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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 September 2021 (23.09.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/188996 A1

(51) International Patent Classification:
C07K 14/005 (2006.01) *C12N 9/50* (2006.01)
C07K 14/47 (2006.01) *C12N 15/11* (2006.01)
C12N 7/00 (2006.01)

(21) International Application Number:
PCT/US2021/023318

(22) International Filing Date:
19 March 2021 (19.03.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/992,400 20 March 2020 (20.03.2020) US

(71) Applicants: **THE BROAD INSTITUTE, INC.** [US/US]; 415 Main Street, Cambridge, Massachusetts 02142 (US). **MASSACHUSETTS INSTITUTE OF TECHNOLOGY** [US/US]; 77 Massachusetts Avenue, Cambridge, Massachusetts 02139 (US).

(72) Inventors: **ZHANG, Feng**; c/o 415 Main Street, Cambridge, Massachusetts 02142 (US). **LADHA, Alim**; c/o 77 Massachusetts Avenue, Cambridge, Massachusetts 02139 (US).

(74) Agent: **MILLER, Carin R.** et al.; Johnson, Marcou, Isaacs & Nix LLC, Customer No. 95280, PO Box 691, Hoschton, Georgia 30548 (US).

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

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

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: COMPOSITIONS AND METHODS FOR ENHANCED LENTIVIRAL PRODUCTION

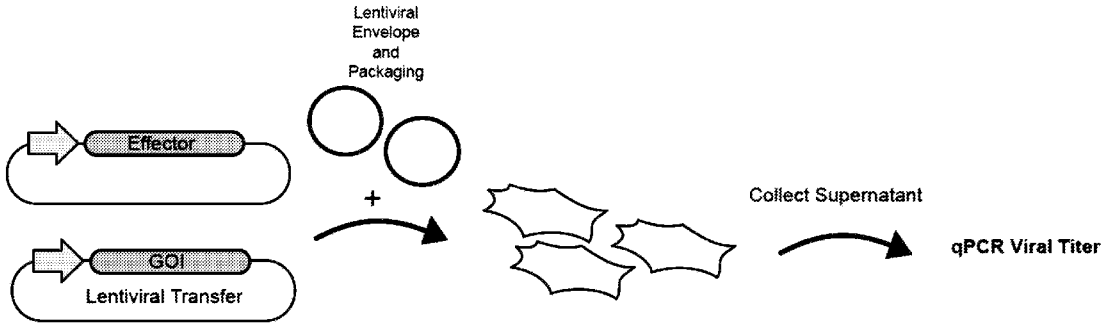

FIG. 1A

(57) Abstract: Described herein are engineered retroviral delivery vesicle generation compositions, systems, and methods to deliver a cargo to a cell. The engineered compositions, systems, and method include one or more polynucleotides encoding one or more elements for forming a delivery vesicle and optionally one or more cargos.

WO 2021/188996 A1

COMPOSITIONS AND METHODS FOR ENHANCED LENTIVIRAL PRODUCTION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 62/992,400, filed March 20, 2020. The entire contents of the above-identified applications are hereby fully incorporated herein by reference.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with government support under Grant No. HL141201 awarded by National Institutes of Health. The government has certain rights in the invention.

SEQUENCE LISTING

[0003] This application contains a sequence listing filed in electronic form as an ASCII.txt file entitled BROD-5075WP_ST25.txt, created on March 19, 2021 and having a size of 90,459 bytes. The content of the sequence listing is incorporated herein in its entirety.

TECHNICAL FIELD

[0004] The subject matter disclosed herein is generally directed to engineered delivery vesicles.

BACKGROUND

[0005] Recombinant viral vectors are widely used gene delivery tools for cells, animal models, and clinical applications, due to their propensity to infect most cells and tissues, and for enhanced safety. Lentiviral vectors offer several advantages over other gene delivery vectors. Their potential for the integration and long-term expression of therapeutic genes renders them an interesting tool for gene and cell therapy interventions.

[0006] While the concept of using viral vectors to deliver a heterologous gene to a recipient cell is well known (Verma and Somia (1997) Nature 389: 239-242), it is widely accepted that there are limits to the size of the heterologous gene that can be (see, for example, page 446, Chapter 9 of Coffin et al., "Retroviruses"; 1997 Cold Spring Harbor Laboratory Press). If the incorporation of the heterologous gene and the associated regulatory elements drastically increases the size of the viral genome, then there is a significant risk of it being unable to be

successfully packaged or at least that the packaging efficiency is significantly reduced. Accordingly, large scale and/or high yield production of lentiviral particles remains an important challenge for the translation of lentiviral-based therapeutic strategies to the clinic. The development of robust and scalable processes for mass and high yield production of lentiviruses is needed.

[0007] Citation or identification of any document in this application is not an admission that such a document is available as prior art to the present invention.

SUMMARY

[0008] Described in certain example embodiments herein are engineered retroviral delivery vesicle generation system comprising (a) one or more polynucleotides encoding one or more effectors; (b) one or more polynucleotides encoding a cargo; and (c) one or more polynucleotides encoding one or more packaging elements, one or more vesicle elements, or both.

[0009] In certain example embodiments, the system is capable of generating 1 to 50 or more fold more lentivirus particles as compared to a system lacking the one or more polynucleotides encoding one or more effectors.

[0010] In certain example embodiments, the system is capable of generating 1 to 50 or more fold more vesicles as compared to a system lacking the one or more polynucleotides encoding one or more effectors.

[0011] In certain example embodiments, the one or more effectors are retrotransposon-derived genes.

[0012] In certain example embodiments, the one or more effectors are PEG10, RTL1, or both.

[0013] In certain example embodiments, the one or more polynucleotides encoding PEG10 is 80 percent to 100 percent identical to SEQ ID NO: 1.

[0014] In certain example embodiments, the one or more polynucleotides encoding RTL1 is 80 percent to 100 percent identical to SEQ ID NO: 2.

[0015] In certain example embodiments, the system is a retroviral system.

[0016] In certain example embodiments, the retroviral system is a lentiviral system.

[0017] In certain example embodiments, (a), (b), and (c) are included in one or more vectors comprising one or more regulatory elements, wherein each of the one or more polynucleotides of (a), (b), and (c) are optionally operably coupled a regulatory element.

[0018] In certain example embodiments, (a), (b), (c), or any combination thereof are included on the same vector or are included in different vectors, or any permissible combination thereof.

[0019] Described in certain example embodiments herein are engineered retroviral delivery vesicle generation systems comprising (a) one or more polynucleotides encoding one or more retroviral polypeptides capable of forming a delivery vesicle and encapsulating one or more cargos therein, wherein at least one of the one or more polypeptides is an effector; and (b) one or more cargos, wherein the one or more cargos are optionally polynucleotide cargos.

[0020] In certain example embodiments, the system is capable of generating 1 to 50 or more fold more delivery vesicles containing one or more cargos as compared to a system lacking the one or more polynucleotides encoding one or more effectors.

[0021] In certain example embodiments, the effector is a retrotransposon-derived effector.

[0022] In certain example embodiments, the effector is PEG10, RTL1, or both.

[0023] In certain example embodiments, PEG10 is encoded by a polynucleotide that is to 80 to 100 percent identical to SEQ ID NO: 1.

[0024] In certain example embodiments, RTL1 is encoded by a polynucleotide that is to 80 to 100 percent identical to SEQ ID NO: 2.

[0025] In certain example embodiments, the one or more retroviral peptides are one or more lentiviral peptides.

[0026] In certain example embodiments, the one or more retroviral peptides comprises one or more packaging elements, vesicle elements, or both.

[0027] In certain example embodiments, (a), (b), optionally (c), or a combination thereof are included in one or more vectors comprising one or more regulatory elements, wherein each of the one or more polynucleotides of (a) and (b)are optionally operably coupled a regulatory element.

[0028] In certain example embodiments, (a), (b), optionally (c), or a combination thereof are included on the same vector or are included on different vectors.

[0029] In certain example embodiments, the one or more cargos comprise one or more packaging elements.

[0030] Described in certain exemplary embodiments herein are methods of generating engineered retroviral delivery vesicles loaded with one or more cargos comprising expressing one or more components of the engineered retroviral delivery vesicle systems as described anywhere herein in one or more suitable bioreactors under conditions such that engineered retroviral delivery vesicles are formed and wherein one or more of the engineered retroviral delivery vesicles contains one or more cargos.

[0031] In certain example embodiments, the one or more suitable bioreactors are cells.

[0032] Described in certain exemplary embodiments herein are engineered retroviral delivery vesicles with one or more cargos generated according a method of generating engineered retroviral delivery vesicles described herein.

[0033] In certain example embodiments, the engineered retroviral delivery vesicle is a lentiviral particle.

[0034] In certain example embodiments, the one or more cargos is/are polynucleotide cargos.

[0035] Described in certain example embodiments herein are co-culture systems comprising two or more cell types wherein at least one, all, or a sub-combination of cell types comprise an engineered retroviral vesicle generation system described herein.

[0036] Described in certain example embodiments herein are methods of cellular delivery of one or more cargos comprising (a) delivering an engineered retroviral delivery vesicle generation system described herein to a donor cell type, wherein expression of the engineered retroviral delivery vesicle generation system in the donor cell type results in the generation of one or more retroviral delivery vesicles comprising one or more cargos; and

(b) delivery to or uptake of one or more engineered retroviral delivery vesicles generated in (a) to or by a recipient cell.

[0037] In certain example embodiments, the method occurs in vivo, in vitro, or ex vivo.

[0038] In certain example embodiments, the recipient cell is diseased or pathogenic.

[0039] In certain example embodiments, the recipient cell is a eukaryote or a prokaryote.

[0040] Described in certain example embodiments herein are cells comprising and capable of expressing an engineered retroviral delivery system as described herein.

[0041] Described in certain example embodiments herein are pharmaceutical formulations comprising an engineered retroviral delivery system as described herein; a engineered retroviral delivery vesicle as described herein; a cell as described herein; or any combination thereof; and a pharmaceutically acceptable carrier.

[0042] Described in certain example embodiments herein are kits comprising an engineered retroviral delivery system as described herein; a delivery vesicle as described herein; a cell as described herein; a pharmaceutical formulation as described herein or any combination thereof

[0043] These and other aspects, objects, features, and advantages of the example embodiments will become apparent to those having ordinary skill in the art upon consideration of the following detailed description of example embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0044] An understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention may be utilized, and the accompanying drawings of which:

[0045] **FIGS. 1A-1B** show **(FIG. 1A)** a schematic for a system for producing lentiviruses and **(FIG. 1B)** a graph showing improvement of yield of lentiviruses from cells using effector genes PEG10 and RTL1.

[0046] The figures herein are for illustrative purposes only and are not necessarily drawn to scale.

DETAILED DESCRIPTION OF THE EXAMPLE EMBODIMENTS

General Definitions

[0047] Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure pertains. Definitions of common terms and techniques in molecular biology may be found in Molecular Cloning: A Laboratory Manual, 2nd edition (1989) (Sambrook, Fritsch, and Maniatis); Molecular Cloning: A Laboratory Manual, 4th edition (2012) (Green and Sambrook); Current Protocols in Molecular Biology (1987) (F.M. Ausubel et al. eds.); the series Methods in Enzymology (Academic Press, Inc.): PCR 2: A Practical Approach (1995)

(M.J. MacPherson, B.D. Hames, and G.R. Taylor eds.): Antibodies, A Laboratory Manual (1988) (Harlow and Lane, eds.): Antibodies A Laboratory Manual, 2nd edition 2013 (E.A. Greenfield ed.); Animal Cell Culture (1987) (R.I. Freshney, ed.); Benjamin Lewin, Genes IX, published by Jones and Bartlet, 2008 (ISBN 0763752223); Kendrew *et al.* (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0632021829); Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 9780471185710); Singleton *et al.*, Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, N.Y. 1992); and Marten H. Hofker and Jan van Deursen, Transgenic Mouse Methods and Protocols, 2nd edition (2011)

.

[0048] As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

[0049] The term "optional" or "optionally" means that the subsequent described event, circumstance or substituent may or may not occur, and that the description includes instances where the event or circumstance occurs and instances where it does not.

[0050] The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

[0051] The terms "about" or "approximately" as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, are meant to encompass variations of and from the specified value, such as variations of +/-10% or less, +/-5% or less, +/-1% or less, and +/-0.1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. It is to be understood that the value to which the modifier "about" or "approximately" refers is itself also specifically, and preferably, disclosed.

[0052] As used herein, a "biological sample" may contain whole cells and/or live cells and/or cell debris. The biological sample may contain (or be derived from) a "bodily fluid". The present invention encompasses embodiments wherein the bodily fluid is selected from amniotic fluid, aqueous humour, vitreous humour, bile, blood serum, breast milk, cerebrospinal fluid, cerumen (earwax), chyle, chyme, endolymph, perilymph, exudates, feces, female ejaculate, gastric acid, gastric juice, lymph, mucus (including nasal drainage and phlegm),

pericardial fluid, peritoneal fluid, pleural fluid, pus, rheum, saliva, sebum (skin oil), semen, sputum, synovial fluid, sweat, tears, urine, vaginal secretion, vomit and mixtures of one or more thereof. Biological samples include cell cultures, bodily fluids, cell cultures from bodily fluids. Bodily fluids may be obtained from a mammal organism, for example by puncture, or other collecting or sampling procedures.

[0053] The terms "subject," "individual," and "patient" are used interchangeably herein to refer to a vertebrate, preferably a mammal, more preferably a human. Mammals include, but are not limited to, murines, simians, humans, farm animals, sport animals, and pets. Tissues, cells and their progeny of a biological entity obtained in vivo or cultured in vitro are also encompassed.

[0054] Various embodiments are described hereinafter. It should be noted that the specific embodiments are not intended as an exhaustive description or as a limitation to the broader aspects discussed herein. One aspect described in conjunction with a particular embodiment is not necessarily limited to that embodiment and can be practiced with any other embodiment(s). Reference throughout this specification to "one embodiment", "an embodiment," "an example 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, appearances of the phrases "in one embodiment," "in an embodiment," or "an example embodiment" in various places throughout this specification are not necessarily all referring to the same embodiment, but may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to a person skilled in the art from this disclosure, in one or more embodiments. Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the invention. For example, in the appended claims, any of the claimed embodiments can be used in any combination.

[0055] All publications, published patent documents, and patent applications cited herein are hereby incorporated by reference to the same extent as though each individual publication, published patent document, or patent application was specifically and individually indicated as being incorporated by reference.

OVERVIEW

[0056] Embodiments disclosed herein provide delivery vesicle generation systems capable of generating delivery vesicles for delivering a cargo.

ENGINEERED RETROVIRAL DELIVERY VESICLE GENERATION SYSTEMS

[0057] Described herein are engineered retroviral delivery vesicle generation systems that are capable of generating engineered retroviral delivery vesicles that can be loaded with a cargo for delivery to a cell. In some embodiments, the engineered retroviral delivery vesicle generation system can be composed of one or more polynucleotides that encode one or more effectors, optionally one or more polynucleotides that encode a cargo, and one or more polynucleotides that encode one or more packaging elements, one or more vesicle elements, or both. In some embodiments, the engineered system can include a polynucleotide encoding an endogenous retroviral polypeptide. For example, the polynucleotide may be a vector suitable for delivery to a cell or other bioreactor system that can facilitate expression of the retroviral polypeptide leading to delivery vesicle formation. As provided in further detail below, a variety of cargo molecules may be packaged within the delivery vesicles disclosed herein. The cargo molecule may be modified with one or more packaging elements that complex or bind to the retroviral polypeptide and facilitate packaging of the cargo molecule into the delivery vesicle. While the term "cargo molecule" or simply "cargo" is referred to in the singular, it is contemplated that multiple copies of a cargo molecule, depending on type and other readily recognizable size constraints of the delivery vesicle, may be packaged within a single delivery vesicle. In some embodiments, the engineered delivery vesicle generation system is a vector system.

[0058] In some embodiments, the invention provides an engineered retroviral system comprising a first polynucleotide comprising a gene of interest; a second polynucleotide comprising one or more effector genes; a third polynucleotide comprising genes encoding packaging proteins; and a fourth polynucleotide encoding an envelope protein. In some embodiments, the retroviral system may be a lentiviral system. Lentiviruses are becoming an increasingly popular choice of gene transfer vehicle for use in the treatment of a variety of genetic and acquired human diseases. As research progresses from basic studies into pre-clinical and clinical phases, there is a growing demand for large volumes of high purity, concentrated vector, and accordingly, the means to produce such quantities. Unlike other viral vectors, lentiviruses are difficult to produce using stable cell lines, therefore transient

transfection of adherent cell lines is conventionally used, and this method has proven challenging to up-scale. A major obstacle in obtaining usable lentivirus is producing a high enough concentration, or titer. This is often due to inefficient packaging of large inserts. As is demonstrated in the Working Examples, addition of the one or more effector genes can, in some embodiments, increase the efficiency of packaging and thus increase yield of viral particles produced by the system. In certain embodiments, the packaging proteins are encoded in more than one polynucleotide, such as in two, three, or more polynucleotides. In certain embodiments, the viral transactivator, tat, is not needed.

[0059] The engineered retroviral delivery vesicle generation systems can result in an increased engineered retroviral delivery vesicle production as compared to conventional systems as is described in greater detail elsewhere herein. In some embodiments, the engineered retroviral delivery vesicle system is capable of generating 1 to 50 or more fold more engineered retroviral delivery vesicles, such as engineered lentiviral particles (or vesicles), as compared to a system lacking the one or more polynucleotides encoding one or more effectors. Thus, incorporation of the effector(s) in the engineered delivery generation system can provide a method of improving yield of retroviral delivery vesicles produced therefrom. As used herein, the term "improved yield" means an effective amount of a retroviral vector or particle that is capable of transducing a target site, such as a cell. As used herein, the term "effective amount" means an amount of a vector or regulated retroviral or lentiviral vector particle which is sufficient to induce gene of interest expression at a target site.

[0060] In some embodiments, yield of engineered retroviral delivery vesicles is increased 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9, 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10, 10.1, 10.2, 10.3, 10.4, 10.5, 10.6, 10.7, 10.8, 10.9, 11, 11.1, 11.2, 11.3, 11.4, 11.5, 11.6, 11.7, 11.8, 11.9, 12, 12.1, 12.2, 12.3, 12.4, 12.5, 12.6, 12.7, 12.8, 12.9, 13, 13.1, 13.2, 13.3, 13.4, 13.5, 13.6, 13.7, 13.8, 13.9, 14, 14.1, 14.2, 14.3, 14.4, 14.5, 14.6, 14.7, 14.8, 14.9, 15, 15.1, 15.2, 15.3, 15.4, 15.5, 15.6, 15.7, 15.8, 15.9, 16, 16.1, 16.2, 16.3, 16.4, 16.5, 16.6, 16.7, 16.8, 16.9, 17, 17.1, 17.2, 17.3, 17.4, 17.5, 17.6, 17.7, 17.8, 17.9, 18, 18.1, 18.2, 18.3, 18.4, 18.5, 18.6, 18.7, 18.8, 18.9, 19, 19.1, 19.2, 19.3, 19.4, 19.5, 19.6, 19.7, 19.8, 19.9, 20, 20.1, 20.2, 20.3, 20.4, 20.5, 20.6, 20.7, 20.8, 20.9, 21, 21.1, 21.2, 21.3, 21.4, 21.5, 21.6, 21.7, 21.8,

21.9, 22, 22.1, 22.2, 22.3, 22.4, 22.5, 22.6, 22.7, 22.8, 22.9, 23, 23.1, 23.2, 23.3, 23.4, 23.5, 23.6, 23.7, 23.8, 23.9, 24, 24.1, 24.2, 24.3, 24.4, 24.5, 24.6, 24.7, 24.8, 24.9, 25, 25.1, 25.2, 25.3, 25.4, 25.5, 25.6, 25.7, 25.8, 25.9, 26, 26.1, 26.2, 26.3, 26.4, 26.5, 26.6, 26.7, 26.8, 26.9, 27, 27.1, 27.2, 27.3, 27.4, 27.5, 27.6, 27.7, 27.8, 27.9, 28, 28.1, 28.2, 28.3, 28.4, 28.5, 28.6, 28.7, 28.8, 28.9, 29, 29.1, 29.2, 29.3, 29.4, 29.5, 29.6, 29.7, 29.8, 29.9, 30, 30.1, 30.2, 30.3, 30.4, 30.5, 30.6, 30.7, 30.8, 30.9, 31, 31.1, 31.2, 31.3, 31.4, 31.5, 31.6, 31.7, 31.8, 31.9, 32, 32.1, 32.2, 32.3, 32.4, 32.5, 32.6, 32.7, 32.8, 32.9, 33, 33.1, 33.2, 33.3, 33.4, 33.5, 33.6, 33.7, 33.8, 33.9, 34, 34.1, 34.2, 34.3, 34.4, 34.5, 34.6, 34.7, 34.8, 34.9, 35, 35.1, 35.2, 35.3, 35.4, 35.5, 35.6, 35.7, 35.8, 35.9, 36, 36.1, 36.2, 36.3, 36.4, 36.5, 36.6, 36.7, 36.8, 36.9, 37, 37.1, 37.2, 37.3, 37.4, 37.5, 37.6, 37.7, 37.8, 37.9, 38, 38.1, 38.2, 38.3, 38.4, 38.5, 38.6, 38.7, 38.8, 38.9, 39, 39.1, 39.2, 39.3, 39.4, 39.5, 39.6, 39.7, 39.8, 39.9, 40, 40.1, 40.2, 40.3, 40.4, 40.5, 40.6, 40.7, 40.8, 40.9, 41, 41.1, 41.2, 41.3, 41.4, 41.5, 41.6, 41.7, 41.8, 41.9, 42, 42.1, 42.2, 42.3, 42.4, 42.5, 42.6, 42.7, 42.8, 42.9, 43, 43.1, 43.2, 43.3, 43.4, 43.5, 43.6, 43.7, 43.8, 43.9, 44, 44.1, 44.2, 44.3, 44.4, 44.5, 44.6, 44.7, 44.8, 44.9, 45, 45.1, 45.2, 45.3, 45.4, 45.5, 45.6, 45.7, 45.8, 45.9, 46, 46.1, 46.2, 46.3, 46.4, 46.5, 46.6, 46.7, 46.8, 46.9, 47, 47.1, 47.2, 47.3, 47.4, 47.5, 47.6, 47.7, 47.8, 47.9, 48, 48.1, 48.2, 48.3, 48.4, 48.5, 48.6, 48.7, 48.8, 48.9, 49, 49.1, 49.2, 49.3, 49.4, 49.5, 49.6, 49.7, 49.8, 49.9 to/or 50 fold as compared to a system lacking the one or more polynucleotides encoding one or more effectors.

[0061] Retroviral vector systems have been proposed as a delivery system for the transfer of a gene of interest to one or more sites of interest. The transfer may occur in vitro, ex vivo, in vivo or combinations thereof. Retroviral vector systems were even exploited to study various aspects of the retrovirus life cycle, including receptor use, reverse transcription and RNA packaging (reviewed by Miller, 1992 Curr Top Microbiol Immunol 158: 1-24).

[0062] For the present application, the term “retrovirus” includes: murine leukemia virus (MLV), human immunodeficiency virus (HIV), equine infectious anemia virus (EIAV), mouse mammary tumor virus (MMTV), Rous sarcoma virus (RSV), sarcoma virus Moloney murine sarcoma virus (Mo-MSV), Abelson murine leukemia virus (A-MLV), Moloney murine leukemia virus (Mo-MLV), Moloney murine leukemia virus (Mo- Avian-29 myelocitomatosis virus (MC29) and Avian erythroblastosis virus (AEV) and all other retroviruses including lentiviruses. A detailed list of retroviruses can be found in Coffin et al. (1997 Cold Spring Harbor Laboratory Ed: Coffin JM, SM Hughes, HE Varmus pp 758-763).

[0063] Lentiviruses also belong to the retrovirus family, but they may infect both dividing and non-dividing cells (Lewis et al. (1992) EMBO J. 3053-3058). The lentivirus group can be divided into "primate" and "non-primate". Examples of primate lentiviruses include human immunodeficiency virus (HIV), the causative agent of human acquired immunodeficiency syndrome (AIDS) and simian immunodeficiency virus (SIV). The non-primate lentiviral group includes the slow virus visna/maedi virus (VMV) prototype, as well as related caprine arthritis encephalitis virus (CAEV), equine infectious anemia virus (EIAV) and virus of the most recently described feline immunodeficiency virus (FIV) and bovine immunodeficiency virus (BIV).

[0064] Details of the genomic structure of some lentiviruses can be found in the art. By way of example, details on HIV and EIAV (i.e., respectively, Genome Accession No. AF033819 and AF033820) can be found in the NCBI Genbank database. Details of HIV variants can also be found at http://hiv-web.lanl.gov. Details of EIAV variants can be found at http://www.ncbi.nlm.nih.gov.

Effectors

[0065] The engineered delivery vesicle generation system includes one or more polynucleotides that encode one or more effectors. In this context herein, "effector" refers to a polynucleotide that encodes for a molecule that selectively binds to a protein or nucleic acid and regulates its biological activity. As such, effector molecules act as ligands that can increase or decrease enzyme activity, gene expression, or cell signaling. In the context of the present invention, the effector gene enhances the packaging efficiency of lentiviral particles and in some embodiments increases the yield of viral particles produced by the system as previously described. Effector molecules can also directly regulate the activity of some mRNA molecules. In some embodiments, effector genes used herein may include, but are not necessarily limited to, the Arc family of genes, the PNM family of genes, the RTL family of genes, the PEG family of genes, the ZCC family of genes, the ZCH family of genes, the MOAP family of genes, or the CCDC family of genes. Specifically, such genes may include, but are not necessarily limited to, ZCC18, ZCH12, PNM8B, PNM8B, PNM6A, PMA6F, PMA6E, PNMA2, PNM8A, PNMA3, PNMA5, PNMA1, MOAP1, and CCDC8. In specific embodiments, the one or more effector genes comprise PEG10, RTL1, or a combination thereof. In some embodiments, the effector is Arc. In some embodiments, the Arc is hARC or dARC1.

[0066] PEG10 is a paternally expressed imprinted gene that is expressed in adult and embryonic tissues. Most notable expression occurs in the placenta. This gene is highly conserved across mammalian species and retains the heptanucleotide (GGGAAAC). PEG10 has been reported to play a role in cell proliferation, differentiation and apoptosis. This gene includes two overlapping reading frames of the same transcript encoding distinct isoforms. The shorter isoform has a CCHC-type zinc finger motif containing a sequence characteristic of gag proteins of most retroviruses and some retrotransposons, and it functions in part by interacting with members of the TGF-beta receptor family. The longer isoform has the active-site DSG consensus sequence of the protease domain of pol proteins. The longer isoform is the result of -1 translational frameshifting that is also seen in some retroviruses. Expression of these two isoforms only comes from the paternal allele due to imprinting. Increased gene expression (as observed by an increase in mRNA levels) is associated with hepatocellular carcinomas.

[0067] RTL1 is a retrotransposon derived protein coding gene. It is also known as PEG11 and is a paternally expressed imprinted gene, part of genomic imprinting. RTL1 plays an important role in the maintenance of fetal capillaries and is expressed in high quantities during late stage of fetal development. The expression of this gene is important for the development of the placenta, the fetus-maternal interface. Because the placenta is the first organ to form during the development of an embryo, problems in its establishment and biological role lead to complications during gestation. This organ maintains the fetus throughout the pregnancy and is therefore sensitive to disruptions. Studies in mice suggest that disruption of the RTL1 concentration, whether increasing or decreasing the amount of this protein coding gene, can lead to serious errors in the conservation of placental fetal capillaries. RTL1 knockout mice have shown obstruction in fetal development along with late fetal/neonatal death. Studies from sheep homologs suggest that high expression levels of RTL1 can lead to skeletal muscle hypertrophy, due to over-expression patterns in the paternal allele specific gene.

[0068] In some embodiments, the one or more effector genes is a retrotransposon-derived gene. Retrotransposons are genetic elements that can amplify themselves in a genome and are ubiquitous components of the DNA of many eukaryotic organisms. These DNA sequences use a "copy-and-paste" mechanism, whereby they are first transcribed into RNA, then converted back into identical DNA sequences using reverse transcription, and these sequences are then inserted into the genome at target sites. As such, retrotransposons sometimes originate from endogenous retroviral elements that are integrated into a host genome and are expressed from

that host genome. In some embodiments, the endogenous retroviral element is an endogenous retroviral gag protein. In some embodiments, the endogenous retroviral element is an endogenous retroviral envelope protein. In some embodiments, the endogenous retroviral element is a retroviral reverse transcriptase. In some embodiments, one or more retroviral elements may be endogenous. In some embodiments, two or more retroviral elements may be endogenous.

[0069] In some embodiments, the PEG10 can have a sequence that is about 80-100 percent, such as 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, to/or 100 percent identical the following DNA sequence:

GCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT
TCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTC
CCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCT
GGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGG
GATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCA
CGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACAC
TTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTT
CCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGAC
CCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGC
CCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAAC
CCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATG
AGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAA
GTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGT
GTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCA
ACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCA
GAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATC
CATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGC
ACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATC
GGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACC
GACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGAC
GGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGG
CGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGC
TGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAAC
ATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAA
GAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGA
GGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTC
TGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCG

TGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGC
TCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGT
TCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCT
ATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGAT
CTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGC
AATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC
TCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTC
ATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCAT
AAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC
CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAG
GCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCT
GCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACG
CAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT
GGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT
CATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTA
AGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG
CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT
CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAAC
CACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGAT
TTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA
ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACC
TATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACG
ATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGC
TCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTT
TATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATA
GTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTC
ATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGT
TAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATG
GCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT
CAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG
ATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGA
TCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA
AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGA
AGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA

ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCTCCC
GATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTC
CCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGG
CTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACG
GGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA
CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAT
GGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG
TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG
TCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA
GCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACT
ATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGCCACCATGGCTGCTGCGGGTGGTTC
CTCCAACTGCCCGCCCCCTCCCCCTCCCCCTCCTCCCAACAACAACAACAACAACAACACCCCAAA
GAGCCCAGGCGTGCCTGACGCCGAAGATGATGATGAACGCAGACACGATGAGCTCCCTGAAGACA
TCAACAACTTTGACGAAGACATGAACAGGCAGTTTGAGAATATGAACCTGCTGGATCAGGTGGAG
TTGCTTGCACAGAGCTACAGTCTGCTGGATCATTTAGATGACTTTGATGATGATGATGAAGACGAT
GACTTTGATCCAGAACCTGACCAGGATGAGCTCCCTGAGTACAGTGACGATGATGACCTGGAGCT
TCAGGGTGCTGCAGCAGCCCCTATCCCAAACTTTTTCTCCGATGATGACTGCCTTGAAGACCTTCCT
GAGAAGTTCGATGGCAACCCTGACATGCTGGGTCCTTTCATGTATCAGTGCCAGCTCTTCATGGAA
AAGAGCACCAGAGATTTCTCAGTTGACCGCATCCGTGTGTGCTTCGTGACAAGCATGCTGATCGGC
CGTGCCGCCCGCTGGGCTACTGCCAAGCTGCAAAGATGTACTTACCTGATGCACAACTACACTGCC
TTTATGATGGAGCTGAAGCATGTCTTTGAAGACCCTCAGAGACGTGAAGCTGCCAAACGCAAGAT
CAGACGTCTGCGCCAGGGCCCTGGGCCTGTTGTGGACTACTCCAATGCATTCCAGATGATTGCCCA
GGACCTGGATTGGACTGAGCCTGCCCTGATGGATCAGTTCCAGGAAGGTCTCAACCCAGACATTC
GCGCAGAGCTGTCTCGCCAGGAGGCCCCCAAGACCCTGGCTGCTCTGATTACTGCCTGTATTCACA
TCGAGAGAAGGCTGGCTCGTGACGCTGCTGCAAAGCCCGATCCTTCACCCAGAGCCTTGGTGATG
CCTCCAAACAGCCAGACCGATCCCACCGAGCCTGTGGGAGGTGCCCGCATGCGCCTGTCCAAGGA
AGAAAAGGAGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCATTTCGCCG
ACACGTGTCCAGCGAAAGCCTCCAAGAATTCGCCGCCGGGAAACTCCCCGGCCCCGCTGTAAGGG
GACCTTCAGCGACAGGGCCAGAACGAATAAGGTCCCCACCCTCCGAGGCTTCGACTCAGCACCTG
CAAGTGATGCTCCAGATTCATATGCCGGGCAGACCCACCCTGTTTGTCCGAGCTATGATTGATTCT
GGTGCATCTGGCAACTTCATTGATCAAGACTTTGTCATACAAAATGCAATTCCTCTCAGAATCAAA
GACTGGCCAGTGATGGTGGAAGCTATTGATGGGCATCCAATTGCCTCGGGCCCAATCATTTTGGAA
ACCCACCACCTGATAGTTGATCTGGGAGACCACCGTGAGATACTGTCATTTGATGTGACTCAGTCT
CCATTCTTTCCTATTGTCCTAGGAATTCGTTGGCTGAGCACGCATGACCCTCACATTACCTGGAGTA
CCCGCTCCATTGTCTTCAACTCTGATTACTGCCGACTTCGCTGCCGGATGTTTGCACAGATACCTTC

TAACTTACTGTTTACAGTGCCACAACCGAATTTGCATCCGTATCTACTTCATCATGTGCATCCGCAT
GTCCATCCGCATATGCATCAGCATCTGCATCAGCATCTGCATCAGTTTCTGCATCCAGATCCGCAT
CAGTATCCGCATCCGGATCCGCATTATCATCATCATCAGCAGGCGGATATGCAGCACCAACTGCAG
CAGTATCTATATCAGTATTTGTATTACCATCTGTATCCGGTTATGCACCACCATCTGCCTCCAGATC
AGCATGAGCATCTGCATGAGTATCTGCATCAGTATCTGCATCAGTATCTGCATCAGTTTCTGCATC
ACCATCTGCATCCGGATCTGCATCAGTATCTGTATCAGTATCTGCATAACCATATGAATCCGGATC
CACATCACCATCCTCATCCAGATCCCCCTCAGGATCCACATCACCCTCCACATCAGGATCCACATC
AGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCACATCAGCATCCGGATCCCCATCAGG
ATCCTCCACATCAGGATCCACATCAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCAC
ATCAGGATGCACATCAGCATCAGGATCCCCATCAGGATGCACATCAGGATCCACATCAGGATGCA
CATCAGGATCCCCATATGGATCCACACCTGCATCAGCACCAGCATCCGCAGCCGCAGCCGCATCC
ACAACAGCATCCTAACCATCCTCAGCAGCCACCATTCTTCTACCACATGGCTGGATTCAGAATTTA
CCACCCTGTAAGGTATTACTATATTCAGAATGTGTATACACCTGTTGATGAGCATGTCTATCCGGG
TCACCGGGTGGTTGACCCTAACATTGAGATGATTCCTGGAGCGCACAGCCTGCCCAGTGGACATTT
GTACTCAATGTCTGAGTCTGAAATGAATGCTCTGCGAAATTTCGTGGACAGGAATGTTAAAGATGG
GCTCATGACTCCCACTGTGGCGCCCAATGGAGCCCAAGTCCTGCAAGTGAAAAGAGGGTGGAAAC
TCCAAGTCACTTACAATTGCCGAGCTCCACAGAGTGGCACCATCCAAAATCAGTACCTACGCATGT
CTCTTCCAAATATGGGAGACCCTGCACACCTGGCAAGCTATGGTGAATTTGTCCAAGTTCCTGGCT
ACCCATATCCAGCCTATGTTTACTATACAAGCCCGCATATGATGACTGCGTGGTACCCAGTAGGAC
GAGATGTACATGGACGAATAATCGTTGTGCCTGTTGTAATCACCTGGTCTCAAAATACGAACCGCC
AGCCTCCGGTGCCCCAGTATCCTCCTCCGCAGCCACCTCCACCACCACCACCACCTCCACCGCCAC
CACCACCTCCACCAGCATCATCCTGCAGTGCTGCGTAGGAATTCTGCAGATATCCA (SEQ ID NO:1)

[0070] In some embodiments, the RTL1 can have a sequence that is about 80-100 percent, such as 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, to/or 100 percent identical the following DNA sequence:

[0071] GCACAGTGGCGGCCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCG
ACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAG
GTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTC
ATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG
GCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGG
GGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTG
ACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTT
CGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGG
CACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGAC
GGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACA
ACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGT
TAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGG

GTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAG
CAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAAT
TAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCC
CATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTG
AGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGA
GCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAA
GATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACA
ACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTT
TGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGC
TGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGG
CTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTA
TCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCAC
CAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGA
TCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGC
CCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAAT
GGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCG
TTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTAC
GGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGG
GACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCA
CCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCC
AGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTG
GTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGC
GTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACG
AGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGT
TGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAAC
GCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT
CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAA
TCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAA
AGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT
CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCC
CTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA
GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGA
ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGA
AAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA

CTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTT
AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTA
GTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGT
CGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGT
TGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGT
TATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTC
TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTG
CCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAA
AACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG
GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTC
CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTA
TTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGAC
GGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGT
TAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAG
CTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCG
CTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAA
TCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAAT
GGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA
GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTG
GCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCC
GCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGAC
TCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA
CGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACG
GTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGA
AATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGCCACCATG
ATAGAACCCTCTGAAGACTCATTTGAGACGATGATGGAGCTTAAGAATCCATCATCAAAACAAAT
GGAGTCCTCCGAGGGCTCATCCAACACCGTTGAAGAGACACCAGGCAGCAGTGGAGCCCAGGCGG
GGGCCCAGGCAGGGGCCCAGGCGGAGGCCCAGGCGGAGACCCAGGTGGAGGCCCAGGCGGAGGC
TCAGGCTGAAGCCCAGGTTGAGGCCCAGGTTGAGGCCCAGGCGGGCTCAGATAGTGGCCCAGCCC
AGGAAGAAAAGGAGCCACCCAGTGGCCCCCTCAAGGAAATGCAAGAGCTGCCCACTAATCTACTC
CAAGAAGTGGAGGAGCCATCCAGTGGCCCCCACCAGGAAATGCAAGAGCTGCCCACTGATCTACT
CCAAGAAGTGGAGGAGCCATCCAGTGGCCCCCACCAGGAAATGCAAGAGCTACCCACTGATCTAC
TCCGGGAAGTGGAGGAGCCATCCAGTGGCCCCTACCAGGAAATGCAAGAGCTACCCACTGATCTA

