication ensures that at least one lentiviral genome is present in each cell, even after cell division, maximizing the safety and therapeutic efficacy of the vector.

Examples of retroviral-derived env genes which can be employed in the invention include, but are not limited to type C retroviral envelope proteins, such as those from Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), and Rous Sarcoma Virus (RSV). Other viral env genes which can be used include, for example, env genes from immunodeficiency viruses (HIV-1, HIV-2, FIV, SIV and EIV), human T cell leukemia viruses (HTLV-1 and HTLV-3), and Vesicular stomatitis virus (VSV) (Protein G). When producing recombinant retroviruses of the invention (*e.g.,* recombinant lentiviruses), the wild-type retroviral (*e.g.,* lentiviral) env gene can be used, or can be substituted with any other viral env gene, such those listed above. Methods of pseudotyping recombinant viruses with envelope proteins from other viruses in this manner are well known in the art. As referred to herein, a "pseudotype envelope" is an envelope protein other than the one that naturally occurs with the retroviral core virion, which encapsidates the retroviral core virion (resulting in a phenotypically mixed virus).

Additional exemplary envelope proteins that may be used include ecotropic or amphotropic MLV envelopes, 10A1 envelope, BAEV, FeLV-B, RD114, SSAV, Ebola, Sendai, FPV (Fowl plague virus), and influenza virus envelopes.

Similarly, genes encoding envelopes from RNA viruses (*e.g.,* RNA virus families of Picornaviridae, Calciviridae, Astroviridae, Togaviridae, Flaviviridae, Coronaviridae, Paramyxoviridae, Rhabdoviridae, Filoviridae, Orthomyxoviridae, Bunyaviridae, Arenaviridae, Reoviridae, Birnaviridae, Retroviridae) as well as from the DNA viruses (families of Hepadnaviridae, Circoviridae, Parvoviridae, Papovaviridae, Adenoviridae, Herpesviridae, Poxyiridae, and Iridoviridae) may be utilized. Representative examples include , FeLV, VEE, HFVW, WDSV, SFV, Rabies, ALV, BIV, BLV, EBV, CAEV, SNV, ChTLV, STLV, MPMV, SMRV, RAV, FuSV, MH2, AEV, AMV, CT10, EIAV.

In other embodiments, envelope proteins for pseudotyping a virus of present invention include, but are not limited to any of the following virus: Influenza A such as H1N1, H1N2, H3N2 and H5N1 (bird flu), Influenza B, Influenza C virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rotavirus, any virus of the Norwalk virus group, enteric adenoviruses, parvovirus, Dengue fever virus, Monkey pox, Mononegavirales, Lyssavirus such as rabies virus, Lagos bat virus, Mokola virus, Duvenhage virus, European bat virus 1 & 2 and Australian bat virus, Ephemerovirus, Vesiculovirus, Vesicular Stomatitis Virus (VSV), Herpesviruses such as Herpes simplex virus types 1 and 2, varicella zoster, cytomegalovirus, Epstein-Bar virus (EBV), human herpesviruses (HHV), human herpesvirus type 6 and 8, Human immunodeficiency virus (HIV), papilloma virus, murine gammaherpesvirus, Arenaviruses such as Argentine hemorrhagic fever virus, Bolivian hemorrhagic fever virus, Sabia-associated hemorrhagic fever virus, Venezuelan hemorrhagic fever virus, Lassa fever virus, Machupo virus, Lymphocytic choriomeningitis virus (LCMV), Bunyaviridiae such as Crimean-Congo hemorrhagic fever virus, Hantavirus, hemorrhagic fever with renal syndrome causing virus, Rift Valley fever virus, Filoviridae (filovirus) including Ebola hemorrhagic fever and Marburg hemorrhagic fever, Flaviviridae including Kaysanur Forest disease virus, Omsk hemorrhagic fever virus, Tick-borne encephalitis causing virus and Paramyxoviridae such as Hendra virus and Nipah virus, variola major and variola minor (smallpox), alphaviruses such as Venezuelan equine encephalitis virus, eastern equine encephalitis virus, western equine encephalitis virus, SARS-associated coronavirus (SARS-CoV), West Nile virus, any encephaliltis causing virus.

In one embodiment, the invention provides packaging cells which produce recombinant retrovirus, *e.g.,* lentivirus, pseudotyped with the VSV-G glycoprotein.

In particular embodiments, stably or transiently expressing a functional tat gene, rev gene, and/or one or functional viral accessory genes, *e.g.,* vif, vpr, vpu, vpx, nef may increase the efficiency of the packaging cell. As used herein, the term "producer cell line" refers to a cell line which is capable of producing recombinant retroviral particles, comprising a packaging cell line and a transfer vector construct comprising a packaging signal.

Any suitable cell line can be employed to prepare packaging cells of the invention. Generally, the cells are mammalian cells. In a particular embodiment, the cells used to produce the packaging cell line are human cells. Suitable cell lines which can be used include, for example, CHO cells, BHK cells, MDCK cells, C3H 10T1/2 cells, FLY cells, Psi-2 cells, BOSC 23 cells, PA317 cells, WEHI cells, COS cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, W138 cells, MRC5 cells, A549 cells, HT1080 cells, 293 cells, 293T cells, B-50 cells, 3T3 cells, NIH3T3 cells, HepG2 cells, Saos-2 cells, Huh7 cells, HeLa cells, W163 cells, 211 cells, and 211A cells. In preferred embodiments, the packaging cells are 293 cells, 293T cells, or A549 cells. In another preferred embodiment, the cells are A549 cells.

Typically, the packaging vectors are included in a packaging cell, and are introduced into the cell via transfection, transduction or infection. Methods for transfection, transduction or infection are well known by those of skill in the art. A retroviral/lentiviral transfer vector of the present invention can be introduced into a packaging cell line, via transfection, transduction or infection, to generate a producer cell or cell line. The packaging vectors of the present invention can be introduced into human cells or cell lines by standard methods including, *e.g.,* calcium phosphate transfection, lipofection or electroporation. In some embodiments, the packaging vectors are introduced into the cells together with a dominant selectable marker, such as neomycin, hygromycin, puromycin, blastocidin, zeocin, thymidine kinase, DHFR, Gln synthetase or ADA, followed by selection in the presence of the appropriate drug and

isolation of clones. A selectable marker gene can be linked physically to genes encoding by the packaging vector, *e.g.,* by IRES or self cleaving viral peptides.

In one embodiment, the packaging system used to generate retroviral vectors of the invention comprises at least two packaging vectors, a first packaging vector which comprises a first nucleotide sequence comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP of a plus-strand or minus strand ssRNA virus; and a second packaging vector comprising a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to a plus-strand ssRNA virus RDRP promoter operably linked to a polynucleotide encoding one or more of the viral structural genes gag, pol, and env, and a CSE for production of a minus strand RNA. In one embodiment, the second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide encoding one or more of the viral structural genes gag, pol, and env, operably linked to a plus-strand or minus-strand ssRNA virus RDRP promoter in the reverse orientation.

In another embodiment, the packaging system used to generate retroviral vectors of the invention comprises at least two packaging vectors, a first packaging vector which comprises a first nucleotide sequence comprising a ubiquitous, tissue specific, or inducible promoter operably linked to one or more alphavirus nonstructural proteins (NSP1-NSP4). In a particular embodiment, the first packaging vector comprises a ubiquitous, tissue specific, or inducible promoter operably linked to one, two, three, or four NSPs. In a certain particular embodiment, the first packaging vector comprises a ubiquitous, tissue specific, or inducible promoter operably linked to NSP4 or NSP2 and NSP4. The polynucleotide sequences encoding NSPs can be contiguous or they can be separated by one or more IRES sequences and or polynucleotide sequences that include self-cleaving polypeptide sites, *e.g.,* a 2A or 2A polypeptides as disclosed elsewhere herein. In preferred embodiments, the first packaging vector comprises an inducible promoter, *e.g.,* a tetracyline or cumate inducible promoter.

In a particular embodiment, second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked an RDRP promoter operably linked to a polynucleotide encoding one or more of

the viral structural genes gag, pol, and env, and a CSE for production of a minus strand RNA. In one embodiment, the second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement sequence of a polynucleotide encoding one or more of the viral structural genes gag, pol, and env, operably linked to an RDRP promoter in the reverse orientation, *e.g.,* alphavirus 26S promoter.

In a particular embodiment, the packaging system used to generate retroviral vectors of the invention comprises at least three packaging vectors, a first packaging vector which comprises a first nucleotide sequence comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP; one, two, three, or four NSPs; NSP2 and NSP4; or NSP4. In preferred embodiments, the first packaging vector comprises an inducible promoter, *e.g.,* a tetracyline or cumate inducible promoter.

In a particular embodiment, the second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP promoter, *e.g.,* alphavirus 26S promoter, operably linked to a polynucleotide encoding gag, pol, or env, and a CSE for production of a minus strand RNA and the third packaging vector comprises a third polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP promoter, *e.g.,* alphavirus 26S promoter, operably linked to a polynucleotide encoding the two viral structural genes not encoded by the second packaging construct, *e.g.,* gag and pol, pol and env, or env and gag, and a CSE for production of a minus strand RNA.

In a particular embodiment, the second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding gag, pol, or env, operably linked to an RDRP promoter in the reverse orientation, *e.g.,* a plus-strand or minus-strand ssRNA virus RDRP promoter; and the third packaging vector comprises a third polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding the two viral structural genes not encoded by the second packaging construct, *e.g.,* gag and pol, pol and env, or env and gag, operably linked to an RNA dependent RNA

polymerase promoter in the reverse orientation, *e.g.,* a plus-strand or minus-strand ssRNA virus RDRP promoter.

In a certain embodiment, the packaging system used to generate retroviral vectors of the invention comprises at least four packaging vectors, a first packaging vector which comprises a first nucleotide sequence comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP; one, two, three, or four NSPs; NSP2 and NSP4; or NSP4. In preferred embodiments, the first packaging vector comprises an inducible promoter, *e.g.,* a tetracyline or cumate inducible promoter.

In a certain embodiment, the second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP promoter, *e.g.,* alphavirus 26S promoter, operably linked to a polynucleotide encoding gag and a CSE for production of a minus strand RNA; the third packaging vector comprises a third polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP promoter, *e.g.,* alphavirus 26S promoter, operably linked to a polynucleotide encoding pol, and a CSE for production of a minus strand RNA; and the fourth packaging vector comprises a fourth polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP promoter, *e.g.,* alphavirus 26S promoter, operably linked to a polynucleotide encoding env, and a CSE for production of a minus strand RNA.

In a certain embodiment, the second packaging vector comprises a second polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding gag, operably linked to an RDRP promoter in the reverse orientation, *e.g.,* a plus-strand or minus-strand ssRNA virus RDRP promoter; the third packaging vector comprises a third polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding pol, operably linked to an RDRP promoter in the reverse orientation, *e.g.,* a plus-strand or minus-strand ssRNA virus RDRP promoter; and the fourth packaging vector comprises a fourth polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding env, operably linked to

an RDRP promoter in the reverse orientation, *e.g.,* a plus-strand or minus-strand ssRNA virus RDRP promoter.

The present invention further contemplates, packaging cell lines comprising at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 or more packaging vectors. In various embodiments, 5, 6, 7, 8, 9, or 10 packaging vector cell lines comprise a 5^{th}, 6^{th}, 7^{th}, 8^{th}, 9^{th}, or 10^{th} polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to an RDRP promoter, *e.g.,* alphavirus 26S promoter, operably linked to a polynucleotide encoding a viral gene selected from the group consisting of tat, rev, vif, vpr, vpu, vpx, and nef, and a CSE for production of a minus strand RNA.

In various other embodiments, 5, 6, 7, 8, 9, or 10 packaging vector cell lines comprise a 5^{th}, 6^{th}, 7^{th}, 8^{th}, 9^{th}, or 10^{th} polynucleotide comprising a ubiquitous, tissue specific, or inducible promoter operably linked to the reverse complement of a polynucleotide encoding a viral gene selected from the group consisting of tat, rev, vif, vpr, vpu, vpx, and nef, operably linked to an RDRP promoter in the reverse orientation, *e.g.,* a plus-strand or minus-strand ssRNA virus RDRP promoter.

One having ordinary skill in the art would appreciate that any number of, or all of the 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 packaging vectors can be designed in the sense or reverse complement orientation, in any combination.

In addition, the present invention contemplates that any number or, or all of the 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 packaging vectors can be expression vectors, *e.g.,* commercial plasmid expression vectors, or retroviral vectors. Accordingly, the invention contemplates providing packaging cell lines by both transient transfection and viral transduction.

In particular embodiments, wherein the packaging vectors are viral vectors, *e.g.,* retroviral or lentiviral transfer vectors, the vectors may comprise one or more of each of the following: 5' and 3' LTRs, optionally modified as discussed elsewhere herein, packaging sequences, posttranscriptional recognition elements (*e.g.,* WPRE, HPRE), insulator elements, cPPT/FLAP sequences, selectable markers, a retroviral export element, *e.g.,* a lentiviral REV response element (RRE), and other lentiviral vector features as disclosed elsewhere herein. In addition, although the

preceding transfer vectors recite an ordered list of features, one having ordinary skill in the art would appreciate that other orders of the features are possible and in some cases preferred.

1. Producer Cell Lines

In various embodiments, the present invention contemplates a producer cell comprising a packaging cell and one or more viral transfer vectors (*e.g.,* plasmid) each comprising viral packaging sequences (Ψs) and LTRs and one or more polynucleotides.

In one embodiment, a transfer vector (*e.g.* plasmid) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to a plus-strand ssRNA virus RDRP promoter sequence, *e.g.,* an alphavirus subgenomic promoter, operably linked to a left (3') retroviral LTR; a packaging sequence (Ψ), another ubiquitous, tissue specific, or inducible promoter operably linked operably to a polynucleotide of interest; a right (3') retroviral LTR; and a CSE for generation of a minus strand RNA. In this arrangement, once the vector has integrated into the genome, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA comprising the entire retroviral genome ready for packaging. In addition, this arrangement allows for controlled amplification of the retroviral genome of the transfer plasmid that will be packaged. When present, the RDRP binds to the 3′ cis-acting sequence, *e.g.,* CSE, at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerizes a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, the RDRP binds 5′ cis-acting sequence or subgenomic promoter and amplifies sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies comprise the entire retroviral genome that will be packaged. Accordingly, this design allows for packaging of genomic RNA from both the ubiquitous, tissue specific, or inducible promoter and the 5′ cis-acting sequence or subgenomic promoter, albeit using different RNA templates. In this manner, the retroviral genomic RNA can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system. Furthermore, this system provides advantages in that it

increases the amount full-length retroviral genomic RNA, thereby increasing viral titer 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In one embodiment, a transfer vector (*e.g.* plasmid) comprises a ubiquitous, tissue specific, or inducible promoter operably linked to an RNA dependent RNA polymerase promoter sequence, *e.g.,* the alphavirus 26S subgenomic promoter, operably linked to a left (3') retroviral LTR; a packaging sequence (Ψ), another ubiquitous, tissue specific, or inducible promoter operably linked operably to a polynucleotide of interest; a right (3') retroviral LTR; and a CSE for generation of a minus strand RNA. In this arrangement, once the vector has integrated into the genome, the ubiquitous, tissue specific, or inducible promoter transcribes a sense (plus-strand) mRNA comprising the entire retroviral genome ready for packaging. In addition, this arrangement allows for controlled amplification of the retroviral genome of the transfer plasmid that will be packaged. When, present, one or more NSPs bind to the CSE at the 3' end of the plus-stranded RNA and using the plus strand as a template, polymerize a non-sense (minus-strand) RNA. Once the minus strand RNA has been formed, one or more NSPs bind to the alphavirus 26S subgenomic promoter and amplify sense (plus-strand) RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA copies comprise the entire retroviral genome that will be packaged. Accordingly, this design allows for packaging of genomic RNA from both the ubiquitous, tissue specific, or inducible promoter and the 26S promoter, albeit using different RNA templates. In this manner, the retroviral genomic RNA can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system. Furthermore, this system provides advantages in that it increases the amount full-length retroviral genomic RNA, thereby increasing viral titer 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In one embodiment, a transfer vector (*e.g.* plasmid) comprises a constitutive, *e.g.,* ubiquitous or tissue specific, or an inducible promoter operably linked to the reverse complement of: a right (3') retroviral LTR; a polynucleotide of interest operably linked to another ubiquitous, tissue specific, or inducible promoter; a

packaging sequence (Ψ); a left (3') retroviral LTR; and a plus-strand or minus-strand ssRNA virus RDRP promoter sequence in the reverse orientation. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the entire retroviral genome that will be packaged the polypeptide of interest and an RDRP 5′ cis-acting sequence or subgenomic promoter in the reverse orientation: the reverse complement of the viral vector cannot be directly packaged into the retroviral particle. However, the transcribed RNA can be used a minus strand RNA template in the 3′ to 5′ orientation. When present, the RDRP binds to the 5′ cis-acting sequence or subgenomic promoter sequence on the transcribed strand and amplifies translatable RNA copies using the transcribed strand as a template. These amplified sense (plus-strand) RNA comprise the entire retroviral genome that will be packaged. Accordingly, the CSE is not needed to generate a minus strand because the RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, the retroviral genomic RNA can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system. Furthermore, this system provides advantages in that it increases the amount full-length retroviral genomic RNA, thereby increasing viral titer 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In one embodiment, the transfer vector (*e.g.* plasmid) comprises a constitutive, *e.g.,* ubiquitous or tissue specific, or an inducible promoter operably linked to the reverse complement of: a right (3') retroviral LTR; a polynucleotide of interest operably linked to another ubiquitous, tissue specific, or inducible promoter; a packaging sequence (Ψ); a left (3') retroviral LTR; and an RNA dependent RNA polymerase promoter sequence in the reverse orientation, *e.g.,* the alphavirus 26S subgenomic promoter. In this arrangement, the ubiquitous, tissue specific, or inducible promoter transcribes a reverse complement of the mRNA that encodes the entire retroviral genome that will be packaged, the polypeptide of interest, and the alphavirus 26S subgenomic promoter in reverse orientation: the reverse complement cannot be directly packaged into the retroviral particle. However, thetranscribed RNA can be used a minus strand RNA template in the 3′ to 5′ orientation. When present, one or