CTCCGGGAAGTGGAGGAGCCATCCAGTGGCCCCTACCAGGAAATGCAAGAGCTACCCACTGATCT
ACTCCGGGAAGTGGAGGAGCCATCCAGTGGCCCCTACCAGGAAATGCAAGAGCTACCCACTGATC
TACTCCGGGAAGTGGAGGAGCCATCCAGTGGCCCCTACCAGGAAATGCAAGAGCTACCCACTGAT
CTACTCCGGGAAGTGGAGGAGCCATCCAGCGACCCCTGTGAGGCATCATCTAACGATCTACCCCA
AGACATGGAGGAATCAAGTGATGATGGTTCGAACCAGGAGTCGAGTGATGGTTCAAACCACGAAT
TAAGTAATGGTTCAAACCATGAGTCGAGCTTTGGTTCAAACCCAGAGTCGAGCGATGTTTCCAACC
TGGAGTCAAGCGGTGGTTCAAACCAGGAGTCCAGCGATGGTTCACAAAAGGAGTCAAGCTATGAT
TCAAACCCGGAGTTAAGTGATAATTCAAACCAGGAGTTAAGCGATAATTCGAACCAGGAGTCAAG
CGACAGTTCAAACCAGTCAAGTGATATTTCAAACCAGGAAGGGTCGGAACCACTCAGTGAGGCAT
CCGATTATAGCATGGATGAAACAATCAACTCATCCGAGACCCAGTCAGACCAAGATGATACTGAC
CTGGGAGATGACGAGGAAGAGGAGGAAGAGGAAGGTGGTGAAGAGGAAGGTCAGCCCAAGAAC
TCTCCAGAAGAAGTCGTGGCCACCATGGGGAACGTCATCTCGCTGTTCCTCCGGATGCAAGACCTA
AAAGAACAACAGAGAGTGGCAGAAAGGCTCATGATGCAGGCTATCAACGAAGGTCGCCTGCCCTC
CCTCAGGCCCTTCTCTGGAGATCGCAGAGACTACCATGAGTTTGTCGTGCTCTGCCAGATGACAAT
GCAGAACTACCCAAGTATGCTATACAATGATGAGCTGCGGGTGAAATTCGTCATCCGCCACCTGA
CTGACCTAGCCTTGGAATGGGCCAACGACCTCGTAGAGCAGAACAGCCCAGTGATAAACAACTTC
TCAGCCTTCCTGGAGGCCATGTCAGAAAAGTTTGAGTATCGGCAGACCCTGCGTGTGGCGGAAGA
CGCCATGTTCAACATCAGGCAGGGGAACCGCTGTGCGGCCGATTACATCAATGAGTTCCGGGGTC
TGATACCCACCTTAGGCTGGCCAGATGAAGTCCTGCAGGCCCACCTGTGCCAGGGGCTCAACGAA
GAGATCAGGCACTATCTGTTCCGGATCCCTCAGCCGAATTCGTGGACTCAGGAGCAGAAGGGAAC
CTGCCTAGACAACCTGATTGTACTTGTCCTGCAACTTGAAGAAAAGCTGGCAGAGAGAAGAGCAT
TACTCAGGCTGCCACCTGAGTCGCGCCCAAGAAGCGTAGCCTGGATGGATGCTCCCGCCCCAGAG
AAGTGGAGGGTAAGCAGCTGGCTGCCCAATGAGTTCCACCCAGACATTGATCGCGATCACCTCTTC
TTACTGCTCTTGGTGAGAGTGGACCCCTACCACAGTGTCGCAGTGCGGGCCCTGGTGGACTCAGGA
GCAGAAGGGAACTACATGGATGAGAGGTTTGCCCAGGAACACTATGTGGAACTCTATGAGAAGCC
TTATCCACAGATAATCCAAGGTGTGGATGGCATCCCAATTGGCAATGAACCCGTCTGGCTGTGCAC
CGAGCCCCTGGTATGTGTCCACCAGAAGCATTATGAGTACATCGAATTTGACATCTTACCTTCGCC
CAACTTCTCCATCGTCTTAGGCATGAAGTGGCTCCGAACCCATGCCCCTGAAGTCGACTGGATGAG
AGGCCGGTGTACCTTCCACTCTCCCTACTGTCTGCGGAATTGCTTCACCCCACCTCCACCATGCATC
GCACTAGAGACATACAGCATAAGCCTACTGCCCGGATTGCCACACACATACTCAGACCTGGCCGA
CGTGTTTAACCCGAGGGAGGCAGATGATGAGACTTCCGACCAGCCAAGCTCAGACGGATCCGATG
ATCTTTCTGAATCAGAGCCCTCTGAGCTTCAGCAGGCTGGAGACAGTGATCAAAGCGGCGTCTTTT
ACGAGTCTGGGGCCCGAGAGACTTTGGAACCTGTGAGCGCCAGGATGCAAGAGAAAGCCAGGCA
GCAAGAGAAAGCCAGGGAGCAAGAGGAATACTGGATACTGTATGACATGCTGACTGACAGACAA
GACTACACGCAGATGGTACCAGAGCTATTCGACCAGCTACATGGAGCGGCATGGTTCACAAAACT
GGAGCTGCTTGGGATTAAGGAAAGTGAGATGAGGCACACGGTCACACACACGGAAGACACTTGG
AGAGCATCGTTCGGTTTCGGCCTTCACCAGATGAGATGCTACCGGCCCTTCACAATGAACTCATAC
TCTGACGAAGGGAACAACATTGTTCACTTCATCCTAAAAGACATCCTAGGGTTGTTTGTGATTTGC
CATGGCAGAGAGGTCCTGGTCTACTCAATGAGCCAGGAGGAACACTCCCAGCATGTCCGCCAAGT

CCTGGTCCGCTTTCGGTATCACAACATCTACTGTTCACTGGACAAAACTCAGTTCCATCGCCAGAC
CGCTGAAATCTTGGGCTTCAACATATCCCCCAAAGGGGTGAAACTGAACAAGAACCTTATGAACC
TCATCGTGGGGTGCCCTGTCCCTGGCAGCAGGAGGTGTCTTCAAAGTGTGATTGACCTTGTCTATC
CCTATCGCCACTTCGTGGAGAACTTCGCTGTCATCGCAGCACCCCTAGTGAGACAACTGCTGAGTT
CAGAGCCCTACTACTGGGGAGAAGAGGAGCAGGAGGCCTTGGAATCCCTGAAGAGGGCTTTCCGC
AAGTCGCCCGTTCTCTATCACCCTAAGCCTCAGAACCCATTCTACTTGGAAACAGACATCACTGGG
TCATTCCTGTCTGCCTCCCTGGTCCAAACCGACGACGAAACTGGCAAGAAATCTACCTGTGCTTTC
TATTCAAGACCCCTCTCTACAATGGAGGTGGAGTACCCTCGAGTGGAGATGAGGATCCTTCCAATC
CGGGCTGCCTTCATGGTGTGGTGCCGCTACCTGGAGAACACCGAGGAGCCCATCATGATCCTTCTC
AACACAGAGGATCTAGCCTCTCTGAATAATGACAGGCTCACCGTACTTCTCCCCGGGCATTGGGTC
TTCTTCTTCTCACACTTCAATTTTGGTGTTATGGAGATGCCAGCTGAAGGTGACACACAAGCACTCT
TCCGCAGATGTTGGAACCAGAGAGGCTTCAGAGCAAGGTTTCTGCGCCCACTGCTGCTTATGTCCA
TTAGGGCAAATCTAAGGTACTTTGACAGATCCTCTGAGACTGAGGATAAAGAGGATGATGAAGAA
GAGGAGGAGGAAGATGGTGAAGAAGAAGAAGGGGAGGAGGAGGAAGATGGTGAAGAAGAAGAA
GGGGAGGAGGAGGAAGATGGTGAAGAAGAAGAGGAGGAGGAGGAAGATGATGAAGAAGAAGAA
GGGGAGGAGGAGGAAGATGGTGAAGAAGAAGAAGGGGAGGAGGAGGAAGATGGTGAAGAAGAA
GAAGGGGAGGAGGAGGAAGATGGTGAAGAAGAAGAAGGGGAGGAGGAGGGAGAGGAGGAGGA
AGAAGGAGAGGAGGAAGAGGAGGAAGAAGAGGATGAGGAGGAAGAGGAGGAGGAAGAGGAAG
AGGAGGAAGAAGAGGAGGAAGAAGAGGAGGAGGAGGAAGAGGAGGAAGAAGAGGAGGAGGAA
GATGAGGAAGAAGAGGATGAGGAGGAAGAAGATGAGGAGGTGCCCAGCATGGTGCGGGAGCTAC
TGGCTGCCATCCCGATGGACCATATCCTCAATGGCTTACTGGCTCACTTCAGTGTGGCCCAGATCC
GGGCCGTTGTCCTAAACTTCTTCCGTGGCCTGCTCTACTGGAAGAGTCTCCTGGGTGTGGCAGCTG
TCCTGGTGATGTTGAGAGCAAGGCAGCCCCTCTCCCCAGTGCCTGCCCCTAATTTGGAAGTGGCCA
GGCCTCAGCACCGGCACACACTCCGACTCATCCTGGACTCGACCCTCATTGCGAGTAGTGGCATGG
CCACAGCCATCGCTCAGTTGCTCAGCCAGATGCCCCCTCTAGTGGGTGCCAACACCCTCCCAGCTC
GAGAGCTGGCTGAGCTTTTCCTGGGCCCAAGGTGCTGGCATCGGAATGCTCTGCATTCCCAGCCTC
CACGGGGGATGCGCTTCACCCCTGGGTTCTGGTTGACCCTGTGTGAGTTCTTTGGGGTCAGAGTCA
ACCCTGAGGATGACGTCTTTCCTGACCCTTACCAGCACCGGTACTTGGAACTGCACGTCGTTGGGG
ATGAAGATGTCGTCTTGCGAGAAGCCCTGCAAGACGACCTGCAACGTTACCGTCAGTGTGGCCTG
CATGACGGCCTGCAAGACACCTCGCAGGACGCGCAGGACAACGACGTGCAAGAGGACCTGTTTGG
CGACCAGGAAGCGGTCACCTTCAGACCACGAAATCTTTTGGACCCAGAGGTCCTAGATTTCCTCAA
CAACCGGCTACTCTACACCCTGGGCACCGATGGCCGCCTTACTCTGCTTAGCAGAGATCAGGTGGC
CCAGGCCCTCACACGATTCCTGGCCATGGCCTCCAGAATGGCGCTGCCTTCCCCAGCCAGGGAACA
GGCAAGGCTAGAAGAGCTATCAGACTCAGATGATGAACTTGACTGAGAATTCTGCAGATATCCA
(SEQ ID NO:2)

[0072] As used herein, "identity," refers to a relationship between two or more nucleotide or polypeptide sequences, as determined by comparing the sequences. In the art, "identity" also refers to the degree of sequence relatedness between polynucleotide or polypeptide

sequences as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including, but not limited to, those described in (Computational Molecular Biology, Lesk, A. M., Ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., Ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., Eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., Eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math. 1988, 48: 1073. Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity are codified in publicly available computer programs. The percent identity between two sequences can be determined by using analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, Madison Wis.) that incorporates the Needelman and Wunsch, (J. Mol. Biol., 1970, 48: 443-453) algorithm (e.g., NBLAST, and XBLAST). The default parameters are used to determine the identity for the polypeptides or polynucleotides of the present disclosure, unless stated otherwise.

Packaging and Vesicle Elements

[0073] The engineered retroviral delivery system can include viral packaging elements, vesicle elements, or both. In some embodiments, the engineered retroviral delivery system includes one or more polynucleotides that encode one or more packaging elements, one or more polynucleotides that encode one or more vesicle elements, or both. Packaging elements are genes/proteins that are involved in viral packaging of a cargo, such as a cargo polynucleotide, and are further described elsewhere herein. Vesicle elements are genes/proteins that are capable of generating a viral particle (or vesicle) and include envelope, capsid, and other structural gene/proteins that form the structural viral particle or vesicle that encapsulates a cargo, for example. Such elements are further described elsewhere herein.

[0074] In some embodiments, the systems described herein comprise a third polynucleotide comprising genes encoding packaging proteins. In some embodiments, packaging proteins may comprise lentiviral gag, pol, tat, and rev genes.

[0075] Each retroviral genome comprises genes designated gag, pol and env which encode for virion proteins and enzymes. These genes are flanked at both ends by regions termed long terminal repeats (LTRs). LTRs are responsible for integration and proviral transcription. These

also serve as enhancer-promoter sequences. In other words, LTRs can control the expression of viral genes. Retroviral RNA encapsidation occurs as a consequence of a psi sequence located at the 5' end of the viral genome.

[0076] The LTRs themselves are identical sequences that can be divided into three elements, which are designated U3, R, and U5. U3 is derived from the single 3 'end sequence of RNA. R is derived from a sequence repeated at both ends of the RNA and U5 is derived from the single 5 'end sequence of the RNA. The sizes of the three elements can vary considerably between different retroviruses.

[0077] In some embodiments, the gene of interest is under the control of an LTR sequence.

[0078] For the viral genome, the site of the transcription initiation lies in the boundary between U3 and R in the LTR on the left side and the poly (A) (termination) addition site lies in the boundary between R and U5 in the LTR of the right side. U3 contains most of the transcriptional control elements of the provirus, which includes the promoter and multiple enhancer sequences corresponding to cellular and in some cases viral transcriptional activating proteins. Some retroviruses have one or more of the following genes encoding proteins that are involved in the regulation of gene expression: tat, rev, tax and rex.

[0079] As regards the structural genes themselves gag, pol and env, gag (group-specific antigen) encodes the internal structural protein of the virus. Gag protein is proteolytically processed into mature MA (matrix), CA (capsid) and NC (nucleocapsid) proteins. The HIV p17 matrix protein (MA) is a 17 kDa protein, of 132 amino acids, which comprises the N-terminus of the Gag polyprotein. It is responsible for targeting Gag polyprotein to the plasma membrane but also makes contacts with the HIV trans-membrane glycoprotein gp41 in the assembled virus and may play a critical role in recruiting Env glycoproteins to viral budding sites.

[0080] Once Gag is translated, Gag polyproteins are myristoylated at their N-terminal glycine residues by N-myristoyltransferase 1, a modification that is critical for plasma membrane targeting. In the membrane-unbound form, the MA myristoyl fatty acid tail is sequestered in a hydrophobic pocket in the core of the MA protein. Recognition of plasma membrane proteins by MA activates a "myristoyl switch", wherein the myristoyl group is extruded from its hydrophobic pocket in MA and embedded in the plasma membrane.

[0081] The HIV nucleocapsid protein (NC) is a 7 kDa zinc finger protein in the Gag polyprotein and which, after viral maturation, forms the viral nucleocapsid. NC recruits full-length viral genomic RNA to nascent virions.

[0082] The neuronal gene Arc bears homology to the Gag component of Ty3/gypsy retrotransposons and exhibits biochemical properties that are reminiscent of retroviral Gag proteins. The Arc protein assembles into virus-like capsids both in cells and when recombinantly expressed in bacteria. Arc capsids are able to encapsulate their own mRNA, mediating their intercellular transfer in extracellular vesicles. Purified Arc proteins may be used to reconstitute capsids with different DNA or RNA or proteins or some mixture thereof and can be packaged into the capsid for delivery into cells. In some embodiments, capsids may be assembled using lipids to aid uptake by cells. Various embodiments may utilize different Arc orthologs.

[0083] In some embodiments, the polynucleotides described herein may comprise a Gag-homology protein or functional domain thereof. The term "functional domain" refers to a polypeptide sequence that has an activity other than binding to the nucleic acid sequence recognized by the nucleic acid binding domain. By combining a nucleic acid binding domain with one or more effector domains, the polynucleotides of the invention may be used to target the one or more functions or activities mediated by the effector domain to a particular target DNA sequence to which the nucleic acid binding domain specifically binds.

[0084] Genes encoding viral polypeptides capable of self-assembly into defective, non-self-propagating viral particles can be obtained from the genomic DNA of a DNA virus or the genomic cDNA of an RNA virus or from available subgenomic clones containing the genes. These genes will include those encoding viral capsid proteins (i.e., proteins that comprise the viral protein shell) and, in the case of enveloped viruses, such as retroviruses, the genes encoding viral envelope glycoproteins. Additional viral genes may also be required for capsid protein maturation and particle self-assembly. These may encode viral proteases responsible for processing of capsid protein or envelope glycoproteins. As an example, the genomic structure of picornaviruses has been well characterized, and the patterns of protein synthesis leading to virion assembly are clear. Rueckert, R. in Virology (1985), B. N. Fields et al. (eds.) Raven Press, New York, pp 705-738. In picornaviruses, the viral capsid proteins are encoded by an RNA genome containing a single long reading frame, and are synthesized as part of a polyprotein which is processed to yield the mature capsid proteins by a combination of cellular

and viral proteases. Thus, the picornavirus genes required for capsid self-assembly include both the capsid structural genes and the viral proteases required for their maturation. Another virus class from which genes encoding self-assembling capsid proteins can be isolated is the lentiviruses, of which HIV is an example. Like the picornaviral capsid proteins, the HIV gag protein is synthesized as a precursor polypeptide that is subsequently processed, by a viral protease, into the mature capsid polypeptides. However, the gag precursor polypeptide can self-assemble into virus-like particles in the absence of protein processing. Gheysen et al., Cell 59:103 (1989); Delchambre et al., The EMBO J. 8:2653-2660 (1989). Unlike picornavirus capsids, HIV capsids are surrounded by a loose membranous envelope that contains the viral glycoproteins. These are encoded by the viral env gene.

[0085] Molecular and genetic determinants of Gag-mediated intercellular communication may be determined by characterizing the mechanisms of capsid-mediated intercellular mRNA transfer, with particular focus on features that could enable use of this system for programmable delivery of genes of interest. Different Gag proteins evolved diverse RNA-binding domains for mediating specific encapsidation of their RNA genomes. The RNA binding sequence specificity of the human Gag homology proteins can be tested through protein pull-down and sequencing of associated RNA and/or through sequencing of the extracellular vesicle fraction from HEK293 cells that over-express each protein.

[0086] The pol gene encodes the reverse transcriptase (RT), which contains the associated DNA polymerase, RNase H and integrase (IN), which mediate genome replication. The env gene encodes the surface glycoprotein (SU) and the 11 (TM) transmembrane protein of the virion, which forms a complex that interacts specifically with cellular receptor proteins. This interaction ultimately leads to fusion infection of the viral membrane with the cell membrane. Retroviruses may also contain additional genes, which encode proteins in addition to gag, pol and env. Examples of additional genes include in HIV, one or more of vif, vpr, vpx, vpu, tat, rev and nef. The EIAV has, for example, additional genes S2 and dUTPase.

[0087] Proteins encoded by additional genes serve several functions, some of which may be duplicators of a function provided by a cellular protein. In EIAV, for example, tat acts as a transcriptional activator of viral LTR. This binds to a stable, double-stranded RNA structure designated as TAR. Rev regulates and coordinates the expression of viral genes through response elements to rev (RRE). The mechanisms of action of these two proteins are thought to be generally similar to the analogous mechanisms in primate viruses. The function of S2 is

unknown. In addition, an EIAV protein, Ttm, has been identified which is encoded by the first exon of tat processed for the env coding sequence at the beginning of the transmembrane protein.

[0088] In some embodiments, a fourth polynucleotide encodes an envelope protein. In certain embodiments, the envelope protein is lentiviral env protein.

[0089] Env is a retroviral gene that encodes the protein that forms the viral envelope. The expression of the *env* gene allows retroviruses to target and attach to specific cell types, and to infiltrate the target cell membrane. The structure and sequence of several different env genes suggests that Env proteins are type 1 fusion machines. Type 1 fusion machines initially bind a receptor on the target cell surface, which triggers a conformational change, allowing for binding of the fusion protein. The fusion peptide inserts itself in the host cell membrane and brings the host cell membrane very close to the viral membrane, allowing for membrane fusion. The sequence of the env gene may differ significantly between retroviruses, however, the gene is always located downstream of gag, pro, and pol. The env mRNA has to be spliced to be expressed.

[0090] Env not only mediates virus entry into cells, but is also a major target for both cellular and antibody responses. It is synthesized as a precursor molecule, gp160, which is subsequently processed into the surface subunit (SU) gp120 and the transmembrane subunit (TM) gp41 by a cellular protease, and exists as a trimer of gp120-gp41 heterodimers on viral or cell membranes. The SU protein domain determines the tropism of the virus because it is responsible for the receptor-binding function of the virus. The SU domain therefore determines the specificity of the virus for a single receptor molecule. gp120 interacts with receptor and coreceptor molecules for HIV and mediates virus attachment to the cell, while gp41 causes subsequent fusion between viral and cell membranes for releasing viral core components into the cell during the initial infection process. The TM protein consists of three distinct domains: the extracellular domain, the transmembrane domain, and the cytoplasmic domain.

[0091] In some embodiments, the envelope protein may be selected from, but is not necessarily limited to, envH1, envH2, envH3, envK1, envK2_1, envK2_2, envK3, envK4, envK5, envK6, envT, envW, envW1, envfrd, envR(b), envR, envF(c)2, or envF(c)1.

[0092] Envelopes from various retrovirus sources can be used for pseudotyping a vector. The exact rules for pseudotyping (i.e., which envelope proteins will interact with the nascent vector particle at the cytoplasmic side of the cell membrane to give a viable viral particle (Tato,

Virology 88:71, 1978) and which will not (Vana, Nature 336:36, 1988), are not well characterized. However, since a piece of cell membrane buds off to form the viral envelope, molecules normally in the membrane are carried along on the viral envelope. Thus, a number of different potential ligands can be put on the surface of viral vectors by manipulating the cell line making gag and pol in which the vectors are produced or choosing various types of cell lines with particular surface markers. One type of surface marker that can be expressed in helper cells and that can give a useful vector-cell interaction is the receptor for another potentially pathogenic virus. The pathogenic virus displays on the infected cell surface its virally specific protein (e.g., env) that normally interacts with the cell surface marker or receptor to give viral infection. This reverses the specificity of the infection of the vector with respect to the potentially pathogenic virus by using the same viral protein-receptor interaction, but with the receptors on the vector and the viral protein on the cell.

[0093] In some embodiments, the env protein is an endogenous retroviral protein. In some embodiments, the env protein is a lentiviral env protein.

[0094] One virus known to participate in pseudotype formation is vesicular stomatitis virus (VSV), the prototypic member of the rhabdovirus family. It is an enveloped virus with a negative stranded RNA genome that causes a self-limiting disease in live-stock and is essentially non-pathogenic in humans. Balachandran and Barber (2000, IUBMB Life 50: 135-8). Rhabdoviruses have single, negative-strand RNA genomes of 11,000 to 12,000 nucleotides (Rose and Schubert, 1987, Rhabdovirus genomes and their products, in The Viruses: The Rhabdoviruses, Plenum Publishing Corp., NY, pp. 129-166). The virus particles contain a helical, nucleocapsid core composed of the genomic RNA and protein. Generally, three proteins, termed N (nucleocapsid, which encases the genome tightly), P (formerly termed NS, originally indicating nonstructural), and L (large) are found to be associated with the nucleocapsid. An additional matrix (M) protein lies within the membrane envelope, perhaps interacting both with the membrane and the nucleocapsid core. A single glycoprotein (G) species spans the membrane and forms the spikes on the surface of the virus particle.

[0095] In some embodiments, the packaging proteins and/or vesicle proteins are encoded in one polynucleotide, in other embodiments, they are encoded in two separate polynucleotides.

Cargos

[0096] The delivery vesicles described herein may be used and further comprise a number of different cargo molecules for delivery. Representative cargo molecules may include, but are

not limited to, nucleic acids, polynucleotides, proteins, polypeptides, polynucleotide/polypeptide complexes, small molecules, sugars, or a combination thereof. Cargos that can be delivered in accordance with the systems and methods described herein include, but are not necessarily limited to, biologically active agents, including, but not limited to, therapeutic agents, imaging agents, and monitoring agents. A cargo may be an exogenous material or an endogenous material. In some embodiments, the cargo can be a "gene of interest".

Polynucleotides

[0097] In some embodiments, the cargo is a cargo polynucleotide. As used herein, "nucleic acid," "nucleotide sequence," and "polynucleotide" can be used interchangeably herein and can generally refer to a string of at least two base-sugar-phosphate combinations and refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single-and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein can refer to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions can be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. "Polynucleotide" and "nucleic acids" also encompasses such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, inter alia. For instance, the term polynucleotide as used herein can include DNAs or RNAs as described herein that contain one or more modified bases. Thus, DNAs or RNAs including unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. "Polynucleotide", "nucleotide sequences" and "nucleic acids" also includes PNAs (peptide nucleic acids), phosphorothioates, and other variants of the phosphate backbone of native nucleic acids. Natural nucleic acids have a phosphate backbone, artificial nucleic acids can contain other types of backbones, but contain the same bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "nucleic acids" or "polynucleotides"

as that term is intended herein. As used herein, "nucleic acid sequence" and "oligonucleotide" also encompasses a nucleic acid and polynucleotide as defined elsewhere herein.

[0098] As used herein, "deoxyribonucleic acid (DNA)" and "ribonucleic acid (RNA)" can generally refer to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. RNA can be in the form of non-coding RNA, including but not limited to, tRNA (transfer RNA), snRNA (small nuclear RNA), rRNA (ribosomal RNA), anti-sense RNA, RNAi (RNA interference construct), siRNA (short interfering RNA), microRNA (miRNA), or ribozymes, aptamers, guide RNA (gRNA), or coding mRNA (messenger RNA).

[0099] In some embodiments, the cargo polynucleotide is DNA. In some embodiments, the cargo polynucleotide is RNA. In some embodiments, the cargo polynucleotide is a polynucleotide (a DNA or an RNA) that encodes an RNA and/or a polypeptide. As used herein with reference to the relationship between DNA, cDNA, cRNA, RNA, protein/peptides, and the like "corresponding to" or "encoding" (used interchangeably herein) refers to the underlying biological relationship between these different molecules. As such, one of skill in the art would understand that operatively "corresponding to" can direct them to determine the possible underlying and/or resulting sequences of other molecules given the sequence of any other molecule which has a similar biological relationship with these molecules. For example, from a DNA sequence an RNA sequence can be determined and from an RNA sequence a cDNA sequence can be determined.

Genes of Interest

[0100] In some embodiments, the systems described herein comprise a polynucleotide encoding a gene of interest. As used herein, the term "gene of interest" refers to the gene selected for a particular purpose and being desired of delivery by a system or vesicle of the present invention. A gene of interest inserted into one or more regions a vector, such as an expression vector (including one or more of the engineered delivery vesicle generation system vectors) such that when expressed in a target cell or recipient cell it can be expressed and produce a desired gene product and/or be packaged as cargo in an engineered delivery vesicle of the present invention. It will be appreciated that other cargos specifically identified can also be genes of interest. For example, a polynucleotide encoding a Cas effector can be a gene of interest in this context where it is desired to deliver a Cas effector to a cell, for example.

[0101] In one embodiment, the gene of interest encodes a gene that provides a therapeutic function for the treatment of a disease. In some embodiments, the gene of interest can also be a vaccinating gene, that is to say a gene encoding an antigenic peptide that is capable of generating an immune response in humans or animals. This may include, but is not necessarily limited to, peptide antigens specific for viral and bacterial infections, or may be tumor-specific. In some embodiments, a gene of interest is a gene which confers a desired phenotype. As the embodiments described herein focus on improved methods for packaging and delivery of a gene of interest, the particular gene of interest is not limiting and the technology can generally be used to deliver any gene of interest generally recognized by one of ordinary skill in the art as deliverable using a lentiviral system. One skilled in the art can design a construct containing any gene that they are interested in. Designing a construct containing a known gene of interest can be performed without undue experimentation. One of ordinary skill in the art routinely selects genes of interest. For example, the GenBank public database has existed since 1982 and is routinely used by persons of ordinary skill in the art relevant to the presently claimed method. As of June 2019, GenBank contains 2013,383,758 loci, 329,835,282,370 bases, from 213,383,758 reported sequences. The nucleotide sequences are from more than 300,000 organisms with supporting bibliographic and biological annotation. GenBank is only example, as there are many other known repositories of sequence information.

[0102] In some embodiments, the gene of interest may be, for example, a synthetic RNA/DNA sequence, a codon optimized RNA/DNA sequence, a recombinant RNA/DNA sequence (i.e. prepared by use of recombinant DNA techniques), a cDNA sequence or a partial genomic DNA sequence, including combinations thereof. Preferably, this is in the sense orientation. Preferably, the sequence is, comprises, or is transcribed from cDNA. The gene(s) of interest may also be referred to herein as “heterologous sequence(s)” “heterologous gene(s)” or “transgene(s)”.

[0103] In some embodiments, the gene of interest may confer some therapeutic benefit. The terms “therapeutic agent”, “therapeutic capable agent” or “treatment agent” are used interchangeably and refer to a molecule or compound that confers some beneficial effect upon administration to a subject. The beneficial effect includes enablement of diagnostic determinations; amelioration of a disease, symptom, disorder, or pathological condition; reducing or preventing the onset of a disease, symptom, disorder or condition; and generally counteracting a disease, symptom, disorder or pathological condition.

[0104] Preferably, the therapeutic agent may be administered in a therapeutically effective amount of the active components. The term "therapeutically effective amount" refers to an amount which can elicit a biological or medicinal response in a tissue, system, animal or human that is being sought by a researcher, veterinarian, medical doctor or other clinician, and in particular can prevent or alleviate one or more of the local or systemic symptoms or features of a disease or condition being treated.

[0105] For example, in methods for treating cancer in a subject, an effective amount of a combination of inhibitors targeting epigenetic genes is any amount that provides an anti-cancer effect, such as reduces or prevents proliferation of a cancer cell or is cytotoxic towards a cancer cell.

[0106] In one aspect, the invention provides for methods and compositions for treating cancer and for targeting mammalian cells that are cancer cells.

[0107] The cancer may include, without limitation, liquid tumors such as leukemia (e.g., acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, acute myeloblastic leukemia, acute promyelocytic leukemia, acute myelomonocytic leukemia, acute monocytic leukemia, acute erythroleukemia, chronic leukemia, chronic myelocytic leukemia, chronic lymphocytic leukemia), polycythemia vera, lymphoma (e.g., Hodgkin's disease, non-Hodgkin's disease), Waldenstrom's macroglobulinemia, heavy chain disease, or multiple myeloma.

[0108] The cancer may include, without limitation, solid tumors such as sarcomas and carcinomas. Examples of solid tumors include, but are not limited to fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, epithelial carcinoma, bronchogenic carcinoma, hepatoma, colorectal cancer (e.g., colon cancer, rectal cancer), anal cancer, pancreatic cancer (e.g., pancreatic adenocarcinoma, islet cell carcinoma, neuroendocrine tumors), breast cancer (e.g., ductal carcinoma, lobular carcinoma, inflammatory breast cancer, clear cell carcinoma, mucinous carcinoma), ovarian carcinoma (e.g., ovarian epithelial carcinoma or surface epithelial-stromal tumour including serous tumour, endometrioid tumor and mucinous cystadenocarcinoma, sex-cord-stromal

tumor), prostate cancer, liver and bile duct carcinoma (e.g., hepatocellular carcinoma, cholangiocarcinoma, hemangioma), choriocarcinoma, seminoma, embryonal carcinoma, kidney cancer (e.g., renal cell carcinoma, clear cell carcinoma, Wilm's tumor, nephroblastoma), cervical cancer, uterine cancer (e.g., endometrial adenocarcinoma, uterine papillary serous carcinoma, uterine clear-cell carcinoma, uterine sarcomas and leiomyosarcomas, mixed mullerian tumors), testicular cancer, germ cell tumor, lung cancer (e.g., lung adenocarcinoma, squamous cell carcinoma, large cell carcinoma, bronchioloalveolar carcinoma, non-small-cell carcinoma, small cell carcinoma, mesothelioma), bladder carcinoma, signet ring cell carcinoma, cancer of the head and neck (e.g., squamous cell carcinomas), esophageal carcinoma (e.g., esophageal adenocarcinoma), tumors of the brain (e.g., glioma, glioblastoma, medullablastoma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, schwannoma, meningioma), neuroblastoma, retinoblastoma, neuroendocrine tumor, melanoma, cancer of the stomach (e.g., stomach adenocarcinoma, gastrointestinal stromal tumor), or carcinoids. Lymphoproliferative disorders are also considered to be proliferative diseases.

[0109] As such, the methods, compositions, and systems described herein may be used to target mammalian cells of any of the tumors or cancers described above.

[0110] In some embodiments, the mammalian cells to be targeted for therapeutic treatment may be infected with one or more pathogens.

[0111] In some embodiments, the one or more pathogens may include, but is not necessarily limited to, one or more viruses. The virus may be a DNA virus, a RNA virus, or a retrovirus. Non-limiting examples of viruses useful with the present invention include, but are not limited to Ebola, measles, SARS, Chikungunya, hepatitis, Marburg, yellow fever, MERS, Dengue, Lassa, influenza, rhabdovirus or HIV. A hepatitis virus may include hepatitis A, hepatitis B, or hepatitis C. An influenza virus may include, for example, influenza A or influenza B. An HIV may include HIV 1 or HIV 2. In certain example embodiments, the viral sequence may be a human respiratory syncytial virus, Sudan ebola virus, Bundibugyo virus, Tai Forest ebola virus, Reston ebola virus, Achimota, Aedes flavivirus, Aguacate virus, Akabane virus, Alethinophid reptarenavirus, Allpahuayo mammarenavirus, Amapari mammarenavirus, Andes virus, Apoi virus, Aravan virus, Aroa virus, Arumwot virus, Atlantic salmon paramyxovirus, Australian bat lyssavirus, Avian bornavirus, Avian metapneumovirus, Avian paramyxoviruses, penguin or Falkland Islandsvirus, BK polyomavirus, Bagaza virus,

Banna virus, Bat hepevirus, Bat sapovirus, Bear Canon mammarenavirus, Beilong virus, Betacoronoavirus, Betapapillomavirus 1-6, Bhanja virus, Bokeloh bat lyssavirus, Borna disease virus, Bourbon virus, Bovine hepacivirus, Bovine parainfluenza virus 3, Bovine respiratory syncytial virus, Brazoran virus, Bunyamwere virus, Caliciviridae virus. California encephalitis virus, Candiru virus, Canine distemper virus, Canaine pneumovirus, Cedar virus, Cell fusing agent virus, Cetacean morbillivirus, Chandipura virus, Chaoyang virus, Chapare mammarenavirus, Chikungunya virus, Colobus monkey papillomavirus, Colorado tick fever virus, Cowpox virus, Crimean-Congo hemorrhagic fever virus, Culex flavivirus, Cupixi mammarenavirus, Dengue virus, Dobrava-Belgrade virus, Donggang virus, Dugbe virus, Duvenhage virus, Eastern equine encephalitis virus, Entebbe bat virus, Enterovirus A-D, European bat lyssavirus 1-2, Eyach virus, Feline morbillivirus, Fer-de-Lance paramyxovirus, Fitzroy River virus, Flaviviridae virus, Flexal mammarenavirus, GB virus C, Gairo virus, Gemycircularvirus, Goose paramyxovirus SF02, Great Island virus, Guanarito mammarenavirus, Hantaan virus, Hantavirus Z10, Heartland virus, Hendra virus, Hepatitis A/B/C/E, Hepatitis delta virus, Human bocavirus, Human coronavirus, Human endogenous retrovirus K, Human enteric coronavirus, Human genital-associated circular DNA virus-1, Human herpesvirus 1-8, Human immunodeficiency virus 1/2, Huan mastadenovirus A-G, Human papillomavirus, Human parainfluenza virus 1-4, Human paraechovirus, Human picobirnavirus, Human smacovirus, Ikoma lyssavirus, Ilheus virus, Influenza A-C, Ippy mammarenavirus, Irkut virus, J-virus, JC polyomavirus, Japanese encephalitis virus, Junin mammarenavirus, KI polyomavirus, Kadipiro virus, Kamiti River virus, Kedougou virus, Khujand virus, Kokobera virus, Kyasanur forest disease virus, Lagos bat virus, Langat virus, Lassa mammarenavirus, Latino mammarenavirus, Leopards Hill virus, Liao ning virus, Ljungan virus, Lloviu virus, Louping ill virus, Lujo mammarenavirus, Luna mammarenavirus, Lunk virus, Lymphocytic choriomeningitis mammarenavirus, Lyssavirus Ozernoe, MSSI2\.225 virus, Machupo mammarenavirus, Mamastrovirus 1, Manzanilla virus, Mapuera virus, Marburg virus, Mayaro virus, Measles virus, Menangle virus, Mercadeo virus, Merkel cell polyomavirus, Middle East respiratory syndrome coronavirus, Mobala mammarenavirus, Modoc virus, Moijang virus, Mokolo virus, Monkeypox virus, Montana myotis leukoenchalitis virus, Mopeia lassa virus reassortant 29, Mopeia mammarenavirus, Morogoro virus, Mossman virus, Mumps virus, Murine pneumonia virus, Murray Valley encephalitis virus, Nariva virus, Newcastle disease virus, Nipah virus, Norwalk virus, Norway rat hepacivirus, Ntaya virus,

O'nyong-nyong virus, Oliveros mammarenavirus, Omsk hemorrhagic fever virus, Oropouche virus, Parainfluenza virus 5, Parana mammarenavirus, Parramatta River virus, Peste-des-petits-ruminants virus, Pichande mammarenavirus, Picornaviridae virus, Pirital mammarenavirus, Piscihepevirus A, Porcine parainfluenza virus 1, porcine rubulavirus, Powassan virus, Primate T-lymphotropic virus 1-2, Primate erythroparvovirus 1, Punta Toro virus, Puumala virus, Quang Binh virus, Rabies virus, Razdan virus, Reptile bornavirus 1, Rhinovirus A-B, Rift Valley fever virus, Rinderpest virus, Rio Bravo virus, Rodent Torque Teno virus, Rodent hepacivirus, Ross River virus, Rotavirus A-I, Royal Farm virus, Rubella virus, Sabia mammarenavirus, Salem virus, Sandfly fever Naples virus, Sandfly fever Sicilian virus, Sapporo virus, Sathuperi virus, Seal anellovirus, Semliki Forest virus, Sendai virus, Seoul virus, Sepik virus, Severe acute respiratory syndrome-related coronavirus, Severe fever with thrombocytopenia syndrome virus, Shamonda virus, Shimoni bat virus, Shuni virus, Simbu virus, Simian torque teno virus, Simian virus 40-41, Sin Nombre virus, Sindbis virus, Small anellovirus, Sosuga virus, Spanish goat encephalitis virus, Spondweni virus, St. Louis encephalitis virus, Sunshine virus, TTV-like mini virus, Tacaribe mammarenavirus, Taila virus, Tamana bat virus, Tamiami mammarenavirus, Tembusu virus, Thogoto virus, Thottapalayam virus, Tick-borne encephalitis virus, Tioman virus, Togaviridae virus, Torque teno canis virus, Torque teno douroucouli virus, Torque teno felis virus, Torque teno midi virus, Torque teno sus virus, Torque teno tamarin virus, Torque teno virus, Torque teno zalophus virus, Tuhoko virus, Tula virus, Tupaia paramyxovirus, Usutu virus, Uukuniemi virus, Vaccinia virus, Variola virus, Venezuelan equine encephalitis virus, Vesicular stomatitis Indiana virus, WU Polyomavirus, Wesselsbron virus, West Caucasian bat virus, West Nile virus, Western equine encephalitis virus, Whitewater Arroyo mammarenavirus, Yellow fever virus, Yokose virus, Yug Bogdanovac virus, Zaire ebolavirus, Zika virus, or Zygosaccharomyces bailii virus Z viral sequence. Examples of RNA viruses that may be detected include one or more of (or any combination of) Coronaviridae virus, a Picornaviridae virus, a Caliciviridae virus, a Flaviviridae virus, a Togaviridae virus, a Bornaviridae, a Filoviridae, a Paramyxoviridae, a Pneumoviridae, a Rhabdoviridae, an Arenaviridae, a Bunyaviridae, an Orthomyxoviridae, or a Deltavirus. In certain example embodiments, the virus is Coronavirus, SARS, Poliovirus, Rhinovirus, Hepatitis A, Norwalk virus, Yellow fever virus, West Nile virus, Hepatitis C virus, Dengue fever virus, Zika virus, Rubella virus, Ross River virus, Sindbis virus, Chikungunya virus, Borna disease virus, Ebola virus, Marburg virus,

Measles virus, Mumps virus, Nipah virus, Hendra virus, Newcastle disease virus, Human respiratory syncytial virus, Rabies virus, Lassa virus, Hantavirus, Crimean-Congo hemorrhagic fever virus, Influenza, or Hepatitis D virus.

[0112] In certain example embodiments, the virus may be a retrovirus. Example retroviruses that may be detected using the embodiments disclosed herein include one or more of or any combination of viruses of the Genus Alpharetrovirus, Betaretrovirus, Gammaretrovirus, Deltaretrovirus, Epsilonretrovirus, Lentivirus, Spumavirus, or the Family Metaviridae, Pseudoviridae, and Retroviridae (including HIV), Hepadnaviridae (including Hepatitis B virus), and Caulimoviridae.

[0113] In some embodiments, the one or more pathogens may include, but is not necessarily limited to, one or more bacteria. Examples of bacteria that cause infections that could be prevented or treated with the systems and methods described herein include without limitation any one or more of (or any combination of) *Acinetobacter baumanii*, *Actinobacillus sp.*, *Actinomycetes*, *Actinomyces sp.* (such as *Actinomyces israelii* and *Actinomyces naeslundii*), *Aeromonas sp.* (such as *Aeromonas hydrophila*, *Aeromonas veronii biovar sobria* (*Aeromonas sobria*), and *Aeromonas caviae*), *Anaplasma phagocytophilum*, *Anaplasma marginale Alcaligenes xylosoxidans*, *Acinetobacter baumanii*, *Actinobacillus actinomycetemcomitans*, *Bacillus sp.* (such as *Bacillus anthracis*, *Bacillus cereus*, *Bacillus subtilis*, *Bacillus thuringiensis*, and *Bacillus stearothermophilus*), *Bacteroides sp.* (such as *Bacteroides fragilis*), *Bartonella sp.* (such as *Bartonella bacilliformis* and *Bartonella henselae*, *Bifidobacterium sp.*, *Bordetella sp.* (such as *Bordetella pertussis*, *Bordetella parapertussis*, and *Bordetella bronchiseptica*), *Borrelia sp.* (such as *Borrelia recurrentis*, and *Borrelia burgdorferi*), *Brucella sp.* (such as *Brucella abortus*, *Brucella canis*, *Brucella melintensis* and *Brucella suis*), *Burkholderia sp.* (such as *Burkholderia pseudomallei* and *Burkholderia cepacia*), *Campylobacter sp.* (such as *Campylobacter jejuni*, *Campylobacter coli*, *Campylobacter lari* and *Campylobacter fetus*), *Capnocytophaga sp.*, *Cardiobacterium hominis*, *Chlamydia trachomatis*, *Chlamydophila pneumoniae*, *Chlamydophila psittaci*, *Citrobacter sp. Coxiella burnetii*, *Corynebacterium sp.* (such as, *Corynebacterium diphtheriae*, *Corynebacterium jeikeum* and *Corynebacterium*), *Clostridium sp.* (such as *Clostridium perfringens*, *Clostridium difficile*, *Clostridium botulinum* and *Clostridium tetani*), *Eikenella corrodens*, *Enterobacter sp.* (such as *Enterobacter aerogenes*, *Enterobacter agglomerans*, *Enterobacter cloacae* and *Escherichia coli*, *including opportunistic Escherichia coli*, such as

enterotoxigenic E. coli, *enteroinvasive E. coli*, *enteropathogenic E. coli*, *enterohemorrhagic E. coli*, *enteroaggregative E. coli* and *uropathogenic E. coli*) *Enterococcus sp.* (such as *Enterococcus faecalis* and *Enterococcus faecium*) *Ehrlichia sp.* (such as *Ehrlichia chafeensia* and *Ehrlichia canis*), *Epidermophyton floccosum, Erysipelothrix rhusiopathiae*, *Eubacterium sp.*, *Francisella tularensis*, *Fusobacterium nucleatum*, *Gardnerella vaginalis*, *Gemella morbillorum*, *Haemophilus sp.* (such as *Haemophilus influenzae*, *Haemophilus ducreyi*, *Haemophilus aegyptius*, *Haemophilus parainfluenzae*, *Haemophilus haemolyticus* and *Haemophilus parahaemolyticus*, *Helicobacter sp.* (such as *Helicobacter pylori*, *Helicobacter cinaedi* and *Helicobacter fennelliae*), *Kingella kingii*, *Klebsiella sp.* (such as *Klebsiella pneumoniae*, *Klebsiella granulomatis* and *Klebsiella oxytoca*), *Lactobacillus sp.*, *Listeria monocytogenes*, *Leptospira interrogans*, *Legionella pneumophila*, *Leptospira interrogans*, *Peptostreptococcus sp.*, *Mannheimia hemolytica, Microsporum canis, Moraxella catarrhalis*, *Morganella sp.*, *Mobiluncus sp.*, *Micrococcus sp.*, *Mycobacterium sp.* (such as *Mycobacterium leprae*, *Mycobacterium tuberculosis*, *Mycobacterium paratuberculosis, Mycobacterium intracellulare*, *Mycobacterium avium*, *Mycobacterium bovis*, and *Mycobacterium marinum*), *Mycoplasma sp.* (such as *Mycoplasma pneumoniae*, *Mycoplasma hominis*, and *Mycoplasma genitalium*), *Nocardia sp.* (such as *Nocardia asteroides*, *Nocardia cyriacigeorgica* and *Nocardia brasiliensis*), *Neisseria sp.* (such as *Neisseria gonorrhoeae* and *Neisseria meningitidis*), *Pasteurella multocida*, *Pityrosporum orbiculare* (Malassezia furfur), *Plesiomonas shigelloides. Prevotella sp.*, *Porphyromonas sp.*, *Prevotella melaninogenica*, *Proteus sp.* (such as *Proteus vulgaris* and *Proteus mirabilis*), *Providencia sp.* (such as *Providencia alcalifaciens*, *Providencia rettgeri* and *Providencia stuartii*), *Pseudomonas aeruginosa*, *Propionibacterium acnes*, *Rhodococcus equi*, *Rickettsia sp.* (such as *Rickettsia rickettsii*, *Rickettsia akari* and *Rickettsia prowazekii*, *Orientia tsutsugamushi* (*formerly: Rickettsia tsutsugamushi*) and *Rickettsia typhi*), *Rhodococcus sp.*, *Serratia marcescens*, *Stenotrophomonas maltophilia*, *Salmonella sp.* (such as *Salmonella enterica*, *Salmonella typhi*, *Salmonella paratyphi*, *Salmonella enteritidis*, *Salmonella cholerasuis* and *Salmonella typhimurium*), *Serratia sp.* (such as *Serratia marcesans* and *Serratia liquifaciens*), *Shigella sp.* (such as *Shigella dysenteriae*, *Shigella flexneri*, *Shigella boydii* and *Shigella sonnei*), *Staphylococcus sp.* (such as *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus hemolyticus*, *Staphylococcus saprophyticus*), *Streptococcus sp.* (such as *Streptococcus pneumoniae* (for example *chloramphenicol-resistant serotype 4 Streptococcus*

pneumoniae, *spectinomycin-resistant serotype 6B Streptococcus pneumoniae*, *streptomycin-resistant serotype 9V Streptococcus pneumoniae*, *erythromycin-resistant serotype 14 Streptococcus pneumoniae*, *optochin-resistant serotype 14 Streptococcus pneumoniae*, *rifampicin-resistant serotype 18C Streptococcus pneumoniae*, *tetracycline-resistant serotype 19F Streptococcus pneumoniae*, *penicillin-resistant serotype 19F Streptococcus pneumoniae*, and *trimethoprim-resistant serotype 23F Streptococcus pneumoniae*, *chloramphenicol-resistant serotype 4 Streptococcus pneumoniae*, *spectinomycin-resistant serotype 6B Streptococcus pneumoniae*, *streptomycin-resistant serotype 9V Streptococcus pneumoniae*, *optochin-resistant serotype 14 Streptococcus pneumoniae*, *rifampicin-resistant serotype 18C Streptococcus pneumoniae*, *penicillin-resistant serotype 19F Streptococcus pneumoniae*, *or trimethoprim-resistant serotype 23F Streptococcus pneumoniae*), *Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pyogenes*, *Group A streptococci*, *Streptococcus pyogenes*, *Group B streptococci*, *Streptococcus agalactiae*, *Group C streptococci*, *Streptococcus anginosus*, *Streptococcus equismilis*, *Group D streptococci*, *Streptococcus bovis*, *Group F streptococci*, and *Streptococcus anginosus Group G streptococci*), *Spirillum minus*, *Streptobacillus moniliformi*, *Treponema sp.* (such as *Treponema carateum*, *Treponema petenue*, *Treponema pallidum* and *Treponema endemicum*, *Trichophyton rubrum, T. mentagrophytes, Tropheryma whippelii*, *Ureaplasma urealyticum*, *Veillonella sp.*, *Vibrio sp.* (such as *Vibrio cholerae*, *Vibrio parahemolyticus*, *Vibrio vulnificus*, *Vibrio parahaemolyticus*, *Vibrio vulnificus*, *Vibrio alginolyticus*, *Vibrio mimicus*, *Vibrio hollisae*, *Vibrio fluvialis*, *Vibrio metchnikovii*, *Vibrio damsela* and *Vibrio furnisii*), *Yersinia sp.* (such as *Yersinia enterocolitica*, *Yersinia pestis*, and *Yersinia pseudotuberculosis*) and *Xanthomonas maltophilia* among others.

[0114] In some embodiments, the one or more pathogens may include, but is not necessarily limited to, one or more fungi. Examples of fungi that cause infections that could be prevented or treated with the systems and methods described herein include without limitation any one or more of (or any combination of), Aspergillus, Blastomyces, Candidiasis, Coccidiodomycosis, Cryptococcus neoformans, Cryptococcus gatti, sp. Histoplasma sp. (such as Histoplasma capsulatum), Pneumocystis sp. (such as Pneumocystis jirovecii), Stachybotrys (such as Stachybotrys chartarum), Mucroymcosis, Sporothrix, fungal eye infections ringworm, Exserohilum, Cladosporium.

[0115] In certain example embodiments, the fungus is a yeast. Examples of yeast include without limitation one or more of (or any combination of), Aspergillus species (such as

Aspergillus fumigatus, Aspergillus flavus and Aspergillus clavatus), Cryptococcus sp. (such as Cryptococcus neoformans, Cryptococcus gattii, Cryptococcus laurentii and Cryptococcus albidus), a Geotrichum species, a Saccharomyces species, a Hansenula species, a Candida species (such as Candida albicans), a Kluyveromyces species, a Debaryomyces species, a Pichia species, or combination thereof. In certain example embodiments, the fungi is a mold. Example molds include, but are not limited to, a Penicillium species, a Cladosporium species, a Byssochlamys species, or a combination thereof.

[0116] In some embodiments, the one or more pathogens may include, but is not necessarily limited to, one or more protozoans. Examples of protozoa that cause infections that could be treated with the systems and methods described herein include without limitation any one or more of (or any combination of), Euglenozoa, Heterolobosea, Diplomonadida, Amoebozoa, Blastocystic, and Apicomplexa. Example Euglenoza include, but are not limited to, Trypanosoma cruzi (Chagas disease), T. brucei gambiense, T. brucei rhodesiense, Leishmania braziliensis, L. infantum, L. mexicana, L. major, L. tropica, and L. donovani. Example Heterolobosea include, but are not limited to, Naegleria fowleri. Example Diplomonadids include, but are not limited to, Giardia intestinalis (G. lamblia, G. duodenalis). Example Amoebozoa include, but are not limited to, Acanthamoeba castellanii, Balamuthia madrillaris, Entamoeba histolytica. Example Blastocysts include, but are not limited to, Blastocystic hominis. Example Apicomplexa include, but are not limited to, Babesia microti, Cryptosporidium parvum, Cyclospora cayetanensis, Plasmodium falciparum, P. vivax, P. ovale, P. malariae, and Toxoplasma gondii.

[0117] In some embodiments, the one or more pathogens may include, but is not necessarily limited to, one or more parasites. Examples of parasites that cause infections that could be treated with the systems and methods described herein include without limitation one or more of (or any combination of), an Onchocerca species and a Plasmodium species.

[0118] In some embodiments, the gene of interest may lead to altered expression in the target cell. As used herein the term "altered expression" may particularly denote altered production of the recited gene products by a cell. As used herein, the term "gene product(s)" includes RNA transcribed from a gene (e.g., mRNA), or a polypeptide encoded by a gene or translated from RNA.

[0119] Also, "altered expression" as intended herein may encompass modulating the activity of one or more endogenous gene products. Accordingly, "altered expression", "altering

expression", "modulating expression", or "detecting expression" or similar may be used interchangeably with respectively "altered expression or activity", "altering expression or activity", "modulating expression or activity", or "detecting expression or activity" or similar. As used herein, "modulating" or "to modulate" generally means either reducing or inhibiting the activity of a target or antigen, or alternatively increasing the activity of the target or antigen, as measured using a suitable in vitro, cellular or in vivo assay. In particular, "modulating" or "to modulate" can mean either reducing or inhibiting the (relevant or intended) activity of, or alternatively increasing the (relevant or intended) biological activity of the target or antigen, as measured using a suitable in vitro, cellular or in vivo assay (which will usually depend on the target or antigen involved), by at least 5%, at least 10%, at least 25%, at least 50%, at least 60%, at least 70%, at least 80%, or 90% or more, compared to activity of the target or antigen in the same assay under the same conditions but without the presence of the inhibitor/antagonist agents or activator/agonist agents described herein.

[0120] As will be clear to the skilled person, "modulating" can also involve effecting a change (which can either be an increase or a decrease) in affinity, avidity, specificity and/or selectivity of a target or antigen, for one or more of its targets compared to the same conditions but without the presence of a modulating agent. Again, this can be determined in any suitable manner and/or using any suitable assay known per se, depending on the target. In particular, an action as an inhibitor/antagonist or activator/agonist can be such that an intended biological or physiological activity is increased or decreased, respectively, by at least 5%, at least 10%, at least 25%, at least 50%, at least 60%, at least 70%, at least 80%, or 90% or more, compared to the biological or physiological activity in the same assay under the same conditions but without the presence of the inhibitor/antagonist agent or activator/agonist agent. Modulating can also involve activating the target or antigen or the mechanism or pathway in which it is involved.

Interference RNAs

[0121] In certain example embodiments, the one or more polynucleotides may encode one or more interference RNAs. Iinterference RNAs are RNA molecules capable of suppressing gene expressions. Example types of interference RNAs include small interfering RNA (siRNA), micro RNA (miRNA), and short hairpin RNA (shRNA).

[0122] In certain example embodiments, the interference RNA may be a siRNAs. Small interfering RNA (siRNA) molecules are capable of inhibiting target gene expression by

interfering RNA. siRNAs may be chemically synthesized, or may be obtained by in vitro transcription, or may be synthesized in vivo in target cell. siRNAs may comprise double-stranded RNA from 15 to 40 nucleotides in length and can contain a protuberant region 3' and/or 5' from 1 to 6 nucleotides in length. Length of protuberant region is independent from total length of siRNA molecule. siRNAs may act by post-transcriptional degradation or silencing of target messenger. In some cases, the exogenous polynucleotides encode shRNAs. In shRNAs the antiparallel strands that form siRNA are connected by a loop or hairpin region.

[0123] The interference RNA (e.g., siRNA) may suppress expression of genes to promote long term survival and functionality of cells after transplanted to a subject. In some examples, the interference RNAs suppress genes in TGFβ pathway, e.g., TGFβ, TGFβ receptors, and SMAD proteins. In some examples, the interference RNAs suppress genes in colony-stimulating factor 1 (CSF1) pathway, e.g., CSF1 and CSF1 receptors. In certain embodiments, the one or more interference RNAs suppress genes in both the CSF1 pathway and the TGFβ pathway. TGFβ pathway genes may comprise one or more of ACVR1, ACVR1C, ACVR2A, ACVR2B, ACVRL1, AMH, AMHR2, BMP2, BMP4, BMP5, BMP6, BMP7, BMP8A, BMP8B, BMPR1A, BMPR1B, BMPR2, CDKN2B, CHRD, COMP, CREBBP, CUL1, DCN, E2F4, E2F5, EP300, FST, GDF5, GDF6, GDF7, ID1, ID2, ID3, ID4, IFNG, INHBA, INHBB, INHBC, INHBE, LEFTY1, LEFTY2, LOC728622, LTBP1, MAPK1, MAPK3, MYC, NODAL, NOG, PITX2, PPP2CA, PPP2CB, PPP2R1A, PPP2R1B, RBL1, RBL2, RBX1, RHOA, ROCK1, ROCK2, RPS6KB1, RPS6KB2, SKP1, SMAD1, SMAD2, SMAD3, SMAD4, SMAD5, SMAD6, SMAD7, SMAD9, SMURF1, SMURF2, SP1, TFDP1, TGFB1, TGFB2, TGFB3, TGFBR1, TGFBR2, THBS1, THBS2, THBS3, THBS4, TNF, ZFYVE16, and/or ZFYVE9.

[0124] In some embodiments, the cargo polynucleotide is an RNAi molecule, antisense molecule, and/or a gene silencing oligonucleotide or a polynucleotide that encodes an RNAi molecule, antisense molecule, and/or gene silencing oligonucleotide.

[0125] As used herein, “gene silencing oligonucleotide” refers to any oligonucleotide that can alone or with other gene silencing oligonucleotides utilize a cell’s endogenous mechanisms, molecules, proteins, enzymes, and/or other cell machinery or exogenous molecule, agent, protein, enzyme, and/or polynucleotide to cause a global or specific reduction or elimination in gene expression, RNA level(s), RNA translation, RNA transcription, that can lead to a reduction or effective loss of a protein expression and/or function of a non-coding

RNA as compared to wild-type or a suitable control. This is synonymous with the phrase "gene knockdown" Reduction in gene expression, RNA level(s), RNA translation, RNA transcription, and/or protein expression can range from about 100, 99, 98, 97, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, to 1% or less reduction. "Gene silencing oligonucleotides" include, but are not limited to, any antisense oligonucleotide, ribozyme, any oligonucleotide (single or double stranded) used to stimulate the RNA interference (RNAi) pathway in a cell (collectively RNAi oligonucleotides), small interfering RNA (siRNA), microRNA, and short-hairpin RNA (shRNA). Commercially available programs and tools are available to design the nucleotide sequence of gene silencing oligonucleotides for a desired gene, based on the gene sequence and other information available to one of ordinary skill in the art.

Therapeutic Polynucleotides

[0126] In some embodiments, the cargo molecule is a therapeutic polynucleotide. Therapeutic polynucleotides are those that provide a therapeutic effect when delivered to a recipient cell. The polynucleotide can be a toxic polynucleotide (a polynucleotide that when transcribed or translated results in the death of the cell) or polynucleotide that encodes a lytic peptide or protein. In embodiments, delivery vesicles having a toxic polynucleotide as a cargo molecule can act as an antimicrobial or antibiotic. This is discussed in greater detail elsewhere herein. In some embodiments, the cargo molecule can be exogenous to the producer cell and/or a first cell. In some embodiments, the cargo molecule can be endogenous to the producer cell and/or a first cell. In some embodiments, the cargo molecule can be exogenous to the recipient cell and/or a second cell. In some embodiments, the cargo molecule can be endogenous to the recipient cell and/or second cell.

[0127] As described herein the cargo polynucleotide can be any polynucleotide endogenous or exogenous to the eukaryotic cell. For example, the cargo polynucleotide can be a polynucleotide residing in the nucleus of the eukaryotic cell. The cargo polynucleotide can be a sequence coding a gene product (e.g., a protein) or a non-coding sequence (e.g., a regulatory polynucleotide).

[0128] In some embodiments, the cargo polynucleotide is a DNA or RNA (e.g. a mRNA) vaccine.

Aptamers

[0129] In certain example embodiments, the polynucleotide may be an aptamer. In certain embodiments, the one or more agents is an aptamer. Nucleic acid aptamers are nucleic acid species that have been engineered through repeated rounds of in vitro selection or equivalently, SELEX (systematic evolution of ligands by exponential enrichment) to bind to various molecular targets such as small molecules, proteins, nucleic acids, cells, tissues and organisms. Nucleic acid aptamers have specific binding affinity to molecules through interactions other than classic Watson-Crick base pairing. Aptamers are useful in biotechnological and therapeutic applications as they offer molecular recognition properties similar to antibodies. In addition to their discriminate recognition, aptamers offer advantages over antibodies as they can be engineered completely in a test tube, are readily produced by chemical synthesis, possess desirable storage properties, and elicit little or no immunogenicity in therapeutic applications. In certain embodiments, RNA aptamers may be expressed from a DNA construct. In other embodiments, a nucleic acid aptamer may be linked to another polynucleotide sequence. The polynucleotide sequence may be a double stranded DNA polynucleotide sequence. The aptamer may be covalently linked to one strand of the polynucleotide sequence. The aptamer may be ligated to the polynucleotide sequence. The polynucleotide sequence may be configured, such that the polynucleotide sequence may be linked to a solid support or ligated to another polynucleotide sequence.

[0130] Aptamers, like peptides generated by phage display or monoclonal antibodies ("mAbs"), are capable of specifically binding to selected targets and modulating the target's activity, e.g., through binding, aptamers may block their target's ability to function. A typical aptamer is 10-15 kDa in size (30-45 nucleotides), binds its target with sub-nanomolar affinity, and discriminates against closely related targets (e.g., aptamers will typically not bind other proteins from the same gene family). Structural studies have shown that aptamers are capable of using the same types of binding interactions (e.g., hydrogen bonding, electrostatic complementarity, hydrophobic contacts, steric exclusion) that drives affinity and specificity in antibody-antigen complexes.

[0131] Aptamers have a number of desirable characteristics for use in research and as therapeutics and diagnostics including high specificity and affinity, biological efficacy, and

excellent pharmacokinetic properties. In addition, they offer specific competitive advantages over antibodies and other protein biologics. Aptamers are chemically synthesized and are readily scaled as needed to meet production demand for research, diagnostic or therapeutic applications. Aptamers are chemically robust. They are intrinsically adapted to regain activity following exposure to factors such as heat and denaturants and can be stored for extended periods (>1 yr) at room temperature as lyophilized powders. Not being bound by a theory, aptamers bound to a solid support or beads may be stored for extended periods.

[0132] Oligonucleotides in their phosphodiester form may be quickly degraded by intracellular and extracellular enzymes such as endonucleases and exonucleases. Aptamers can include modified nucleotides conferring improved characteristics on the ligand, such as improved in vivo stability or improved delivery characteristics. Examples of such modifications include chemical substitutions at the ribose and/or phosphate and/or base positions. SELEX identified nucleic acid ligands containing modified nucleotides are described, e.g., in U.S. Pat. No. 5,660,985, which describes oligonucleotides containing nucleotide derivatives chemically modified at the 2' position of ribose, 5 position of pyrimidines, and 8 position of purines, U.S. Pat. No. 5,756,703 which describes oligonucleotides containing various 2' -modified pyrimidines, and U.S. Pat. No. 5,580,737 which describes highly specific nucleic acid ligands containing one or more nucleotides modified with 2'-amino (2'-NH2), 2'-fluoro (2'-F), and/or 2'-0-methyl (2'-OMe) substituents. Modifications of aptamers may also include modifications at exocyclic amines, substitution of 4- thiouridine, substitution of 5-bromo or 5-iodo-uracil; backbone modifications, phosphorothioate or allyl phosphate modifications, methylations, and unusual base-pairing combinations such as the isobases isocytidine and isoguanosine. Modifications can also include 3' and 5' modifications such as capping. As used herein, the term phosphorothioate encompasses one or more non-bridging oxygen atoms in a phosphodiester bond replaced by one or more sulfur atoms. In further embodiments, the oligonucleotides comprise modified sugar groups, for example, one or more of the hydroxyl groups is replaced with halogen, aliphatic groups, or functionalized as ethers or amines. In one embodiment, the 2'-position of the furanose residue is substituted by any of an O-methyl, O-alkyl, O-allyl, S-alkyl, S-allyl, or halo group. Methods of synthesis of 2'-modified sugars are described, e.g., in Sproat, et al., Nucl. Acid Res. 19:733-738 (1991); Cotten, et al, Nucl. Acid Res. 19:2629-2635 (1991); and Hobbs, et al, Biochemistry 12:5138-5145 (1973). Other modifications are known to one of

ordinary skill in the art. In certain embodiments, aptamers include aptamers with improved off-rates as described in International Patent Publication No. WO 2009012418, "Method for generating aptamers with improved off-rates," incorporated herein by reference in its entirety. In certain embodiments aptamers are chosen from a library of aptamers. Such libraries include, but are not limited to those described in Rohloff et al., "Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents," Molecular Therapy Nucleic Acids (2014) 3, e201. Aptamers are also commercially available (see e.g., SomaLogic, Inc., Boulder, Colorado). In certain embodiments, the present invention may utilize any aptamer containing any modification as described herein.

[0133] In certain other example embodiments, the polynucleotide may be a ribozyme or other enzymatically active polynucleotide.

Biologically active agents

[0134] In some embodiments, the cargo is a biologically active agent. Biologically active agents include any molecule that induces, directly or indirectly, an effect in a cell. Biologically active agents may be a protein, a nucleic acid, a small molecule, a carbohydrate, and a lipid. When the cargo is or comprises a nucleic acid, the nucleic acid may be a separate entity from the DNA-based carrier. In these embodiments, the DNA-based carrier is not itself the cargo. In other embodiments, the DNA-based carrier may itself comprise a nucleic acid cargo. Therapeutic agents include, without limitation, chemotherapeutic agents, anti-oncogenic agents, anti-angiogenic agents, tumor suppressor agents, anti-microbial agents, enzyme replacement agents, gene expression modulating agents and expression constructs comprising a nucleic acid encoding a therapeutic protein or nucleic acid, and vaccines. Therapeutic agents may be peptides, proteins (including enzymes, antibodies and peptidic hormones), ligands of cytoskeleton, nucleic acid, small molecules, non-peptidic hormones and the like. To increase affinity for the nucleus, agents may be conjugated to a nuclear localization sequence. Nucleic acids that may be delivered by the method of the invention include synthetic and natural nucleic acid material, including DNA, RNA, transposon DNA, antisense nucleic acids, dsRNA, siRNAs, transcription RNA, messenger RNA, ribosomal RNA, small nucleolar RNA, microRNA, ribozymes, plasmids, expression constructs, etc.

[0135] Imaging agents include contrast agents, such as ferrofluid-based MRI contrast agents and gadolinium agents for PET scans, fluorescein isothiocyanate and 6-TAMARA. Monitoring agents include reporter probes, biosensors, green fluorescent protein and the like.

Reporter probes include photo-emitting compounds, such as phosphors, radioactive moieties and fluorescent moieties, such as rare earth chelates (e.g., europium chelates), Texas Red, rhodamine, fluorescein, FITC, fluo-3, 5 hexadecanoyl fluorescein, Cy2, fluor X, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, dansyl, phycocrytherin, phycocyanin, spectrum orange, spectrum green, and/or derivatives of any one or more of the above. Biosensors are molecules that detect and transmit information regarding a physiological change or process, for instance, by detecting the presence or change in the presence of a chemical. The information obtained by the biosensor typically activates a signal that is detected with a transducer. The transducer typically converts the biological response into an electrical signal. Examples of biosensors include enzymes, antibodies, DNA, receptors and regulator proteins used as recognition elements, which can be used either in whole cells or isolated and used independently (D'Souza, 2001, Biosensors and Bioelectronics 16:337-353).

[0136] One or two or more different cargoes may be delivered by the delivery particles described herein.

[0137] In some embodiments, the cargo may be linked to one or more envelope proteins by a linker, as described elsewhere herein. A suitable linker may include, but is not necessarily limited to, a glycine-serine linker. In some embodiments, the glycine-serine linker is $(GGS)_3$ (SEQ ID NO: 3).

[0138] In some embodiments, the cargo comprises a ribonucleoprotein. In specific embodiments, the cargo comprises a genetic modulating agent.

[0139] As used herein the term "altered expression" may particularly denote altered production of the recited gene products by a cell. As used herein, the term "gene product(s)" includes RNA transcribed from a gene (e.g., mRNA), or a polypeptide encoded by a gene or translated from RNA.

[0140] Also, "altered expression" as intended herein may encompass modulating the activity of one or more endogenous gene products. Accordingly, "altered expression", "altering expression", "modulating expression", or "detecting expression" or similar may be used interchangeably with respectively "altered expression or activity", "altering expression or activity", "modulating expression or activity", or "detecting expression or activity" or similar terms. As used herein, "modulating" or "to modulate" generally means either reducing or inhibiting the activity of a target or antigen, or alternatively increasing the activity of the target or antigen, as measured using a suitable *in vitro*, cellular or *in vivo* assay. In particular,

"modulating" or "to modulate" can mean either reducing or inhibiting the (relevant or intended) activity of, or alternatively increasing the (relevant or intended) biological activity of the target or antigen, as measured using a suitable *in vitro*, cellular or *in vivo* assay (which will usually depend on the target or antigen involved), by at least 5%, at least 10%, at least 25%, at least 50%, at least 60%, at least 70%, at least 80%, or 90% or more, compared to activity of the target or antigen in the same assay under the same conditions but without the presence of the inhibitor/antagonist agents or activator/agonist agents described herein.

[0141] As will be clear to the skilled person, "modulating" can also involve effecting a change (which can either be an increase or a decrease) in affinity, avidity, specificity and/or selectivity of a target or antigen, for one or more of its targets compared to the same conditions but without the presence of a modulating agent. Again, this can be determined in any suitable manner and/or using any suitable assay known per se, depending on the target. In particular, an action as an inhibitor/antagonist or activator/agonist can be such that an intended biological or physiological activity is increased or decreased, respectively, by at least 5%, at least 10%, at least 25%, at least 50%, at least 60%, at least 70%, at least 80%, or 90% or more, compared to the biological or physiological activity in the same assay under the same conditions but without the presence of the inhibitor/antagonist agent or activator/agonist agent. Modulating can also involve activating the target or antigen or the mechanism or pathway in which it is involved.

Genetic Modifying Systems

[0142] In some embodiments, the cargo is a polynucleotide modifying system or component(s) thereof. In some embodiments the polynucleotide modifying system is a gene modifying system. In some embodiments, the gene modifying system is or is composed of a gene modulating agent. In some embodiments, the genetic modulating agent may comprise one or more components of a polynucleotide modification system (e.g., a gene editing system) and/or polynucleotides encoding thereof.

[0143] In some embodiments, the gene editing system may be an RNA-guided system or other programmable nuclease system. In some embodiments, the gene editing system is an IscB system. In some embodiments, the gene editing system may be a CRISPR-Cas system.

CRISPR-Cas Systems

[0144] In general, a CRISPR-Cas or CRISPR system as used in herein and in documents, such as WO 2014/093622 (PCT/US2013/074667), refers collectively to transcripts and other

elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (e.g. tracrRNA or an active partial tracrRNA), a tracr-mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), or "RNA(s)" as that term is herein used (e.g., RNA(s) to guide Cas, such as Cas9, e.g. CRISPR RNA and transactivating (tracr) RNA or a single guide RNA (sgRNA) (chimeric RNA)) or other sequences and transcripts from a CRISPR locus. In general, a CRISPR system is characterized by elements that promote the formation of a CRISPR complex at the site of a target sequence (also referred to as a protospacer in the context of an endogenous CRISPR system). See, e.g., Shmakov et al. (2015) "Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems", Molecular Cell, DOI: dx.doi.org/10.1016/j.molcel.2015.10.008.

Class 1 Systems

[0145] The methods, systems, and tools provided herein may be designed for use with Class 1 CRISPR proteins. In certain example embodiments, the Class 1 system may be Type I, Type III or Type IV Cas proteins as described in Makarova et al. "Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants" Nature Reviews Microbiology, 18:67-81 (Feb 2020)., incorporated in its entirety herein by reference, and particularly as described in Figure 1, p. 326. The Class 1 systems typically use a multi-protein effector complex, which can, in some embodiments, include ancillary proteins, such as one or more proteins in a complex referred to as a CRISPR-associated complex for antiviral defense (Cascade), one or more adaptation proteins (e.g. Cas1, Cas2, RNA nuclease), and/or one or more accessory proteins (e.g. Cas 4, DNA nuclease), CRISPR associated Rossman fold (CARF) domain containing proteins, and/or RNA transcriptase. Although Class 1 systems have limited sequence similarity, Class 1 system proteins can be identified by their similar architectures, including one or more Repeat Associated Mysterious Protein (RAMP) family subunits, e.g. Cas 5, Cas6, Cas7. RAMP proteins are characterized by having one or more RNA recognition motif domains. Large subunits (for example cas8 or cas10) and small subunits (for example, cas11) are also typical of Class 1 systems. See, e.g., Figures 1 and 2. Koonin EV, Makarova KS. 2019 Origins and evolution of CRISPR-Cas systems. Phil. Trans. R. Soc. B 374: 20180087, DOI: 10.1098/rstb.2018.0087. In one aspect, Class 1 systems are

characterized by the signature protein Cas3. The Cascade in particular Class1 proteins can comprise a dedicated complex of multiple Cas proteins that binds pre-crRNA and recruits an additional Cas protein, for example Cas6 or Cas5, which is the nuclease directly responsible for processing pre-crRNA. In one aspect, the Type I CRISPR protein comprises an effector complex comprises one or more Cas5 subunits and two or more Cas7 subunits. Class 1 subtypes include Type I-A, I-B, I-C, I-U, I-D, I-E, and I-F, Type IV-A and IV-B, and Type III-A, III-D, III-C, and III-B. Class 1 systems also include CRISPR-Cas variants, including Type I-A, I-B, I-E, I-F and I-U variants, which can include variants carried by transposons and plasmids, including versions of subtype I-F encoded by a large family of Tn7-like transposon and smaller groups of Tn7-like transposons that encode similarly degraded subtype I-B systems. Peters *et al.*, PNAS 114 (35) (2017); DOI: 10.1073/pnas.1709035114; *see also,* Makarova et al, the CRISPR Journal, v. 1, n5, Figure 5.

Class 2 Systems

[0146] The compositions, systems, and methods described in greater detail elsewhere herein can be designed and adapted for use with Class 2 CRISPR-Cas systems. Thus, in some embodiments, the CRISPR-Cas system is a Class 2 CRISPR-Cas system. Class 2 systems are distinguished from Class 1 systems in that they have a single, large, multi-domain effector protein. In certain example embodiments, the Class 2 system can be a Type II, Type V, or Type VI system, which are described in Makarova et al. "Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants" Nature Reviews Microbiology, 18:67-81 (Feb 2020), incorporated herein by reference. Each type of Class 2 system is further divided into subtypes. *See* Markova et al. 2020, particularly at Figure. 2. Class 2, Type II systems can be divided into 4 subtypes: II-A, II-B, II-C1, and II-C2. Class 2, Type V systems can be divided into 17 subtypes: V-A, V-B1, V-B2, V-C, V-D, V-E, V-F1, V-F1(V-U3), V-F2, V-F3, V-G, V-H, V-I, V-K (V-U5), V-U1, V-U2, and V-U4. Class 2, Type IV systems can be divided into 5 subtypes: VI-A, VI-B1, VI-B2, VI-C, and VI-D.

[0147] The distinguishing feature of these types is that their effector complexes consist of a single, large, multi-domain protein. Type V systems differ from Type II effectors (e.g., Cas9), which contain two nuclear domains that are each responsible for the cleavage of one strand of the target DNA, with the HNH nuclease inserted inside the Ruv-C like nuclease domain sequence. The Type V systems (e.g., Cas12) only contain a RuvC-like nuclease domain that cleaves both strands. Type VI (Cas13) are unrelated to the effectors of Type II and V systems

and contain two HEPN domains and target RNA. Cas13 proteins also display collateral activity that is triggered by target recognition. Some Type V systems have also been found to possess this collateral activity with two single-stranded DNA in in vitro contexts.

[0148] In some embodiments, the Class 2 system is a Type II system. In some embodiments, the Type II CRISPR-Cas system is a II-A CRISPR-Cas system. In some embodiments, the Type II CRISPR-Cas system is a II-B CRISPR-Cas system. In some embodiments, the Type II CRISPR-Cas system is a II-C1 CRISPR-Cas system. In some embodiments, the Type II CRISPR-Cas system is a II-C2 CRISPR-Cas system. In some embodiments, the Type II system is a Cas9 system. In some embodiments, the Type II system includes a Cas9.

[0149] In some embodiments, the Class 2 system is a Type V system. In some embodiments, the Type V CRISPR-Cas system is a V-A CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-B1 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-B2 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-C CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-D CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-E CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-F1 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-F1 (V-U3) CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-F2 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-F3 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-G CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-H CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-I CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-K (V-U5) CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-U1 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-U2 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system is a V-U4 CRISPR-Cas system. In some embodiments, the Type V CRISPR-Cas system includes a Cas12a (Cpf1), Cas12b (C2c1), Cas12c (C2c3), Cas12d (CasY), Cas12e (CasX), Cas14, and/or CasΦ.

[0150] In some embodiments the Class 2 system is a Type VI system. In some embodiments, the Type VI CRISPR-Cas system is a VI-A CRISPR-Cas system. In some

embodiments, the Type VI CRISPR-Cas system is a VI-B1 CRISPR-Cas system. In some embodiments, the Type VI CRISPR-Cas system is a VI-B2 CRISPR-Cas system. In some embodiments, the Type VI CRISPR-Cas system is a VI-C CRISPR-Cas system. In some embodiments, the Type VI CRISPR-Cas system is a VI-D CRISPR-Cas system. In some embodiments, the Type VI CRISPR-Cas system includes a Cas13a (C2c2), Cas13b (Group 29/30), Cas13c, and/or Cas13d.

Guide Molecules

[0151] The CRISPR-Cas or Cas-Based system described herein can, in some embodiments, include one or more guide molecules. The terms guide molecule, guide sequence and guide polynucleotide refer to polynucleotides capable of guiding Cas to a target genomic locus and are used interchangeably as in foregoing cited documents such as International Patent Publication No. WO 2014/093622 (PCT/US2013/074667). In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with a target polynucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of a CRISPR complex to the target sequence. The guide molecule can be a polynucleotide.

[0152] The ability of a guide sequence (within a nucleic acid-targeting guide RNA) to direct sequence-specific binding of a nucleic acid-targeting complex to a target nucleic acid sequence may be assessed by any suitable assay. For example, the components of a nucleic acid-targeting CRISPR system sufficient to form a nucleic acid-targeting complex, including the guide sequence to be tested, may be provided to a host cell having the corresponding target nucleic acid sequence, such as by transfection with vectors encoding the components of the nucleic acid-targeting complex, followed by an assessment of preferential targeting (e.g., cleavage) within the target nucleic acid sequence, such as by Surveyor assay (Qui et al. 2004. BioTechniques. 36(4)702-707). Similarly, cleavage of a target nucleic acid sequence may be evaluated in a test tube by providing the target nucleic acid sequence, components of a nucleic acid-targeting complex, including the guide sequence to be tested and a control guide sequence different from the test guide sequence, and comparing binding or rate of cleavage at the target sequence between the test and control guide sequence reactions. Other assays are possible and will occur to those skilled in the art.

[0153] In some embodiments, the guide molecule is an RNA. The guide molecule(s) (also referred to interchangeably herein as guide polynucleotide and guide sequence) that are included in the CRISPR-Cas or Cas based system can be any polynucleotide sequence having

sufficient complementarity with a target nucleic acid sequence to hybridize with the target nucleic acid sequence and direct sequence-specific binding of a nucleic acid-targeting complex to the target nucleic acid sequence. In some embodiments, the degree of complementarity, when optimally aligned using a suitable alignment algorithm, can be about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting examples of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g., the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; available at www.novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net).

[0154] A guide sequence, and hence a nucleic acid-targeting guide, may be selected to target any target nucleic acid sequence. The target sequence may be DNA. The target sequence may be any RNA sequence. In some embodiments, the target sequence may be a sequence within an RNA molecule selected from the group consisting of messenger RNA (mRNA), pre-mRNA, ribosomal RNA (rRNA), transfer RNA (tRNA), micro-RNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), double stranded RNA (dsRNA), non-coding RNA (ncRNA), long non-coding RNA (lncRNA), and small cytoplasmatic RNA (scRNA). In some preferred embodiments, the target sequence may be a sequence within an RNA molecule selected from the group consisting of mRNA, pre-mRNA, and rRNA. In some preferred embodiments, the target sequence may be a sequence within an RNA molecule selected from the group consisting of ncRNA, and lncRNA. In some more preferred embodiments, the target sequence may be a sequence within an mRNA molecule or a pre-mRNA molecule.

[0155] In some embodiments, a nucleic acid-targeting guide is selected to reduce the degree secondary structure within the nucleic acid-targeting guide. In some embodiments, about or less than about 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, or fewer of the nucleotides of the nucleic acid-targeting guide participate in self-complementary base pairing when optimally folded. Optimal folding may be determined by any suitable polynucleotide folding algorithm. Some programs are based on calculating the minimal Gibbs free energy. An example of one such algorithm is mFold, as described by Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Another example folding algorithm is the online webserver RNAfold,

developed at Institute for Theoretical Chemistry at the University of Vienna, using the centroid structure prediction algorithm (*see* e.g., A.R. Gruber et al., 2008, Cell 106(1): 23-24; and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62).

[0156] In certain embodiments, a guide RNA or crRNA may comprise, consist essentially of, or consist of a direct repeat (DR) sequence and a guide sequence or spacer sequence. In certain embodiments, the guide RNA or crRNA may comprise, consist essentially of, or consist of a direct repeat sequence fused or linked to a guide sequence or spacer sequence. In certain embodiments, the direct repeat sequence may be located upstream (i.e., 5') from the guide sequence or spacer sequence. In other embodiments, the direct repeat sequence may be located downstream (i.e., 3') from the guide sequence or spacer sequence.

[0157] In certain embodiments, the crRNA comprises a stem loop, preferably a single stem loop. In certain embodiments, the direct repeat sequence forms a stem loop, preferably a single stem loop.

[0158] In certain embodiments, the spacer length of the guide RNA is from 15 to 35 nt. In certain embodiments, the spacer length of the guide RNA is at least 15 nucleotides. In certain embodiments, the spacer length is from 15 to 17 nt, e.g., 15, 16, or 17 nt, from 17 to 20 nt, e.g., 17, 18, 19, or 20 nt, from 20 to 24 nt, e.g., 20, 21, 22, 23, or 24 nt, from 23 to 25 nt, e.g., 23, 24, or 25 nt, from 24 to 27 nt, e.g., 24, 25, 26, or 27 nt, from 27 to 30 nt, e.g., 27, 28, 29, or 30 nt, from 30 to 35 nt, e.g., 30, 31, 32, 33, 34, or 35 nt, or 35 nt or longer.

[0159] The "tracrRNA" sequence or analogous terms includes any polynucleotide sequence that has sufficient complementarity with a crRNA sequence to hybridize. In some embodiments, the degree of complementarity between the tracrRNA sequence and crRNA sequence along the length of the shorter of the two when optimally aligned is about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. In some embodiments, the tracr sequence is about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides in length. In some embodiments, the tracr sequence and crRNA sequence are contained within a single transcript, such that hybridization between the two produces a transcript having a secondary structure, such as a hairpin.