more NSPs bind to the alphavirus 26S subgenomic promoter sequence on the transcribed strand and amplify translatable RNA copies using the minus strand as a template. These amplified sense (plus-strand) RNA comprise the entire retroviral genome that will be packaged. Accordingly, the CSE is not needed to generate a minus strand because the RDRP can use the plus-strand RNA in the 3′ to 5′ orientation as a minus-strand template. In this manner, the retroviral genomic RNA can be amplified 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system. Furthermore, this system provides advantages in that it increases the amount full-length retroviral genomic RNA, thereby increasing viral titer 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or 100 fold or more compared to a system that lacks an RDRP/promoter system.

In a particular embodiment, the present invention contemplates, that each of the foregoing transfer vectors (*e.g.,* plasmid) may further comprise recombinase recognition sites that flank the retroviral genome to be packaged, and optionally a selectable marker. In this manner, once the transfer vector bearing the recombinase recognition sites has integrated into the packaging cell genome, other transfer vectors comprising recombinase recognition sites compatible to those flanking the retroviral genome of the integrated transfer vector can be exchanged with the integrated transfer vector in the presence of a suitable recombinase. This process, Recombinase Mediated Cassette Exchange (RMCE) if highly efficient and may be serially repeated 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more times on the integration locus. In preferred embodiments, the recombinase activity is introduced transiently in the producer cell for a time and duration effective to mediate the cassette exchange. The recombinase can be introduced into the packaging cell transiently or permanently and under the control of a ubiquitous, tissue specific, or inducible promoter.

Recombinase mediated cassette exchange (RMCE) is a process in which site-specific recombinases exchange one gene cassette flanked by a pair of incompatible target sites for another cassette flanked by an identical pair of sites. Typically one cassette is present in the host genome, *e.g.,* the packaging cell, whereas the other gene cassette is introduced into the host cell by chemical or biological means, *e.g.,* on a transfer plasmid. In addition, successful RMCE can be accomplished using a retroviral

vector carrying both the transgene cassette to be inserted and the recombinase coding region. RMCE proceeds efficiently in cells in which the vector is able to replicate. Additional transfection of the transgene cassette significantly enhances the RMCE frequency. This demonstrates that an RMCE system in the context of a viral vector allows the site directed insertion of a transgene into a defined genomic site.

In this way, stable packaging cell lines can be developed having tight gene switch regulation of the viral genes necessary for packaging and a placeholder genomic integration for the lentiviral genome to be packaged, flanked by recombinase sites. Once a transfer vector is ready for viral particle production, it can be recombined into the packaging cell. However, although the packaging cell/producer cell has all the genes and elements required for proper viral packaging, the geneswitch of the viral genes required for packaging is "off"; thus, no packaging takes place and the cell is completely stable. The viral production system is switched "on" when the packaging/producer cells are contacted with a polynucleotide encoding an expressible RDRP or an RDRP polypeptide capable of entering the cell. Once the system is switched "on" robust viral particle production takes place. In certain embodiments, the system may be switched "off" again, either by RDRP protein turnover or regulatable expression of the RDRP expression vector.

In particular embodiments, wherein the packaging vectors are transfer vectors, *e.g.,* retroviral or lentiviral transfer vectors (*e.g.,* plasmids), the vectors may comprise one or more of each of the following: 5' and 3' LTRs, optionally modified as discussed elsewhere herein, packaging sequences, posttranscriptional recognition elements (*e.g.,* WPRE, HPRE), insulator elements, cPPT/FLAP sequences, selectable markers, a retroviral export element, *e.g.,* a lentiviral REV response element (RRE), and other lentiviral vector features as disclosed elsewhere herein. In addition, although the preceding transfer vectors recite an ordered list of features, one having ordinary skill in the art would appreciate that other orders of the features are possible and in some cases preferred.

In one embodiment, the invention employs an inducible promoter within the retroviral vectors, so that transcription of selected genes can be turned on and off. This minimizes cellular toxicity caused by expression of cytotoxic viral proteins,

increasing the stability of the packaging cells containing the vectors. For example, high levels of expression of VSV-G (envelope protein) and Vpr can be cytotoxic (Yee, J. -K., *et al.*, 1994. *Proc. Natl. Acad. Sci.*, 91:9654-9568) and, therefore, expression of these proteins in packaging cells of the invention can be controlled by an inducible promoter systems as discussed elsewhere herein.

Viruses may be used to infect cells *in vivo, ex vivo*, or *in vitro* using techniques well known in the art For example, when cells, for instance $CD34^+$ cells, dendritic cells, peripheral blood cells or tumor cells are transduced *ex vivo*, the vector particles may be incubated with the cells using a dose generally in the order of between 1 to 50 multiplicities of infection (MOI) which also corresponds to $1x10^5$ to $50x10^5$ transducing units of the viral vector per 10^5 cells. This, of course, includes amount of vector corresponding to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, and 50 MOI.

Viruses may also be delivered to a subject *in vivo*, by direct injection to the cell, tissue, or organ in need of therapy. Direct injection requires on the order of between 1 to 50 multiplicities of infection (MOI) which also corresponds to $1x10^5$ to $50x10^5$ transducing units of the viral vector per 10^5 cells.

Viruses may also be delivered according to viral titer (TU/mL), which can be measured, for example, by using a commercially available p24 titer assay, which is an ELISA against the p24 viral coat protein. The following formula can be used to calculate the pg/mL of p24: there are approximately 2000 molecules of p24 per physical particle (PP) of lentivirus: (2 x 103) x (24 x 103 Da of p24 per PP), 48 x 106/Avogadro = (48 x 106) / (6 x 1023) = 8 x 10-17 g of p24 per PP, approximately 1 PP per 1 x 10-16 g of p24, 1 x 104 PP per pg of p24. A reasonably well packaged, VSV-G pseudotyped lentiviral vector will have an infectivity index in the range of 1 TU per 1000 physical particles (PP) to 1 TU per 100 PP (or less). Thus, the range is approximately 10 to 100 TU/pg of p24. It is through this conversion that TU/mL is obtained.

Based on previous experience, the amount of lentivirus directly injected is determined by total TU and can vary based on both the volume that could be feasibly injected to the site and the type of tissue to be injected. For example, a brain injection

site may only allow for a very small volume of virus to be injected, so a high titer prep would be preferred, a TU of about 1 x 10^6 to 1 x 10^7, about 1 x 10^6 to 1 x 10^8, 1 x 10^6 to 1 x 10^9, about 1 x 10^7 to 1 x 10^{10}, 1 x 10^8 to 1 x 10^{11}, about 1 x 10^8 to 1 x 10^{12}, or about 1 x 10^{10} to 1 x 10^{12} or more per injection could be used. However, a systemic delivery could accommodate a much larger TU, a load of 1 x 10^8, 1 x 10^9, 1 x 10^{10}, 1 x 10^{11}, 1 x 10^{12}, 1 x 10^{13}, 1 x 10^{14}, or 1 x 10^{15}, could be delivered.

J. Gene Therapy Methods

The retroviral vector geneswitch expression systems of the invention provide improved viral titers and highly controlled strategies for amplification of one or more mammalian polypeptides, *e.g.,* therapeutic polypeptides, in a transduced cell. Thus, the viruses and transduced cells herein are suitable for gene therapy applications. As used herein, the term "gene therapy" refers to the introduction of a gene into a cell's genome. The virus can infect and transduce the cell *in vivo, ex vivo*, or *in vitro*. In *ex vivo* and *in vitro* embodiments, the transduced cells can then be administered to a subject in need of therapy. The present invention contemplates that the vector systems, viral particles, and transduced cells of the invention are broadly applicable to any disease, disorder, or condition. In particular embodiments, the gene therapy systems of the invention are broadly applicable to treat, prevent, and/or ameliorate monogenic diseases, disorders, or conditions.

In various embodiments, the retroviral vectors are administered by direct injection to a cell, tissue, or organ of a subject in need of gene therapy, *in vivo*. In various other embodiments, cells are transduced *in vitro* or *ex vivo* with one or more retroviral, *e.g.,* lentiviral, vectors of the invention depending on whether the single vector system, the binary vector system, or the polypeptide composition/ vector system is used. The transduced cells are then administered to a subject in need of gene therapy.