[0160] In general, degree of complementarity is with reference to the optimal alignment of the sca sequence and tracr sequence, along the length of the shorter of the two sequences. Optimal alignment may be determined by any suitable alignment algorithm and may further account for secondary structures, such as self-complementarity within either the sca sequence

or tracr sequence. In some embodiments, the degree of complementarity between the tracr sequence and sca sequence along the length of the shorter of the two when optimally aligned is about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher.

[0161] In some embodiments, the degree of complementarity between a guide sequence and its corresponding target sequence can be about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or 100%; a guide or RNA or sgRNA can be about or more than about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, or more nucleotides in length; or guide or RNA or sgRNA can be less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length; and tracr RNA can be 30 or 50 nucleotides in length. In some embodiments, the degree of complementarity between a guide sequence and its corresponding target sequence is greater than 94.5% or 95% or 95.5% or 96% or 96.5% or 97% or 97.5% or 98% or 98.5% or 99% or 99.5% or 99.9%, or 100%. Off target is less than 100% or 99.9% or 99.5% or 99% or 99% or 98.5% or 98% or 97.5% or 97% or 96.5% or 96% or 95.5% or 95% or 94.5% or 94% or 93% or 92% or 91% or 90% or 89% or 88% or 87% or 86% or 85% or 84% or 83% or 82% or 81% or 80% complementarity between the sequence and the guide, with it being advantageous that off target is 100% or 99.9% or 99.5% or 99% or 99% or 98.5% or 98% or 97.5% or 97% or 96.5% or 96% or 95.5% or 95% or 94.5% complementarity between the sequence and the guide.

[0162] In some embodiments according to the invention, the guide RNA (capable of guiding Cas to a target locus) may comprise (1) a guide sequence capable of hybridizing to a genomic target locus in the eukaryotic cell; (2) a tracr sequence; and (3) a tracr mate sequence. All (1) to (3) may reside in a single RNA, i.e., an sgRNA (arranged in a 5' to 3' orientation), or the tracr RNA may be a different RNA than the RNA containing the guide and tracr sequence. The tracr hybridizes to the tracr mate sequence and directs the CRISPR/Cas complex to the target sequence. Where the tracr RNA is on a different RNA than the RNA containing the guide and tracr sequence, the length of each RNA may be optimized to be shortened from their respective native lengths, and each may be independently chemically modified to protect from degradation by cellular RNase or otherwise increase stability.

[0163] Many modifications to guide sequences are known in the art and are further contemplated within the context of this invention. Various modifications may be used to increase the specificity of binding to the target sequence and/or increase the activity of the Cas

protein and/or reduce off-target effects. Example guide sequence modifications are described in International Patent Application No. PCT US2019/045582, specifically paragraphs [0178]-[0333]. which is incorporated herein by reference.

Target Sequences, PAMs, and PFSs

[0164] In the context of formation of a CRISPR complex, "target sequence" refers to a sequence to which a guide sequence is designed to have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a CRISPR complex. A target sequence may comprise RNA polynucleotides. The term "target RNA" refers to an RNA polynucleotide being or comprising the target sequence. In other words, the target polynucleotide can be a polynucleotide or a part of a polynucleotide to which a part of the guide sequence is designed to have complementarity with and to which the effector function mediated by the complex comprising the CRISPR effector protein and a guide molecule is to be directed. In some embodiments, a target sequence is located in the nucleus or cytoplasm of a cell.

[0165] The guide sequence can specifically bind a target sequence in a target polynucleotide. The target polynucleotide may be DNA. The target polynucleotide may be RNA. The target polynucleotide can have one or more (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, etc. or more) target sequences. The target polynucleotide can be on a vector. The target polynucleotide can be genomic DNA. The target polynucleotide can be episomal. Other forms of the target polynucleotide are described elsewhere herein.

[0166] The target sequence may be DNA. The target sequence may be any RNA sequence. In some embodiments, the target sequence may be a sequence within an RNA molecule selected from the group consisting of messenger RNA (mRNA), pre-mRNA, ribosomal RNA (rRNA), transfer RNA (tRNA), micro-RNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), double stranded RNA (dsRNA), non-coding RNA (ncRNA), long non-coding RNA (lncRNA), and small cytoplasmatic RNA (scRNA). In some preferred embodiments, the target sequence (also referred to herein as a target polynucleotide) may be a sequence within an RNA molecule selected from the group consisting of mRNA, pre-mRNA, and rRNA. In some preferred embodiments, the target sequence may be a sequence within an RNA molecule selected from the group consisting of ncRNA, and lncRNA. In some more preferred embodiments, the target sequence may be a sequence within an mRNA molecule or a pre-mRNA molecule.

PAM and PFS Elements

[0167] PAM elements are sequences that can be recognized and bound by Cas proteins. Cas proteins/effector complexes can then unwind the dsDNA at a position adjacent to the PAM element. It will be appreciated that Cas proteins and systems that include them that target RNA do not require PAM sequences (Marraffini et al. 2010. Nature. 463:568-571). Instead, many rely on PFSs, which are discussed elsewhere herein. In certain embodiments, the target sequence should be associated with a PAM (protospacer adjacent motif) or PFS (protospacer flanking sequence or site), that is, a short sequence recognized by the CRISPR complex. Depending on the nature of the CRISPR-Cas protein, the target sequence should be selected, such that its complementary sequence in the DNA duplex (also referred to herein as the non-target sequence) is upstream or downstream of the PAM. In the embodiments, the complementary sequence of the target sequence is downstream or 3' of the PAM or upstream or 5' of the PAM. The precise sequence and length requirements for the PAM differ depending on the Cas protein used, but PAMs are typically 2-5 base pair sequences adjacent the protospacer (that is, the target sequence). Examples of the natural PAM sequences for different Cas proteins are provided herein below and the skilled person will be able to identify further PAM sequences for use with a given Cas protein.

[0168] The ability to recognize different PAM sequences depends on the Cas polypeptide(s) included in the system. *See* e.g., Gleditzsch et al. 2019. RNA Biology. 16(4):504-517. **Table 1** (from Gleditzsch et al. 2019) below shows several Cas polypeptides and the PAM sequence they recognize.

Table 1– Example PAM Sequences

Cas Protein	PAM Sequence
SpCas9	NGG/NRG
SaCas9	NGRRT or NGRRN
NmeCas9	NNNNGATT
CjCas9	NNNNRYAC
StCas9	NNAGAAW
Cas12a (Cpf1) (including LbCpf1 and AsCpf1)	TTTV
Cas12b (C2c1)	TTT, TTA, and TTC
Cas12c (C2c3)	TA
Cas12d (CasY)	TA
Cas12e (CasX)	5′-TTCN-3′

[0169] In a preferred embodiment, the CRISPR effector protein may recognize a 3' PAM. In certain embodiments, the CRISPR effector protein may recognize a 3' PAM which is 5'H, wherein H is A, C or U.

[0170] Further, engineering of the PAM Interacting (PI) domain on the Cas protein may allow programing of PAM specificity, improve target site recognition fidelity, and increase the versatility of the CRISPR-Cas protein, for example as described for Cas9 in Kleinstiver BP et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015 Jul 23;523(7561):481-5. doi: 10.1038/nature14592. As further detailed herein, the skilled person will understand that Cas13 proteins may be modified analogously. Gao *et al*, "Engineered Cpf1 Enzymes with Altered PAM Specificities," bioRxiv 091611; doi: http://dx.doi.org/10.1101/091611 (Dec. 4, 2016). Doench et al. created a pool of sgRNAs, tiling across all possible target sites of a panel of six endogenous mouse and three endogenous human genes and quantitatively assessed their ability to produce null alleles of their target gene by antibody staining and flow cytometry. The authors showed that optimization of the PAM improved activity and also provided an on-line tool for designing sgRNAs.

[0171] PAM sequences can be identified in a polynucleotide using an appropriate design tool, which are commercially available as well as online. Such freely available tools include, but are not limited to, CRISPRFinder and CRISPRTarget. Mojica et al. 2009. Microbiol. 155(Pt. 3):733-740; Atschul et al. 1990. J. Mol. Biol. 215:403-410; Biswass et al. 2013 RNA Biol. 10:817-827; and Grissa et al. 2007. Nucleic Acid Res. 35:W52-57. Experimental approaches to PAM identification can include, but are not limited to, plasmid depletion assays (Jiang et al. 2013. Nat. Biotechnol. 31:233-239; Esvelt et al. 2013. Nat. Methods. 10:1116-1121; Kleinstiver et al. 2015. Nature. 523:481-485), screened by a high-throughput *in vivo* model called PAM-SCNAR (Pattanayak et al. 2013. Nat. Biotechnol. 31:839-843 and Leenay et al. 2016.Mol. Cell. 16:253), and negative screening (Zetsche et al. 2015. Cell. 163:759-771).

[0172] As previously mentioned, CRISPR-Cas systems that target RNA do not typically rely on PAM sequences. Instead, such systems typically recognize protospacer flanking sites (PFSs) instead of PAMs Thus, Type VI CRISPR-Cas systems typically recognize protospacer flanking sites (PFSs) instead of PAMs. PFSs represents an analogue to PAMs for RNA targets. Type VI CRISPR-Cas systems employ a Cas13. Some Cas13 proteins analyzed to date, such as Cas13a (C2c2) identified from Leptotrichia shahii (LShCAs13a) have a specific discrimination against G at the 3'end of the target RNA. The presence of a C at the

corresponding crRNA repeat site can indicate that nucleotide pairing at this position is rejected. However, some Cas13 proteins (e.g., LwaCAs13a and PspCas13b) do not seem to have a PFS preference. See e.g., Gleditzsch et al. 2019. RNA Biology. 16(4):504-517.

[0173] Some Type VI proteins, such as subtype B, have 5′-recognition of D (G, T, A) and a 3′-motif requirement of NAN or NNA. One example is the Cas13b protein identified in Bergeyella zoohelcum (BzCas13b). See e.g., Gleditzsch et al. 2019. RNA Biology. 16(4):504-517.

[0174] Overall Type VI CRISPR-Cas systems appear to have less restrictive rules for substrate (e.g., target sequence) recognition than those that target DNA (e.g., Type V and type II).

Sequences related to nucleus targeting and transportation

[0175] In some embodiments, one or more components (e.g., the Cas protein and/or deaminase) in the composition for engineering cells may comprise one or more sequences related to nucleus targeting and transportation. Such sequence may facilitate the one or more components in the composition for targeting a sequence within a cell. In order to improve targeting of the CRISPR-Cas protein and/or the nucleotide deaminase protein or catalytic domain thereof used in the methods of the present disclosure to the nucleus, it may be advantageous to provide one or both of these components with one or more nuclear localization sequences (NLSs).

[0176] In some embodiments, the NLSs used in the context of the present disclosure are heterologous to the proteins. Non-limiting examples of NLSs include an NLS sequence derived from: the NLS of the SV40 virus large T-antigen, having the amino acid sequence PKKKRKV (SEQ ID NO:4) or PKKKRKVEAS (SEQ ID NO:5); the NLS from nucleoplasmin (e.g., the nucleoplasmin bipartite NLS with the sequence KRPAATKKAGQAKKKK (SEQ ID NO:6)); the c-myc NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO:7) or RQRRNELKRSP (SEQ ID NO:8); the hRNPA1 M9 NLS having the sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO:9); the sequence RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO:10) of the IBB domain from importin-alpha; the sequences VSRKRPRP (SEQ ID NO:11) and PPKKARED (SEQ ID NO:12) of the myoma T protein; the sequence PQPKKKPL (SEQ ID NO:13) of human p53; the sequence SALIKKKKKMAP (SEQ ID NO:14) of mouse c-abl IV; the sequences DRLRR (SEQ ID NO:15) and PKQKKRK (SEQ ID NO:16) of the influenza

virus NS1; the sequence RKLKKKIKKL (SEQ ID NO:17) of the Hepatitis virus delta antigen; the sequence REKKKFLKRR (SEQ ID NO:18) of the mouse Mx1 protein; the sequence KRKGDEVDGVDEVAKKKSKK (SEQ ID NO:19) of the human poly(ADP-ribose) polymerase; and the sequence RKCLQAGMNLEARKTKK (SEQ ID NO:20) of the steroid hormone receptors (human) glucocorticoid. In general, the one or more NLSs are of sufficient strength to drive accumulation of the DNA-targeting Cas protein in a detectable amount in the nucleus of a eukaryotic cell. In general, strength of nuclear localization activity may derive from the number of NLSs in the CRISPR-Cas protein, the particular NLS(s) used, or a combination of these factors. Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to the nucleic acid-targeting protein, such that location within a cell may be visualized, such as in combination with a means for detecting the location of the nucleus (e.g., a stain specific for the nucleus such as DAPI). Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly, such as by an assay for the effect of nucleic acid-targeting complex formation (e.g., assay for deaminase activity) at the target sequence, or assay for altered gene expression activity affected by DNA-targeting complex formation and/or DNA-targeting), as compared to a control not exposed to the CRISPR-Cas protein and deaminase protein or exposed to a CRISPR-Cas and/or deaminase protein lacking the one or more NLSs.

[0177] The CRISPR-Cas and/or nucleotide deaminase proteins may be provided with 1 or more, such as with, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more heterologous NLSs. In some embodiments, the proteins comprises about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the amino-terminus, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more NLSs at or near the carboxy-terminus, or a combination of these (e.g., zero or at least one or more NLS at the amino-terminus and zero or at one or more NLS at the carboxy terminus). When more than one NLS is present, each may be selected independently of the others, such that a single NLS may be present in more than one copy and/or in combination with one or more other NLSs present in one or more copies. In some embodiments, an NLS is considered near the N- or C-terminus when the nearest amino acid of the NLS is within about 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, or more amino acids along the polypeptide chain from the N- or C-terminus. In

preferred embodiments of the CRISPR-Cas proteins, an NLS attached to the C-terminal of the protein.

[0178] In certain embodiments, the CRISPR-Cas protein and the deaminase protein are delivered to the cell or expressed within the cell as separate proteins. In these embodiments, each of the CRISPR-Cas and deaminase protein can be provided with one or more NLSs as described herein. In certain embodiments, the CRISPR-Cas and deaminase proteins are delivered to the cell or expressed with the cell as a fusion protein. In these embodiments one or both of the CRISPR-Cas and deaminase protein is provided with one or more NLSs. Where the nucleotide deaminase is fused to an adaptor protein (such as MS2) as described above, the one or more NLS can be provided on the adaptor protein, provided that this does not interfere with aptamer binding. In particular embodiments, the one or more NLS sequences may also function as linker sequences between the nucleotide deaminase and the CRISPR-Cas protein.

[0179] In certain embodiments, guides of the disclosure comprise specific binding sites (e.g., aptamers) for adapter proteins, which may be linked to or fused to a nucleotide deaminase or catalytic domain thereof. When such a guide forms a CRISPR complex (e.g., CRISPR-Cas protein binding to guide and target), the adapter proteins bind and the nucleotide deaminase or catalytic domain thereof associated with the adapter protein is positioned in a spatial orientation which is advantageous for the attributed function to be effective.

[0180] The skilled person will understand that modifications to the guide which allow for binding of the adapter + nucleotide deaminase, but not proper positioning of the adapter + nucleotide deaminase (e.g. due to steric hindrance within the three-dimensional structure of the CRISPR complex) are modifications which are not intended. The one or more modified guide may be modified at the tetra loop, the stem loop 1, stem loop 2, or stem loop 3, as described herein, preferably at either the tetra loop or stem loop 2, and in some cases at both the tetra loop and stem loop 2.

[0181] In some embodiments, a component (e.g., the dead Cas protein, the nucleotide deaminase protein or catalytic domain thereof, or a combination thereof) in the systems may comprise one or more nuclear export signals (NES), one or more nuclear localization signals (NLS), or any combinations thereof. In some cases, the NES may be an HIV Rev NES. In certain cases, the NES may be MAPK NES. When the component is a protein, the NES or NLS may be at the C terminus of component. Alternatively or additionally, the NES or NLS may be at the N terminus of component. In some examples, the Cas protein and optionally said

nucleotide deaminase protein or catalytic domain thereof comprise one or more heterologous nuclear export signal(s) (NES(s)) or nuclear localization signal(s) (NLS(s)), preferably an HIV Rev NES or MAPK NES, preferably C-terminal.

[0182] It will be appreciated that NLS and NES described herein with respect to Cas proteins can be used with other cargos, in particularly, gene modifying agents herein, and other proteins that can benefit from translocation in or out of a nuclease of a cell, such as a target cell.

Donor Templates

[0183] In some embodiments, the composition for engineering cells comprise a template, e.g., a recombination template. A template may be a component of another vector as described herein, contained in a separate vector, or provided as a separate polynucleotide. In some embodiments, a recombination template is designed to serve as a template in homologous recombination, such as within or near a target sequence nicked or cleaved by a nucleic acid-targeting effector protein as a part of a nucleic acid-targeting complex.

[0184] In an embodiment, the template nucleic acid alters the sequence of the target position. In an embodiment, the template nucleic acid results in the incorporation of a modified, or non-naturally occurring base into the target nucleic acid.

[0185] The template sequence may undergo a breakage mediated or catalyzed recombination with the target sequence. In an embodiment, the template nucleic acid may include sequence that corresponds to a site on the target sequence that is cleaved by a Cas protein mediated cleavage event. In an embodiment, the template nucleic acid may include a sequence that corresponds to both, a first site on the target sequence that is cleaved in a first Cas protein mediated event, and a second site on the target sequence that is cleaved in a second Cas protein mediated event.

[0186] In certain embodiments, the template nucleic acid can include a sequence which results in an alteration in the coding sequence of a translated sequence, e.g., one which results in the substitution of one amino acid for another in a protein product, e.g., transforming a mutant allele into a wild type allele, transforming a wild type allele into a mutant allele, and/or introducing a stop codon, insertion of an amino acid residue, deletion of an amino acid residue, or a nonsense mutation. In certain embodiments, the template nucleic acid can include a sequence which results in an alteration in a non-coding sequence, e.g., an alteration in an exon or in a 5' or 3' non-translated or non-transcribed region. Such alterations include an alteration

in a control element, e.g., a promoter, enhancer, and an alteration in a cis-acting or trans-acting control element.

[0187] A template nucleic acid having homology with a target position in a target gene may be used to alter the structure of a target sequence. The template sequence may be used to alter an unwanted structure, e.g., an unwanted or mutant nucleotide. The template nucleic acid may include a sequence which, when integrated, results in decreasing the activity of a positive control element; increasing the activity of a positive control element; decreasing the activity of a negative control element; increasing the activity of a negative control element; decreasing the expression of a gene; increasing the expression of a gene; increasing resistance to a disorder or disease; increasing resistance to viral entry; correcting a mutation or altering an unwanted amino acid residue conferring, increasing, abolishing or decreasing a biological property of a gene product, e.g., increasing the enzymatic activity of an enzyme, or increasing the ability of a gene product to interact with another molecule.

[0188] The template nucleic acid may include a sequence which results in a change in sequence of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 1 1, 12 or more nucleotides of the target sequence.

[0189] A template polynucleotide may be of any suitable length, such as about or more than about 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1000, or more nucleotides in length. In an embodiment, the template nucleic acid may be 20+/- 10, 30+/- 10, 40+/- 10, 50+/- 10, 60+/- 10, 70+/- 10, 80+/- 10, 90+/- 10, 100+/- 10, 1 10+/- 10, 120+/- 10, 130+/- 10, 140+/- 10, 150+/- 10, 160+/- 10, 170+/- 10, 1 80+/- 10, 190+/- 10, 200+/- 10, 210+/-10, of 220+/- 10 nucleotides in length. In an embodiment, the template nucleic acid may be 30+/-20, 40+/-20, 50+/-20, 60+/- 20, 70+/- 20, 80+/-20, 90+/-20, 100+/-20, 1 10+/-20, 120+/-20, 130+/-20, 140+/-20, I 50+/-20, 160+/-20, 170+/-20, 180+/-20, 190+/-20, 200+/-20, 210+/-20, of 220+/-20 nucleotides in length. In an embodiment, the template nucleic acid is 10 to 1 ,000, 20 to 900, 30 to 800, 40 to 700, 50 to 600, 50 to 500, 50 to 400, 50 to300, 50 to 200, or 50 to 100 nucleotides in length.

[0190] In some embodiments, the template polynucleotide is complementary to a portion of a polynucleotide comprising the target sequence. When optimally aligned, a template polynucleotide might overlap with one or more nucleotides of a target sequences (e.g. about or more than about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more nucleotides). In some embodiments, when a template sequence and a polynucleotide comprising a target sequence are optimally aligned, the nearest nucleotide of the template polynucleotide is within

about 1, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1000, 5000, 10000, or more nucleotides from the target sequence.

[0191] The exogenous polynucleotide template comprises a sequence to be integrated (e.g., a mutated gene). The sequence for integration may be a sequence endogenous or exogenous to the cell. Examples of a sequence to be integrated include polynucleotides encoding a protein or a non-coding RNA (e.g., a microRNA). Thus, the sequence for integration may be operably linked to an appropriate control sequence or sequences. Alternatively, the sequence to be integrated may provide a regulatory function.

[0192] An upstream or downstream sequence may comprise from about 20 bp to about 2500 bp, for example, about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, or 2500 bp. In some methods, the exemplary upstream or downstream sequence have about 200 bp to about 2000 bp, about 600 bp to about 1000 bp, or more particularly about 700 bp to about 1000.

[0193] An upstream or downstream sequence may comprise from about 20 bp to about 2500 bp, for example, about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, or 2500 bp. In some methods, the exemplary upstream or downstream sequence have about 200 bp to about 2000 bp, about 600 bp to about 1000 bp, or more particularly about 700 bp to about 1000.

[0194] In certain embodiments, one or both homology arms may be shortened to avoid including certain sequence repeat elements. For example, a 5' homology arm may be shortened to avoid a sequence repeat element. In other embodiments, a 3' homology arm may be shortened to avoid a sequence repeat element. In some embodiments, both the 5' and the 3' homology arms may be shortened to avoid including certain sequence repeat elements.

[0195] In some methods, the exogenous polynucleotide template may further comprise a marker. Such a marker may make it easy to screen for targeted integrations. Examples of suitable markers include restriction sites, fluorescent proteins, or selectable markers. The exogenous polynucleotide template of the disclosure can be constructed using recombinant techniques (see, for example, Sambrook et al., 2001 and Ausubel et al., 1996).

[0196] In certain embodiments, a template nucleic acid for correcting a mutation may designed for use as a single-stranded oligonucleotide. When using a single-stranded oligonucleotide, 5' and 3' homology arms may range up to about 200 base pairs (bp) in length, e.g., at least 25, 50, 75, 100, 125, 150, 175, or 200 bp in length.

[0197] Suzuki et al. describe *in vivo* genome editing via CRISPR/Cas9 mediated homology-independent targeted integration (2016, Nature 540:144–149).

Specialized Cas-based Systems

[0198] In some embodiments, the system is a Cas-based system that is capable of performing a specialized function or activity. For example, the Cas protein may be fused, operably coupled to, or otherwise associated with one or more functionals domains. In certain example embodiments, the Cas protein may be a catalytically dead Cas protein ("dCas") and/or have nickase activity. A nickase is a Cas protein that cuts only one strand of a double stranded target. In such embodiments, the dCas or nickase provide a sequence specific targeting functionality that delivers the functional domain to or proximate a target sequence. Example functional domains that may be fused to, operably coupled to, or otherwise associated with a Cas protein can be or include, but are not limited to a nuclear localization signal (NLS) domain, a nuclear export signal (NES) domain, a translational activation domain, a transcriptional activation domain (e.g. VP64, p65, MyoD1, HSF1, RTA, and SET7/9), a translation initiation domain, a transcriptional repression domain (e.g., a KRAB domain, NuE domain, NcoR domain, and a SID domain such as a SID4X domain), a nuclease domain (e.g., FokI), a histone modification domain (e.g., a histone acetyltransferase), a light inducible/controllable domain, a chemically inducible/controllable domain, a transposase domain, a homologous recombination machinery domain, a recombinase domain, an integrase domain, and combinations thereof. Methods for generating catalytically dead Cas9 or a nickase Cas9 (WO 2014/204725, Ran et al. Cell. 2013 Sept 12; 154(6):1380-1389), Cas12 (Liu et al. Nature Communications, 8, 2095 (2017), and Cas13 (International Patent Publication Nos. WO 2019/005884 and WO2019/060746) are known in the art and incorporated herein by reference.

[0199] In some embodiments, the functional domains can have one or more of the following activities: methylase activity, demethylase activity, translation activation activity, translation initiation activity, translation repression activity, transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity, single-strand RNA cleavage activity, double-strand RNA cleavage activity, single-strand DNA cleavage activity, double-strand DNA cleavage activity, molecular switch activity, chemical inducibility, light inducibility, and nucleic acid binding activity. In some embodiments, the one or more functional domains may comprise epitope tags or reporters. Non-limiting examples of epitope tags include histidine (His) tags, V5 tags, FLAG

tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Examples of reporters include, but are not limited to, glutathione-S-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) beta-galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and auto-fluorescent proteins including blue fluorescent protein (BFP).

[0200] The one or more functional domain(s) may be positioned at, near, and/or in proximity to a terminus of the effector protein (e.g., a Cas protein). In embodiments having two or more functional domains, each of the two can be positioned at or near or in proximity to a terminus of the effector protein (e.g., a Cas protein). In some embodiments, such as those where the functional domain is operably coupled to the effector protein, the one or more functional domains can be tethered or linked via a suitable linker (including, but not limited to, GlySer linkers) to the effector protein (e.g., a Cas protein). When there is more than one functional domain, the functional domains can be same or different. In some embodiments, all the functional domains are the same. In some embodiments, all of the functional domains are different from each other. In some embodiments, at least two of the functional domains are different from each other. In some embodiments, at least two of the functional domains are the same as each other.

[0201] Other suitable functional domains can be found, for example, in International Patent Publication No. WO 2019/018423.

Split CRISPR-Cas systems

[0202] In some embodiments, the CRISPR-Cas system is a split CRISPR-Cas system. *See* e.g., Zetche et al., 2015. Nat. Biotechnol. 33(2): 139-142 and International Patent Publication WO 2019/018423 , the compositions and techniques of which can be used in and/or adapted for use with the present invention. Split CRISPR-Cas proteins are set forth herein and in documents incorporated herein by reference in further detail herein. In certain embodiments, each part of a split CRISPR protein are attached to a member of a specific binding pair, and when bound with each other, the members of the specific binding pair maintain the parts of the CRISPR protein in proximity. In certain embodiments, each part of a split CRISPR protein is associated with an inducible binding pair. An inducible binding pair is one which is capable of being switched "on" or "off" by a protein or small molecule that binds to both members of the inducible binding pair. In some embodiments, CRISPR proteins may preferably split

between domains, leaving domains intact. In particular embodiments, said Cas split domains (e.g., RuvC and HNH domains in the case of Cas9) can be simultaneously or sequentially introduced into the cell such that said split Cas domain(s) process the target nucleic acid sequence in the algae cell. The reduced size of the split Cas compared to the wild type Cas allows other methods of delivery of the systems to the cells, such as the use of cell penetrating peptides as described herein.

DNA and RNA Base Editing

[0203] In some embodiments, a polynucleotide of the present invention described elsewhere herein can be modified using a base editing system. In some embodiments, a Cas protein is connected or fused to a nucleotide deaminase. Thus, in some embodiments the Cas-based system can be a base editing system. As used herein, "base editing" refers generally to the process of polynucleotide modification via a CRISPR-Cas-based or Cas-based system that does not include excising nucleotides to make the modification. Base editing can convert base pairs at precise locations without generating excess undesired editing byproducts that can be made using traditional CRISPR-Cas systems.

[0204] In certain example embodiments, the nucleotide deaminase may be a DNA base editor used in combination with a DNA binding Cas protein such as, but not limited to, Class 2 Type II and Type V systems. Two classes of DNA base editors are generally known: cytosine base editors (CBEs) and adenine base editors (ABEs). CBEs convert a C•G base pair into a T•A base pair (Komor et al. 2016. Nature. 533:420-424; Nishida et al. 2016. Science. 353; and Li et al. Nat. Biotech. 36:324-327) and ABEs convert an A•T base pair to a G•C base pair. Collectively, CBEs and ABEs can mediate all four possible transition mutations (C to T, A to G, T to C, and G to A). Rees and Liu. 2018.Nat. Rev. Genet. 19(12): 770-788, particularly at Figures 1b, 2a-2c, 3a-3f, and Table 1. In some embodiments, the base editing system includes a CBE and/or an ABE. In some embodiments, a polynucleotide of the present invention described elsewhere herein can be modified using a base editing system. Rees and Liu. 2018. Nat. Rev. Gent. 19(12):770-788. Base editors also generally do not need a DNA donor template and/or rely on homology-directed repair. Komor et al. 2016. Nature. 533:420-424; Nishida et al. 2016. Science. 353; and Gaudeli et al. 2017. Nature. 551:464-471. Upon binding to a target locus in the DNA, base pairing between the guide RNA of the system and the target DNA strand leads to displacement of a small segment of ssDNA in an "R-loop". Nishimasu et al. Cell. 156:935-949. DNA bases within the ssDNA bubble are modified by the enzyme

component, such as a deaminase. In some systems, the catalytically disabled Cas protein can be a variant or modified Cas can have nickase functionality and can generate a nick in the non-edited DNA strand to induce cells to repair the non-edited strand using the edited strand as a template. Komor et al. 2016. Nature. 533:420-424; Nishida et al. 2016. Science. 353; and Gaudeli et al. 2017. Nature. 551:464-471.

[0205] Other Example Type V base editing systems are described in International Patent Publication Nos. WO 2018/213708, WO 2018/213726, and International Patent Applications No. PCT/US2018/067207, PCT/US2018/067225, and PCT/US2018/067307, each of which is incorporated herein by reference.

[0206] In certain example embodiments, the base editing system may be an RNA base editing system. As with DNA base editors, a nucleotide deaminase capable of converting nucleotide bases may be fused to a Cas protein. However, in these embodiments, the Cas protein will need to be capable of binding RNA. Example RNA binding Cas proteins include, but are not limited to, RNA-binding Cas9s such as *Francisella novicida* Cas9 ("FnCas9"), and Class 2 Type VI Cas systems. The nucleotide deaminase may be a cytidine deaminase or an adenosine deaminase, or an adenosine deaminase engineered to have cytidine deaminase activity. In certain example embodiments, the RNA base editor may be used to delete or introduce a post-translation modification site in the expressed mRNA. In contrast to DNA base editors, whose edits are permanent in the modified cell, RNA base editors can provide edits where finer, temporal control may be needed, for example in modulating a particular immune response. Example Type VI RNA-base editing systems are described in Cox et al. 2017. Science 358: 1019-1027, International Patent Publication Nos. WO 2019/005884, WO 2019/005886, and WO 2019/071048, and International Patent Application Nos. PCT/US20018/05179 and PCT/US2018/067207, which are incorporated herein by reference. An example FnCas9 system that may be adapted for RNA base editing purposes is described in International Patent Publication No. WO 2016/106236, which is incorporated herein by reference.

[0207] An example method for delivery of base-editing systems, including use of a split-intein approach to divide CBE and ABE into reconstitutable halves, is described in Levy et al. Nature Biomedical Engineering doi.org/10.1038/s41441-019-0505-5 (2019), which is incorporated herein by reference.

Prime Editors

[0208] In some embodiments, a polynucleotide of the present invention described elsewhere herein can be modified using a prime editing system. See e.g. Anzalone et al. 2019. Nature. 576: 149-157. Like base editing systems, prime editing systems can be capable of targeted modification of a polynucleotide without generating double stranded breaks and does not require donor templates. Further prime editing systems can be capable of all 12 possible combination swaps. Prime editing can operate via a "search-and-replace" methodology and can mediate targeted insertions, deletions, all 12 possible base-to-base conversion and combinations thereof. Generally, a prime editing system, as exemplified by PE1, PE2, and PE3 (*Id.*), can include a reverse transcriptase fused or otherwise coupled or associated with an RNA-programmable nickase and a prime-editing extended guide RNA (pegRNA) to facility direct copying of genetic information from the extension on the pegRNA into the target polynucleotide. Embodiments that can be used with the present invention include these and variants thereof. Prime editing can have the advantage of lower off-target activity than traditional CRIPSR-Cas systems along with few byproducts and greater or similar efficiency as compared to traditional CRISPR-Cas systems.

[0209] In some embodiments, the prime editing guide molecule can specify both the target polynucleotide information (e.g., sequence) and contain a new polynucleotide cargo that replaces target polynucleotides. To initiate transfer from the guide molecule to the target polynucleotide, the PE system can nick the target polynucleotide at a target side to expose a 3'hydroxyl group, which can prime reverse transcription of an edit-encoding extension region of the guide molecule (e.g., a prime editing guide molecule or peg guide molecule) directly into the target site in the target polynucleotide. See e.g., Anzalone et al. 2019. Nature. 576: 149-157, particularly at Figures 1b, 1c, related discussion, and Supplementary discussion.

[0210] In some embodiments, a prime editing system can be composed of a Cas polypeptide having nickase activity, a reverse transcriptase, and a guide molecule. The Cas polypeptide can lack nuclease activity. The guide molecule can include a target binding sequence as well as a primer binding sequence and a template containing the edited polynucleotide sequence. The guide molecule, Cas polypeptide, and/or reverse transcriptase can be coupled together or otherwise associate with each other to form an effector complex and edit a target sequence. In some embodiments, the Cas polypeptide is a Class 2, Type V Cas polypeptide. In some embodiments, the Cas polypeptide is a Cas9 polypeptide (e.g. is a Cas9

nickase). In some embodiments, the Cas polypeptide is fused to the reverse transcriptase. In some embodiments, the Cas polypeptide is linked to the reverse transcriptase.

[0211] In some embodiments, the prime editing system can be a PE1 system or variant thereof, a PE2 system or variant thereof, or a PE3 (e.g. PE3, PE3b) system. See e.g., Anzalone et al. 2019. Nature. 576: 149-157, particularly at pgs. 2-3, Figs. 2a, 3a-3f, 4a-4b, Extended data Figs. 3a-3b, 4,

[0212] The peg guide molecule can be about 10 to about 200 or more nucleotides in length, such as 10 to/or 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, or 200 or more nucleotides in length. Optimization of the peg guide molecule can be accomplished as described in Anzalone et al. 2019. Nature. 576: 149-157, particularly at pg. 3, Fig. 2a-2b, and Extended Data Figs. 5a-c.

CRISPR Associated Transposase (CAST) Systems

[0213] In some embodiments, a polynucleotide of the present invention described elsewhere herein can be modified using a CRISPR Associated Transposase ("CAST") system. CAST system can include a Cas protein that is catalytically inactive, or engineered to be catalytically active, and further comprises a transposase (or subunits thereof) that catalyze RNA-guided DNA transposition. Such systems are able to insert DNA sequences at a target site in a DNA molecule without relying on host cell repair machinery. CAST systems can be Class1 or Class 2 CAST systems. An example Class 1 system is described in Klompe *et al.* Nature, doi:10.1038/s41586-019-1323, which is in incorporated herein by reference. An example Class 2 system is described in Strecker et al. Science. 10/1126/science. aax9181 (2019), and PCT/US2019/066835 which are incorporated herein by reference.

IscBs

[0214] In some embodiments, the nucleic acid-guided nucleases herein may be IscB proteins. An IscB protein may comprise an X domain and a Y domain as described herein. In some examples, the IscB proteins may form a complex with one or more guide molecules. In some cases, the IscB proteins may form a complex with one or more hRNA molecules which serve as a scaffold molecule and comprise guide sequences. In some examples, the IscB proteins are CRISPR-associated proteins, e.g., the loci of the nucleases are associated with an CRISPR array. In some examples, the IscB proteins are not CRISPR-associated.

[0215] In some examples, the IscB protein may be homolog or ortholog of IscB proteins described in Kapitonov VV et al., ISC, a Novel Group of Bacterial and Archaeal DNA Transposons That Encode Cas9 Homologs, J Bacteriol. 2015 Dec 28;198(5):797-807. doi: 10.1128/JB.00783-15, which is incorporated by reference herein in its entirety.

[0216] In some embodiments, the IscBs may comprise one or more domains, e.g., one or more of a X domain (e.g., at N-terminus), a RuvC domain, a Bridge Helix domain, and a Y domain (e.g., at C-terminus). In some examples, the nucleic-acid guided nuclease comprises an N-terminal X domain, a RuvC domain (e.g., including a RuvC-I, RuvC-II, and RuvC-III subdomains), a Bridge Helix domain, and a C-terminal Y domain. In some examples, the nucleic-acid guided nuclease comprises In some examples, the nucleic-acid guided nuclease comprises an N-terminal X domain, a RuvC domain (e.g., including a RuvC-I, RuvC-II, and RuvC-III subdomains), a Bridge Helix domain, an HNH domain, and a C-terminal Y domain.

[0217] In some embodiments, the nucleic acid-guided nucleases may have a small size. For example, the nucleic acid-guided nucleases may be no more than 50, no more than 100, no more than 150, no more than 200, no more than 250, no more than 300, no more than 350, no more than 400, no more than 450, no more than 500, no more than 550, no more than 600, no more than 650, no more than 700, no more than 750, no more than 800, no more than 850, no more than 900, no more than 950, or no more than 1000 amino acids in length.

[0218] In some examples, the IscB protein shares at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with a IscB protein selected from **Table 2**.

Table 2.

No.	Proteins	Sequences
1	IscB(-HNH) EFH81386	MSTDATLIRTTPSHAEADATDTLVATPLMPPRRVISPWPGPGEGQSLMRIPVVDIRG MALMPCTPAKARHLLKSGNARPKRNKLGLFYVQLSYEQEPDNQSLVAGVDPGSKF EGLSVVGTKDTVLNLMVEAPDHVKGAVQTRRTMRRARRQRKWRRPKRFHNRLNR

		MQRIPPSTRSRWEAKARIVAHLRTILPFTDVVVEDVQAVTRKGKGGTWNGSFSPVQ VGKEHLYRLLRAMGLTLHLREGWQTKELREQHGLKKTKSKSKQSFESHAVDSWV LAASISGAEHPTCTRLWYMVPAILHRRQLHRLQASKGGVRKPYGGTRSLGVKRGTL VEHKKYGRCTVGGVDRKRNTISLHEYRTNTRLTQAAKVETCRVLTWLSWRSWLLR GKRTSSKGKGSHSS (SEQ ID NO: 21)
2	IscB(+HNH) TAE54104.1	MQPAKQQNWVFQINGDKQPLDMINPGRCRELQNRGKLASFRRFPYVVIQQQTIENP QTKEYILKIDPGSQWTGFAIQCGNDILFRAELNHRGEAIKFDLVKRAWFRRGRRSRN LRYRKKRLNRAKPEGWLAPSIRHRVLTVETWIKRFMRYCPIAWIEIEQVRFDTQKLA NPEIDGVEYQQGELQGYEVREYLLQKWGRKCAYCGTENVPLEVEHIQSKSKGGSSR IGNLTLACHVCNVKKGNLDVRDFL AKSPDILNQV LENSTKPLKD AAAVNSTRYA IVKMAKSICENVKCSSGARTKMNRVRQGLEKTHSLDAACVGESGASIRVLTDRPLLI TCKGHGSRQSIRVNASGFPAVKNAKTVFTHIAAGDVVRFTIGKDRKKAQAGTYTAR VKTPTPKGFEVLIDGAR ISLSTMSNVVFVHRSDGYGY EL (SEQ ID NO: 22)
3	IscB(+HNH) WP_038093640.1	MAVFVIDKHKRPLMPCSEKRARLLLERGRAVVHRQVPFVIRLKDRTVQHSAVQPLR VALDPGSRATGMALVREKNTVDTGTGEVYRERIALNLFELVHRGHRIREQLDQRRN FRRRRRGANLRYRAPRFDNRRRPPGWLAPSLQHRVDTTMAWVRRLCRWAPASAIG IETVRFDTQRLQNPEISGVEYQQGALAGCEVREYLLEKWGRKCAYCGAENVPLEIE HIVPKSRGGSDRVSNLALACRACNQAKGNRDVRAFLADQPERLARILAQAKAPLK DAAAVNATRWALYRALVDTGLPVEAGTGGRTKWNRTRLGLPKTHALDALCVGQ VDQVRHWRVPVLGIRCAGRGSYRRTRLTRHGFPRGYLTRNKSAFGFQTGDLIRAVV TKGKKAGTYLGRIAIRASGSFNIQTPMGVVQGIHHRFCTLLQRADGYGYFVQP KPTEAALSSP RLKAGVSSAG N (SEQ ID NO: 23)
4	IscB(+HNH) WP_052490348.1	MTTNVVFVIDTNQKPLQPCSAAVARKLLLRGKAAMFRRYPAVIILKKEVDSVGKPK IELRIDPGSKYTGFALVDSKDNADFIIWGTELEHRGAAICKELTKRSAIRRSRRNRKT RYRKKRFERRKPEGWLAPSLQHRVDTTLTWVKRICKFVPIMSISVEQVKFDLQKLE NSDIQGIEYQQGTLAGYTLREALLEHWGRKCAYCDVENVFLEIEHIYPKSKGGSDKF SNLTLACHKCNINKGNKSIDEFLLSDHKRLEQIKLHQKKTLKDAAAVNATRKKLVT TLQEKTFLNVLVSDGASTKMTRLSSSLAKRHWIDAGCVNTTLIVILKTLQPLQVKCN GHGNKQFVTMDAYGFPRKSYEPKKVRKDWKAGDIIRVTKKDGTMLMGRVKKAA KKLVYIPFGGKEASFSS ENAKAIHRSDGYRYSFAAID SELLQKMAT (SEQ ID NO: 24)
5	IscB(+HNH) WP_015325818.1	MPNKYAFVLDSKGKLLDPTKSKKAWYLIRKGKASLVEEYPLIIKLKREVPKDQVNS DKLILGIDDGTKKVGFALVQKCQTKNKVLFKAVMEQRQDVSKKMEERRGYRRYR RSHKRYRPARFDNRSSSKRKGRIPPSILQKKQAILRVVNKLKKYIRIDKIVLEDVSIDI RKLTEGRELYNWEYQESNRLDENLRKATLYRDDCTCQLCGTTETMLHAHHIMPRR DGGADSIYNLITLCKACHKDKVDNNEYQYKDQFLAIIDSKELSDLKSASHVMQGKT WLRDKLSKIAQLEITSGGNTANKRIDYEIEKSHSNDAICTTGLLPVDNIDDIKEYYIKP LRKKSKAKIKELKCFRQRDLVKYTKRNGETYTGYITSLRIKNNKYNSKVCNFSTLK GKIFRGYGFRNLTLLNRPKGLMIV (SEQ ID NO: 25)
6	sp|G3ECR1|CAS9_STRTR	MLFNKCIIISINLDFSNKEKCMTKPYSIGLDIGTNSVGWAVITDNYKVPSKKMKVLG NTSKKYIKKNLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRILYLQEIFSTEMATL DDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKVYHDEFPTIYHLRKYLADSTKKA DLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSK QLEEIVKDKISKLEKKDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASL HFSKESYDEDLETLLGYIGDDYSDVFLKAKKLYDAILLSGFLTVTDNETEAPLSSAM IKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLK NLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPFLA KNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFIN RMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIV RLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLLNIINDKEFLD DSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWGKLSAKLI NGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEV VKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSNSQQR LKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDI DRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDFPSLEVVKKRKTFWYQLL KSKLISQRKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKD ENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVIASALLKKY PKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVN EETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSK PKPNSNENLVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQGIS ILDRINYRKDKLNFLLEKGYKDIELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIH KGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEELFYYILEFNENYVGA KKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVK IPRYRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG (SEQ ID NO: 26)

7	sp|J7RUA5|CAS9_STAAU	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGAR RLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAA LLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKD GEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGP GEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVI TRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPE FTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEE IEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQK EIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMIN EMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLL NNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYE TFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGL MNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIA NADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHI KDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDK LKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTK YSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDN GVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKI NGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKK YSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 27)
8	Streptococcus_pyogenes_SF370	KYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDK KHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQS FIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGE QKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQL KRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFK EDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPEN IVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRG KSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFI KRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 28)

Table 2 (continued)

No.	Proteins	Domains and amino acid positions
1	IscB(-HNH) EFH81386	X domain: 51-97 RuvC-I: 104-118 Bridge Helix: 140-160 RuvC-II: 169-212 RuvC-III: 226-278
2	IscB(+HNH) TAE54104.1	X domain: 11-56 RuvC-I: 63-77 Bridge Helix: 100-121 RuvC-II: 129-172 HNH: 211-243 RuvC-III: 279-321
3	IscB(+HNH) WP_038093640.1	X domain: 4-50

		RuvC-I: 57-71 Bridge Helix: 108-129 RuvC-II: 138-181 HNH: 220-252 RuvC-III: 288-330
4	IscB(+HNH) WP_052490348.1	X domain: 7-52 RuvC-I: 59-73 Bridge Helix: 100-121 RuvC-II: 129-172 HNH: 211-243 RuvC-III: 279-322
5	IscB(+HNH) WP_015325818.1	X domain: 7-52 RuvC-I: 61-75 Bridge Helix: 101-121 RuvC-II: 132-175 HNH: 215-247 RuvC-III: 284-327
6	sp|G3ECR1|CAS9_STRTR	RuvC-I: 28-42 Bridge Helix: 85-108 Rec: 118-736 RuvC-II: 750-799 HNH: 864-896 RuvC-III: 957-1019 PAM Interaction (PI): 1119-1409
7	sp|J7RUA5|CAS9_STAAU	RuvC-I: 7-21 Bridge Helix: 49-72 Rec: 80-433 RuvC-II: 445-493 HNH: 553-585 RuvC-III: 654-709 PAM Interaction (PI): 789-1053
8	Streptococcus_pyogenes_SF370	RuvC-I: 4-18 Bridge Helix: 61-84 Rec: 94-718 RuvC-II: 725-774 HNH: 833-865 RuvC-III: 926-988 PAM Interaction (PI): 1099-1365

X domains

[0219] In some embodiments, the IscB proteins comprise an X domain, e.g., at its N-terminal.

[0220] In certain embodiments, the X domain include the X domains in **Table 2**. Examples of the X domains also include any polypeptides a structural similarity and/or sequence similarity to a X domain described in the art. In some examples, the X domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with X domains in **Table 2**.

[0221] In some examples, the X domain may be no more than 10, no more than 20, no more than 30, no more than 40, no more than 50, no more than 60, no more than 70, no more than 80, no more than 90, or no more than 100 amino acids in length. For example, the X

domain may be no more than 50 amino acids in length, such as comprising 2 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

Y domain

[0222] In some embodiments, the IscB proteins comprise a Y domain, e.g., at its C-terminal.

[0223] In certain embodiments, the X domain include Y domains in **Table 2**. Examples of the Y domain also include any polypeptides a structural similarity and/or sequence similarity to a Y domain described in the art. In some examples, the Y domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with Y domains in **Table 2**.

RuvC domain

[0224] In some embodiments, the IscB proteins comprises at least one nuclease domain. In certain embodiments, the IscB proteins comprise at least two nuclease domains. In certain embodiments, the one or more nuclease domains are only active upon presence of a cofactor. In certain embodiments, the cofactor is Magnesium (Mg). In embodiments where more than one nuclease domain is present and the substrate is a double-strand polynucleotide, the nuclease domains each cleave a different strand of the double-strand polynucleotide. In certain embodiments, the nuclease domain is a RuvC domain.

[0225] The IscB proteins may comprise a RuvC domain. The RuvC domain may comprise multiple subdomains, e.g., RuvC-I, RuvC-II and RuvC-III. The subdomains may be separated by interval sequences on the amino acid sequence of the protein.

[0226] In certain embodiments, examples of the RuvC domain include those in **Table 2**. Examples of the RuvC domain also include any polypeptides a structural similarity and/or sequence similarity to a RuvC domain described in the art. For example, the RuvC domain may share a structural similarity and/or sequence similarity to a RuvC of Cas9. In some examples, the RuvC domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with RuvC domains in **Table 2**.

Bridge helix

[0227] The IscB proteins comprise a bridge helix (BH) domain. The bridge helix domain refers to a helix and arginine rich polypeptide. The bridge helix domain may be located next to anyone of the amino acid domains in the nucleic-acid guided nuclease. In some embodiments, the bridge helix domain is next to a RuvC domain, e.g., next to RuvC-I, RuvC-II, or RuvC-III subdomain. In one example, the bridge helix domain is between a RuvC-1 and RuvC2 subdomains.

[0228] The bridge helix domain may be from 10 to 100, from 20 to 60, from 30 to 50, e.g., 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 or 47, 48, 49, or 50 amino acids in length. Examples of bridge helix includes the polypeptide of amino acids 60-93 of the sequence of S. pyogenes Cas9.

[0229] In certain embodiments, examples of the BH domain include those in **Table 2**. Examples of the BH domain also include any polypeptides a structural similarity and/or sequence similarity to a BH domain described in the art. For example, the BH domain may share a structural similarity and/or sequence similarity to a BH domain of Cas9. In some examples, the BH domain may have an amino acid sequence that share at least 50%, at least 55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with BH domains in **Table 2**.

HNH domain

[0230] The IscB proteins comprise an HNH domain. In certain embodiments, at least one nuclease domain shares a substantial structural similarity or sequence similarity to a HNH domain described in the art.

[0231] In some examples, the nucleic acid-guided nuclease comprises a HNH domain and a RuvC domain. In the cases where the RuvC domain comprises RuvC-I, RuvC-II, and RuvC-III domain, the HNH domain may be located between the Ruv C II and RuvC III subdomains of the RuvC domain.

[0232] In certain embodiments, examples of the HNH domain include those in **Table 2**. Examples of the HNH domain also include any polypeptides a structural similarity and/or sequence similarity to a HNH domain described in the art. For example, the HNH domain may share a structural similarity and/or sequence similarity to a HNH domain of Cas9. In some examples, the HNH domain may have an amino acid sequence that share at least 50%, at least

55%, at least 60%, at least 5%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or 100% sequence identity with HNH domains in **Table 2**.

hRNA

[0233] In some examples, the IscB proteins capable of forming a complex with one or more hRNA molecules. The hRNA complex can comprise a guide sequence and a scaffold that interacts with the IscB polypeptide. An hRNA molecules may form a complex with an IscB polypeptide nuclease or IscB polypeptide and direct the complex to bind with a target sequence. In certain example embodiments, the hRNA molecule is a single molecule comprising a scaffold sequence and a spacer sequence. In certain example embodiments, the spacer is 5' of the scaffold sequence. In certain example embodiments, the hRNA molecule may further comprise a conserved nucleic acid sequence between the scaffold and spacer portions.

[0234] As used herein, a heterologous hRNA molecule is an hRNA molecule that is not derived from the same species as the IscB polypeptide nuclease, or comprises a portion of the molecule, e.g. spacer, that is not derived from the same species as the IscB polypeptide nuclease, e.g. IscB protein. For example, a heterologous hRNA molecule of a IscB polypeptide nuclease derived from species A comprises a polynucleotide derived from a species different from species A, or an artificial polynucleotide.

TALE Nucleases

[0235] In some embodiments, a TALE nuclease or TALE nuclease system can be used to modify a polynucleotide. In some embodiments, the methods provided herein use isolated, non-naturally occurring, recombinant or engineered DNA binding proteins that comprise TALE monomers or TALE monomers or half monomers as a part of their organizational structure that enable the targeting of nucleic acid sequences with improved efficiency and expanded specificity.

[0236] Naturally occurring TALEs or "wild type TALEs" are nucleic acid binding proteins secreted by numerous species of proteobacteria. TALE polypeptides contain a nucleic acid binding domain composed of tandem repeats of highly conserved monomer polypeptides that are predominantly 33, 34 or 35 amino acids in length and that differ from each other mainly in amino acid positions 12 and 13. In advantageous embodiments the nucleic acid is DNA. As used herein, the term "polypeptide monomers", "TALE monomers" or "monomers" will be used to refer to the highly conserved repetitive polypeptide sequences within the TALE nucleic acid binding domain and the term "repeat variable di-residues" or "RVD" will be used to refer

to the highly variable amino acids at positions 12 and 13 of the polypeptide monomers. As provided throughout the disclosure, the amino acid residues of the RVD are depicted using the IUPAC single letter code for amino acids. A general representation of a TALE monomer which is comprised within the DNA binding domain is X_{1-11}-($X_{12}X_{13}$)-X_{14-33} or $_{34}$ or $_{35}$, where the subscript indicates the amino acid position and X represents any amino acid. $X_{12}X_{13}$ indicate the RVDs. In some polypeptide monomers, the variable amino acid at position 13 is missing or absent and in such monomers, the RVD consists of a single amino acid. In such cases the RVD may be alternatively represented as X*, where X represents X_{12} and (*) indicates that X_{13} is absent. The DNA binding domain comprises several repeats of TALE monomers and this may be represented as (X_{1-11}-($X_{12}X_{13}$)-X_{14-33} or $_{34}$ or $_{35}$)$_z$, where in an advantageous embodiment, z is at least 5 to 40. In a further advantageous embodiment, z is at least 10 to 26.

[0237] The TALE monomers can have a nucleotide binding affinity that is determined by the identity of the amino acids in its RVD. For example, polypeptide monomers with an RVD of NI can preferentially bind to adenine (A), monomers with an RVD of NG can preferentially bind to thymine (T), monomers with an RVD of HD can preferentially bind to cytosine (C) and monomers with an RVD of NN can preferentially bind to both adenine (A) and guanine (G). In some embodiments, monomers with an RVD of IG can preferentially bind to T. Thus, the number and order of the polypeptide monomer repeats in the nucleic acid binding domain of a TALE determines its nucleic acid target specificity. In some embodiments, monomers with an RVD of NS can recognize all four base pairs and can bind to A, T, G or C. The structure and function of TALEs is further described in, for example, Moscou et al., Science 326:1501 (2009); Boch et al., Science 326:1509-1512 (2009); and Zhang et al., Nature Biotechnology 29:149-153 (2011).

[0238] The polypeptides used in methods of the invention can be isolated, non-naturally occurring, recombinant or engineered nucleic acid-binding proteins that have nucleic acid or DNA binding regions containing polypeptide monomer repeats that are designed to target specific nucleic acid sequences.

[0239] As described herein, polypeptide monomers having an RVD of HN or NH preferentially bind to guanine and thereby allow the generation of TALE polypeptides with high binding specificity for guanine containing target nucleic acid sequences. In some embodiments, polypeptide monomers having RVDs RN, NN, NK, SN, NH, KN, HN, NQ, HH, RG, KH, RH and SS can preferentially bind to guanine. In some embodiments, polypeptide

monomers having RVDs RN, NK, NQ, HH, KH, RH, SS and SN can preferentially bind to guanine and can thus allow the generation of TALE polypeptides with high binding specificity for guanine containing target nucleic acid sequences. In some embodiments, polypeptide monomers having RVDs HH, KH, NH, NK, NQ, RH, RN and SS can preferentially bind to guanine and thereby allow the generation of TALE polypeptides with high binding specificity for guanine containing target nucleic acid sequences. In some embodiments, the RVDs that have high binding specificity for guanine are RN, NH RH and KH. Furthermore, polypeptide monomers having an RVD of NV can preferentially bind to adenine and guanine. In some embodiments, monomers having RVDs of H*, HA, KA, N*, NA, NC, NS, RA, and S* bind to adenine, guanine, cytosine and thymine with comparable affinity.

[0240] The predetermined N-terminal to C-terminal order of the one or more polypeptide monomers of the nucleic acid or DNA binding domain determines the corresponding predetermined target nucleic acid sequence to which the polypeptides of the invention will bind. As used herein the monomers and at least one or more half monomers are "specifically ordered to target" the genomic locus or gene of interest. In plant genomes, the natural TALE-binding sites always begin with a thymine (T), which may be specified by a cryptic signal within the non-repetitive N-terminus of the TALE polypeptide; in some cases, this region may be referred to as repeat 0. In animal genomes, TALE binding sites do not necessarily have to begin with a thymine (T) and polypeptides of the invention may target DNA sequences that begin with T, A, G or C. The tandem repeat of TALE monomers always ends with a half-length repeat or a stretch of sequence that may share identity with only the first 20 amino acids of a repetitive full-length TALE monomer and this half repeat may be referred to as a half-monomer. Therefore, it follows that the length of the nucleic acid or DNA being targeted is equal to the number of full monomers plus two.

[0241] As described in Zhang et al., Nature Biotechnology 29:149-153 (2011), TALE polypeptide binding efficiency may be increased by including amino acid sequences from the "capping regions" that are directly N-terminal or C-terminal of the DNA binding region of naturally occurring TALEs into the engineered TALEs at positions N-terminal or C-terminal of the engineered TALE DNA binding region. Thus, in certain embodiments, the TALE polypeptides described herein further comprise an N-terminal capping region and/or a C-terminal capping region.

[0242] An exemplary amino acid sequence of a N-terminal capping region is:

M D P I R S R T P S P A R E L L S G P Q P D G V Q P T A D R G V S P P A G G P L D G L P A R R T M S R T R L P S P P A P S P A F S A D S F S D L L R Q F D P S L F N T S L F D S L P P F G A H H T E A A T G E W D E V Q S G L R A A D A P P P T M R V A V T A A R P P R A K P A P R R R A A Q P S D A S P A A Q V D L R T L G Y S Q Q Q Q E K I K P K V R S T V A Q H H E A L V G H G F T H A H I V A L S Q H P A A L G T V A V K Y Q D M I A A L P E A T H E A I V G V G K Q W S G A R A L E A L L T V A G E L R G P P L Q L D T G Q L L K I A K R G G V T A V E A V H A W R N A L T G A P L N (SEQ ID NO:29)

[0243] An exemplary amino acid sequence of a C-terminal capping region is: R P A L E S I V A Q L S R P D P A L A A L T N D H L V A L A C L G G R P A L D A V K K G L P H A P A L I K R T N R R I P E R T S H R V A D H A Q V V R V L G F F Q C H S H P A Q A F D D A M T Q F G M S R H G L L Q L F R R V G V T E L E A R S G T L P P A S Q R W D R I L Q A S G M K R A K P S P T S T Q T P D Q A S L H A F A D S L E R D L D A P S P M H E G D Q T R A S (SEQ ID NO:30)

[0244] As used herein the predetermined "N-terminus" to "C terminus" orientation of the N-terminal capping region, the DNA binding domain comprising the repeat TALE monomers and the C-terminal capping region provide structural basis for the organization of different domains in the d-TALEs or polypeptides of the invention.

[0245] The entire N-terminal and/or C-terminal capping regions are not necessary to enhance the binding activity of the DNA binding region. Therefore, in certain embodiments, fragments of the N-terminal and/or C-terminal capping regions are included in the TALE polypeptides described herein.

[0246] In certain embodiments, the TALE polypeptides described herein contain a N-terminal capping region fragment that included at least 10, 20, 30, 40, 50, 54, 60, 70, 80, 87, 90, 94, 100, 102, 110, 117, 120, 130, 140, 147, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260 or 270 amino acids of an N-terminal capping region. In certain embodiments, the N-terminal capping region fragment amino acids are of the C-terminus (the DNA-binding region proximal end) of an N-terminal capping region. As described in Zhang et al., *Nature Biotechnology* 29:149-153 (2011), N-terminal capping region fragments that include the C-terminal 240 amino acids enhance binding activity equal to the full length capping region, while fragments that include the C-terminal 147 amino acids retain greater than 80% of the

efficacy of the full length capping region, and fragments that include the C-terminal 117 amino acids retain greater than 50% of the activity of the full-length capping region.

[0247] In some embodiments, the TALE polypeptides described herein contain a C-terminal capping region fragment that included at least 6, 10, 20, 30, 37, 40, 50, 60, 68, 70, 80, 90, 100, 110, 120, 127, 130, 140, 150, 155, 160, 170, 180 amino acids of a C-terminal capping region. In certain embodiments, the C-terminal capping region fragment amino acids are of the N-terminus (the DNA-binding region proximal end) of a C-terminal capping region. As described in Zhang et al., Nature Biotechnology 29:149-153 (2011), C-terminal capping region fragments that include the C-terminal 68 amino acids enhance binding activity equal to the full-length capping region, while fragments that include the C-terminal 20 amino acids retain greater than 50% of the efficacy of the full-length capping region.

[0248] In certain embodiments, the capping regions of the TALE polypeptides described herein do not need to have identical sequences to the capping region sequences provided herein. Thus, in some embodiments, the capping region of the TALE polypeptides described herein have sequences that are at least 50%, 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical or share identity to the capping region amino acid sequences provided herein. Sequence identity is related to sequence homology. Homology comparisons may be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs may calculate percent (%) homology between two or more sequences and may also calculate the sequence identity shared by two or more amino acid or nucleic acid sequences. In some preferred embodiments, the capping region of the TALE polypeptides described herein have sequences that are at least 95% identical or share identity to the capping region amino acid sequences provided herein.

[0249] Sequence homologies can be generated by any of a number of computer programs known in the art, which include but are not limited to BLAST or FASTA. Suitable computer programs for carrying out alignments like the GCG Wisconsin Bestfit package may also be used. Once the software has produced an optimal alignment, it is possible to calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

[0250] In some embodiments described herein, the TALE polypeptides of the invention include a nucleic acid binding domain linked to the one or more effector domains. The terms

"effector domain" or "regulatory and functional domain" refer to a polypeptide sequence that has an activity other than binding to the nucleic acid sequence recognized by the nucleic acid binding domain. By combining a nucleic acid binding domain with one or more effector domains, the polypeptides of the invention may be used to target the one or more functions or activities mediated by the effector domain to a particular target DNA sequence to which the nucleic acid binding domain specifically binds.

[0251] In some embodiments of the TALE polypeptides described herein, the activity mediated by the effector domain is a biological activity. For example, in some embodiments the effector domain is a transcriptional inhibitor (i.e., a repressor domain), such as an mSin interaction domain (SID). SID4X domain or a Krüppel-associated box (KRAB) or fragments of the KRAB domain. In some embodiments, the effector domain is an enhancer of transcription (i.e., an activation domain), such as the VP16, VP64 or p65 activation domain. In some embodiments, the nucleic acid binding is linked, for example, with an effector domain that includes but is not limited to a transposase, integrase, recombinase, resolvase, invertase, protease, DNA methyltransferase, DNA demethylase, histone acetylase, histone deacetylase, nuclease, transcriptional repressor, transcriptional activator, transcription factor recruiting, protein nuclear-localization signal or cellular uptake signal.

[0252] In some embodiments, the effector domain is a protein domain which exhibits activities which include but are not limited to transposase activity, integrase activity, recombinase activity, resolvase activity, invertase activity, protease activity, DNA methyltransferase activity, DNA demethylase activity, histone acetylase activity, histone deacetylase activity, nuclease activity, nuclear-localization signaling activity, transcriptional repressor activity, transcriptional activator activity, transcription factor recruiting activity, or cellular uptake signaling activity. Other preferred embodiments of the invention may include any combination of the activities described herein.

[0253] Other preferred tools for genome editing for use in the context of this invention include zinc finger systems and TALE systems. One type of programmable DNA-binding domain is provided by artificial zinc-finger (ZF) technology, which involves arrays of ZF modules to target new DNA-binding sites in the genome. Each finger module in a ZF array targets three DNA bases. A customized array of individual zinc finger domains is assembled into a ZF protein (ZFP).

Zinc Finger Nucleases

[0254] Zinc Finger protens can comprise a functional domain. The first synthetic zinc finger nucleases (ZFNs) were developed by fusing a ZF protein to the catalytic domain of the Type IIS restriction enzyme FokI. (Kim, Y. G. et al., 1994, Chimeric restriction endonuclease, Proc. Natl. Acad. Sci. U.S.A. 91, 883−887; Kim, Y. G. et al., 1996, Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc. Natl. Acad. Sci. U.S.A. 93, 1156−1160). Increased cleavage specificity can be attained with decreased off target activity by use of paired ZFN heterodimers, each targeting different nucleotide sequences separated by a short spacer. (Doyon, Y. et al., 2011, Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat. Methods 8, 74−79). ZFPs can also be designed as transcription activators and repressors and have been used to target many genes in a wide variety of organisms. Exemplary methods of genome editing using ZFNs can be found for example in U.S. Patent Nos. 6,534,261, 6,607,882, 6,746,838, 6,794,136, 6,824,978, 6,866,997, 6,933,113, 6,979,539, 7,013,219, 7,030,215, 7,220,719, 7,241,573, 7,241,574, 7,585,849, 7,595,376, 6,903,185, and 6,479,626, all of which are specifically incorporated by reference.

Meganucleases

[0255] In some embodiments, a meganuclease or system thereof can be used to modify a polynucleotide. Meganucleases, which are endodeoxyribonucleases characterized by a large recognition site (double-stranded DNA sequences of 12 to 40 base pairs). Exemplary methods for using meganucleases can be found in US Patent Nos. 8,163,514, 8,133,697, 8,021,867, 8,119,361, 8,119,381, 8,124,369, and 8,129,134, which are specifically incorporated herein by reference.

RNAi

[0256] In certain embodiments, the genetic modifying agent is RNAi (e.g., shRNA). As used herein, “gene silencing” or “gene silenced” in reference to an activity of an RNAi molecule, for example a siRNA or miRNA refers to a decrease in the mRNA level in a cell for a target gene by at least about 5%, about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, about 99%, about 100% of the mRNA level found in the cell without the presence of the miRNA or RNA interference molecule. In one preferred embodiment, the mRNA levels are decreased by at least about 70%, about 80%, about 90%, about 95%, about 99%, about 100%.

[0257] As used herein, the term "RNAi" refers to any type of interfering RNA, including but not limited to, siRNAi, shRNAi, endogenous microRNA and artificial microRNA. For instance, it includes sequences previously identified as siRNA, regardless of the mechanism of down-stream processing of the RNA (i.e. although siRNAs are believed to have a specific method of *in vivo* processing resulting in the cleavage of mRNA, such sequences can be incorporated into the vectors in the context of the flanking sequences described herein). The term "RNAi" can include both gene silencing RNAi molecules, and also RNAi effector molecules which activate the expression of a gene.

[0258] As used herein, a "siRNA" refers to a nucleic acid that forms a double stranded RNA, which double stranded RNA has the ability to reduce or inhibit expression of a gene or target gene when the siRNA is present or expressed in the same cell as the target gene. The double stranded RNA siRNA can be formed by the complementary strands. In one embodiment, a siRNA refers to a nucleic acid that can form a double stranded siRNA. The sequence of the siRNA can correspond to the full-length target gene, or a subsequence thereof. Typically, the siRNA is at least about 15-50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is about 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, preferably about 19-30 base nucleotides, preferably about 20-25 nucleotides in length, e.g., 20, 21 , 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length).

[0259] As used herein "shRNA" or "small hairpin RNA" (also called stem loop) is a type of siRNA. In one embodiment, these shRNAs are composed of a short, e.g. about 19 to about 25 nucleotide, antisense strand, followed by a nucleotide loop of about 5 to about 9 nucleotides, and the analogous sense strand. Alternatively, the sense strand can precede the nucleotide loop structure and the antisense strand can follow.

[0260] The terms "microRNA" or "miRNA" are used interchangeably herein are endogenous RNAs, some of which are known to regulate the expression of protein-coding genes at the posttranscriptional level. Endogenous microRNAs are small RNAs naturally present in the genome that are capable of modulating the productive utilization of mRNA. The term artificial microRNA includes any type of RNA sequence, other than endogenous microRNA, which is capable of modulating the productive utilization of mRNA. MicroRNA sequences have been described in publications such as Lim, et al., Genes & Development, 17, p. 991 - 1008 (2003), Lim et al Science 299, 1540 (2003), Lee and Ambros Science, 294, 862

(2001), Lau et al., Science 294, 858-861 (2001), Lagos-Quintana et al, Current Biology, 12, 735-739 (2002), Lagos Quintana et al, Science 294, 853- 857 (2001), and Lagos-Quintana et al, RNA, 9, 175- 179 (2003), which are incorporated herein by reference. Multiple microRNAs can also be incorporated into a precursor molecule. Furthermore, miRNA-like stem-loops can be expressed in cells as a vehicle to deliver artificial miRNAs and short interfering RNAs (siRNAs) for the purpose of modulating the expression of endogenous genes through the miRNA and or RNAi pathways.

[0261] As used herein, "double stranded RNA" or "dsRNA" refers to RNA molecules that are comprised of two strands. Double-stranded molecules include those comprised of a single RNA molecule that doubles back on itself to form a two-stranded structure. For example, the stem loop structure of the progenitor molecules from which the single-stranded miRNA is derived, called the pre-miRNA (Bartel et al. 2004. Cell 1 16:281 -297), comprises a dsRNA molecule.

Polypeptides

[0262] In certain example embodiments, the cargo molecule may one or more polypeptides. The polypeptide may be a full-length protein or a functional fragment or functional domain thereof, that is a fragment or domain that maintains the desired functionality of the full-length protein. As used within this section "protein" is meant to refer to full-length proteins and functional fragments and domains thereof. A wide array of polypeptides may be delivered using the engineered delivery vesicles described herein, including but not limited to, secretory proteins, immunomodulatory proteins, anti-fibrotic proteins, proteins that promote tissue regeneration and/or transplant survival functions, hormones, anti-microbial proteins, anti-fibrillating polypeptides, and antibodies. The one or more polypeptides may also comprise combinations of the aforementioned example classes of polypeptides. It will be appreciated that any of the polypeptides described herein can also be delivered via the engineered delivery vesicles and systems described herein via delivery of the corresponding encoding polynucleotide.

Secretory Proteins

[0263] In certain example embodiments, the one or more polypeptides may comprise one or more secretory proteins. A secretory is a protein that is actively transported out of the cell, for example, the protein, whether it be endocrine or exocrine, is secreted by a cell. Secretory pathways have been shown conserved from yeast to mammals, and both conventional and

unconventional protein secretion pathways have been demonstrated in plants. Chung et al., "An Overview of Protein Secretion in Plant Cells," MIMB, 1662:19-32, September 1, 2017. Accordingly, identification of secretory proteins in which one or more polynucleotides may be inserted can be identified for particular cells and applications. In embodiments, one of skill in the art can identify secretory proteins based on the presence of a signal peptide, which consists of a short hydrophobic N-terminal sequence.

[0264] In embodiments, the protein is secreted by the secretory pathway. In embodiments, the proteins are exocrine secretion proteins or peptides, comprising enzymes in the digestive tract. In embodiments the protein is endocrine secretion protein or peptide, for example, insulin and other hormones released into the blood stream. In other embodiments, the protein is involved in signaling between or within cells via secreted signaling molecules, for example, paracrine, autocrine, endocrine or neuroendocrine. In embodiments, the secretory protein is selected from the group of cytokines, kinases, hormones and growth factors that bind to receptors on the surface of target cells.

[0265] As described, secretory proteins include hormones, enzymes, toxins, and antimicrobial peptides. Examples of secretory proteins include serine proteases (e.g., pepsins, trypsin, chymotrypsin, elastase and plasminogen activators), amylases, lipases, nucleases (e.g. deoxyribonucleases and ribonucleases), peptidases enzyme inhibitors such as serpins (e.g., α1-antitrypsin and plasminogen activator inhibitors), cell attachment proteins such as collagen, fibronectin and laminin, hormones and growth factors such as insulin, growth hormone, prolactin platelet-derived growth factor, epidermal growth factor, fibroblast growth factors, interleukins, interferons, apolipoproteins, and carrier proteins such as transferrin and albumins. In some examples, the secretory protein is insulin or a fragment thereof. In one example, the secretory protein is a precursor of insulin or a fragment thereof. In certain examples, the secretory protein is c-peptide. In a preferred embodiment, the one or more polynucleotides is inserted in the middle of the c-peptide. In some aspects, the secretory protein is GLP-1, glucagon, betatrophin, pancreatic amylase, pancreatic lipase, carboxypeptidase, secretin, CCK, a PPAR (e.g. PPAR-alpha, PPAR-gamma, PPAR-delta or a precursor thereof (e.g. preprotein or preproprotein). In aspects, the secretory protein is fibronectin, a clotting factor protein (e.g. Factor VII, VIII, IX, etc.), α2-macroglobulin, α1-antitrypsin, antithrombin III, protein S, protein C, plasminogen, α2-antiplasmin, complement components (e.g. complement component C1-9), albumin, ceruloplasmin, transcortin, haptoglobin, hemopexin, IGF binding

protein, retinol binding protein, transferrin, vitamin-D binding protein, transthyretin, IGF-1, thrombopoietin, hepcidin, angiotensinogen, or a precursor protein thereof. In aspects, the secretory protein is pepsinogen, gastric lipase, sucrase, gastrin, lactase, maltase, peptidase, or a precursor thereof. In aspects, the secretory protein is renin, erythropoietin, angiotensin, adrenocorticotropic hormone (ACTH), amylin, atrial natriuretic peptide (ANP), calcitonin, ghrelin, growth hormone (GH), leptin, melanocyte-stimulating hormone (MSH), oxytocin, prolactin, follicle-stimulating hormone (FSH), thyroid stimulating hormone (TSH), thyrotropin-releasing hormone (TRH), vasopressin, vasoactive intestinal peptide, or a precursor thereof.

Immunomodulatory Polypeptides

[0266] In certain example embodiments, the one or more polypeptides may comprise one or more immunomodulatory protein. In certain embodiments, the present invention provides for modulating immune states. The immune state can be modulated by modulating T cell function or dysfunction. In particular embodiments, the immune state is modulated by expression and secretion of IL-10 and/or other cytokines as described elsewhere herein. In certain embodiments, T cells can affect the overall immune state, such as other immune cells in proximity.

[0267] The polynucleotides may encode one or more immunomodulatory proteins, including immunosuppressive proteins. The term "immunosuppressive" means that immune response in an organism is reduced or depressed. An immunosuppressive protein may suppress, reduce, or mask the immune system or degree of response of the subject being treated. For example, an immunosuppressive protein may suppress cytokine production, downregulate or suppress self-antigen expression, or mask the MHC antigens. As used herein, the term “immune response” refers to a response by a cell of the immune system, such as a B cell, T cell (CD4+ or CD8+), regulatory T cell, antigen-presenting cell, dendritic cell, monocyte, macrophage, NKT cell, NK cell, basophil, eosinophil, or neutrophil, to a stimulus. In some embodiments, the response is specific for a particular antigen (an “antigen-specific response”), and refers to a response by a CD4 T cell, CD8 T cell, or B cell via their antigen-specific receptor. In some embodiments, an immune response is a T cell response, such as a CD4+ response or a CD8+ response. Such responses by these cells can include, for example, cytotoxicity, proliferation, cytokine or chemokine production, trafficking, or phagocytosis, and can be dependent on the nature of the immune cell undergoing the response. In some cases,

the immunosuppressive proteins may exert pleiotropic functions. In some cases, the immunomodulatory proteins may maintain proper regulatory T cells versus effector T cells (Treg/Teff) balance. For examples, the immunomodulatory proteins may expand and/or activate the Tregs and blocks the actions of Teffs, thus providing immunoregulation without global immunosuppression. Target genes associated with immune suppression include, for example, checkpoint inhibitors such PD1, Tim3, Lag3, TIGIT, CTLA-4, and combinations thereof.

[0268] The term "immune cell" as used throughout this specification generally encompasses any cell derived from a hematopoietic stem cell that plays a role in the immune response. The term is intended to encompass immune cells both of the innate or adaptive immune system. The immune cell as referred to herein may be a leukocyte, at any stage of differentiation (e.g., a stem cell, a progenitor cell, a mature cell) or any activation stage. Immune cells include lymphocytes (such as natural killer cells, T-cells (including, e.g., thymocytes, Th or Tc; Th1, Th2, Th17, Thαβ, CD4+, CD8+, effector Th, memory Th, regulatory Th, CD4+/CD8+ thymocytes, CD4–/CD8– thymocytes, γδ T cells, etc.) or B-cells (including, e.g., pro-B cells, early pro-B cells, late pro-B cells, pre-B cells, large pre-B cells, small pre-B cells, immature or mature B-cells, producing antibodies of any isotype, T1 B-cells, T2, B-cells, naïve B-cells, GC B-cells, plasmablasts, memory B-cells, plasma cells, follicular B-cells, marginal zone B-cells, B-1 cells, B-2 cells, regulatory B cells, etc.), such as for instance, monocytes (including, e.g., classical, non-classical, or intermediate monocytes), (segmented or banded) neutrophils, eosinophils, basophils, mast cells, histiocytes, microglia, including various subtypes, maturation, differentiation, or activation stages, such as for instance hematopoietic stem cells, myeloid progenitors, lymphoid progenitors, myeloblasts, promyelocytes, myelocytes, metamyelocytes, monoblasts, promonocytes, lymphoblasts, prolymphocytes, small lymphocytes, macrophages (including, e.g., Kupffer cells, stellate macrophages, M1 or M2 macrophages), (myeloid or lymphoid) dendritic cells (including, e.g., Langerhans cells, conventional or myeloid dendritic cells, plasmacytoid dendritic cells, mDC-1, mDC-2, Mo-DC, HP-DC, veiled cells), granulocytes, polymorphonuclear cells, antigen-presenting cells (APC), etc.

[0269] T cell response refers more specifically to an immune response in which T cells directly or indirectly mediate or otherwise contribute to an immune response in a subject. T cell-mediated response may be associated with cell mediated effects, cytokine mediated

effects, and even effects associated with B cells if the B cells are stimulated, for example, by cytokines secreted by T cells. By means of an example but without limitation, effector functions of MHC class I restricted Cytotoxic T lymphocytes (CTLs), may include cytokine and/or cytolytic capabilities, such as lysis of target cells presenting an antigen peptide recognized by the T cell receptor (naturally-occurring TCR or genetically engineered TCR, e.g., chimeric antigen receptor, CAR), secretion of cytokines, preferably IFN gamma, TNF alpha and/or or more immunostimulatory cytokines, such as IL-2, and/or antigen peptide-induced secretion of cytotoxic effector molecules, such as granzymes, perforins or granulysin. By means of example but without limitation, for MHC class II restricted T helper (Th) cells, effector functions may be antigen peptide-induced secretion of cytokines, preferably, IFN gamma, TNF alpha, IL-4, IL5, IL-10, and/or IL-2. By means of example but without limitation, for T regulatory (Treg) cells, effector functions may be antigen peptide-induced secretion of cytokines, preferably, IL-10, IL-35, and/or TGF-beta. B cell response refers more specifically to an immune response in which B cells directly or indirectly mediate or otherwise contribute to an immune response in a subject. Effector functions of B cells may include in particular production and secretion of antigen-specific antibodies by B cells (e.g., polyclonal B cell response to a plurality of the epitopes of an antigen (antigen-specific antibody response)), antigen presentation, and/or cytokine secretion.

[0270] During persistent immune activation, such as during uncontrolled tumor growth or chronic infections, subpopulations of immune cells, particularly of CD8+ or CD4+ T cells, become compromised to different extents with respect to their cytokine and/or cytolytic capabilities. Such immune cells, particularly CD8+ or CD4+ T cells, are commonly referred to as “dysfunctional” or as “functionally exhausted” or “exhausted”. As used herein, the term “dysfunctional” or “functional exhaustion” refer to a state of a cell where the cell does not perform its usual function or activity in response to normal input signals, and includes refractivity of immune cells to stimulation, such as stimulation via an activating receptor or a cytokine. Such a function or activity includes, but is not limited to, proliferation (e.g., in response to a cytokine, such as IFN-gamma) or cell division, entrance into the cell cycle, cytokine production, cytotoxicity, migration and trafficking, phagocytotic activity, or any combination thereof. Normal input signals can include, but are not limited to, stimulation via a receptor (e.g., T cell receptor, B cell receptor, co-stimulatory receptor). Unresponsive immune cells can have a reduction of at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%,

90%, 95%, or even 100% in cytotoxic activity, cytokine production, proliferation, trafficking, phagocytotic activity, or any combination thereof, relative to a corresponding control immune cell of the same type. In some particular embodiments of the aspects described herein, a cell that is dysfunctional is a CD8+ T cell that expresses the CD8+ cell surface marker. Such CD8+ cells normally proliferate and produce cell killing enzymes, e.g., they can release the cytotoxins perforin, granzymes, and granulysin. However, exhausted/dysfunctional T cells do not respond adequately to TCR stimulation, and display poor effector function, sustained expression of inhibitory receptors and a transcriptional state distinct from that of functional effector or memory T cells. Dysfunction/exhaustion of T cells thus prevents optimal control of infection and tumors. Exhausted/dysfunctional immune cells, such as T cells, such as CD8+ T cells, may produce reduced amounts of IFN-gamma, TNF-alpha and/or one or more immunostimulatory cytokines, such as IL-2, compared to functional immune cells. Exhausted/dysfunctional immune cells, such as T cells, such as CD8+ T cells, may further produce (increased amounts of) one or more immunosuppressive transcription factors or cytokines, such as IL-10 and/or Foxp3, compared to functional immune cells, thereby contributing to local immunosuppression. Dysfunctional CD8+ T cells can be both protective and detrimental against disease control. As used herein, a "dysfunctional immune state" refers to an overall suppressive immune state in a subject or microenvironment of the subject (e.g., tumor microenvironment). For example, increased IL-10 production leads to suppression of other immune cells in a population of immune cells.