In methods comprising use of the binary vector system, the virus of the first or second vectors may be directly administered to the subject to transduced the desired cells, in any order or together, or the cells may be transduced *ex vivo* or *in vitro,* with either one of, or both of, the viruses from the first and second vectors, and administered to the subject. If the cells is transduced with only one of the vectors, the

cells can be administered to the subject and the second vector can be administered as a virus or expression vector to the transduced cells of the subject to control and/or amplify therapeutic polypeptide expression.

In methods comprising use of the polypeptide composition/ vector system, the virus may be directly administered to the subject to transduce the desired cells or the cells may be transduced *ex vivo* or *in vitro* and administered to the subject. Once the transduced cells are provided to the subject, a composition comprising RDRP or polynucleotides encoding an expressible RDRP, *e.g.,* NSP1-NSP4, NSP2-2A-NSP4, or NSP4, can be administered to the subject to control and/or amplify therapeutic polypeptide expression.

Cells suitable for transduction and administration in the gene therapy methods of the invention include, but are not limited to stem cells, progenitor cells, and differentiated cells.

In various embodiments, the use of stem cells is preferred because they have the ability to differentiate into the appropriate cell types when administered to a particular biological niche, *in vivo*. The term "stem cell" refers to a cell which is an undifferentiated cell capable of (1) long term self -renewal, or the ability to generate at least one identical copy of the original cell, (2) differentiation at the single cell level into multiple, and in some instance only one, specialized cell type and (3) of *in vivo* functional regeneration of tissues. Stem cells are subclassified according to their developmental potential as totipotent, pluripotent, multipotent and oligo/unipotent. "Self-renewal" refers a cell with a unique capacity to produce unaltered daughter cells and to generate specialized cell types (potency). Self-renewal can be achieved in two ways. Asymmetric cell division produces one daughter cell that is identical to the parental cell and one daughter cell that is different from the parental cell and is a progenitor or differentiated cell. Asymmetric cell division does not increase the number of cells. Symmetric cell division produces two identical daughter cells. "Proliferation" or "expansion" of cells refers to symmetrically dividing cells.

As used herein, the term "totipotent" means the ability of a cell to form all cell lineages of an organism. For example, in mammals, only the zygote and the first cleavage stage blastomeres are totipotent. As used herein, the term "pluripotent"

means the ability of a cell to form all lineages of the body or soma (*i.e.*, the embryo proper). For example, embryonic stem cells are a type of pluripotent stem cells that are able to form cells from each of the three germs layers, the ectoderm, the mesoderm, and the endoderm. As used herein, the term "multipotent" refers to the ability of an adult stem cell to form multiple cell types of one lineage. For example, hematopoietic stem cells are capable of forming all cells of the blood cell lineage, *e.g.,* lymphoid and myeloid cells. As used herein, the term "oligopotent" refers to the ability of an adult stem cell to differentiate into only a few different cell types. For example, lymphoid or myeloid stem cells are capable of forming cells of either the lymphoid or myeloid lineages, respectively. As used herein, the term "unipotent" means the ability of a cell to form a single cell type. For example, spermatogonial stem cells are only capable of forming sperm cells.

As used herein, the term "progenitor" or "progenitor cells" refers to cells have the capacity to self-renew and to differentiate into more mature cells. Many progenitor cells differentiate along a single lineage, but may have quite extensive proliferative capacity.

In particular embodiments, the population or source of cells used in the methods contemplated herein comprises mesenchymal stem and/or progenitor cells, mesodermal stem and/or progenitor cells, endodermal stem and/or progenitor cells, or ectodermal stem and/or progenitor cells. In certain embodiments, the population or source of cells used in the methods contemplated herein comprises bone marrow stem cells, umbilical cord blood stem and/or progenitor cells, bone stem and/or progenitor cells, muscle stem and/or progenitor cells, hematopoietic stem and/or progenitor cells, fat stem and/or progenitor cells, cartilage stem and/or progenitor cells, neural stem and/or progenitor cells, skin stem and/or progenitor cells, liver stem and/or progenitor cells, pancreas stem and/or progenitor cells, kidney stem and/or progenitor cells, gastric stem and/or progenitor cells, and intestinal stem and/or progenitor cells.

Hematopoietic stem cells (HSCs) give rise to committed hematopoietic progenitor cells (HPCs) that are capable of generating the entire repertoire of mature blood cells over the lifetime of an organism. The term "hematopoietic stem cell" or "HSC" refers to multipotent stem cells that give rise to the all the blood cell types of an

organism, including myeloid (*e.g.,* monocytes and macrophages, neutrophils, basophils, eosinophils, erythrocytes, megakaryocytes/platelets, dendritic cells), and lymphoid lineages (*e.g.,* T-cells, B-cells, NK-cells), and others known in the art (See Fei, R., *et al.*, U.S. Patent No. 5,635,387; McGlave, *et al.*, U.S. Patent No. 5,460,964; Simmons, P., et al, U.S. Patent No. 5,677,136; Tsukamoto, *et al.*, U.S. Patent No. 5,750,397; Schwartz, *et al.*, U.S. Patent No. 5,759,793; DiGuisto, *et al.*, U.S. Patent No. 5,681,599; Tsukamoto, *et al.*, U.S. Patent No. 5,716,827). When transplanted into lethally irradiated animals or humans, hematopoietic stem and progenitor cells can repopulate the erythroid, neutrophil-macrophage, megakaryocyte and lymphoid hematopoietic cell pool.

Cells of the invention can be autologous/autogeneic ("self") or non-autologous ("non-self," *e.g.,* allogeneic, syngeneic or xenogeneic). "Autologous," as used herein, refers to cells from the same subject. "Allogeneic," as used herein, refers to cells of the same species that differ genetically to the cell in comparison. "Syngeneic," as used herein, refers to cells of a different subject that are genetically identical to the cell in comparison. "Xenogeneic," as used herein, refers to cells of a different species to the cell in comparison. In preferred embodiments, the cells of the invention are allogeneic.

A "subject," as used herein, includes any animal that exhibits a symptom that can be treated with the gene therapy vectors, cell-based therapeutics, and methods disclosed elsewhere herein. Suitable subjects (*e.g.,* patients) include laboratory animals (such as mouse, rat, rabbit, or guinea pig), farm animals, and domestic animals or pets (such as a cat or dog). Non-human primates and, preferably, human patients, are included. Typical subjects include animals that exhibit aberrant amounts (lower or higher amounts than a "normal" or "healthy" subject) of one or more physiological activities that can be modulated by gene therapy.

As used herein "treatment" or "treating," includes any beneficial or desirable effect on the symptoms or pathology of a disease or pathological condition, and may include even minimal reductions in one or more measurable markers of the disease or condition being treated. Treatment can involve optionally either the reduction or amelioration of symptoms of the disease or condition, or the delaying of

the progression of the disease or condition. "Treatment" does not necessarily indicate complete eradication or cure of the disease or condition, or associated symptoms thereof.

As used herein, "prevent," and similar words such as "prevented," "preventing" *etc.*, indicate an approach for preventing, inhibiting, or reducing the likelihood of the occurrence or recurrence of, a disease or condition. It also refers to delaying the onset or recurrence of a disease or condition or delaying the occurrence or recurrence of the symptoms of a disease or condition. As used herein, "prevention" and similar words also includes reducing the intensity, effect, symptoms and/or burden of a disease or condition prior to onset or recurrence of the disease or condition.

As used herein, the term "amount" refers to "an amount effective" or "an effective amount" of a virus or transduced therapeutic cell to achieve a beneficial or desired prophylactic or therapeutic result, including clinical results.

A "prophylactically effective amount" refers to an amount of a virus or transduced therapeutic cell effective to achieve the desired prophylactic result. Typically but not necessarily, since a prophylactic dose is used in subjects prior to or at an earlier stage of disease, the prophylactically effective amount is less than the therapeutically effective amount.

A "therapeutically effective amount" of a virus or transduced therapeutic cell may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of the stem and progenitor cells to elicit a desired response in the individual. A therapeutically effective amount is also one in which any toxic or detrimental effects of the virus or transduced therapeutic cells are outweighed by the therapeutically beneficial effects. The term "therapeutically effective amount" includes an amount that is effective to "treat" a subject (*e.g.,* a patient).

Viral vectors of the present invention may be used to control the expression of one or more mammalian polypeptides useful in the treatment of a disease, disorder, or condition that is amenable to treatment with gene therapy, including, but not limited to monogenic diseases, disorders, and conditions.

1. Monogenic Diseases

As used herein, the term "monogenic disease" refers to a disease in which modifications of a single gene is associated with a disorder, disease, or condition in a subject. Though relatively rare, monogenic diseases affect millions of people worldwide. Scientists currently estimate that over 10,000 human diseases are known to be monogenic. Pure genetic diseases are caused by a single error in a single gene in the human DNA. The nature of disease depends on the functions performed by the modified gene. The single-gene or monogenic diseases can be classified into three main categories: Dominant, Recessive, and X-linked. All human beings have two sets or copies of each gene called "allele"; one copy on each side of the chromosome pair. Recessive diseases are monogenic disorders that occur due to damages in both copies or allele. Dominant diseases are monogenic disorders that involve damage to only one gene copy. X linked diseases are monogenic disorders that are linked to defective genes on the X chromosome which is the sex chromosome. The X linked alleles can also be dominant or recessive. These alleles are expressed equally in men and women, more so in men as they carry only one copy of X chromosome (XY) whereas women carry two (XX).

Monogenic diseases are responsible for a heavy loss of life. The global prevalence of all single gene diseases at birth is approximately 1/100. In Canada, it has been estimated that taken together, monogenic diseases may account for up to 40% of the work of hospital based pediatric practice (Scriver, 1995).

Illustrative embodiments of monogenic disorders, diseases, and conditions that can be treated, prevented, or ameliorated with compositions and methods of the present invention include, but are not limited to: 11-hydroxylase deficiency; 17,20-desmolase deficiency; 17-hydroxylase deficiency; 3-hydroxyisobutyrate aciduria; 3-hydroxysteroid dehydrogenase deficiency; 46,XY gonadal dysgenesis; AAA syndrome; ABCA3 deficiency; ABCC8-associated hyperinsulinism; aceruloplasminemia; achondrogenesis type 2; acral peeling skin syndrome; acrodermatitis enteropathica; adrenocortical micronodular hyperplasia; adrenoleukodystrophies; adrenomyeloneuropathies; Aicardi-Goutieres syndrome; Alagille disease; Alpers syndrome; alpha-mannosidosis; Alstrom syndrome; Alzheimer

disease; amelogenesis imperfecta; amish type microcephaly; amyotrophic lateral sclerosis; anauxetic dysplasia; androgen insentivity syndrome; Antley-Bixler syndrome; APECED, Apert syndrome, aplasia of lacrimal and salivary glands, argininemia, arrhythmogenic right ventricular dysplasia, Arts syndrome, ARVD2, arylsulfatase deficiency type metachromatic leokodystrophy, ataxia telangiectasia, autoimmune lymphoproliferative syndrome; autoimmune polyglandular syndrome type 1; autosomal dominant anhidrotic ectodermal dysplasia; autosomal dominant polycystic kidney disease; autosomal recessive microtia; autosomal recessive renal glucosuria; autosomal visceral heterotaxy; Bardet-Biedl syndrome; Bartter syndrome; basal cell nevus syndrome; Batten disease; benign recurrent intrahepatic cholestasis; beta-mannosidosis; Bethlem myopathy; Blackfan-Diamond anemia; blepharophimosis; Byler disease; C syndrome; CADASIL; carbamyl phosphate synthetase deficiency; cardiofaciocutaneous syndrome; Carney triad; carnitine palmitoyltransferase deficiencies; cartilage-hair hypoplasia; cblC type of combined methylmalonic aciduria; CD18 deficiency; CD3Z-associated primary T-cell immunodeficiency; CD40L deficiency; CDAGS syndrome; CDG1A; CDG1B; CDG1M; CDG2C; CEDNIK syndrome; central core disease; centronuclear myopathy; cerebral capillary malformation; cerebrooculofacioskeletal syndrome type 4; cerebrooculogacioskeletal syndrome; cerebrotendinous xanthomatosis; CHARGE association; cherubism; CHILD syndrome; chronic granulomatous disease; chronic recurrent multifocal osteomyelitis; citrin deficiency; classic hemochromatosis; CNPPB syndrome; cobalamin C disease; Cockayne syndrome; coenzyme Q10 deficiency; Coffin-Lowry syndrome; Cohen syndrome; combined deficiency of coagulation factors V; common variable immune deficiency; complete androgen insentivity; cone rod dystrophies; conformational diseases; congenital bile adid synthesis defect type 1; congenital bile adid synthesis defect type 2; congenital defect in bile acid synthesis type; congenital erythropoietic porphyria; congenital generalized osteosclerosis; Cornelia de Lange syndrome; Cousin syndrome; Cowden disease; COX deficiency; Crigler-Najjar disease; Crigler-Najjar syndrome type 1; Crisponi syndrome; Currarino syndrome; Curth-Macklin type ichthyosis hystrix; cutis laxa; cystinosis; d-2-hydroxyglutaric aciduria; DDP syndrome; Dejerine-Sottas disease; Denys-Drash syndrome; desmin cardiomyopathy; desmin myopathy; DGUOK-

associated mitochondrial DNA depletion; disorders of glutamate metabolism; distal spinal muscular atrophy type 5; DNA repair diseases; dominant optic atrophy; Doyne honeycomb retinal dystrophy; Duchenne muscular dystrophy; dyskeratosis congenita; Ehlers-Danlos syndrome type 4; Ehlers-Danlos syndromes; Elejalde disease; Ellis-van Creveld disease; Emery-Dreifuss muscular dystrophies; encephalomyopathic mtDNA depletion syndrome; enzymatic diseases; EPCAM-associated congenital tufting enteropathy; epidermolysis bullosa with pyloric atresia; exercise-induced hypoglycemia; facioscapulohumeral muscular dystrophy; Faisalabad histiocytosis; familial atypical mycobacteriosis; familial capillary malformation-arteriovenous; familial esophageal achalasia; familial glomuvenous malformation; familial hemophagocytic lymphohistiocytosis; familial mediterranean fever; familial megacalyces; familial schwannomatosisl; familial spina bifida; familial splenic asplenia/hypoplasia; familial thrombotic thrombocytopenic purpura; Fanconi disease; Feingold syndrome; FENIB; fibrodysplasia ossificans progressiva; FKTN; Francois-Neetens fleck corneal dystrophy; Frasier syndrome; Friedreich ataxia; FTDP-17; fucosidosis; G6PD deficiency; galactosialidosis; Galloway syndrome; Gardner syndrome; Gaucher disease; Gitelman syndrome; GLUT1 deficiency; glycogen storage disease type 1b; glycogen storage disease type 2; glycogen storage disease type 3; glycogen storage disease type 4; glycogen storage disease type 9a; glycogen storage diseases; GM1-gangliosidosis; Greenberg syndrome; Greig cephalopolysyndactyly syndrome; hair genetic diseases; HANAC syndrome; harlequin type ichtyosis congenita; HDR syndrome; hemochromatosis type 3; hemochromatosis type 4; hemophilia A; hereditary angioedema type 3; hereditary angioedemas; hereditary hemorrhagic telangiectasia; hereditary hypofibrinogenemia; hereditary intraosseous vascular malformation; hereditary leiomyomatosis and renal cell cancer; hereditary neuralgic amyotrophy; hereditary sensory and autonomic neuropathy type; Hermansky-Pudlak disease; HHH syndrome; HHT2; hidrotic ectodermal dysplasia type 1; hidrotic ectodermal dysplasias; HNF4A-associated hyperinsulinism; HNPCC; human immunodeficiency with microcephaly; Huntington disease; hyper-IgD syndrome; hyperinsulinism-hyperammonemia syndrome; hypertrophy of the retinal pigment epithelium; hypochondrogenesis; hypohidrotic ectodermal dysplasia; ICF syndrome;