[0271] CD8+ T cell function is associated with their cytokine profiles. It has been reported that effector CD8+ T cells with the ability to simultaneously produce multiple cytokines (polyfunctional CD8+ T cells) are associated with protective immunity in patients with controlled chronic viral infections as well as cancer patients responsive to immune therapy (Spranger et al., 2014, J. Immunother. Cancer, vol. 2, 3). In the presence of persistent antigen CD8+ T cells were found to have lost cytolytic activity completely over time (Moskophidis et al., 1993, Nature, vol. 362, 758–761). It was subsequently found that dysfunctional T cells can differentially produce IL-2, TNFa and IFNg in a hierarchical order (Wherry et al., 2003, J. Virol., vol. 77, 4911–4927). Decoupled dysfunctional and activated CD8+ cell states have also been described (see, e.g., Singer, et al. (2016). A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. Cell 166, 1500-1511 e1509; WO/2017/075478; and WO/2018/049025).

[0272] The invention provides compositions and methods for modulating T cell balance. The invention provides T cell modulating agents that modulate T cell balance. For example, in some embodiments, the invention provides T cell modulating agents and methods of using these T cell modulating agents to regulate, influence or otherwise impact the level of and/or balance between T cell types, e.g., between Th17 and other T cell types, for example, Th1-like cells. For example, in some embodiments, the invention provides T cell modulating agents and methods of using these T cell modulating agents to regulate, influence or otherwise impact the level of and/or balance between Th17 activity and inflammatory potential. As used herein, terms such as "Th17 cell" and/or "Th17 phenotype" and all grammatical variations thereof refer to a differentiated T helper cell that expresses one or more cytokines selected from the group the consisting of interleukin 17A (IL-17A), interleukin 17F (IL-17F), and interleukin 17A/F heterodimer (IL17-AF). As used herein, terms such as "Th1 cell" and/or "Th1 phenotype" and all grammatical variations thereof refer to a differentiated T helper cell that expresses interferon gamma (IFNγ). As used herein, terms such as "Th2 cell" and/or "Th2 phenotype" and all grammatical variations thereof refer to a differentiated T helper cell that expresses one or more cytokines selected from the group the consisting of interleukin 4 (IL-4), interleukin 5 (IL-5) and interleukin 13 (IL-13). As used herein, terms such as "Treg cell" and/or "Treg phenotype" and all grammatical variations thereof refer to a differentiated T cell that expresses Foxp3.

[0273] In some examples, immunomodulatory proteins may be immunosuppressive cytokines. In general, cytokines are small proteins and include interleukins, lymphokines and cell signal molecules, such as tumor necrosis factor and the interferons, which regulate inflammation, hematopoiesis, and response to infections. Examples of immunosuppressive cytokines include interleukin 10 (IL-10), TGF-β, IL-Ra, IL-18Ra, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-34, IL-35, IL-36, IL-37, PGE2, SCF, G-CSF, CSF-1R, M-CSF, GM-CSF, IFN-α, IFN-β, IFN-γ, IFN-λ, bFGF, CCL2, CXCL1, CXCL8, CXCL12, CX3CL1, CXCR4, TNF- α and VEGF. Examples of immunosuppressive proteins may further include FOXP3, AHR, TRP53, IKZF3, IRF4, IRF1, and SMAD3. In one example, the immunosuppressive protein is IL-10. In one example, the immunosuppressive protein is IL-6. In one example, the immunosuppressive protein is IL-2.

Anti-fibrotic proteins

[0274] In certain example embodiments, the one or more polypeptides may comprise an anti-fibrotic protein. Examples of anti-fibrotic proteins include any protein that reduces or inhibits the production of extracellular matrix components, fibronectin, proteoglycan, collagen, elastin, TGIFs, and SMAD7. In embodiments, the anti-fibrotic protein is a peroxisome proliferator-activated receptor (PPAR), or may include one or more PPARs. In some embodiments, the protein is PPARα, PPAR γ is a dual PPARα/γ. Derosa et al., "The role of various peroxisome proliferator-activated receptors and their ligands in clinical practice" January 18, 2017 J. Cell. Phys. 223:1 153-161.

Proteins that promote tissue regeneration and/or transplant survival functions

[0275] In certain example embodiments, the one or more polypeptides may comprise an proteins that proteins that promote tissue regeneration and/or transplant survival functions. In some cases, such proteins may induce and/or up-regulate the expression of genes for pancreatic β cell regeneration. In some cases, the proteins that promote transplant survival and functions include the products of genes for pancreatic β cell regeneration. Such genes may include proislet peptides that are proteins or peptides derived from such proteins that stimulate islet cell neogenesis. Examples of genes for pancreatic β cell regeneration include Reg1, Reg2, Reg3, Reg4, human proislet peptide, parathyroid hormone-related peptide (1-36), glucagon-like peptide-1 (GLP-1), extendin-4, prolactin, Hgf, Igf-1, Gip-1, adipsin, resistin, leptin, IL-6, IL-10, Pdx1, Ptfa1, Mafa, Pax6, Pax4, Nkx6.1, Nkx2.2, PDGF, vglycin, placental lactogens (somatomammotropins, e.g. CSH1, CHS2), isoforms thereof, homologs thereof, and orthologs thereof. In certain embodiments, the protein promoting pancreatic B cell regeneration is a cytokine, myokine, and/or adipokine.

Hormones

[0276] In certain embodiments, the one or mor polynucleotides may comprise one or more hormones. The term "hormone" refers to polypeptide hormones, which are generally secreted by glandular organs with ducts. Hormones include proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence hormone, including synthetically produced small-molecule entities and pharmaceutically acceptable derivatives and salts thereof. Included among the hormones are, for example, growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin;

glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); prolactin, placental lactogen, mouse gonadotropin-associated peptide, inhibin; activin; mullerian-inhibiting substance; and thrombopoietin, growth hormone (GH), adrenocorticotropic hormone (ACTH), dehydroepiandrosterone (DHEA), cortisol, epinephrine, thyroid hormone, estrogen, progesterone, placental lactogens (somatomammotropins, e.g. CSH1, CHS2), testosterone. and neuroendocrine hormones. In certain examples, the hormone is secreted from pancreas, e.g., insulin, glucagon, somatostatin, pancreatic polypeptide and ghrelin. In some examples, the hormone is insulin.

[0277] Hormones herein may also include growth factors, e.g., fibroblast growth factor (FGF) family, bone morphogenic protein (BMP) family, platelet derived growth factor (PDGF) family, transforming growth factor beta (TGFbeta) family, nerve growth factor (NGF) family, epidermal growth factor (EGF) family, insulin related growth factor (IGF) family, hepatocyte growth factor (HGF) family, hematopoietic growth factors (HeGFs), platelet-derived endothelial cell growth factor (PD-ECGF), angiopoietin, vascular endothelial growth factor (VEGF) family, and glucocorticoids. In a particular embodiment, the hormone is insulin or incretins such as exenatide, GLP-1.

Neurohormones

[0278] In embodiments, the secreted peptide is a neurohormone, a hormone produced and released by neuroendocrine cells. Example neurohormones include Thyrotropin-releasing hormone, Corticotropin-releasing hormone, Histamine, Growth hormone-releasing hormone, Somatostatin, Gonadotropin-releasing hormone, Serotonin, Dopamine, Neurotensin, Oxytocin, Vasopressin, Epinephrine, and Norepinephrine.

Anti-microbial Proteins

[0279] In some embodiments, the one or more polypeptides may comprise one or more anti-microbial proteins. In embodiments where the cell is mammalian cell, human host defense antimicrobial peptides and proteins (AMPs) play a critical role in warding off invading microbial pathogens. In certain embodiments, the anti-microbial is α-defensin HD-6, HNP-1 and β-defensin hBD-3, lysozyme, cathelcidin LL-37, C-type lectin RegIIIalpha, for example. See, e.g. Wang, "Human Antimicrobial Peptide and Proteins" *Pharma*, May 2014, 7(5): 545-594, incorporated herein by reference.

Anti-fibrillating Proteins

[0280] In certain example embodiments, the one or more polypeptides may comprise one or more anti-fibrillating polypeptides. The anti-fibrillating polypeptide can be the secreted polypeptide. In some embodiments, the anti-fibrillating polypeptide is co-expressed with one or more other polynucleotides and/or polypeptides described elsewhere herein. The anti-fibrillating agent can be secreted and act to inhibit the fibrillation and/or aggregation of endogenous proteins and/or exogenous proteins that it may be co-expressed therewith. In some embodiments, the anti-fibrillating agent is P4 (VITYF (SEQ ID NO:31)), P5 (VVVVV (SEQ ID NO:32)), KR7 (KPWWPRR (SEQ ID NO:33)), NK9 (NIVNVSLVK (SEQ ID NO:34)), iAb5p (Leu-Pro-Phe-Phe-Asp (SEQ ID NO:35)), KLVF (SEQ ID NO:36) and derivatives thereof, indolicidin, carnosine, a hexapeptide as set forth in Wang et al. 2014. ACS Chem Neurosci. 5:972-981, alpha sheet peptides having alternating D-amino acids and L-amino acids as set forth in Hopping et al. 2014. Elife 3:e01681, D-(PGKLVYA (SEQ ID NO:37)), RI-OR2-TAT, cyclo(17, 21)-(Lys17, Asp21)A_(1-28), SEN304, SEN1576, D3, R8-Aβ(25-35), human yD-crystallin (HGD), poly-lysine, heparin, poly-Asp, polyGl, poly-L-lysine, poly-L-glutamic acid, LVEALYL (SEQ ID NO:38), RGFFYT (SEQ ID NO:39), a peptide set forth or as designed/generated by the method set forth in US Pat. No. 8,754,034, and combinations thereof. In aspects, the anti-fibrillating agent is a D-peptide. In aspects, the anti-fibrillating agent is an L-peptide. In aspects, the anti-fibrillating agent is a retro-inverso modified peptide. Retro-inverso modified peptides are derived from peptides by substituting the L-amino acids for their D-counterparts and reversing the sequence to mimic the original peptide since they retain the same spatial positioning of the side chains and 3D structure. In aspects, the retro-inverso modified peptide is derived from a natural or synthetic Aβ peptide. In some embodiments, the polynucleotide encodes a fibrillation resistant protein. In some embodiments, the fibrillation resistant protein is a modified insulin, see e.g. U.S. Pat. No.: 8,343,914.

Antibodies

[0281] In certain embodiments, the one or more polypeptides may comprise one or more antibodies. The term "antibody" is used interchangeably with the term "immunoglobulin" herein, and includes intact antibodies, fragments of antibodies, e.g., Fab, F(ab')2 fragments, and intact antibodies and fragments that have been mutated either in their constant and/or variable region (e.g., mutations to produce chimeric, partially humanized, or fully humanized

antibodies, as well as to produce antibodies with a desired trait, e.g., enhanced binding and/or reduced FcR binding). The term "fragment" refers to a part or portion of an antibody or antibody chain comprising fewer amino acid residues than an intact or complete antibody or antibody chain. Fragments can be obtained via chemical or enzymatic treatment of an intact or complete antibody or antibody chain. Fragments can also be obtained by recombinant means. Exemplary fragments include Fab, Fab', F(ab')2, Fabc, Fd, dAb, V_{HH} and scFv and/or Fv fragments.

[0282] As used herein, a preparation of antibody protein having less than about 50% of non-antibody protein (also referred to herein as a "contaminating protein"), or of chemical precursors, is considered to be "substantially free." 40%, 30%, 20%, 10% and more preferably 5% (by dry weight), of non-antibody protein, or of chemical precursors is considered to be substantially free. When the antibody protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 30%, preferably less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume or mass of the protein preparation.

[0283] The term "antigen-binding fragment" refers to a polypeptide fragment of an immunoglobulin or antibody that binds antigen or competes with intact antibody (i.e., with the intact antibody from which they were derived) for antigen binding (i.e., specific binding). As such these antibodies or fragments thereof are included in the scope of the invention, provided that the antibody or fragment binds specifically to a target molecule.

[0284] It is intended that the term "antibody" encompass any Ig class or any Ig subclass (e.g. the IgG1, IgG2, IgG3, and IgG4 subclassess of IgG) obtained from any source (e.g., humans and non-human primates, and in rodents, lagomorphs, caprines, bovines, equines, ovines, etc.).

[0285] The term "Ig class" or "immunoglobulin class", as used herein, refers to the five classes of immunoglobulin that have been identified in humans and higher mammals, IgG, IgM, IgA, IgD, and IgE. The term "Ig subclass" refers to the two subclasses of IgM (H and L), three subclasses of IgA (IgA1, IgA2, and secretory IgA), and four subclasses of IgG (IgG1, IgG2, IgG3, and IgG4) that have been identified in humans and higher mammals. The antibodies can exist in monomeric or polymeric form; for example, IgM antibodies exist in pentameric form, and IgA antibodies exist in monomeric, dimeric or multimeric form.

[0286] The term "IgG subclass" refers to the four subclasses of immunoglobulin class IgG - IgG1, IgG2, IgG3, and IgG4 that have been identified in humans and higher mammals by the heavy chains of the immunoglobulins, V1 - γ4, respectively. The term "single-chain immunoglobulin" or "single-chain antibody" (used interchangeably herein) refers to a protein having a two-polypeptide chain structure consisting of a heavy and a light chain, said chains being stabilized, for example, by interchain peptide linkers, which has the ability to specifically bind antigen. The term "domain" refers to a globular region of a heavy or light chain polypeptide comprising peptide loops (e.g., comprising 3 to 4 peptide loops) stabilized, for example, by β pleated sheet and/or intrachain disulfide bond. Domains are further referred to herein as "constant" or "variable", based on the relative lack of sequence variation within the domains of various class members in the case of a "constant" domain, or the significant variation within the domains of various class members in the case of a "variable" domain. Antibody or polypeptide "domains" are often referred to interchangeably in the art as antibody or polypeptide "regions". The "constant" domains of an antibody light chain are referred to interchangeably as "light chain constant regions", "light chain constant domains", "CL" regions or "CL" domains. The "constant" domains of an antibody heavy chain are referred to interchangeably as "heavy chain constant regions", "heavy chain constant domains", "CH" regions or "CH" domains). The "variable" domains of an antibody light chain are referred to interchangeably as "light chain variable regions", "light chain variable domains", "VL" regions or "VL" domains). The "variable" domains of an antibody heavy chain are referred to interchangeably as "heavy chain constant regions", "heavy chain constant domains", "VH" regions or "VH" domains).

[0287] The term "region" can also refer to a part or portion of an antibody chain or antibody chain domain (e.g., a part or portion of a heavy or light chain or a part or portion of a constant or variable domain, as defined herein), as well as more discrete parts or portions of said chains or domains. For example, light and heavy chains or light and heavy chain variable domains include "complementarity determining regions" or "CDRs" interspersed among "framework regions" or "FRs", as defined herein.

[0288] The term "conformation" refers to the tertiary structure of a protein or polypeptide (e.g., an antibody, antibody chain, domain or region thereof). For example, the phrase "light (or heavy) chain conformation" refers to the tertiary structure of a light (or heavy) chain

variable region, and the phrase "antibody conformation" or "antibody fragment conformation" refers to the tertiary structure of an antibody or fragment thereof.

[0289] The term “antibody-like protein scaffolds” or “engineered protein scaffolds” broadly encompasses proteinaceous non-immunoglobulin specific-binding agents, typically obtained by combinatorial engineering (such as site-directed random mutagenesis in combination with phage display or other molecular selection techniques). Usually, such scaffolds are derived from robust and small soluble monomeric proteins (such as Kunitz inhibitors or lipocalins) or from a stably folded extra-membrane domain of a cell surface receptor (such as protein A, fibronectin or the ankyrin repeat).

[0290] Such scaffolds have been extensively reviewed in Binz et al. (Engineering novel binding proteins from nonimmunoglobulin domains. Nat Biotechnol 2005, 23:1257-1268), Gebauer and Skerra (Engineered protein scaffolds as next-generation antibody therapeutics. Curr Opin Chem Biol. 2009, 13:245-55), Gill and Damle (Biopharmaceutical drug discovery using novel protein scaffolds. Curr Opin Biotechnol 2006, 17:653-658), Skerra (Engineered protein scaffolds for molecular recognition. J Mol Recognit 2000, 13:167-187), and Skerra (Alternative non-antibody scaffolds for molecular recognition. Curr Opin Biotechnol 2007, 18:295-304), and include without limitation affibodies, based on the Z-domain of staphylococcal protein A, a three-helix bundle of 58 residues providing an interface on two of its alpha-helices (Nygren, Alternative binding proteins: Affibody binding proteins developed from a small three-helix bundle scaffold. FEBS J 2008, 275:2668-2676); engineered Kunitz domains based on a small (ca. 58 residues) and robust, disulphide-crosslinked serine protease inhibitor, typically of human origin (e.g. LACI-D1), which can be engineered for different protease specificities (Nixon and Wood, Engineered protein inhibitors of proteases. Curr Opin Drug Discov Dev 2006, 9:261-268); monobodies or adnectins based on the 10th extracellular domain of human fibronectin III (10Fn3), which adopts an Ig-like beta-sandwich fold (94 residues) with 2–3 exposed loops, but lacks the central disulphide bridge (Koide and Koide, Monobodies: antibody mimics based on the scaffold of the fibronectin type III domain. Methods Mol Biol 2007, 352:95-109); anticalins derived from the lipocalins, a diverse family of eight-stranded beta-barrel proteins (ca. 180 residues) that naturally form binding sites for small ligands by means of four structurally variable loops at the open end, which are abundant in humans, insects, and many other organisms (Skerra, Alternative binding proteins: Anticalins—harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel

binding activities. FEBS J 2008, 275:2677-2683); DARPins, designed ankyrin repeat domains (166 residues), which provide a rigid interface arising from typically three repeated beta-turns (Stumpp et al., DARPins: a new generation of protein therapeutics. Drug Discov Today 2008, 13:695-701); avimers (multimerized LDLR-A module) (Silverman et al., Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains. Nat Biotechnol 2005, 23:1556-1561); and cysteine-rich knottin peptides (Kolmar, Alternative binding proteins: biological activity and therapeutic potential of cystine-knot miniproteins. FEBS J 2008, 275:2684-2690).

[0291] "Specific binding" of an antibody means that the antibody exhibits appreciable affinity for a particular antigen or epitope and, generally, does not exhibit significant cross reactivity. "Appreciable" binding includes binding with an affinity of at least 25 μM. Antibodies with affinities greater than 1×10^7 M^{-1} (or a dissociation coefficient of 1μM or less or a dissociation coefficient of 1nm or less) typically bind with correspondingly greater specificity. Values intermediate of those set forth herein are also intended to be within the scope of the present invention and antibodies of the invention bind with a range of affinities, for example, 100nM or less, 75nM or less, 50nM or less, 25nM or less, for example 10nM or less, 5nM or less, 1nM or less, or in embodiments 500pM or less, 100pM or less, 50pM or less or 25pM or less. An antibody that "does not exhibit significant crossreactivity" is one that will not appreciably bind to an entity other than its target (e.g., a different epitope or a different molecule). For example, an antibody that specifically binds to a target molecule will appreciably bind the target molecule but will not significantly react with non-target molecules or peptides. An antibody specific for a particular epitope will, for example, not significantly crossreact with remote epitopes on the same protein or peptide. Specific binding can be determined according to any art-recognized means for determining such binding. Preferably, specific binding is determined according to Scatchard analysis and/or competitive binding assays.

[0292] As used herein, the term "affinity" refers to the strength of the binding of a single antigen-combining site with an antigenic determinant. Affinity depends on the closeness of stereochemical fit between antibody combining sites and antigen determinants, on the size of the area of contact between them, on the distribution of charged and hydrophobic groups, etc. Antibody affinity can be measured by equilibrium dialysis or by the kinetic BIACORE™

method. The dissociation constant, Kd, and the association constant, Ka, are quantitative measures of affinity.

[0293] As used herein, the term "monoclonal antibody" refers to an antibody derived from a clonal population of antibody-producing cells (e.g., B lymphocytes or B cells) which is homogeneous in structure and antigen specificity. The term "polyclonal antibody" refers to a plurality of antibodies originating from different clonal populations of antibody-producing cells which are heterogeneous in their structure and epitope specificity but which recognize a common antigen. Monoclonal and polyclonal antibodies may exist within bodily fluids, as crude preparations, or may be purified, as described herein.

[0294] The term "binding portion" of an antibody (or "antibody portion") includes one or more complete domains, e.g., a pair of complete domains, as well as fragments of an antibody that retain the ability to specifically bind to a target molecule. It has been shown that the binding function of an antibody can be performed by fragments of a full-length antibody. Binding fragments are produced by recombinant DNA techniques, or by enzymatic or chemical cleavage of intact immunoglobulins. Binding fragments include Fab, Fab', F(ab')2, Fabc, Fd, dAb, Fv, single chains, single-chain antibodies, e.g., scFv, and single domain antibodies.

[0295] "Humanized" forms of non-human (e.g., murine) antibodies are chimeric antibodies that contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hypervariable region of the recipient are replaced by residues from a hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit or nonhuman primate having the desired specificity, affinity, and capacity. In some instances, FR residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin.

[0296] Examples of portions of antibodies or epitope-binding proteins encompassed by the present definition include: (i) the Fab fragment, having V_L, C_L, V_H and C_H1 domains; (ii) the

Fab' fragment, which is a Fab fragment having one or more cysteine residues at the C-terminus of the C_H1 domain; (iii) the Fd fragment having V_H and C_H1 domains; (iv) the Fd' fragment having V_H and C_H1 domains and one or more cysteine residues at the C-terminus of the CHI domain; (v) the Fv fragment having the V_L and V_H domains of a single arm of an antibody; (vi) the dAb fragment (Ward et al., 341 Nature 544 (1989)) which consists of a V_H domain or a V_L domain that binds antigen; (vii) isolated CDR regions or isolated CDR regions presented in a functional framework; (viii) $F(ab')_2$ fragments which are bivalent fragments including two Fab' fragments linked by a disulphide bridge at the hinge region; (ix) single chain antibody molecules (e.g., single chain Fv; scFv) (Bird et al., 242 Science 423 (1988); and Huston et al., 85 PNAS 5879 (1988)); (x) "diabodies" with two antigen binding sites, comprising a heavy chain variable domain (V_H) connected to a light chain variable domain (V_L) in the same polypeptide chain (see, e.g., EP 404,097; WO 93/11161; Hollinger et al., 90 PNAS 6444 (1993)); (xi) "linear antibodies" comprising a pair of tandem Fd segments (V_H-C_h1-V_H-C_h1) which, together with complementary light chain polypeptides, form a pair of antigen binding regions (Zapata et al., Protein Eng. 8(10):1057-62 (1995); and U.S. Patent No. 5,641,870).

[0297] As used herein, a "blocking" antibody or an antibody "antagonist" is one which inhibits or reduces biological activity of the antigen(s) it binds. In certain embodiments, the blocking antibodies or antagonist antibodies or portions thereof described herein completely inhibit the biological activity of the antigen(s).

[0298] Antibodies may act as agonists or antagonists of the recognized polypeptides. For example, the present invention includes antibodies which disrupt receptor/ligand interactions either partially or fully. The invention features both receptor-specific antibodies and ligand-specific antibodies. The invention also features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or of one of its down-stream substrates by immunoprecipitation followed by western blot analysis. In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

[0299] The invention also features receptor-specific antibodies which both prevent ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex. Likewise, encompassed by the invention are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included in the invention are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the receptor. The antibodies may be specified as agonists, antagonists or inverse agonists for biological activities comprising the specific biological activities of the peptides disclosed herein. The antibody agonists and antagonists can be made using methods known in the art. See, e.g., PCT publication WO 96/40281; U.S. Pat. No. 5,811,097; Deng et al., Blood 92(6):1981-1988 (1998); Chen et al., Cancer Res. 58(16):3668-3678 (1998); Harrop et al., J. Immunol. 161(4):1786-1794 (1998); Zhu et al., Cancer Res. 58(15):3209-3214 (1998); Yoon et al., J. Immunol. 160(7):3170-3179 (1998); Prat et al., J. Cell. Sci. III (Pt2):237-247 (1998); Pitard et al., J. Immunol. Methods 205(2):177-190 (1997); Liautard et al., Cytokine 9(4):233-241 (1997); Carlson et al., J. Biol. Chem. 272(17):11295-11301 (1997); Taryman et al., Neuron 14(4):755-762 (1995); Muller et al., Structure 6(9):1153-1167 (1998); Bartunek et al., Cytokine 8(1):14-20 (1996).

[0300] The antibodies as defined for the present invention include derivatives that are modified, i.e., by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified, e.g., by glycosylation, acetylation, pegylation, phosphylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

Protease Cleavage Sites

[0301] The one or more cargo polypeptides, as exemplified above, may comprise one or more protease cleavage sites, i.e., amino acid sequences that can be recognized and cleaved by

a protease. The protease cleavage sites may be used for generating desired gene products (e.g., intact gene products without any tags or portion of other proteins). The protease cleavage site may be one end or both ends of the protein. Examples of protease cleavage sites that can be used herein include an enterokinase cleavage site, a thrombin cleavage site, a Factor Xa cleavage site, a human rhinovirus 3C protease cleavage site, a tobacco etch virus (TEV) protease cleavage site, a dipeptidyl aminopeptidase cleavage site and a small ubiquitin-like modifier (SUMO)/ubiquitin-like protein-1(ULP-1) protease cleavage site. In certain examples, the protease cleavage site comprises Lys-Arg.

Small Molecules

[0302] In some embodiments, the engineered delivery vesicle can deliver one or more small molecule compounds. Thus, in some embodiments, the cargo molecule is a small molecule. In some embodiments, the small molecule compound(s) can be linked or directly attached to a polynucleotide that can bind a polynucleotide binding protein that can be included in the engineered delivery system polynucleotide. In some embodiments, the engineered delivery system polynucleotide can include a small molecule binding protein (e.g. a receptor for the small molecule) that, like the polynucleotide binding protein discussed elsewhere herein, can be incorporated in to the engineered delivery vesicle.

[0303] In some embodiments, the small molecule compound(s) can be linked or directly attached to a polynucleotide that can bind a polynucleotide binding protein that can be included in the engineered delivery system polynucleotide or delivery vesicle. In some embodiments, the engineered delivery system polynucleotide or delivery vesicle can include a small molecule binding protein (e.g. a receptor for the small molecule) that, like the polynucleotide binding protein discussed elsewhere herein, can be incorporated in to the engineered delivery system polynucleotide or delivery vesicle.

[0304] Suitable hormones include, but are not limited to, amino-acid derived hormones (e.g. melatonin and thyroxine), small peptide hormones and protein hormones (e.g. thyrotropin-releasing hormone, vasopressin, insulin, growth hormone, luteinizing hormone, follicle-stimulating hormone, and thyroid-stimulating hormone), eicosanoids (e.g. arachidonic acid, lipoxins, and prostaglandins), and steroid hormones (e.g. estradiol, testosterone, tetrahydro testosteron Cortisol). Suitable immunomodulators include, but are not limited to, prednisone, azathioprine, 6-MP, cyclosporine, tacrolimus, methotrexate, interleukins (e.g. IL-2, IL-7, and IL-12), cytokines (e.g. interferons (e.g. IFN-a, IFN-β, IFN-ε, IFN-K, IFN-ω, and IFN-γ),

granulocyte colony-stimulating factor, and imiquimod), chemokines (e.g. CCL3, CCL26 and CXCL7) , cytosine phosphate-guanosine, oligodeoxynucleotides, glucans, antibodies, and aptamers).

[0305] Suitable antipyretics include, but are not limited to, non-steroidal anti-inflammants (e.g. ibuprofen, naproxen, ketoprofen, and nimesulide), aspirin and related salicylates (e.g. choline salicylate, magnesium salicylae, and sodium salicaylate), paracetamol/acetaminophen, metamizole, nabumetone, phenazone, and quinine.

[0306] Suitable anxiolytics include, but are not limited to, benzodiazepines (e.g. alprazolam, bromazepam, chlordiazepoxide, clonazepam, clorazepate, diazepam, flurazepam, lorazepam, oxazepam, temazepam, triazolam, and tofisopam), serotenergic antidepressants (e.g. selective serotonin reuptake inhibitors, tricyclic antidepresents, and monoamine oxidase inhibitors), mebicar, afobazole, selank, bromantane, emoxypine, azapirones, barbiturates, hydroxyzine, pregabalin, validol, and beta blockers.

[0307] Suitable antipsychotics include, but are not limited to, benperidol, bromoperidol, droperidol, haloperidol, moperone, pipaperone, timiperone, fluspirilene, penfluridol, pimozide, acepromazine, chlorpromazine, cyamemazine, dizyrazine, fluphenazine, levomepromazine, mesoridazine, perazine, pericyazine, perphenazine, pipotiazine, prochlorperazine, promazine, promethazine, prothipendyl, thioproperazine, thioridazine, trifluoperazine, triflupromazine, chlorprothixene, clopenthixol, flupentixol, tiotixene, zuclopenthixol, clotiapine, loxapine, prothipendyl, carpipramine, clocapramine, molindone, mosapramine, sulpiride, veralipride, amisulpride, amoxapine, aripiprazole, asenapine, clozapine, blonanserin, iloperidone, lurasidone, melperone, nemonapride, olanzapine, paliperidone, perospirone, quetiapine, remoxipride, risperidone, sertindole, trimipramine, ziprasidone, zotepine, alstonie, befeprunox, bitopertin, brexpiprazole, cannabidiol, cariprazine, pimavanserin, pomaglumetad methionil, vabicaserin, xanomeline, and zicronapine.

[0308] Suitable analgesics include, but are not limited to, paracetamol/acetaminophen, nonsteroidal anti-inflammants (e.g. ibuprofen, naproxen, ketoprofen, and nimesulide), COX-2 inhibitors (e.g. rofecoxib, celecoxib, and etoricoxib), opioids (e.g. morphine, codeine, oxycodone, hydrocodone, dihydromorphine, pethidine, buprenorphine), tramadol, norepinephrine, flupiretine, nefopam, orphenadrine, pregabalin, gabapentin, cyclobenzaprine, scopolamine, methadone, ketobemidone, piritramide, and aspirin and related salicylates (e.g. choline salicylate, magnesium salicylate, and sodium salicylate).

[0309] Suitable antispasmodics include, but are not limited to, mebeverine, papverine, cyclobenzaprine, carisoprodol, orphenadrine, tizanidine, metaxalone, methodcarbamol, chlorzoxazone, baclofen, dantrolene, baclofen, tizanidine, and dantrolene. Suitable anti-inflammatories include, but are not limited to, prednisone, non-steroidal anti-inflammants (e.g. ibuprofen, naproxen, ketoprofen, and nimesulide), COX-2 inhibitors (e.g. rofecoxib, celecoxib, and etoricoxib), and immune selective anti-inflammatory derivatives (e.g. submandibular gland peptide-T and its derivatives).

[0310] Suitable anti-histamines include, but are not limited to, H1 -receptor antagonists (e.g. acrivastine, azelastine, bilastine, brompheniramine, buclizine, bromodiphenhydramine, carbinoxamine, cetirizine, chlorpromazine, cyclizine, chlorpheniramine, clemastine, cyproheptadine, desloratadine, dexbromapheniramine, dexchlorpheniramine, dimenhydrinate, dimetindene, diphenhydramine, doxylamine, ebasine, embramine, fexofenadine, hydroxyzine, levocetirzine, loratadine, meclozine, mirtazapine, olopatadine, orphenadrine, phenindamine, pheniramine, phenyltoloxamine, promethazine, pyrilamine, quetiapine, rupatadine, tripelennamine, and triprolidine), H2-receptor antagonists (e.g. cimetidine, famotidine, lafutidine, nizatidine, rafitidine, and roxatidine), tritoqualine, catechin, cromoglicate, nedocromil, and p2-adrenergic agonists.

[0311] Suitable anti-infectives include, but are not limited to, amebicides (e.g. nitazoxanide, paromomycin, metronidazole, tinidazole, chloroquine, miltefosine, amphotericin b, and iodoquinol), aminoglycosides (e.g. paromomycin, tobramycin, gentamicin, amikacin, kanamycin, and neomycin), anthelmintics (e.g. pyrantel, mebendazole, ivermectin, praziquantel, abendazole, thiabendazole, oxamniquine), antifungals (e.g. azole antifungals (e.g. itraconazole, fluconazole, posaconazole, ketoconazole, clotrimazole, miconazole, and voriconazole), echinocandins (e.g. caspofungin, anidulafungin, and micafungin), griseofulvin, terbinafine, flucytosine, and polyenes (e.g. nystatin, and amphotericin b), antimalarial agents (e.g. pyrimethamine/sulfadoxine, artemether/lumefantrine, atovaquone/proquanil, quinine, hydroxychloroquine, mefloquine, chloroquine, doxycycline, pyrimethamine, and halofantrine), antituberculosis agents (e.g. aminosalicylates (e.g. aminosalicylic acid), isoniazid/rifampin, isoniazid/pyrazinamide/rifampin, bedaquiline, isoniazid, ethambutol, rifampin, rifabutin, rifapentine, capreomycin, and cycloserine), antivirals (e.g. amantadine, rimantadine, abacavir/lamivudine, emtricitabine/tenofovir, cobicistat/elvitegravir/emtricitabine/tenofovir, efavirenz/emtricitabine/tenofovir,

avacavir/lamivudine/zidovudine, lamivudine/zidovudine, emtricitabine/tenofovir, emtricitabine/opinavir/ritonavir/tenofovir, interferon alfa-2v/ribavirin, peginterferon alfa-2b, maraviroc, raltegravir, dolutegravir, enfuvirtide, foscarnet, fomivirsen, oseltamivir, zanamivir, nevirapine, efavirenz, etravirine, rilpivirine, delaviridine, nevirapine, entecavir, lamivudine, adefovir, sofosbuvir, didanosine, tenofovir, avacivr, zidovudine, stavudine, emtricitabine, xalcitabine, telbivudine, simeprevir, boceprevir, telaprevir, lopinavir/ritonavir, fosamprenvir, dranuavir, ritonavir, tipranavir, atazanavir, nelfinavir, amprenavir, indinavir, sawuinavir, ribavirin, valcyclovir, acyclovir, famciclovir, ganciclovir, and valganciclovir), carbapenems (e.g. doripenem, meropenem, ertapenem, and cilastatin/imipenem), cephalosporins (e.g. cefadroxil, cephradine, cefazolin, cephalexin, cefepime, ceflaroline, loracarbef, cefotetan, cefuroxime, cefprozil, loracarbef, cefoxitin, cefaclor, ceftibuten, ceftriaxone, cefotaxime, cefpodoxime, cefdinir, cefixime, cefditoren, cefizoxime, and ceftazidime), glycopeptide antibiotics (e.g. vancomycin, dalbavancin, oritavancin, and telvancin), glycylcyclines (e.g. tigecycline), leprostatics (e.g. clofazimine and thalidomide), lincomycin and derivatives thereof (e.g. clindamycin and lincomycin), macrolides and derivatives thereof (e.g. telithromycin, fidaxomicin, erthromycin, azithromycin, clarithromycin, dirithromycin, and troleandomycin), linezolid, sulfamethoxazole/trimethoprim, rifaximin, chloramphenicol, fosfomycin, metronidazole, aztreonam, bacitracin, penicillins (amoxicillin, ampicillin, bacampicillin, carbenicillin, piperacillin, ticarcillin, amoxicillin/clavulanate, ampicillin/sulbactam, piperacillin/tazobactam, clavulanate/ticarcillin, penicillin, procaine penicillin, oxaxillin, dicloxacillin, and nafcillin), quinolones (e.g. lomefloxacin, norfloxacin, ofloxacin, qatifloxacin, moxifloxacin, ciprofloxacin, levofloxacin, gemifloxacin, moxifloxacin, cinoxacin, nalidixic acid, enoxacin, grepafloxacin, gatifloxacin, trovafloxacin, and sparfloxacin), sulfonamides (e.g. sulfamethoxazole/trimethoprim, sulfasalazine, and sulfasoxazole), tetracyclines (e.g. doxycycline, demeclocycline, minocycline, doxycycline/salicyclic acid, doxycycline/omega-3 polyunsaturated fatty acids, and tetracycline), and urinary anti-infectives (e.g. nitrofurantoin, methenamine, fosfomycin, cinoxacin, nalidixic acid, trimethoprim, and methylene blue).

[0312] Suitable chemotherapeutics include, but are not limited to, paclitaxel, brentuximab vedotin, doxorubicin, 5-FU (fluorouracil), everolimus, pemetrexed, melphalan, pamidronate, anastrozole, exemestane, nelarabine, ofatumumab, bevacizumab, belinostat, tositumomab, carmustine, bleomycin, bosutinib, busulfan, alemtuzumab, irinotecan, vandetanib,

bicalutamide, lomustine, daunorubicin, clofarabine, cabozantinib, dactinomycin, ramucirumab, cytarabine, Cytoxan, cyclophosphamide, decitabine, dexamethasone, docetaxel, hydroxyurea, decarbazine, leuprolide, epirubicin, oxaliplatin, asparaginase, estramustine, cetuximab, vismodegib, asparginase Erwinia chrysanthemi, amifostine, etoposide, flutamide, toremifene, fulvestrant, letrozole, degarelix, pralatrexate, methotrexate, floxuridine, obinutuzumab, gemcitabine, afatinib, imatinib mesylatem, carmustine, eribulin, trastuzumab, altretamine, topotecan, ponatinib, idarubicin, ifosfamide, ibrutinib, axitinib, interferon alfa-2a, gefitinib, romidepsin, ixabepilone, ruxolitinib, cabazitaxel, ado-trastuzumab emtansine, carfilzomib, chlorambucil, sargramostim, cladribine, mitotane, vincristine, procarbazine, megestrol, trametinib, mesna, strontium-89 chloride, mechlorethamine, mitomycin, busulfan, gemtuzumab ozogamicin, vinorelbine, filgrastim, pegfilgrastim, sorafenib, nilutamide, pentostatin, tamoxifen, mitoxantrone, pegaspargase, denileukin diftitox, alitretinoin, carboplatin, pertuzumab, cisplatin, pomalidomide, prednisone, aldesleukin, mercaptopurine, zoledronic acid, lenalidomide, rituximab, octretide, dasatinib, regorafenib, histrelin, sunitinib, siltuximab, omacetaxine, thioguanine (tioguanine), dabrafenib, erlotinib, bexarotene, temozolomide, thiotepa, thalidomide, BCG, temsirolimus, bendamustine hydrochloride, triptorelin, aresnic trioxide, lapatinib, valrubicin, panitumumab, vinblastine, bortezomib, tretinoin, azacitidine, pazopanib, teniposide, leucovorin, crizotinib, capecitabine, enzalutamide, ipilimumab, goserelin, vorinostat, idelalisib, ceritinib, abiraterone, epothilone, tafluposide, azathioprine, doxifluridine, vindesine, and all-trans retinoic acid.

Engineered Delivery Vesicle Generation Vector Systems

[0313] The engineered delivery vesicle generation system can be an engineered vector system. In some embodiments, the engineered delivery vesicle generation system is an engineered viral vector system. In some embodiments, the engineered delivery vesicle generation system is an engineered retroviral vector system. In some embodiments, the engineered delivery vesicle generation system is an engineered lentiviral vector system.

[0314] The term of art “viral vector” and as used herein in this context refers to polynucleotide based vectors that contain one or more elements from or based upon one or more elements of a virus that can be capable of expressing and packaging a polynucleotide, such as a cargo s polynucleotide described elsewhere herein, into a virus particle and producing said virus particle when used alone or with one or more other viral vectors (such as in a viral vector system). Viral vectors and systems thereof can be used for producing viral particles for

delivery of and/or expression of one or more cargos described herein. The viral vector can be part of a viral vector system involving multiple vectors. In some embodiments, systems incorporating multiple viral vectors can increase the safety of these systems. Suitable viral vectors can include retroviral-based vectors, lentiviral-based vectors, adenoviral-based vectors, adeno associated vectors, helper-dependent adenoviral (HdAd) vectors, hybrid adenoviral vectors, herpes simplex virus-based vectors, poxvirus-based vectors, and Epstein-Barr virus-based vectors. Other embodiments of viral vectors and viral particles produce therefrom are described elsewhere herein. In some embodiments, the viral vectors are configured to produce replication incompetent viral particles for improved safety of these systems.