idiopathic congenital intestinal pseudo-obstruction; immunodeficiency with hyper-IgM type 1; immunodeficiency with hyper-IgM type 3; immunodeficiency with hyper-IgM type 4; immunodeficiency with hyper-IgM type 5; inborm errors of thyroid metabolism; infantile visceral myopathy; infantile X-linked spinal muscular atrophy; intrahepatic cholestasis of pregnancy; IPEX syndrome; IRAK4 deficiency; isolated congenital asplenia; Jeune syndrome Imag; Johanson-Blizzard syndrome; Joubert syndrome; JP-HHT syndrome; juvenile hemochromatosis; juvenile hyalin fibromatosis; juvenile nephronophthisis; Kabuki mask syndrome; Kallmann syndromes; Kartagener syndrome; KCNJ11-associated hyperinsulinism; Kearns-Sayre syndrome; Kostmann disease; Kozlowski type of spondylometaphyseal dysplasia; Krabbe disease; LADD syndrome; late infantile-onset neuronal ceroid lipofuscinosis; LCK deficiency; LDHCP syndrome; Legius syndrome; Leigh syndrome; lethal congenital contracture syndrome 2; lethal congenital contracture syndromes; lethal contractural syndrome type 3; lethal neonatal CPT deficiency type 2; lethal osteosclerotic bone dysplasia; LIG4 syndrome; lissencephaly type 1 Imag; lissencephaly type 3; Loeys-Dietz syndrome; low phospholipid-associated cholelithiasis; lysinuric protein intolerance; Maffucci syndrome; Majeed syndrome; mannose-binding protein deficiency; Marfan disease; Marshall syndrome; MASA syndrome; MCAD deficiency; McCune-Albright syndrome; MCKD2; Meckel syndrome; Meesmann corneal dystrophy; megacystis-microcolon-intestinal hypoperistalsis; megaloblastic anemia type 1; MEHMO; MELAS; Melnick-Needles syndrome; MEN2s; Menkes disease; metachromatic leukodystrophies; methylmalonic acidurias; methylvalonic aciduria; microcoria-congenital nephrosis syndrome; microvillous atrophy; mitochondrial neurogastrointestinal encephalomyopathy; monilethrix; monosomy X; mosaic trisomy 9 syndrome; Mowat-Wilson syndrome; mucolipidosis type 2; mucolipidosis type IIIa; mucolipidosis type IV; mucopolysaccharidoses; mucopolysaccharidosis type 3A; mucopolysaccharidosis type 3C; mucopolysaccharidosis type 4B; multiminicore disease; multiple acyl-CoA dehydrogenation deficiency; multiple cutaneous and mucosal venous malformations; multiple endocrine neoplasia type 1; multiple sulfatase deficiency; NAIC; nail-patella syndrome; nemaline myopathies; neonatal diabetes mellitus; neonatal surfactant deficiency; nephronophtisis; Netherton disease;

neurofibromatoses; neurofibromatosis type 1; Niemann-Pick disease type A; Niemann-Pick disease type B; Niemann-Pick disease type C; NKX2E; Noonan syndrome; North American Indian childhood cirrhosis; NROB1 duplication-associated DSD; ocular genetic diseases; oculo-auricular syndrome; OLEDAID; oligomeganephronia; oligomeganephronic renal hypolasia; Ollier disease; Opitz-Kaveggia syndrome; orofaciodigital syndrome type 1; orofaciodigital syndrome type 2; osseous Paget disease; otopalatodigital syndrome type 2; OXPHOS diseases; palmoplantar hyperkeratosis; panlobar nephroblastomatosis; Parkes-Weber syndrome; Parkinson disease; partial deletion of 21q22.2-q22.3; Pearson syndrome; Pelizaeus-Merzbacher disease; Pendred syndrome; pentalogy of Cantrell; peroxisomal acyl-CoA-oxidase deficiency; Peutz-Jeghers syndrome; Pfeiffer syndrome; Pierson syndrome; pigmented nodular adrenocortical disease; pipecolic acidemia; Pitt-Hopkins syndrome; plasmalogens deficiency; pleuropulmonary blastoma and cystic nephroma; polycystic lipomembranous osteodysplasia; porphyrias; premature ovarian failure; primary erythermalgia; primary hemochromatoses; primary hyperoxaluria; progressive familial intrahepatic cholestasis; propionic acidemia; pyruvate decarboxylase deficiency; RAPADILINO syndrome; renal cystinosis; rhabdoid tumor predisposition syndrome; Rieger syndrome; ring chromosome 4; Roberts syndrome; Robinow-Sorauf syndrome; Rothmund-Thomson syndrome; SCID; Saethre-Chotzen syndrome; Sandhoff disease; SC phocomelia syndrome; SCA5; Schinzel phocomelia syndrome; short rib-polydactyly syndrome type 1; short rib-polydactyly syndrome type 4; short-rib polydactyly syndrome type 2; short-rib polydactyly syndrome type 3; Shwachman disease; Shwachman-Diamond disease; sickle cell anemia; Silver-Russell syndrome; Simpson-Golabi-Behmel syndrome; Smith-Lemli-Opitz syndrome; SPG7-associated hereditary spastic paraplegia; spherocytosis; split-hand/foot malformation with long bone deficiencies; spondylocostal dysostosis; sporadic visceral myopathy with inclusion bodies; storage diseases; STRA6-associated syndrome; Tay-Sachs disease; thanatophoric dysplasia; thyroid metabolism diseases; Tourette syndrome; transthyretin-associated amyloidosis; trisomy 13; trisomy 22; trisomy 2p syndrome; tuberous sclerosis; tufting enteropathy; urea cycle diseases; Van Den Ende-Gupta syndrome; Van der Woude syndrome; variegated mosaic aneuploidy syndrome;

VLCAD deficiency; von Hippel-Lindau disease; Waardenburg syndrome; WAGR syndrome; Walker-Warburg syndrome; Werner syndrome; Wilson disease; Wolcott-Rallison syndrome; Wolfram syndrome; X-linked agammaglobulinemia; X-linked chronic idiopathic intestinal pseudo-obstruction; X-linked cleft palate with ankyloglossia; X-linked dominant chondrodysplasia punctata ; X-linked ectodermal dysplasia; X-linked Emery-Dreifuss muscular dystrophy; X-linked lissencephaly; X-linked lymphoproliferative disease; X-linked visceral heterotaxy; xanthinuria type 1; xanthinuria type 2; xeroderma pigmentosum; XPV; and Zellweger disease.

Other illustrative embodiments of diseases, disorders, and conditions that can be treated, prevented, or ameliorated using methods and compositions of the present invention include, but not limited to treating macrophage inhibitory and/or T cell inhibitory activity and thus, anti-inflammatory activity; anti-immune activity, *i.e.* inhibitory effects against a cellular and/or humoral immune response, including a response not associated with inflammation; inhibit the ability of macrophages and T cells to adhere to extracellular matrix components and fibronectin, as well as up-regulated fas receptor expression in T cells; inhibit unwanted immune reaction and inflammation including arthritis, including rheumatoid arthritis, inflammation associated with hypersensitivity, allergic reactions, asthma, systemic lupus erythematosus, collagen diseases and other autoimmune diseases, inflammation associated with atherosclerosis, arteriosclerosis, atherosclerotic heart disease, reperfusion injury, cardiac arrest, myocardial infarction, vascular inflammatory disorders, respiratory distress syndrome or other cardiopulmonary diseases, inflammation associated with peptic ulcer, ulcerative colitis and other diseases of the gastrointestinal tract, hepatic fibrosis, liver cirrhosis or other hepatic diseases, thyroiditis or other glandular diseases, glomerulonephritis or other renal and urologic diseases, otitis or other oto-rhino-laryngological diseases, dermatitis or other dermal diseases, periodontal diseases or other dental diseases, orchitis or epididimo-orchitis, infertility, orchidal trauma or other immune-related testicular diseases, placental dysfunction, placental insufficiency, habitual abortion, eclampsia, pre-eclampsia and other immune and/or inflammatory-related gynecological diseases, posterior uveitis, intermediate uveitis, anterior uveitis, conjunctivitis, chorioretinitis, uveoretinitis, optic

neuritis, intraocular inflammation, *e.g.* retinitis or cystoid macular oedema, sympathetic ophthalmia, scleritis, retinitis pigmentosa, immune and inflammatory components of degenerative fondus disease, inflammatory components of ocular trauma, ocular inflammation caused by infection, proliferative vitreo-retinopathies, acute ischaemic optic neuropathy, excessive scarring, *e.g.* following glaucoma filtration operation, immune and/or inflammation reaction against ocular implants and other immune and inflammatory-related ophthalmic diseases, inflammation associated with autoimmune diseases or conditions or disorders where, both in the central nervous system (CNS) or in any other organ, immune and/or inflammation suppression would be beneficial, Parkinson's disease, complication and/or side effects from treatment of Parkinson's disease, AIDS-related dementia complex HIV-related encephalopathy, Devic's disease, Sydenham chorea, Alzheimer's disease and other degenerative diseases, conditions or disorders of the CNS, inflammatory components of strokes, post-polio syndrome, immune and inflammatory components of psychiatric disorders, myelitis, encephalitis, subacute sclerosing pan-encephalitis, encephalomyelitis, acute neuropathy, subacute neuropathy, chronic neuropathy, Guillaim-Barre syndrome, Sydenham chora, myasthenia gravis, pseudo-tumour cerebri, Down's Syndrome, Huntington's disease, amyotrophic lateral sclerosis, inflammatory components of CNS compression or CNS trauma or infections of the CNS, inflammatory components of muscular atrophies and dystrophies, and immune and inflammatory related diseases, conditions or disorders of the central and peripheral nervous systems, post-traumatic inflammation, septic shock, infectious diseases, inflammatory complications or side effects of surgery, bone marrow transplantation or other transplantation complications and/or side effects, inflammatory and/or immune complications and side effects of gene therapy, *e.g.* due to infection with a viral carrier, or inflammation associated with AIDS, to suppress or inhibit a humoral and/or cellular immune response, to treat or ameliorate monocyte or leukocyte proliferative diseases, *e.g.* leukaemia, by reducing the amount of monocytes or lymphocytes, for the prevention and/or treatment of graft rejection in cases of transplantation of natural or artificial cells, tissue and organs such as cornea, bone marrow, organs, lenses, pacemakers, natural or artificial skin tissue.

In one preferred embodiment, the invention provides improved viral vector systems optimized to express high levels of one or more therapeutic proteins in erythroid cells or erythroid precursor cells. Retroviral vectors, including lentiviral vectors, of the invention further comprise a polynucleotide-of-interest, including, for example, a globin gene or a gene which encodes an antisickling protein. In one embodiment, the globin gene expressed in the retroviral vector of the invention is β-globin, δ-globin, or γ-globin. In another embodiment, the human β-globin gene is the wild type human β-globin gene or human β^A-globin gene. In another embodiment, the human β-globin gene comprises one or more deletions of intron sequences or is a mutated human β-globin gene encoding at least one antisickling amino acid residue. Antisickling amino acids can be derived from human δ-globin or human γ-globin. In another embodiment, the mutated human β-globin gene encodes a threonine to glutamine mutation at codon 87 (βA-T87Q).

Retroviral vectors, including lentiviral vectors, of the invention can be used in gene therapy, including for the treatment of hemoglobinopathies. In particular embodiments, the invention provides methods for using the foregoing vectors to achieve stable, high levels of gene expression in erythroid cells, *e.g.,* in order to treat erythroid-specific diseases. In a particular embodiment, the gene therapy vectors are used to treat hemoglobinopathies, including, for example, sickle cell disease (SCD). In another preferred embodiment, the gene therapy vectors are used for treatment of thalassemias, including, but not limited to, β-thalassemia.

In another preferred embodiment, vectors of the invention comprise an ABCD1 gene for treatment of adrenoleukodystropies and/or adrenomyeloneuropathies.

The various embodiments described above can be combined to provide further embodiments. All of the U.S. patents, U.S. patent application publications, U.S. patent applications, foreign patents, foreign patent applications and non-patent publications referred to in this specification and/or listed in the Application Data Sheet are incorporated herein by reference, in their entirety. Aspects of the embodiments can be modified, if necessary to employ concepts of the various patents, applications and publications to provide yet further embodiments.

These and other changes can be made to the embodiments in light of the above-detailed description. In general, in the following claims, the terms used should not be construed to limit the claims to the specific embodiments disclosed in the specification and the claims, but should be construed to include all possible embodiments along with the full scope of equivalents to which such claims are entitled. Accordingly, the claims are not limited by the disclosure.

CLAIMS

1. A mammalian packaging cell comprising:

a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to the reverse complement of one or more viral structural genes selected from the group consisting of gag, pol, and env; operably linked to an RNA dependent RNA polymerase (RDRP) promoter.

2. The mammalian packaging cell claim 1, wherein the single stranded RNA virus is a minus-strand single stranded RNA virus.

3. The mammalian packaging cell claim 1, wherein the single stranded RNA virus is a plus-strand single stranded RNA virus.

4. A mammalian packaging cell comprising:

a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to an RNA dependent RNA polymerase promoter of a single stranded RNA virus operably linked to one or more viral structural genes selected from the group consisting of gag, pol, and env; and an RNA polymerase recognition site (CSE).

5. The mammalian packaging cell claim 4, wherein the single stranded RNA virus is a plus-strand single stranded RNA virus.

6. The mammalian packaging cell of claim 3 or claim 5, wherein the plus-strand single stranded RNA virus is a narnavirus or an alphavirus.

7. The mammalian packaging cell of claim 6, wherein the alphavirus is selected from the group consisting of: Aura virus, Bebaru virus, Cabassou virus,

Chikungunya virus, Eastern equine encephalomyelitis virus, Fort Morgan virus, Getah virus, Kyzylagach virus, Mayaro virus, Middleburg virus, Mucambo virus, Ndumu virus, Pixuna virus, Ross River virus, Semliki Forest virus, Sindbis virus, Tonate virus, Triniti virus, Una virus, Venezuelan equine encephalomyelitis virus, Western equine encephalomyelitis virus, Whataroa virus, and Y-62-33 virus.

8. The mammalian packaging cell of any one of claims 1-7, wherein the first polynucleotide comprises any one of a gag gene, a pol gene, or an env gene.

9. The mammalian packaging cell of any one of claims 1-7, wherein the first polynucleotide comprises any two of a gag gene, a pol gene, or an env gene.

10. The mammalian packaging cell of any one of claims 1-7, wherein the first polynucleotide comprises a gag gene, a pol gene, and an env gene.

11. The mammalian packaging cell of any one of claims 1-7, wherein the cell comprises a second polynucleotide comprising a second promoter operably linked to a viral RNA dependent RNA polymerase promoter operably linked to a viral structural gene not encoded by the first polynucleotide and a CSE.

12. The mammalian packaging cell of any one of claims 1-7, wherein the cell comprises a second polynucleotide comprising a second promoter operably linked to the reverse complement of a viral structural gene not encoded by the first polynucleotide; operably linked to a viral RNA dependent RNA polymerase promoter.

13. The mammalian packaging cell of any one of claims 11-12, wherein the cell comprises a third polynucleotide comprising a third promoter operably linked to a viral RNA dependent RNA polymerase promoter operably linked to a viral structural gene not encoded by the first or second polynucleotides and a CSE.

14. The mammalian packaging cell of any one of claims 11-12, wherein the cell comprises a third polynucleotide comprising a third promoter operably linked to the reverse complement of a viral structural gene not encoded by the first or second polynucleotides; operably linked to a viral RNA dependent RNA polymerase promoter.

15. The mammalian packaging cell according to any one of claims 1-14, wherein the first, second, third, and fourth promoters are independently selected from the group consisting of: an constitutive promoter and an inducible promoter.

16. The mammalian packaging cell of claim 15, wherein the first promoter is an inducible promoter and one or more of the second, third, and fourth promoters are the same inducible promoter or a different inducible promoter.

17. The mammalian packaging cell of claim 15, wherein the first promoter is an constitutive promoter and one or more of the second, third, and fourth promoters are the same constitutive promoter or a different constitutive promoter.

18. The mammalian packaging cell of any one of claims 1-17, wherein the cell further comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more isolated non-structural protein genes of the single-stranded RNA virus.

19. The mammalian packaging cell of claim 14-17, wherein the constitutive promoter is selected from the group consisting of: a cytomegalovirus immediate early gene promoter (CMV), an elongation factor 1 alpha promoter (EF1-α), a phosphoglycerate kinase-1 promoter (PGK), a ubiquitin-C promoter (UBQ-C), a cytomegalovirus enhancer/chicken beta-actin promoter (CAG), polyoma enhancer/herpes simplex thymidine kinase promoter (MC1), a beta actin promoter (β-ACT), and a simian virus 40 promoter (SV40).

20. The mammalian packaging cell of claim 14-17, wherein the inducible promoter is selected from the group consisting of: a tetracycline responsive promoter, an ecdysone responsive promoter, a cumate responsive promoter, a glucocorticoid responsive promoter, and estrogen responsive promoter, and an RU-486 responsive promoter.

21. The mammalian packaging cell of any one of claims 1-20, wherein the cell is contacted with a composition comprising one or more isolated non-structural protein polypeptides of the single-stranded RNA virus.

22. The mammalian packaging cell of claim 21, wherein the one or more non-structural protein polypeptides comprise a cell permeable peptide domain.

23. The mammalian packaging cell of claim 22, wherein the cell permeable peptide domain is selected from the group consisting of: RKKRRQRRR, KKRRQRRR, RKKRRQRR, RRRRRRRRR, KKKKKKKKK, RQIKIWFQNRRMKWKK, RQIKIWFQNRRMKSKK, RQIKIWFQNKRAKIKK, RQIKIWFQNRRMKWKK, and RVIRVWFQNKRCKDKK.

24. The mammalian packaging cell of any one of claims 1-23, wherein the one or more viral structural genes are independently isolated from a virus selected from the group consisting of: human immunodeficiency virus type 1 (HIV-1), human immunodeficiency virus type 2 (HIV-2), visna virus, caprine arthritis-encephalitis virus (CAEV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), bovine immune deficiency virus (BIV), and simian immunodeficiency virus (SIV).

25. The mammalian packaging cell of any one of claims 1-23, wherein the wherein the gag and pol viral structural genes are independently selected from the group consisting of: human immunodeficiency virus type 1 (HIV-1), human

immunodeficiency virus type 2 (HIV-2), visna virus, caprine arthritis-encephalitis virus (CAEV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), bovine immune deficiency virus (BIV), and simian immunodeficiency virus (SIV).

26. The mammalian packaging cell of any one of claims 1-25, wherein the wherein the pol viral structural gene is isolated from the caprine arthritis-encephalitis virus (CAEV).