[0315] In certain embodiments, the virus structural component, which can be encoded by one or more polynucleotides in a viral vector or vector system, comprises one or more capsid proteins including an entire capsid. In certain embodiments, such as wherein a viral capsid comprises multiple copies of different proteins, the delivery system can provide one or more of the same protein or a mixture of such proteins. For example, AAV comprises 3 capsid proteins, VP1, VP2, and VP3, thus delivery systems of the invention can comprise one or more of VP1, and/or one or more of VP2, and/or one or more of VP3. Accordingly, the present invention is applicable to a virus within the family Adenoviridae, such as Atadenovirus, e.g., Ovine atadenovirus D, Aviadenovirus, e.g., Fowl aviadenovirus A, Ichtadenovirus, e.g., Sturgeon ichtadenovirus A, Mastadenovirus (which includes adenoviruses such as all human adenoviruses), e.g., Human mastadenovirus C, and Siadenovirus, e.g., Frog siadenovirus A. Thus, a virus of within the family Adenoviridae is contemplated as within the invention with discussion herein as to adenovirus applicable to other family members. Target-specific AAV capsid variants can be used or selected. Non-limiting examples include capsid variants selected to bind to chronic myelogenous leukemia cells, human CD34 PBPC cells, breast cancer cells, cells of lung, heart, dermal fibroblasts, melanoma cells, stem cell, glioblastoma cells, coronary artery endothelial cells and keratinocytes. See, e.g., Buning et al, 2015, Current Opinion in Pharmacology 24, 94-104. From teachings herein and knowledge in the art as to modifications of adenovirus (see, e.g., US Patents 9,410,129, 7,344,872, 7,256,036, 6,911,199, 6,740,525; Matthews, "Capsid-Incorporation of Antigens into Adenovirus Capsid Proteins for a Vaccine Approach," Mol Pharm, 8(1): 3-11 (2011)), as well as regarding modifications of AAV, the skilled person can readily obtain a modified adenovirus that has a large cargo, despite that heretofore it was not expected that such a large protein could be provided on an adenovirus.

And as to the viruses related to adenovirus mentioned herein, as well as to the viruses related to AAV mentioned elsewhere herein, the teachings herein as to modifying adenovirus and AAV, respectively, can be applied to those viruses without undue experimentation from this disclosure and the knowledge in the art.

[0316] In some embodiments, the viral vector is configured such that when the cargo is packaged the cargo(s), is external to the capsid or virus particle. In the sense that it is not inside the capsid (enveloped or encompassed with the capsid), but is externally exposed so that it can contact the target genomic DNA. In some embodiments, the viral vector is configured such that all the cargo(s) are contained within the capsid after packaging.

Split Viral Vector Systems

[0317] When the engineered delivery vesicle generation viral vector or vector system (be it a retroviral (e.g. AAV) or lentiviral vector) is designed so as to position the cargo(s) at the internal surface of the capsid once formed, the cargo(s) will fill most or all of internal volume of the capsid. In other embodiments, the cargo can be modified or divided so as to occupy a less of the capsid internal volume. Accordingly, in certain embodiments, the cargo can be divided in two portions, one portion comprises in one viral particle or capsid and the second portion comprised in a second viral particle or capsid. In certain embodiments, by splitting the cargo in two portions, space is made available to link one or more heterologous domains to one or both cargo portions. Such systems can be referred to as "split vector systems" or in the context of the present disclosure a "split cargo system" a "split protein", and the like. When the concept is applied to a vector system, it thus describes putting pieces of the split proteins on different vectors thus reducing the payload of any one vector. This approach can facilitate delivery of systems where the total system size is close to or exceeds the packaging capacity of the vector. This is independent of any regulation of the cargo system that can be achieved with a split system or split protein design.

Retroviral and Lentiviral Vectors

[0318] Retroviral vectors can be composed of cis-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum cis-acting LTRs are sufficient for replication and packaging of the vectors, which are then used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Suitable retroviral vectors for the engineered delivery vesicle generation systems of the present invention can include those based upon murine leukemia virus (MuLV), gibbon ape leukemia

virus (GaLV), Simian immunodeficiency virus (SIV), human immunodeficiency virus (HIV), and combinations thereof (see, e.g., Buchscher et al., J. Virol. 66:2731-2739 (1992); Johann et al., J. Virol. 66:1635-1640 (1992); Sommnerfelt et al., Virol. 176:58-59 (1990); Wilson et al., J. Virol. 63:2374-2378 (1989); Miller et al., J. Virol. 65:2220-2224 (1991); PCT/US94/05700). Selection of a retroviral gene transfer system may therefore depend on the target tissue.

[0319] The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors that are able to transduce or infect non-dividing cells and are described in greater detail elsewhere herein. A retrovirus can also be engineered to allow for conditional expression of the inserted transgene, such that only certain cell types are infected by the lentivirus.

[0320] Lentiviruses are complex retroviruses that have the ability to infect and express their genes in both mitotic and post-mitotic cells. Advantages of using a lentiviral approach can include the ability to transduce or infect non-dividing cells and their ability to typically produce high viral titers, which can increase efficiency or efficacy of production and delivery. Suitable lentiviral vectors include, but are not limited to, human immunodeficiency virus (HIV)-based lentiviral vectors, feline immunodeficiency virus (FIV)-based lentiviral vectors, simian immunodeficiency virus (SIV)-based lentiviral vectors, Moloney Murine Leukaemia Virus (Mo-MLV), Visna.maedi virus (VMV)-based lentiviral vector, carpine arthritis-encephalitis virus (CAEV)-based lentiviral vector, bovine immune deficiency virus (BIV)-based lentiviral vector, and Equine infectious anemia (EIAV)-based lentiviral vector. In some embodiments, an HIV-based lentiviral vector system can be used. In some embodiments, a FIV-based lentiviral vector system can be used.

[0321] In some embodiments, the lentiviral vector is an EIAV-based lentiviral vector or vector system. EIAV vectors have been used to mediate expression, packaging, and/or delivery in other contexts, such as for ocular gene therapy (see, e.g., Balagaan, J Gene Med 2006; 8: 275 – 285). In another embodiment, RetinoStat®, (see, e.g., Binley et al., HUMAN GENE THERAPY 23:980–991 (September 2012)), which describes RetinoStat®, an equine infectious anemia virus-based lentiviral gene therapy vector that expresses angiostatic proteins endostatin and angiostatin that is delivered via a subretinal injection for the treatment of the wet form of age-related macular degeneration. Any of these vectors described in these publications can be modified for the elements of the engineered delivery vesicle generation systems.

[0322] In some embodiments, the lentiviral vector or vector system thereof can be a first-generation lentiviral vector or vector system thereof. First-generation lentiviral vectors can contain a large portion of the lentivirus genome, including the gag and pol genes, other additional viral proteins (e.g. VSV-G) and other accessory genes (e.g. vif, vprm vpu, nef, and combinations thereof), regulatory genes (e.g. tat and/or rev) as well as the gene of interest between the LTRs. First generation lentiviral vectors can result in the production of virus particles that can be capable of replication in vivo, which may not be appropriate for some instances or applications.

[0323] In some embodiments, the lentiviral vector or vector system thereof can be a second-generation lentiviral vector or vector system thereof. Second-generation lentiviral vectors do not contain one or more accessory virulence factors and do not contain all components necessary for virus particle production on the same lentiviral vector. This can result in the production of a replication-incompetent virus particle and thus increase the safety of these systems over first-generation lentiviral vectors. In some embodiments, the second-generation vector lacks one or more accessory virulence factors (e.g. vif, vprm, vpu, nef, and combinations thereof). Unlike the first-generation lentiviral vectors, no single second generation lentiviral vector includes all features necessary to express and package a polynucleotide into a virus particle. In some embodiments, the envelope and packaging components are split between two different vectors with the gag, pol, rev, and tat genes being contained on one vector and the envelope protein (e.g. VSV-G) are contained on a second vector. The gene of interest, its promoter, and LTRs can be included on a third vector that can be used in conjunction with the other two vectors (packaging and envelope vectors) to generate a replication-incompetent virus particle.

[0324] In some embodiments, the lentiviral vector or vector system thereof can be a third-generation lentiviral vector or vector system thereof. Third-generation lentiviral vectors and vector systems thereof have increased safety over first- and second-generation lentiviral vectors and systems thereof because, for example, the various components of the viral genome are split between two or more different vectors but used together in vitro to make virus particles, they can lack the tat gene (when a constitutively active promoter is included up-stream of the LTRs), and they can include one or more deletions in the 3'LTR to create self-inactivating (SIN) vectors having disrupted promoter/enhancer activity of the LTR. In some embodiments, a third-generation lentiviral vector system can include (i) a vector plasmid that

contains the polynucleotide of interest and upstream promoter that are flanked by the 5' and 3' LTRs, which can optionally include one or more deletions present in one or both of the LTRs to render the vector self-inactivating; (ii) a "packaging vector(s)" that can contain one or more genes involved in packaging a polynucleotide into a virus particle that is produced by the system (e.g. gag, pol, and rev) and upstream regulatory sequences (e.g. promoter(s)) to drive expression of the features present on the packaging vector, and (iii) an "envelope vector" that contains one or more envelope protein genes and upstream promoters. In certain embodiments, the third-generation lentiviral vector system can include at least two packaging vectors, with the gag-pol being present on a different vector than the rev gene.

[0325] In some embodiments, self-inactivating lentiviral vectors with an siRNA targeting a common exon shared by HIV tat/rev, a nucleolar-localizing TAR decoy, and an anti–CCR5-specific hammerhead ribozyme (see, e.g., DiGiusto et al. (2010) Sci Transl Med 2:36ra43) can be used/and or adapted to the engineered delivery vesicle generation systems of the present invention.

[0326] In some embodiments, the pseudotype and infectivity or tropism of a lentivirus particle can be tuned by altering the type of envelope protein(s) included in the lentiviral vector or system thereof. As used herein, an "envelope protein" or "outer protein" means a protein exposed at the surface of a viral particle that is not a capsid protein. For example, envelope or outer proteins typically comprise proteins embedded in the envelope of the virus. In some embodiments, a lentiviral vector or vector system thereof can include a VSV-G envelope protein. VSV-G mediates viral attachment to an LDL receptor (LDLR) or an LDLR family member present on a host cell, which triggers endocytosis of the viral particle by the host cell. Because LDLR is expressed by a wide variety of cells, viral particles expressing the VSV-G envelope protein can infect or transduce a wide variety of cell types. Other suitable envelope proteins can be incorporated based on the host cell that a user desires to be infected by a virus particle produced from a lentiviral vector or system thereof described herein and can include, but are not limited to, feline endogenous virus envelope protein (RD114) (see e.g. Hanawa et al. Molec. Ther. 2002 5(3) 242-251), modified Sindbis virus envelope proteins (see e.g. Morizono et al. 2010. J. Virol. 84(14) 6923-6934; Morizono et al. 2001. J. Virol. 75:8016-8020; Morizono et al. 2009. J. Gene Med. 11:549-558; Morizono et al. 2006 Virology 355:71-81; Morizono et al J. Gene Med. 11:655-663, Morizono et al. 2005 Nat. Med. 11:346-352), baboon retroviral envelope protein (see e.g. Girard-Gagnepain et al. 2014. Blood. 124: 1221-

1231); Tupaia paramyxovirus glycoproteins (see e.g. Enkirch T. et al., 2013. Gene Ther. 20:16-23); measles virus glycoproteins (see e.g. Funke et al. 2008. Molec. Ther. 16(8): 1427-1436), rabies virus envelope proteins, MLV envelope proteins, Ebola envelope proteins, baculovirus envelope proteins, filovirus envelope proteins, hepatitis E1 and E2 envelope proteins, gp41 and gp120 of HIV, hemagglutinin, neuraminidase, M2 proteins of influenza virus, and combinations thereof.

[0327] In some embodiments, the tropism of the resulting lentiviral particle can be tuned by incorporating cell targeting peptides into a lentiviral vector such that the cell targeting peptides are expressed on the surface of the resulting lentiviral particle. In some embodiments, a lentiviral vector can contain an envelope protein that is fused to a cell targeting protein (see e.g. Buchholz et al. 2015. Trends Biotechnol. 33:777-790; Bender et al. 2016. PLoS Pathog. 12(e1005461); and Friedrich et al. 2013. Mol. Ther. 2013. 21: 849-859.

[0328] In some embodiments, a split-intein-mediated approach to target lentiviral particles to a specific cell type can be used (see e.g. Chamoun-Emaneulli et al. 2015. Biotechnol. Bioeng. 112:2611-2617, Ramirez et al. 2013. Protein. Eng. Des. Sel. 26:215-233. In these embodiments, a lentiviral vector can contain one half of a splicing-deficient variant of the naturally split intein from Nostoc punctiforme fused to a cell targeting peptide and the same or different lentiviral vector can contain the other half of the split intein fused to an envelope protein, such as a binding-deficient, fusion-competent virus envelope protein. This can result in production of a virus particle from the lentiviral vector or vector system that includes a split intein that can function as a molecular Velcro linker to link the cell-binding protein to the pseudotyped lentivirus particle. This approach can be advantageous for use where surface-incompatibilities can restrict the use of, e.g., cell targeting peptides.

[0329] In some embodiments, a covalent-bond-forming protein-peptide pair can be incorporated into one or more of the lentiviral vectors described herein to conjugate a cell targeting peptide to the virus particle (see e.g. Kasaraneni et al. 2018. Sci. Reports (8) No. 10990). In some embodiments, a lentiviral vector can include an N-terminal PDZ domain of InaD protein (PDZ1) and its pentapeptide ligand (TEFCA (SEQ ID NO:40)) from NorpA, which can conjugate the cell targeting peptide to the virus particle via a covalent bond (e.g. a disulfide bond). In some embodiments, the PDZ1 protein can be fused to an envelope protein, which can optionally be binding deficient and/or fusion competent virus envelope protein and included in a lentiviral vector. In some embodiments, the TEFCA (SEQ ID NO:40) can be

fused to a cell targeting peptide and the TEFCA-CPT (SEQ ID NO:41) fusion construct can be incorporated into the same or a different lentiviral vector as the PDZ1-envenlope protein construct. During virus production, specific interaction between the PDZ1 and TEFCA (SEQ ID NO:42) facilitates producing virus particles covalently functionalized with the cell targeting peptide and thus capable of targeting a specific cell-type based upon a specific interaction between the cell targeting peptide and cells expressing its binding partner. This approach can be advantageous for use where surface-incompatibilities can restrict the use of, e.g., cell targeting peptides.

[0330] Lentiviral vectors have been disclosed as in the treatment for Parkinson's Disease, see, e.g., US Patent Publication No. 20120295960 and US Patent Nos. 7303910 and 7351585. Lentiviral vectors have also been disclosed for the treatment of ocular diseases, see e.g., US Patent Publication Nos. 20060281180, 20090007284, US20110117189; US20090017543; US20070054961, US20100317109. Lentiviral vectors have also been disclosed for delivery to the brain, see, e.g., US Patent Publication Nos. US20110293571; US20110293571, US20040013648, US20070025970, US20090111106 and US Patent No. US7259015. Any of these systems or a variant thereof can be adapted for use with the engineered delivery vesicle generation systems of the present invention.

[0331] In some embodiments, an engineered lentiviral vector system of the present invention can include one or more transfer plasmids. Transfer plasmids can be generated from various other vector backbones and can include one or more features that can work with other retroviral and/or lentiviral vectors in the system that can, for example, improve safety of the vector and/or vector system, increase virial titers, and/or increase or otherwise enhance expression of the desired insert to be expressed and/or packaged into the viral particle. Suitable features that can be included in a transfer plasmid can include, but are not limited to, 5'LTR, 3'LTR, SIN/LTR, origin of replication (Ori), selectable marker genes (e.g. antibiotic resistance genes), Psi (Ψ), RRE (rev response element), cPPT (central polypurine tract), promoters, WPRE (woodchuck hepatitis post-transcriptional regulatory element), SV40 polyadenylation signal, pUC origin, SV40 origin, F1 origin, and combinations thereof.

[0332] In another embodiment, Cocal vesiculovirus envelope pseudotyped retroviral or lentiviral vector particles are contemplated (see, e.g., US Patent Publication No. 20120164118 assigned to the Fred Hutchinson Cancer Research Center). Cocal virus is in the Vesiculovirus genus, and is a causative agent of vesicular stomatitis in mammals. Cocal virus was originally

isolated from mites in Trinidad (Jonkers et al., Am. J. Vet. Res. 25:236-242 (1964)), and infections have been identified in Trinidad, Brazil, and Argentina from insects, cattle, and horses. Many of the vesiculoviruses that infect mammals have been isolated from naturally infected arthropods, suggesting that they are vector-borne. Antibodies to vesiculoviruses are common among people living in rural areas where the viruses are endemic and laboratory-acquired; infections in humans usually result in influenza-like symptoms. The Cocal virus envelope glycoprotein shares 71.5% identity at the amino acid level with VSV-G Indiana, and phylogenetic comparison of the envelope gene of vesiculoviruses shows that Cocal virus is serologically distinct from, but most closely related to, VSV-G Indiana strains among the vesiculoviruses. Jonkers et al., Am. J. Vet. Res. 25:236-242 (1964) and Travassos da Rosa et al., Am. J. Tropical Med. & Hygiene 33:999-1006 (1984). The Cocal vesiculovirus envelope pseudotyped retroviral vector particles may include for example, lentiviral, alpharetroviral, betaretroviral, gammaretroviral, deltaretroviral, and epsilonretroviral vector particles that may comprise retroviral Gag, Pol, and/or one or more accessory protein(s) and a Cocal vesiculovirus envelope protein. In certain embodiments of these embodiments, the Gag, Pol, and accessory proteins are lentiviral and/or gammaretroviral. In some embodiments, a retroviral vector can contain encoding polypeptides for one or more Cocal vesiculovirus envelope proteins such that the resulting viral or pseudoviral particles are Cocal vesiculovirus envelope pseudotyped.

Vector Features

[0333] The vectors can include additional features that can confer one or more functionalities to the vector, the polynucleotide to be delivered, a virus particle produced there from, or polypeptide expressed thereof. Such features include, but are not limited to, regulatory elements, selectable markers, molecular identifiers (e.g. molecular barcodes), stabilizing elements, and the like. It will be appreciated by those skilled in the art that the design of the expression vector and additional features included can depend on such factors as the choice of the host cell to be transformed, the level of expression desired, etc.

Regulatory Elements

[0334] In certain embodiments, the polynucleotides and/or vectors thereof described herein can include one or more regulatory elements that can be operatively linked to the one or more polynucleotides of the vectors, such as effector polynucleotide(s), packaging polynuclotide(s), vesicle element polynucleotide(s), and/or cargo polynucleotide(s). The term "regulatory element" is intended to include promoters, enhancers, internal ribosomal entry sites (IRES),

other expression control elements (e.g., transcription termination signals, such as polyadenylation signals and poly-U sequences) and cellular localization signals (e.g. nuclear localization signals). Such regulatory elements are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory elements include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). A tissue-specific promoter can direct expression primarily in a desired tissue of interest, such as muscle, neuron, bone, skin, blood, specific organs (e.g., liver, pancreas), or particular cell types (e.g., lymphocytes). Regulatory elements may also direct expression in a temporal-dependent manner, such as in a cell-cycle dependent or developmental stage-dependent manner, which may or may not also be tissue or cell-type specific. In some embodiments, a vector comprises one or more pol III promoter (e.g., 1, 2, 3, 4, 5, or more pol III promoters), one or more pol II promoters (e.g., 1, 2, 3, 4, 5, or more pol II promoters), one or more pol I promoters (e.g., 1, 2, 3, 4, 5, or more pol I promoters), or combinations thereof. Examples of pol III promoters include, but are not limited to, U6 and H1 promoters. Examples of pol II promoters include, but are not limited to, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) (see, e.g., Boshart et al, Cell, 41:521-530 (1985)), the SV40 promoter, the dihydrofolate reductase promoter, the β-actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1α promoter. Also encompassed by the term "regulatory element" are enhancer elements, such as WPRE; CMV enhancers; the R-U5' segment in LTR of HTLV-I (Mol. Cell. Biol., Vol. 8(1), p. 466-472, 1988); SV40 enhancer; and the intron sequence between exons 2 and 3 of rabbit β-globin (Proc. Natl. Acad. Sci. USA., Vol. 78(3), p. 1527-31, 1981).

[0335] In some embodiments, the regulatory sequence can be a regulatory sequence described in U.S. Pat. No. 7,776,321, U.S. Pat. Pub. No. 2011/0027239, and International Patent Publication No. WO 2011/028929, the contents of which are incorporated by reference herein in their entirety. In some embodiments, the vector can contain a minimal promoter. In some embodiments, the minimal promoter is the Mecp2 promoter, tRNA promoter, H1, T7, or U6. In a further embodiment, the minimal promoter is tissue specific. In some embodiments, the length of the vector polynucleotide the minimal promoters and polynucleotide sequences is less than 4.4Kb.

[0336] To express a polynucleotide, the vector can include one or more transcriptional and/or translational initiation regulatory sequences, e.g. promoters, that direct the transcription of the gene and/or translation of the encoded protein in a cell. In some embodiments a constitutive promoter may be employed. Suitable constitutive promoters for mammalian cells are generally known in the art and include, but are not limited to SV40, CAG, CMV, EF-1α, β-actin, retroviral Rous sarcoma virus (RSV) LTR promoter, RSV, dihydrofolate reductase promoter, and phosphoglycerol kinase (PGK). Suitable constitutive promoters for bacterial cells, yeast cells, and fungal cells are generally known in the art, such as a T-7 promoter for bacterial expression and an alcohol dehydrogenase promoter for expression in yeast.

[0337] In some embodiments, the regulatory element can be a regulated promoter. "Regulated promoter" refers to promoters that direct gene expression not constitutively, but in a temporally- and/or spatially-regulated manner, and includes tissue-specific, tissue-preferred and inducible promoters. Regulated promoters include conditional promoters and inducible promoters. In some embodiments, conditional promoters can be employed to direct expression of a polynucleotide in a specific cell type, under certain environmental conditions, and/or during a specific state of development. Suitable tissue specific promoters can include, but are not limited to, liver specific promoters (e.g. APOA2, SERPIN A1 (hAAT), CYP3A4, and MIR122), pancreatic cell promoters (e.g. INS, IRS2, Pdx1, Alx3, Ppy), cardiac specific promoters (e.g. Myh6 (alpha MHC), MYL2 (MLC-2v), TNI3 (cTnl), NPPA (ANF), Slc8a1 (Ncx1)), central nervous system cell promoters (SYN1, GFAP, INA, NES, MOBP, MBP, TH, FOXA2 (HNF3 beta)), skin cell specific promoters (e.g. FLG, K14, TGM3), immune cell specific promoters, (e.g. ITGAM, CD43 promoter, CD14 promoter, CD45 promoter, CD68 promoter), urogenital cell specific promoters (e.g. Pbsn, Upk2, Sbp, Fer1l4), endothelial cell specific promoters (e.g. ENG), pluripotent and embryonic germ layer cell specific promoters (e.g. Oct4, NANOG, Synthetic Oct4, T brachyury, NES, SOX17, FOXA2, MIR122), and muscle cell specific promoter (e.g. Desmin). Other tissue and/or cell specific promoters are generally known in the art and are within the scope of this disclosure.

[0338] Inducible/conditional promoters can be positively inducible/conditional promoters (e.g. a promoter that activates transcription of the polynucleotide upon appropriate interaction with an activated activator, or an inducer (compound, environmental condition, or other stimulus) or a negative/conditional inducible promoter (e.g. a promoter that is repressed (e.g. bound by a repressor) until the repressor condition of the promotor is removed (e.g. inducer

binds a repressor bound to the promoter stimulating release of the promoter by the repressor or removal of a chemical repressor from the promoter environment). The inducer can be a compound, environmental condition, or other stimulus. Thus, inducible/conditional promoters can be responsive to any suitable stimuli such as chemical, biological, or other molecular agents, temperature, light, and/or pH. Suitable inducible/conditional promoters include, but are not limited to, Tet-On, Tet-Off, Lac promoter, pBad, AlcA, LexA, Hsp70 promoter, Hsp90 promoter, pDawn, XVE/OlexA, GVG, and pOp/LhGR.

[0339] Where expression in a plant cell is desired, the cargo or other polynuclotide of the system described herein are typically placed under control of a plant promoter, i.e. a promoter operable in plant cells. The use of different types of promoters is envisaged.

[0340] A constitutive plant promoter is a promoter that is able to express the open reading frame (ORF) that it controls in all or nearly all of the plant tissues during all or nearly all developmental stages of the plant (referred to as "constitutive expression"). One non-limiting example of a constitutive promoter is the cauliflower mosaic virus 35S promoter. Different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. In particular embodiments, one or more of the cargos are expressed under the control of a constitutive promoter, such as the cauliflower mosaic virus 35S promoter issue-preferred promoters can be utilized to target enhanced expression in certain cell types within a particular plant tissue, for instance vascular cells in leaves or roots or in specific cells of the seed. Examples of particular promoters for use in the in the present invention are found in Kawamata et al., (1997) Plant Cell Physiol 38:792-803; Yamamoto et al., (1997) Plant J 12:255-65; Hire et al, (1992) Plant Mol Biol 20:207-18, Kuster et al, (1995) Plant Mol Biol 29:759-72, and Capana et al., (1994) Plant Mol Biol 25:681 -91.

[0341] Examples of promoters that are inducible and that can allow for spatiotemporal control of gene editing or gene expression may use a form of energy. The form of energy may include but is not limited to sound energy, electromagnetic radiation, chemical energy and/or thermal energy. Examples of inducible systems include tetracycline inducible promoters (Tet-On or Tet-Off), small molecule two-hybrid transcription activations systems (FKBP, ABA, etc.), or light inducible systems (Phytochrome, LOV domains, or cryptochrome)., such as a Light Inducible Transcriptional Effector (LITE) that direct changes in transcriptional activity in a sequence-specific manner. The components of a light inducible system may include one

or more elements of the engineered delivery vesicle generation system described herein, a light-responsive cytochrome heterodimer (e.g. from Arabidopsis thaliana), and a transcriptional activation/repression domain. In some embodiments, the vector can include one or more of the inducible DNA binding proteins provided in International Patent Publication No. WO 2014/018423 and US Patent Publication Nos., 2015/0291966, 2017/0166903, 2019/0203212, which describe e.g. embodiments of inducible DNA binding proteins and methods of use and can be adapted for use with the present invention.

[0342] In some embodiments, transient or inducible expression can be achieved by including, for example, chemical-regulated promotors, i.e. whereby the application of an exogenous chemical induces gene expression. Modulation of gene expression can also be obtained by including a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters include, but are not limited to, the maize ln2-2 promoter, activated by benzene sulfonamide herbicide safeners (De Veylder et al., (1997) Plant Cell Physiol 38:568-77), the maize GST promoter (GST-ll-27, WO93/01294), activated by hydrophobic electrophilic compounds used as pre-emergent herbicides, and the tobacco PR-1 a promoter (Ono et al., (2004) Biosci Biotechnol Biochem 68:803-7) activated by salicylic acid. Promoters which are regulated by antibiotics, such as tetracycline-inducible and tetracycline-repressible promoters (Gatz et al., (1991) Mol Gen Genet 227:229-37; U.S. Patent Nos. 5,814,618 and 5,789,156) can also be used herein.

[0343] In some embodiments, the polynucleotide, vector or system thereof can include one or more elements capable of translocating and/or expressing a cargo polynucleotide to/in a specific cell component or organelle. Such organelles can include, but are not limited to, nucleus, ribosome, endoplasmic reticulum, Golgi apparatus, chloroplast, mitochondria, vacuole, lysosome, cytoskeleton, plasma membrane, cell wall, peroxisome, centrioles, etc. Such regulatory elements can include, but are not limited to, nuclear localization signals (examples of which are described in greater detail elsewhere herein), any such as those that are annotated in the LocSigDB database (see e.g. http://genome.unmc.edu/LocSigDB/ and Negi et al., 2015. Database. 2015: bav003; doi: 10.1093/database/bav003), nuclear export signals (e.g. LXXXLXXLXL (SEQ ID NO:42) and others described elsewhere herein), endoplasmic reticulum localization/retention signals (e.g. KDEL (SEQ ID NO:43), KDXX (SEQ ID NO:44), KKXX (SEQ ID NO:45), KXX, and others described elsewhere herein; and see e.g. Liu et al. 2007 Mol. Biol. Cell. 18(3):1073-1082 and Gorleku et al., 2011. J. Biol. Chem.

286:39573-39584), mitochondria (see e.g. Cell Reports. 22:2818-2826, particularly at Fig. 2; Doyle et al. 2013. PLoS ONE 8, e67938; Funes et al. 2002. J. Biol. Chem. 277:6051-6058; Matouschek et al. 1997. PNAS USA 85:2091-2095; Oca-Cossio et al., 2003. 165:707-720; Waltner et al., 1996. J. Biol. Chem. 271:21226-21230; Wilcox et al., 2005. PNAS USA 102:15435-15440; Galanis et al., 1991. FEBS Lett 282:425-430, peroxisome (e.g. (S/A/C)-(K/R/H)-(L/A), SLK, (R/K)-(L/V/I)-XXXXX-(H/Q)-(L/A/F) (SEQ ID NO: 51). Suitable protein targeting motifs can also be designed or identified using any suitable database or prediction tool, including but not limited to Minimotif Miner (http:minimotifminer.org, http://mitominer.mrc-mbu.cam.ac.uk/release-4.0/embodiment.do?name=Protein%20MTS), LocDB (see above), PTSs predictor, TargetP-2.0 (http://www.cbs.dtu.dk/services/TargetP/), ChloroP (http://www.cbs.dtu.dk/services/ChloroP/); NetNES (http://www.cbs.dtu.dk/services/NetNES/), Predotar (https://urgi.versailles.inra.fr/predotar/), and SignalP (http://www.cbs.dtu.dk/services/SignalP/).

Selectable Markers and Tags

[0344] One or more of the engineered delivery vesicle generation system polynucleotides can be operably linked, fused to, or otherwise modified to include a polynucleotide that encodes or is a selectable marker or tag, which can be a polynucleotide or polypeptide. In some embodiments, the polynucleotide encoding a polypeptide selectable marker can be incorporated in the engineered delivery vesicle generation system such that the selectable marker polypeptide, when translated, is inserted between two amino acids between the N- and C- terminus of the cargo polypeptide or at the N- and/or C-terminus of the cargo polypeptide. In some embodiments, the selectable marker or tag is a polynucleotide barcode or unique molecular identifier (UMI).

[0345] It will be appreciated that the polynucleotide encoding such selectable markers or tags can be incorporated into a polynucleotide encoding one or more components of the engineered delivery vesicle generation system described herein in an appropriate manner to allow expression of the selectable marker or tag. Such techniques and methods are described elsewhere herein and will be instantly appreciated by one of ordinary skill in the art in view of this disclosure. Many such selectable markers and tags are generally known in the art and are intended to be within the scope of this disclosure.

[0346] Suitable selectable markers and tags include, but are not limited to, affinity tags, such as chitin binding protein (CBP), maltose binding protein (MBP), glutathione-S-

transferase (GST), poly(His) tag; solubilization tags such as thioredoxin (TRX) and poly(NANP), MBP, and GST; chromatography tags such as those consisting of polyanionic amino acids, such as FLAG-tag; epitope tags such as V5-tag, Myc-tag, HA-tag and NE-tag; protein tags that can allow specific enzymatic modification (such as biotinylation by biotin ligase) or chemical modification (such as reaction with FlAsH-EDT2 for fluorescence imaging), DNA and/or RNA segments that contain restriction enzyme or other enzyme cleavage sites; DNA segments that encode products that provide resistance against otherwise toxic compounds including antibiotics, such as, spectinomycin, ampicillin, kanamycin, tetracycline, Basta, neomycin phosphotransferase II (NEO), hygromycin phosphotransferase (HPT)) and the like; DNA and/or RNA segments that encode products that are otherwise lacking in the recipient cell (e.g., tRNA genes, auxotrophic markers); DNA and/or RNA segments that encode products which can be readily identified (e.g., phenotypic markers such as β-galactosidase, GUS; fluorescent proteins such as green fluorescent protein (GFP), cyan (CFP), yellow (YFP), red (RFP), luciferase, and cell surface proteins); polynucleotides that can generate one or more new primer sites for PCR (e.g., the juxtaposition of two DNA sequences not previously juxtaposed), DNA sequences not acted upon or acted upon by a restriction endonuclease or other DNA modifying enzyme, chemical, etc.; epitope tags (e.g. GFP, FLAG- and His-tags), and, DNA sequences that make a molecular barcode or unique molecular identifier (UMI), DNA sequences required for a specific modification (e.g., methylation) that allows its identification. Other suitable markers will be appreciated by those of skill in the art.

[0347] Selectable markers and tags can be operably linked to one or more components of the engineered delivery vesicle generation system herein via suitable linker, such as a glycine or glycine serine linkers as short as GS or GG up to $(GGGGG)_3$ (SEQ ID NO:46) or $(GGGGS)_3$ (SEQ ID NO:47). Other suitable linkers are described elsewhere herein.

[0348] The vector or vector system can include one or more polynucleotides encoding one or more targeting moieties. In some embodiments, the targeting moiety encoding polynucleotides can be included in the vector or vector system, such as a viral vector system, such that they are expressed within and/or on the virus particle(s) produced such that the virus particles can be targeted to specific cells, tissues, organs, etc. In some embodiments, the targeting moiety encoding polynucleotides can be included in the vector or vector system such that the engineered delivery vesicle generation system and/or cargo polynucleotide(s) and/or products expressed therefrom include the targeting moiety and can be targeted to specific cells,

tissues, organs, etc. In some embodiments, such as non-viral carriers, the targeting moiety can be attached to the carrier (e.g. polymer, lipid, inorganic molecule etc.) and can be capable of targeting the carrier and any attached or associated engineered delivery vesicle generation system polynucleotide(s) to specific cells, tissues, organs, etc.

Codon Optimization of Vector Polynucleotides

[0349] As described elsewhere herein, the polynucleotide encoding one or more embodiments of the engineered delivery vesicle generation system and/or cargos described herein can be codon optimized. In some embodiments, one or more polynucleotides contained in a vector ("vector polynucleotides") described herein that are in addition to an optionally codon optimized polynucleotide encoding embodiments of the engineered delivery vesicle generation system and/or cargos described herein can be codon optimized. In general, codon optimization refers to a process of modifying a nucleic acid sequence for enhanced expression in the host cells of interest by replacing at least one codon (e.g., about or more than about 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more codons) of the native sequence with codons that are more frequently or most frequently used in the genes of that host cell while maintaining the native amino acid sequence. Various species exhibit particular bias for certain codons of a particular amino acid. Codon bias (differences in codon usage between organisms) often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, among other things, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization. Codon usage tables are readily available, for example, at the "Codon Usage Database" available at www.kazusa.orjp/codon/ and these tables can be adapted in a number of ways. See Nakamura, Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000). Computer algorithms for codon optimizing a particular sequence for expression in a particular host cell are also available, such as Gene Forge (Aptagen; Jacobus, PA), are also available. In some embodiments, one or more codons (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, or more, or all codons) in a sequence encoding a DNA/RNA-targeting Cas protein corresponds to the most frequently used codon for a particular amino acid. As to codon usage in yeast, reference is made to the online Yeast Genome database available at

http://www.yeastgenome.org/community/codon_usage.shtml, or *Codon selection in yeast*, Bennetzen and Hall, J Biol Chem. 1982 Mar 25;257(6):3026-31. As to codon usage in plants including algae, reference is made to *Codon usage in higher plants, green algae, and cyanobacteria*, Campbell and Gowri, Plant Physiol. 1990 Jan; 92(1): 1–11.; as well as *Codon usage in plant genes*, Murray et al, Nucleic Acids Res. 1989 Jan 25;17(2):477-98; or *Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages*, Morton BR, J Mol Evol. 1998 Apr;46(4):449-59.

[0350] The vector polynucleotide can be codon optimized for expression in a specific cell-type, tissue type, organ type, and/or subject type. In some embodiments, a codon optimized sequence is a sequence optimized for expression in a eukaryote, e.g., humans (i.e. being optimized for expression in a human or human cell), or for another eukaryote, such as another animal (e.g. a mammal or avian) as is described elsewhere herein. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein. In some embodiments, the polynucleotide is codon optimized for a specific cell type. Such cell types can include, but are not limited to, epithelial cells (including skin cells, cells lining the gastrointestinal tract, cells lining other hollow organs), nerve cells (nerves, brain cells, spinal column cells, nerve support cells (e.g. astrocytes, glial cells, Schwann cells etc.) , muscle cells (e.g. cardiac muscle, smooth muscle cells, and skeletal muscle cells), connective tissue cells (fat and other soft tissue padding cells, bone cells, tendon cells, cartilage cells), blood cells, stem cells and other progenitor cells, immune system cells, germ cells, and combinations thereof. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein. In some embodiments, the polynucleotide is codon optimized for a specific tissue type. Such tissue types can include, but are not limited to, muscle tissue, connective tissue, connective tissue, nervous tissue, and epithelial tissue. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein. In some embodiments, the polynucleotide is codon optimized for a specific organ. Such organs include, but are not limited to, muscles, skin, intestines, liver, spleen, brain, lungs, stomach, heart, kidneys, gallbladder, pancreas, bladder, thyroid, bone, blood vessels, blood, and combinations thereof. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein.

[0351] In some embodiments, a vector polynucleotide is codon optimized for expression in particular cells, such as prokaryotic or eukaryotic cells. The eukaryotic cells may be those

of or derived from a particular organism, such as a plant or a mammal, including but not limited to human, or non-human eukaryote or animal or mammal as discussed herein, e.g., mouse, rat, rabbit, dog, livestock, or non-human mammal or primate.

Vector Construction

[0352] The vectors described herein can be constructed using any suitable process or technique. In some embodiments, one or more suitable recombination and/or cloning methods or techniques can be used to the vector(s) described herein. Suitable recombination and/or cloning techniques and/or methods can include, but not limited to, those described in U.S. Patent Publication No. US 2004/0171156 A1. Other suitable methods and techniques are described elsewhere herein.

[0353] Construction of recombinant AAV vectors are described in a number of publications, including U.S. Pat. No. 5,173,414; Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, et al., Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); and Samulski et al., J. Virol. 63:03822-3828 (1989). Any of the techniques and/or methods can be used and/or adapted for constructing an AAV or other vector described herein. nAAV vectors are discussed elsewhere herein.

[0354] In some embodiments, a vector comprises one or more insertion sites, such as a restriction endonuclease recognition sequence (also referred to as a "cloning site"). In some embodiments, one or more insertion sites (e.g., about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more insertion sites) are located upstream and/or downstream of one or more sequence elements of one or more vectors. When multiple different guide polynucleotides are used, a single expression construct may be used to target nucleic acid-targeting activity to multiple different, corresponding target sequences within a cell. For example, a single vector may comprise about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more guide s polynucleotides. In some embodiments, about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more such guide-polynucleotide-containing vectors may be provided, and optionally delivered to a cell.

[0355] Delivery vehicles, vectors, particles, nanoparticles, formulations and components thereof for expression of one or more elements of an engineered delivery vesicle generation system and/or cargo described herein are as used in the foregoing documents, such as International Patent Publication No. WO 2014/093622 (PCT/US2013/074667) and are discussed in greater detail herein.

DELIVERY VESICLES

[0356] Also envisioned within the scope of the invention are engineered retroviral delivery vesicles generated from the engineered retroviral delivery generation compositions and systems described elsewhere herein. In some embodiments, the engineered retroviral delivery vesicles includes one or more viral polypeptides and optionally a non-heterologous cargo molecule, the endogenous viral (e.g., a retroviral or lentiviral) polypeptide forming the engineered delivery vesicle and encapsulating the non-heterologous cargo molecule. As used herein "non-heterologous" is used to refer to cargo molecules not normally packaged by the delivery vesicle. For example, in the context of PEG10 which can package its own mRNA, a non-heterologous cargo molecule would exclude a naturally occurring PEG10 delivery vesicle comprising its own naturally occurring mRNA. In some embodiments, the delivery vesicle elicits a poor immune response, as described elsewhere herein. In some embodiments, the engineered delivery vesicle generation systems can produce an increased amount of delivery vesicles as compared to other viral particle production systems lacking PEG10 and/or RTL1.

[0357] The engineereddelivery vesicles produced by the engineered delivery vesicle generation system of the present invention are viral particles. In some embodiments, the engineered delivery vesicles produced by the engineered delivery vesicle generation system of the present invention are engineered retroviral particles. In some embodiments, the engineered delivery vesicles produced by the engineered delivery vesicle generation system of the present invention are engineered lentiviral particles.

ENGINEERED CELLS AND ORGANISMS

[0358] Described herein are engineered cells, cell populations, tissues, and organisms that can be generated using any of the engineered delivery vesicle generation compositions, systems, formulations, and/or generated delivery vesicles described herein. In some embodiments, the engineered cells contain one or more of the engineered compositions, systems, or formulations thereof and are capable of expressing one or more of the engineered compositions, systems, or formulations thereof and producing one or more engineered delivery vesicles that can optionally contain one or more cargos. Such cells are also referred to herein as "producer cells" or donor cells, depending on the context. It will be appreciated that these engineered cells are different from "modified cells" described elsewhere herein in that the modified cells are not necessarily producer or donor cells (e.g., they do not make engineered delivery vesicles) unless they include one or more of the engineered delivery system molecules

or vectors described herein that render the cells capable of producing an engineered delivery vesicle. Modified cells can be recipient cells of an engineered delivery vesicle and can, in some embodiments, be said to be modified by the engineered delivery vesicles and/or a cargo present in the engineered delivery vesicle that is delivered to the recipient cell. The term "modification" can be used in connection with modification of a cell that is not dependent on being a recipient cell. For example, isolated cells can be modified prior to receiving an engineered delivery system or engineered delivery vesicle and/or cargo. It will be appreciated that populations of producer cells can be included or part of a tissue, organ, or organism, such as a non-human animal or non-animal organism.