27. The mammalian packaging cell of any one of claims 1-26, wherein the wherein the env viral structural gene encodes an envelope polypeptide selected from a virus selected from the group consisting of FIV, HIV, VSV, MoMLV, GALV, JSRV, LCMV, Marburg virus, Mokola virus, Ebola virus, Rabies virus, Ross River Virus, Sendai virus, foul plague virus, influenza virus, and Lagos-bat virus.

28. The mammalian packaging cell of any one of claims 1-27, wherein the cell is selected from the group consisting of: CHO, BHK, MDCK, 10T1/2, WEHI cells, COS, BSC 1, BSC 40, BMT 10, VERO, W138, MRC5, A549, HT1080, 293, 293T, B-50, 3T3, NIH3T3, HepG2, Saos-2, Huh7, HeLa cells W163, HeLa, Vero, 211, and 211A.

29. A mammalian packaging cell comprising:

a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to an alphavirus 26S subgenomic promoter operably linked to one or more viral structural genes selected from the group consisting of gag, pol, and env; and a CSE.

30. A mammalian packaging cell comprising:

a cell having a first polynucleotide comprising a first promoter operable in mammalian cells operably linked to the reverse complement one or more viral

structural genes selected from the group consisting of gag, pol, and env; operably linked to an alphavirus 26S subgenomic promoter.

31. The mammalian packaging cell of any one of claims 29-30, wherein the cell further comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more alphavirus non-structural protein genes.

32. The mammalian packaging cell of any one of claims 29-30, wherein the cell is contacted with a composition comprising one or more alphavirus non-structural protein polypeptides.

33. The mammalian packaging cell of claim 32, wherein the one or more non-structural protein polypeptides comprise a cell permeable peptide domain.

34. The mammalian packaging cell of claim 33, wherein the cell permeable peptide domain is selected from the group consisting of: RKKRRQRRR, KKRRQRRR, RKKRRQRR, RRRRRRRRR, KKKKKKKKK, RQIKIWFQNRRMKWKK, RQIKIWFQNRRMKSKK, RQIKIWFQNKRAKIKK, RQIKIWFQNRRMKWKK, and RVIRVWFQNKRCKDKK.

35. The mammalian packaging cell of any one of claims 31-34, wherein the one or more non-structural proteins is selected from the group consisting of: NSP1, NSP2, NSP3, and NSP4.

36. The mammalian packaging cell of any one of claims 31-35, wherein the one or more non-structural proteins are NSP2 and NSP4.

37. The mammalian packaging cell of any one of claims 31-36, wherein the non-structural protein is NSP4.

38. A producer cell comprising:

a) the mammalian packaging cell of any one of claims 1-28,

wherein the cell is contacted with a composition comprising one or more non-structural proteins of the single-stranded RNA virus or,

wherein the cell comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more non-structural genes isolated from the single-stranded RNA virus; and

b) a transfer vector comprising a replication defective lentiviral genome, a packaging signal, and a mammalian promoter operably linked to a mammalian polypeptide;

wherein the mammalian cell is cultured for a time sufficient to package the lentiviral genome and produce lentiviral particles.

39. A producer cell comprising:

a) the mammalian packaging cell of any one of claims 29-37,

wherein the cell is contacted with a composition comprising one or more non-structural proteins of the alphavirus or,

wherein the cell comprises a polynucleotide comprising a constitutive or inducible promoter operably linked to one or more non-structural genes isolated from the alphavirus; and

b) a transfer vector comprising a replication defective lentiviral genome, a packaging signal, and a mammalian promoter operably linked to a mammalian polypeptide;

wherein the mammalian cell is cultured for a time sufficient to package the lentiviral genome and produce lentiviral particles.

40. A method of producing a lentiviral particle comprising:

a) culturing a mammalian cell that expresses lentiviral gag and pol genes, a heterologous envelope gene,

wherein expression of one or more of the gag, pol, and env genes is dependent upon expression of one or more non-structural proteins of a single stranded RNA virus in the cell;

b) introducing a polynucleotide into the mammalian cell, the polynucleotide comprising a replication defective lentiviral genome; and

c) culturing the mammalian cell for a time sufficient to package the lentiviral genome and produce lentiviral particles.

41. A lentiviral vector comprising:

a) a 5' LTR;

b) a packaging signal;

c) a promoter operative in mammalian cells;

d) an RNA dependent RNA polymerase promoter of a single stranded RNA virus operably linked to a polynucleotide encoding a mammalian polypeptide; and

e) a 3' LTR.

42. The vector of claim 41, wherein single stranded RNA virus is a plus-strand single stranded RNA virus.

43. A lentiviral vector comprising

a) a 5' LTR;

b) packaging signal;

c) a mammalian promoter;

d) the reverse complement sequence of a polynucleotide encoding a mammalian polypeptide operably linked to an RNA dependent RNA polymerase promoter of a single stranded RNA virus; and

e) a 3' LTR.

44. A lentiviral vector comprising:

a) a 5' LTR;

b) packaging signal;

c) a mammalian promoter;

d) an RNA dependent RNA polymerase promoter of a single stranded RNA virus operably linked to a polynucleotide encoding a mammalian polypeptide;

e) an RNA polymerase recognition site (CSE); and

f) a 3′ LTR.

45. The vector of claim 43 or 44, wherein single stranded virus is a plus-strand single stranded RNA virus or a minus-strand single stranded RNA virus.

46. The vector of claim 45, wherein the RNA dependent RNA polymerase promoter comprises the alphavirus 26S subgenomic promoter.

47. The vector of any one of claims 41-46, wherein the U3 region of the 5' LTR comprises a heterologous transcriptional regulatory element.

48. The vector of claim 47, wherein the heterologous transcriptional regulatory element is selected from the group consisting of: a cytomegalovirus enhancer, a cytomegalovirus promoter, a cytomegalovirus enhancer and promoter, a Rous sarcoma virus enhancer, a Rous sarcoma virus promoter, and a Rous sarcoma virus enhancer and promoter.

49. The vector of any one of claims 41-46, wherein promoter operative in mammalian cells is selected from the group consisting of: a ubiquitous promoter, a tissue specific promoter, and an inducible promoter.

50. The vector of claim 49, wherein the ubiquitous promoter is selected from the group consisting of: a cytomegalovirus immediate early gene promoter (CMV), an elongation factor 1 alpha promoter (EF1-α), a phosphoglycerate kinase-1 promoter (PGK), a ubiquitin-C promoter (UBQ-C), a cytomegalovirus enhancer/chicken beta-actin promoter (CAG), polyoma enhancer/herpes simplex thymidine kinase promoter (MC1), a beta actin promoter (β-ACT), and a simian virus 40 promoter (SV40).

51. The vector of claim 49, wherein the tissue specific promoter expresses the polynucleotide encoding the mammalian polypeptide in ectodermal cells, endodermal cells, mesodermal cells, or mesenchymal cells.

52. The vector of claim 49, wherein the inducible promoter is selected from the group consisting of: a tetracycline responsive promoter, an ecdysone responsive promoter, a cumate responsive promoter, a glucocorticoid responsive promoter, and estrogen responsive promoter, and an RU-486 promoter.

53. The vector of any one of claims 41-52, wherein the lentivirus is selected from the group consisting of: human immunodeficiency virus type 1 (HIV-1), human immunodeficiency virus type 2 (HIV-2), visna virus, caprine arthritis-encephalitis virus (CAEV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), bovine immune deficiency virus (BIV), and simian immunodeficiency virus (SIV).

54. The vector of claim 46, wherein the alphavirus is selected from the group consisting of: Aura virus, Bebaru virus, Cabassou virus, Chikungunya virus, Eastern equine encephalomyelitis virus, Fort Morgan virus, Getah virus, Kyzylagach virus, Mayaro virus, Middleburg virus, Mucambo virus, Ndumu virus, Pixuna virus, Ross River virus, Semliki Forest virus, Sindbis virus, Tonate virus, Triniti virus, Una virus, Venezuelan equine encephalomyelitis virus, Western equine encephalomyelitis virus, Whataroa virus, and Y-62-33 virus.

55. The vector of any one of claims 41-54, wherein the mammalian polypeptide is adrenoleukodystrophy protein (ABCD1).

56. The vector of any one of claims 41-54, wherein the mammalian polypeptide is β-globin.

57. The vector of any one of claims 41-56, wherein one or more nucleotides of the U3 region of the 3' LTR are deleted.

58. The vector of any one of claims 41-57, wherein vector comprises one or more vector sequences selected from the group consisting of: a central polypurine tract and the central termination sequence (cPPT/CTS), a woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), hepatitis B virus posttranscriptional regulatory element (HPRE), and mammalian polyadenylation sequence 3' of the 3' R region.

59. The vector of any one of claims 41-58, wherein vector comprises one or more recombinase recognition sites selected from the group consisting of: LoxP, Lox511, Lox5171, Lox2272, m2, Lox71, Lox66, FRT, F_1, F_2, F_3, F_4, F_5, FRT(LE), FRT(RE), attB, attP, attL, and attR.