[0359] In some embodiments, a cargo of the engineered delivery vesicle generation compositions, systems, formulations thereof is a suitable polynucleotide and/or genome modifying agent(s) and/or systems for modifying a polynucleotide and/or genome of a cell so as to produce the engineered cells, tissues, and organisms. The modified cells, cell populations, and organisms can have an insertion of one or more polynucleotides, deletion of one or more polynucleotides, mutation of one or more polynucleotides, or a combination thereof. The modification can result in activation of one or more genes, inactivation of one or more genes, modulation of one or more genes, or a combination thereof. In some embodiments, one or more exogenous and/or heterologous genes or portions thereof are inserted into a cell so as to modify the cell. Cells, including cells in an organism, can be modified in vitro, in situ, ex vivo, or in vivo. In some embodiments, the modification is an insertion and/or deletion of a polynucleotide, gene, or allele of interest. Exemplary genes that can be modified or inserted are described in greater detail elsewhere herein.

[0360] In some embodiments, the engineered organism is a human, non-human animal, or non-animal organism (e.g., plant, fungi, prokaryote, and the like).

Cells

[0361] The cells to which the engineered compositions, systems, formulations, vesicles, and optionally cargos can be delivered to so as to produce a delivery vesicle producer cell or modified cell (collectively referred to in this context as "engineered cells") or population thereof can be any suitable eukaryotic or prokaryotic cell or population thereof. The engineered cell can be any eukaryotic cell, including but not limited to, human, non-human animal, plant, algae, and the like.

[0362] As used herein, a "population" of cells is any number of cells greater than 1, but is preferably at least $1X10^3$ cells, at least $1X10^4$ cells, at least at least $1X10^5$ cells, at least $1X10^6$ cells, at least $1X10^7$ cells, at least $1X10^8$ cells, at least $1X10^9$ cells, or at least $1X10^{10}$ cells.

[0363] In certain embodiments, the cell population can be composed of a single cell type or subtype. In some embodiments, the cell population may comprise several cell types and the combination of cell types and/or subtypes may comprise an immune cell, intestinal cell, liver cell, kidney cell, lung cell, brain cell, epithelial cell, endoderm cell, neuron, ectoderm cell, islet cell, acinar cell, oocyte, sperm, hematopoietic cell, hepatocyte, skin/keratinocyte, melanocyte, bone/osteocyte, hair/dermal papilla cell, cartilage/chondrocyte, fat cell/adipocyte, skeletal muscular cell, endothelium cell, cardiac muscle/cardiomyocyte, trophoblast, tumor cell, or tumor microenvironment (TME) cell.

[0364] Cell lines suitable for use in the present invention can be found at any public cell line database, including, but not necessarily limited to, ThermoFisher Scientific or Cellosaurus-ExPASy, all of which are incorporated by reference herein. Other cell lines suitable for biopharmaceutical manufacturing are discussed in Dumont et al. *Crit Rev Biotechnol* 36(6):1110-1122 (2016) and Ulrich et al. Encyclopedia of Genetics (2001), each of which is incorporated by reference herein. In some embodiments, the population of cells comprises 293T or 293FT cells. Other exemplary cells are described and elsewhere herein.

[0365] Cell lines suitable for use in the present invention can be found at any public cell line database, including, but not necessarily limited to, ThermoFisher Scientific or Cellosaurus-ExPASy, all of which are incorporated by reference herein. Other cell lines suitable for biopharmaceutical manufacturing are discussed in Dumont et al. *Crit Rev Biotechnol* 36(6):1110-1122 (2016) and Ulrich et al. Encyclopedia of Genetics (2001), each of which is incorporated by reference herein. In some embodiments, the population of cells comprises 293T or 293FT cells.

[0366] The engineered cell can be a prokaryotic cell. The prokaryotic cell can be bacterial cell. The prokaryotic cell can be an archaea cell. The bacterial cell can be any suitable bacterial cell. Suitable bacterial cells can be from the genus Escherichia, Bacillus, Lactobacillus, Rhodococcus, Rodhobacter, Synechococcus, Synechoystis, Pseudomonas, Psedoaltermonas, Stenotrophamonas, and Streptomyces Suitable bacterial cells include, but are not limited to Escherichia coli cells, Caulobacter crescentus cells, Rodhobacter sphaeroides cells,

Psedoaltermonas haloplanktis cells. Suitable strains of bacterial include, but are not limited to BL21(DE3), DL21(DE3)-pLysS, BL21 Star-pLysS, BL21-SI, BL21-AI, Tuner, Tuner pLysS, Origami, Origami B pLysS, Rosetta, Rosetta pLysS, Rosetta-gami-pLysS, BL21 CodonPlus, AD494, BL2trxB, HMS174, NovaBlue(DE3), BLR, C41(DE3), C43(DE3), Lemo21(DE3), Shuffle T7, ArcticExpress and ArticExpress (DE3).

[0367] The engineered cell can be a eukaryotic cell. The eukaryotic cells may be those of or derived from a particular organism, such as a plant or a mammal, including but not limited to human, or non-human eukaryote or animal or mammal as herein discussed, e.g., mouse, rat, rabbit, dog, livestock, or non-human mammal or primate. In some embodiments, the engineered cell can be a cell line. Examples of cell lines include, but are not limited to, HEK293 and variants (e.g. HEK293T) C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panc1, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 monkey kidney epithelial, BALB/ 3T3 mouse embryo fibroblast, 3T3 Swiss, 3T3-L1, 132-d5 human fetal fibroblasts; 10.1 mouse fibroblasts, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 cells, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr -/-, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, JY cells, K562 cells, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK II, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN / OPCT cell lines, Peer, PNT-1A / PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, THP1 cell line, U373, U87, U937, VCaP, Vero cells, WM39, WT-49, X63, YAC-1, YAR, and transgenic varieties thereof. Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassus, Va.)).

[0368] In some embodiments, the cell is obtained from a subject or a tumor therein. The subject can be a human, non-human animal, or plant.

[0369] Further, the engineered cell may be a fungus cell. As used herein, a "fungal cell" refers to any type of eukaryotic cell within the kingdom of fungi. Phyla within the kingdom of fungi include Ascomycota, Basidiomycota, Blastocladiomycota, Chytridiomycota, Glomeromycota, Microsporidia, and Neocallimastigomycota. Fungal cells may include yeasts, molds, and filamentous fungi. In some embodiments, the fungal cell is a yeast cell.

[0370] As used herein, the term "yeast cell" refers to any fungal cell within the phyla Ascomycota and Basidiomycota. Yeast cells may include budding yeast cells, fission yeast cells, and mold cells. Without being limited to these organisms, many types of yeast used in laboratory and industrial settings are part of the phylum Ascomycota. In some embodiments, the yeast cell is an S. cerevisiae, Kluyveromyces marxianus, or Issatchenkia orientalis cell. Other yeast cells may include without limitation Candida spp. (e.g., Candida albicans), Yarrowia spp. (e.g., Yarrowia lipolytica), Pichia spp. (e.g., Pichia pastoris), Kluyveromyces spp. (e.g., Kluyveromyces lactis and Kluyveromyces marxianus), Neurospora spp. (e.g., Neurospora crassa), Fusarium spp. (e.g., Fusarium oxysporum), and Issatchenkia spp. (e.g., Issatchenkia orientalis, a.k.a. Pichia kudriavzevii and Candida acidothermophilum). In some embodiments, the fungal cell is a filamentous fungal cell. As used herein, the term "filamentous fungal cell" refers to any type of fungal cell that grows in filaments, i.e., hyphae or mycelia. Examples of filamentous fungal cells may include without limitation Aspergillus spp. (e.g., Aspergillus niger), Trichoderma spp. (e.g., Trichoderma reesei), Rhizopus spp. (e.g., Rhizopus oryzae), and Mortierella spp. (e.g., Mortierella isabellina).

[0371] In some embodiments, the fungal cell is an industrial strain. As used herein, "industrial strain" refers to any strain of fungal cell used in or isolated from an industrial process, e.g., production of a product on a commercial or industrial scale. Industrial strain may refer to a fungal species that is typically used in an industrial process, or it may refer to an isolate of a fungal species that may be also used for non-industrial purposes (e.g., laboratory research). Examples of industrial processes may include fermentation (e.g., in production of food or beverage products), distillation, biofuel production, production of a compound, and production of a polypeptide. Examples of industrial strains can include, without limitation, JAY270 and ATCC4124.

[0372] In some embodiments, the fungal cell is a polyploid cell. As used herein, a "polyploid" cell may refer to any cell whose genome is present in more than one copy. A polyploid cell may refer to a type of cell that is naturally found in a polyploid state, or it may refer to a cell that has been induced to exist in a polyploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). A polyploid cell may refer to a cell whose entire genome is polyploid, or it may refer to a cell that is polyploid in a particular genomic locus of interest.

[0373] In some embodiments, the fungal cell is a diploid cell. As used herein, a "diploid" cell may refer to any cell whose genome is present in two copies. A diploid cell may refer to a type of cell that is naturally found in a diploid state, or it may refer to a cell that has been induced to exist in a diploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). For example, the S. cerevisiae strain S228C may be maintained in a haploid or diploid state. A diploid cell may refer to a cell whose entire genome is diploid, or it may refer to a cell that is diploid in a particular genomic locus of interest. In some embodiments, the fungal cell is a haploid cell. As used herein, a "haploid" cell may refer to any cell whose genome is present in one copy. A haploid cell may refer to a type of cell that is naturally found in a haploid state, or it may refer to a cell that has been induced to exist in a haploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). For example, the S. cerevisiae strain S228C may be maintained in a haploid or diploid state. A haploid cell may refer to a cell whose entire genome is haploid, or it may refer to a cell that is haploid in a particular genomic locus of interest.

[0374] In some embodiments, the cell is an insect cell, such as Sf9, or those derived from Bombyx mori, Mamestra brassicae, Spodoptera fruiperda, Trichoplusi ni, and Drosophilia melanogaster. See also e.g., Drugmand et al., 2011. Biotechnology Advances. 30(5):1140-1157.

[0375] Conventional viral and non-viral based gene transfer methods can be used to introduce nucleic acids (e.g. such as one or more of the polynucleotides of the engineered delivery system described herein) in cells or target tissues. Such methods can be used to administer nucleic acids encoding components of a nucleic acid-targeting system to cells in culture, or in a host organism. In some embodiments, a delivery is via a polynucleotide molecule (e.g. a DNA or RNA molecule) not contained in a vector. In some embodiments,

delivery is via a vector. In some embodiments, delivery, is via viral particles. In aspects delivery is via a particle, (e.g. a nanoparticle) carrying one or more engineered delivery system polynucleotides, vectors, or viral particles. Particles, including nanoparticles, are discussed in greater detail elsewhere herein.

[0376] Vector delivery can be appropriate in some embodiments, where in vivo expression is envisaged. It will be appreciated that the engineered cells can be generated in vitro, ex vivo, in situ, or in vivo by delivery of one or more components of the engineered delivery systems as described elsewhere herein.

[0377] Suitable conventional viral and non-viral based methods of engineering cells to contain and/or express the engineered delivery system polynucleotides and/or vectors described herein are generally known in the art and/or described elsewhere herein.

Engineered Delivery Vesicle Producing Cells

[0378] In some embodiments, the engineered cell is a cell obtained from a subject, such as a human, non-human animal, or plant subject. In some embodiments, the subject is a healthy or non-diseased subject. In some embodiments, the subject is a subject with a desired physiological and/or biological characteristic such that when an engineered delivery vesicle is produced it can package one or more molecules that are within the producer cell that can be related to the desired physiological and/or biological characteristic. In this context, the cargo molecules incorporated into the delivery vesicles can be capable of transferring the desired characteristic to a recipient cell.

[0379] In some embodiments, a cell can be obtained from a subject, modified such that it is an engineered delivery vesicle producer cell, and administered back to the subject from which it was obtained (autologous) or delivered to an allogenic subject. In other words, a producer cell described herein can be used in an autologous or allogenic context, such as in a cell therapy. In these embodiments, the cells can deliver a cargo, such as a therapeutic cargo or a cargo that can manipulate a cellular microenvironment within the subject.

Modified Cells

[0380] Also described herein are modified cells and cell populations that can be modified by an embodiment of engineered composition, system, formulation, delivery vesicle, and/or cargo of the present invention. In some embodiments, a cell is modified by a programmable nuclease-based system such as a TALEN, Zinc-finger nuclease, or an RNA guided nuclease

system (such as a CRISPR-Cas or IscB system). In some embodiments, the cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is a mammalian cell. In some embodiments, the eukaryotic cell is a non-human mammalian cell. In some embodiments, the cell is a human cell. In some embodiments, the cell is a plant cell. In some embodiments, the cell is a fungal cell. In some embodiments, the cell is a prokaryotic cell. The cells can be modified in vitro, ex vivo, or in vivo. The cells can be modified by delivering a polynucleotide modifying agent or system described in greater detail elsewhere herein or a component thereof into a cell by a suitable delivery mechanism. Suitable delivery methods and techniques include but are not limited to, transfection via a vector, transduction with viral particles, electroporation, endocytic methods, and others, which are described elsewhere herein and will be appreciated by those of ordinary skill in the art in view of this disclosure.

[0381] The modified cells can be further optionally cultured and/or expanded in vitro or ex vivo using any suitable cell culture techniques or conditions, which unless specified otherwise herein, will be appreciated by one of ordinary skill in the art in view of this disclosure. In some embodiments, the cells can be modified, optionally cultured and/or expanded, and administered to a subject in need thereof. In some embodiments, cells can be isolated from a subject, subsequently modified and optionally cultured and/or expanded, and administered back to the subject. Such administration can be referred to as autologous administration. In some embodiments, cells can be isolated from a first subject, subsequently modified, optionally cultured and/or expanded, and administered to a second subject, where the first subject and the second subject are different. Such administration can be referred to as non-autologous administration.

Organisms

[0382] Also described herein are modified organisms. In some embodiments, the modified organisms can include one or more modified cells as are described elsewhere herein. In some embodiments, the modified organism is a non-human mammal. In some embodiments, the modified organism is a modified plant. In some embodiments, the modified organism is an insect. In some embodiments, the modified organism is a fungus. In some embodiments, the modified organism is a fungus. The modified organisms can be generated using a that can be modified by an embodiment of the engineered or non-natural guided excision-transposition system described herein. Methods of making modified organisms are described in greater detail elsewhere herein.

[0383] The systems and methods described herein can be used in non-animal organisms, e.g., plants, fungi to generated modified non-animal organisms. The system and methods described can be used to generate non-human animal organisms. The system and methods described herein can be used to modify non-germline cells in a human. In some embodiments, the modification is expression of a polynucleotide of interest, gene of interest, and/or allele of interest. Exemplary genes and polynucleotides that can be modified are described in greater detail elsewhere herein

Non-Animal Organisms

[0384] In some embodiments, the modified non-animal organisms are plants, yeasts, etc. In general, the term "plant" relates to any various photosynthetic, eukaryotic, unicellular or multicellular organism of the kingdom Plantae characteristically growing by cell division, containing chloroplasts, and having cell walls comprised of cellulose. The term plant encompasses monocotyledonous and dicotyledonous plants. The term plant encompasses plant parts, clippings, grafts, and progeny thereof. Specifically, the plants are intended to comprise without limitation angiosperm and gymnosperm plants such as acacia, alfalfa, amaranth, apple, apricot, artichoke, ash tree, asparagus, avocado, banana, barley, beans, beet, birch, beech, blackberry, blueberry, broccoli, Brussel's sprouts, cabbage, canola, cantaloupe, carrot, cassava, cauliflower, cedar, a cereal, celery, chestnut, cherry, Chinese cabbage, citrus, clementine, clover, coffee, corn, cotton, cowpea, cucumber, cypress, eggplant, elm, endive, eucalyptus, fennel, figs, fir, geranium, grape, grapefruit, groundnuts, ground cherry, gum hemlock, hickory, kale, kiwifruit, kohlrabi, larch, lettuce, leek, lemon, lime, locust, pine, maidenhair, maize, mango, maple, melon, millet, mushroom, mustard, nuts, oak, oats, oil palm, okra, onion, orange, an ornamental plant or flower or tree, papaya, palm, parsley, parsnip, pea, peach, peanut, pear, peat, pepper, persimmon, pigeon pea, pine, pineapple, plantain, plum, pomegranate, potato, pumpkin, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, safflower, sallow, soybean, spinach, spruce, squash, strawberry, sugar beet, sugarcane, sunflower, sweet potato, sweet corn, tangerine, tea, tobacco, tomato, trees, triticale, turf grasses, turnips, vine, walnut, watercress, watermelon, wheat, yams, yew, and zucchini. The term plant also encompasses Algae, which are mainly photoautotrophs unified primarily by their lack of roots, leaves and other organs that characterize higher plants.

[0385] The modifications can be used to confer desired traits on essentially any plant. A wide variety of plants and plant cell systems may be engineered for the desired physiological

and agronomic characteristics described herein using the nucleic acid constructs of the present disclosure and the various transformation methods mentioned above. In some embodiments, target plants and plant cells for engineering include, but are not limited to, those monocotyledonous and dicotyledonous plants, such as crops including grain crops (e.g., wheat, maize, rice, millet, barley), fruit crops (e.g., tomato, apple, pear, strawberry, orange), forage crops (e.g., alfalfa), root vegetable crops (e.g., carrot, potato, sugar beets, yam), leafy vegetable crops (e.g., lettuce, spinach); flowering plants (e.g., petunia, rose, chrysanthemum), conifers and pine trees (e.g., pine fir, spruce); plants used in phytoremediation (e.g., heavy metal accumulating plants); oil crops (e.g., sunflower, rape seed) and plants used for experimental purposes (e.g., Arabidopsis). Plant cells and tissues for engineering include, without limitation, roots, stems, leaves, flowers, and reproductive structures, undifferentiated meristematic cells, parenchyma, collenchyma, sclerenchyma, xylem, phloem, epidermis, and germplasm. Thus, the methods and engineered compositions, systems, and delivery vesicles described herein can be used over a broad range of plants, such as for example, with dicotyledonous plants belonging to the orders Magniolales, Illiciales, Laurales, Piperales, Aristochiales, Nymphaeales, Ranunculales, Papeverales, Sarraceniaceae, Trochodendrales, Hamamelidales, Eucomiales, Leitneriales, Myricales, Fagales, Casuarinales, Caryophyllales, Batales, Polygonales, Plumbaginales, Dilleniales, Theales, Malvales, Urticales, Lecythidales, Violales, Salicales, Capparales, Ericales, Diapensales, Ebenales, Primulales, Rosales, Fabales, Podostemales, Haloragales, Myrtales, Cornales, Proteales, San tales, Rafflesiales, Celastrales, Euphorbiales, Rhamnales, Sapindales, Juglandales, Geraniales, Polygalales, Umbellales, Gentianales, Polemoniales, Lamiales, Plantaginales, Scrophulariales, Campanulales, Rubiales, Dipsacales, and Asterales. In some embodiments, the plants that are to be or are modified are monocotyledonous plants, such as those belonging to the orders Alismatales, Hydrocharitales, Najadales, Triuridales, Commelinales, Eriocaulales, Restionales, Poales, Juncales, Cyperales, Typhales, Bromeliales, Zingiberales, Arecales, Cyclanthales, Pandanales, Arales, Lilliales, and Orchid ales, or with plants belonging to Gymnospermae, e.g those belonging to the orders Pinales, Ginkgoales, Cycadales, Araucariales, Cupressales and Gnetales.

[0386] The engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention can be used over a broad range of plant species, included in the non-limitative list of dicot, monocot or gymnosperm genera hereunder: Atropa, Alseodaphne, Anacardium, Arachis, Beilschmiedia, Brassica, Carthamus, Cocculus,

Croton, Cucumis, Citrus, Citrullus, Capsicum, Catharanthus, Cocos, Coffea, Cucurbita, Daucus, Duguetia, Eschscholzia, Ficus, Fragaria, Glaucium, Glycine, Gossypium, Helianthus, Hevea, Hyoscyamus, Lactuca, Landolphia, Linum, Litsea, Lycopersicon, Lupinus, Manihot, Majorana, Malus, Medicago, Nicotiana, Olea, Parthenium, Papaver, Persea, Phaseolus, Pistacia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Senecio, Sinomenium, Stephania, Sinapis, Solanum, Theobroma, Trifolium, Trigonella, Vicia, Vinca, Vilis, and Vigna; and the genera Allium, Andropogon, Aragrostis, Asparagus, Avena, Cynodon, Elaeis, Festuca, Festulolium, Heterocallis, Hordeum, Lemna, Lolium, Musa, Oryza, Panicum, Pannesetum, Phleum, Poa, Secale, Sorghum, Triticum, Zea, Abies, Cunninghamia, Ephedra, Picea, Pinus, and Pseudotsuga.

[0387] The engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom can also be used over a broad range of "algae" or "algae cells"; including for example algae selected from several eukaryotic phyla, including the Rhodophyta (red algae), Chlorophyta (green algae), Phaeophyta (brown algae), Bacillariophyta (diatoms), Eustigmatophyta and dinoflagellates as well as the prokaryotic phylum Cyanobacteria (blue-green algae). The term "algae" includes for example algae selected from : Amphora, Anabaena, Anikstrodesmis, Botryococcus, Chaetoceros, Chlamydomonas, Chlorella, Chlorococcum, Cyclotella, Cylindrotheca, Dunaliella, Emiliana, Euglena, Hematococcus, Isochrysis, Monochrysis, Monoraphidium, Nannochloris, Nannnochloropsis, Navicula, Nephrochloris, Nephroselmis, Nitzschia, Nodularia, Nostoc, Oochromonas, Oocystis, Oscillartoria, Pavlova, Phaeodactylum, Playtmonas, Pleurochrysis, Porhyra, Pseudoanabaena, Pyramimonas, Stichococcus, Synechococcus, Synechocystis, Tetraselmis, Thalassiosira, and Trichodesmium.

[0388] A part of a plant, e.g., a "plant tissue" may be treated according to the methods of the present invention to produce an improved plant. Plant tissue also encompasses plant cells. The term “plant cell” as used herein refers to individual units of a living plant, either in an intact whole plant or in an isolated form grown in in vitro tissue cultures, on media or agar, in suspension in a growth media or buffer or as a part of higher organized unites, such as, for example, plant tissue, a plant organ, or a whole plant.

[0389] A “protoplast” refers to a plant cell that has had its protective cell wall completely or partially removed using, for example, mechanical or enzymatic means resulting in an intact biochemical competent unit of living plant that can reform their cell wall, proliferate and regenerate grow into a whole plant under proper growing conditions.

[0390] The term "transformation" broadly refers to the process by which a plant host is genetically modified by the introduction of DNA by means of Agrobacteria or one of a variety of chemical or physical methods. As used herein, the term "plant host" refers to plants, including any cells, tissues, organs, or progeny of the plants. Many suitable plant tissues or plant cells can be transformed and include, but are not limited to, protoplasts, somatic embryos, pollen, leaves, seedlings, stems, calli, stolons, microtubers, and shoots. A plant tissue also refers to any clone of such a plant, seed, progeny, propagule whether generated sexually or asexually, and descendants of any of these, such as cuttings or seed.

[0391] The term "transformed" as used herein, refers to a cell, tissue, organ, or organism into which a foreign DNA molecule, such as a construct, has been introduced. The introduced DNA molecule may be integrated into the genomic DNA of the recipient cell, tissue, organ, or organism such that the introduced DNA molecule is transmitted to the subsequent progeny. In these embodiments, the "transformed" or “transgenic” cell or plant may also include progeny of the cell or plant and progeny produced from a breeding program employing such a transformed plant as a parent in a cross and exhibiting an altered phenotype resulting from the presence of the introduced DNA molecule. Preferably, the transgenic plant is fertile and capable of transmitting the introduced DNA to progeny through sexual reproduction.

[0392] The term “progeny”, such as the progeny of a transgenic plant, is one that is born of, begotten by, or derived from a plant or the transgenic plant. The introduced DNA molecule may also be transiently introduced into the recipient cell such that the introduced DNA molecule is not inherited by subsequent progeny and thus not considered “transgenic”. Accordingly, as used herein, a “non-transgenic” plant or plant cell is a plant which does not contain a foreign DNA stably integrated into its genome.

[0393] The term “plant promoter” as used herein is a promoter capable of initiating transcription in plant cells, whether or not its origin is a plant cell. Exemplary suitable plant promoters include, but are not limited to, those that are obtained from plants, plant viruses, and bacteria such as Agrobacterium or Rhizobium which comprise genes expressed in plant cells.

[0394] As used herein, a "fungal cell" refers to any type of eukaryotic cell within the kingdom of fungi. Phyla within the kingdom of fungi include Ascomycota, Basidiomycota, Blastocladiomycota, Chytridiomycota, Glomeromycota, Microsporidia, and Neocallimastigomycota. Fungal cells may include yeasts, molds, and filamentous fungi. In some embodiments, the fungal cell is a yeast cell.

[0395] As used herein, the term "yeast cell" refers to any fungal cell within the phyla Ascomycota and Basidiomycota. Yeast cells may include budding yeast cells, fission yeast cells, and mold cells. Without being limited to these organisms, many types of yeast used in laboratory and industrial settings are part of the phylum Ascomycota. In some embodiments, the yeast cell is an S. cerervisiae, Kluyveromyces marxianus, or Issatchenkia orientalis cell. Other yeast cells may include without limitation Candida spp. (e.g., Candida albicans), Yarrowia spp. (e.g., Yarrowia lipolytica), Pichia spp. (e.g., Pichia pastoris), Kluyveromyces spp. (e.g., Kluyveromyces lactis and Kluyveromyces marxianus), Neurospora spp. (e.g., Neurospora crassa), Fusarium spp. (e.g., Fusarium oxysporum), and Issatchenkia spp. (e.g., Issatchenkia orientalis, a.k.a. Pichia kudriavzevii and Candida acidothermophilum). In some embodiments, the fungal cell is a filamentous fungal cell. As used herein, the term "filamentous fungal cell" refers to any type of fungal cell that grows in filaments, i.e., hyphae or mycelia. Examples of filamentous fungal cells may include without limitation Aspergillus spp. (e.g., Aspergillus niger), Trichoderma spp. (e.g., Trichoderma reesei), Rhizopus spp. (e.g., Rhizopus oryzae), and Mortierella spp. (e.g., Mortierella isabellina).

[0396] In some embodiments, the fungal cell is an industrial strain. As used herein, "industrial strain" refers to any strain of fungal cell used in or isolated from an industrial process, e.g., production of a product on a commercial or industrial scale. Industrial strain may refer to a fungal species that is typically used in an industrial process, or it may refer to an isolate of a fungal species that may be also used for non-industrial purposes (e.g., laboratory research). Examples of industrial processes may include fermentation (e.g., in production of food or beverage products), distillation, biofuel production, production of a compound, and production of a polypeptide. Examples of industrial strains may include, without limitation, JAY270 and ATCC4124.

[0397] In some embodiments, the fungal cell is a polyploid cell. As used herein, a "polyploid" cell may refer to any cell whose genome is present in more than one copy. A polyploid cell may refer to a type of cell that is naturally found in a polyploid state, or it may refer to a cell that has been induced to exist in a polyploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). A polyploid cell may refer to a cell whose entire genome is polyploid, or it may refer to a cell that is polyploid in a particular genomic locus of interest. Without wishing to be bound to theory, it is thought that the abundance of guideRNA may more often be a rate-

limiting component in genome engineering of polyploidy cells than in haploid cells, and thus the methods using the systems described herein may take advantage of using a certain fungal cell type.

[0398] In some embodiments, the fungal cell is a diploid cell. As used herein, a "diploid" cell may refer to any cell whose genome is present in two copies. A diploid cell may refer to a type of cell that is naturally found in a diploid state, or it may refer to a cell that has been induced to exist in a diploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). For example, the S. cerevisiae strain S228C may be maintained in a haploid or diploid state. A diploid cell may refer to a cell whose entire genome is diploid, or it may refer to a cell that is diploid in a particular genomic locus of interest. In some embodiments, the fungal cell is a haploid cell. As used herein, a "haploid" cell may refer to any cell whose genome is present in one copy. A haploid cell may refer to a type of cell that is naturally found in a haploid state, or it may refer to a cell that has been induced to exist in a haploid state (e.g., through specific regulation, alteration, inactivation, activation, or modification of meiosis, cytokinesis, or DNA replication). For example, the S. cerevisiae strain S228C may be maintained in a haploid or diploid state. A haploid cell may refer to a cell whose entire genome is haploid, or it may refer to a cell that is haploid in a particular genomic locus of interest.

[0399] As used herein, a "yeast expression vector" refers to a nucleic acid that contains one or more sequences encoding an RNA and/or polypeptide and may further contain any desired elements that control the expression of the nucleic acid(s), as well as any elements that enable the replication and maintenance of the expression vector inside the yeast cell. Many suitable yeast expression vectors and features thereof are known in the art; for example, various vectors and techniques are illustrated in in Yeast Protocols, 2nd edition, Xiao, W., ed. (Humana Press, New York, 2007) and Buckholz, R.G. and Gleeson, M.A. (1991) Biotechnology (NY) 9(11): 1067-72. Yeast vectors may contain, without limitation, a centromeric (CEN) sequence, an autonomous replication sequence (ARS), a promoter, such as an RNA Polymerase III promoter, operably linked to a sequence or gene of interest, a terminator such as an RNA polymerase III terminator, an origin of replication, and a marker gene (e.g., auxotrophic, antibiotic, or other selectable markers). Examples of expression vectors for use in yeast may include plasmids, yeast artificial chromosomes, 2μ plasmids, yeast integrative plasmids, yeast replicative plasmids, shuttle vectors, and episomal plasmids.

[0400] Described herein are plants and/or plant cells that can be produced by one or more of the methods described herein, or a progeny thereof. The progeny may be a clone of the produced plant or animal, or may result from sexual reproduction by crossing with other individuals of the same species to introgress further desirable traits into their offspring. The cell may be in vivo or ex vivo in the cases of multicellular organisms, particularly plant. This is described in greater detail herein.

[0401] Also described herein are gametes, seeds, germplasm, embryos, either zygotic or somatic, progeny or hybrids of plants comprising the genetic modification, which are produced by traditional breeding methods, are also included within the scope of the present invention. Such plants may contain a heterologous or foreign DNA sequence inserted at or instead of a target sequence. Alternatively, such plants may contain only an alteration (mutation, deletion, insertion, substitution) in one or more nucleotides. As such, such plants will only be different from their progenitor plants by the presence of the particular modification.

[0402] The polynucleotide modifying agent(s) and/or systems described herein can be used to confer desired traits on essentially any plant, algae, fungus, yeast, etc. A wide variety of plants, algae, fungus, yeast, etc. and plant algae, fungus, yeast cell or tissue systems may be engineered for the desired physiological and agronomic characteristics described herein using the nucleic acid constructs of the present disclosure and the various transformation methods mentioned above.

[0403] In particular embodiments, the methods described herein are used to modify endogenous genes or to modify their expression without the permanent introduction into the genome of the plant, algae, fungus, yeast, etc. of any foreign gene, including those encoding CRISPR components, so as to avoid the presence of foreign DNA in the genome of the plant. This can be of interest as the regulatory requirements for non-transgenic plants are less rigorous.

[0404] Also described herein are modified non-animal organisms (plants, yeast, algae, and other microorganisms) that can express one or more polynucleotides, genes or alleles of interest. In some embodiments the gene of interest are one or more engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom. In some embodiments, the gene of interest is a cargo that can be included in the engineered compositions, systems, formulations, delivery vesicles, etc. Additional exemplary genes of interest are described in greater detail elsewhere herein.

Stable integration in the genome of plants and plant cells

[0405] In particular embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention are introduced for stable integration into the genome of a plant cell. In these embodiments, the design of the transformation vector or the expression system can be adjusted depending on for when, where and under what conditions the polynucleotide modifying agents or systems thereof are expressed. Suitable vectors and delivery are described in greater detail elsewhere herein.

[0406] In particular embodiments, the p engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention are stably introduced into the genomic DNA of a plant cell. In particular embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention are introduced for stable integration into the DNA of a plant organelle such as, but not limited to a plastid, e mitochondrion or a chloroplast. In some embodiments, an expression system for stable integration into the genome of a plant cell can contain one or more of the following elements: a promoter element that can be used to express a polynucleotide modifying agent(s) or a system thereof in a plant cell; a 5' untranslated region to enhance expression; an intron element to further enhance expression in certain cells, such as monocot cells; a multiple-cloning site to provide convenient restriction sites for inserting the polynucleotide modifying agent(s) or a system thereof and other desired elements; and a 3' untranslated region to provide for efficient termination of the expressed transcript. The elements of the expression system can be on one or more expression constructs which are either circular such as a plasmid or transformation vector, or non-circular such as linear double stranded DNA.

[0407] DNA construct(s) containing the components of the engineered delivery vesicle generation systems of the present invention, and, where applicable, template sequence may be introduced into the genome of a plant, plant part, or plant cell by a variety of conventional techniques. The process generally comprises the steps of selecting a suitable host cell or host tissue, introducing the construct(s) into the host cell or host tissue.

[0408] In particular embodiments, the DNA construct may be introduced into the plant cell using techniques such as but not limited to electroporation, microinjection, aerosol beam injection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using biolistic methods, such as DNA particle bombardment (see also e.g., Fu et al.,

Transgenic Res. 2000 Feb;9(1):11-9). The basis of particle bombardment is the acceleration of particles coated with gene/s of interest toward cells, resulting in the penetration of the protoplasm by the particles and typically stable integration into the genome. (see e.g., Klein et al, Nature (1987), Klein et ah, Bio/Technology (1992), Casas et ah, Proc. Natl. Acad. Sci. USA (1993).).

[0409] In particular embodiments, the DNA constructs containing components of the engineered delivery vesicle generation systems of the present invention can be introduced into the plant by Agrobacterium-mediated transformation. The DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional Agrobacterium tumefaciens host vector. The foreign DNA can be incorporated into the genome of plants by infecting the plants or by incubating plant protoplasts with Agrobacterium bacteria, containing one or more Ti (tumor-inducing) plasmids. See e.g., Fraley et al., (1985), Rogers et al., (1987) and U.S. Pat. No. 5,563,055).

Transient expression of in plants and plant cells

[0410] In some embodiments, one or more of the engineered compositions, systems, formulations thereof, vesicles, and/or cargo delivered therefrom is/are transiently expressed in the plant cell. In these embodiments, the system can ensure modification of a target gene only when all the required components of the system (e.g., in the context of a typical CRISPR-Cas system, the Cas enzyme(s) and guide RNA(s)) are present in a cell, such that polynucleotide modification can further be controlled. As the expression of the necessary components of the modification agent and/or system is transient, plants regenerated from such plant cells typically contain no foreign DNA. It will be appreciated that not all components must be expressed transiently for modification to be controlled by transient expression. In some embodiments where multiple components are necessary for modification to occur, one or more components of the modification system are expressed transiently and one or more components of the system are stably expressed. In some embodiments where an RNA guided nuclease system (such as a CRISPR-Cas or IscB system) is employed or delivered and/or generated by one or more of the engineered compositions, systems, and formulations thereof of the present invention, the effector nuclease is stably expressed by the plant cell and the guide sequence is transiently expressed. In some embodiments where a CRISPR-Cas or IscB system is employed, the Cas or IscB enzyme is transiently expressed by the plant cell and the guide sequence is stably expressed.

[0411] In particular embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom can be transiently introduced in the plant cells using a plant viral vector (Scholthof et al. 1996, Annu Rev Phytopathol. 1996;34:299-323). In further particular embodiments, said viral vector is a vector from, derived from, or based at least in part on a DNA virus. For example, geminivirus (e.g., cabbage leaf curl virus, bean yellow dwarf virus, wheat dwarf virus, tomato leaf curl virus, maize streak virus, tobacco leaf curl virus, or tomato golden mosaic virus) or nanovirus (e.g., Faba bean necrotic yellow virus). In other particular embodiments, said viral vector is a vector from, derived from, or based at least in part on an RNA virus. For example, tobravirus (e.g., tobacco rattle virus, tobacco mosaic virus), potexvirus (e.g., potato virus X), or hordeivirus (e.g., barley stripe mosaic virus). The replicating genomes of plant viruses are non-integrative vectors.

[0412] In particular embodiments, the vector used for transient expression of constructs is for instance a pEAQ vector (or based upon, such as in the context of the present invention), which is tailored for Agrobacterium-mediated transient expression (Sainsbury F. et al., Plant Biotechnol J. 2009 Sep;7(7):682-93) in the protoplast. Precise targeting of genomic locations was demonstrated using a modified Cabbage Leaf Curl virus (CaLCuV) vector to express gRNAs in stable transgenic plants expressing a CRISPR enzyme (Scientific Reports 5, Article number: 14926 (2015), doi:10.1038/srep14926). Such an approach can be applied and adapted for use with the engineered compositions and systems of the present invention.

[0413] In particular embodiments, double-stranded DNA fragments encoding the polynucleotide modifying agent(s) and/or system component(s) (e.g., where a CRISPR-Cas system is employed, a guide RNA and/or the Cas gene) can be transiently introduced into the plant cell. In such embodiments, the introduced double-stranded DNA fragments are provided in sufficient quantity to modify the cell but do not persist after a contemplated period of time has passed or after one or more cell divisions. Methods for direct DNA transfer in plants are known by the skilled artisan (see for instance Davey et al. Plant Mol Biol. 1989 Sep;13(3):273-85.)

[0414] In other embodiments, an RNA polynucleotide encoding an effector protein (such as an RNA guided nuclease) of a modifying agent or system component (e.g., where a CRISPR-Cas system is employed, a Cas protein) is introduced into the plant cell, which is then translated and processed by the host cell generating the protein in sufficient quantity to modify the cell (in the presence of at least one guide RNA) but which does not persist after a contemplated

period of time has passed or after one or more cell divisions. In some embodiments, the RNA polynucleotide is a cargo that can be delivered by one or more of the engineered compositions, systems, and vesicles of the present invention. Methods for introducing mRNA to plant protoplasts for transient expression are known by the skilled artisan (see e.g., in Gallie, Plant Cell Reports (1993), 13;119-122).

[0415] In some embodiments, a combination of the different methods described above can be used.

Plant promoters

[0416] In some embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom described elsewhere herein can be placed under control of a suitable plant promoter, i.e., a promoter operable in plant cells. The use of different types of promoters is envisaged. Plant promoters can be constitutive, inducible, and/or tissue specific.

[0417] A constitutive plant promoter is a promoter that is able to express the open reading frame (ORF) that it controls in all or nearly all of the plant tissues during all or nearly all developmental stages of the plant (referred to as "constitutive expression"). One non-limiting example of a constitutive promoter is the cauliflower mosaic virus 35S promoter. "Regulated promoter" refers to promoters that direct gene expression not constitutively, but in a temporally- and/or spatially-regulated manner, and includes tissue-specific, tissue-preferred and inducible promoters. Different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. In particular embodiments, one or more of the gene modifying agents are expressed under the control of a constitutive promoter, such as the cauliflower mosaic virus 35S promoter issue-preferred promoters can be utilized to target enhanced expression in certain cell types within a particular plant tissue, for instance vascular cells in leaves or roots or in specific cells of the seed. Examples of particular promoters for use in the system are found in Kawamata et al., (1997) Plant Cell Physiol 38:792-803; Yamamoto et al., (1997) Plant J 12:255-65; Hire et al, (1992) Plant Mol Biol 20:207-18, Kuster et al, (1995) Plant Mol Biol 29:759-72, and Capana et al., (1994) Plant Mol Biol 25:681 -91.

[0418] Examples of promoters that are inducible and that allow for spatiotemporal control of gene editing or gene expression may use a form of energy. The form of energy may include but is not limited to sound energy, electromagnetic radiation, chemical energy and/or thermal

energy. Examples of inducible systems include tetracycline inducible promoters (Tet-On or Tet-Off), small molecule two-hybrid transcription activations systems (FKBP, ABA, etc), or light inducible systems (Phytochrome, LOV domains, or cryptochrome)., such as a Light Inducible Transcriptional Effector (LITE) that direct changes in transcriptional activity in a sequence-specific manner. The components of a light inducible system may include one or more gene modifying agents, a light-responsive cytochrome heterodimer (e.g., from Arabidopsis thaliana), and a transcriptional activation/repression domain. Further examples of inducible DNA binding proteins and methods for their use are provided in US 61/736465 and US 61/721,283, which is hereby incorporated by reference in its entirety.

[0419] In particular embodiments, transient or inducible expression can be achieved by using, for example, chemical-regulated promotors, i.e., whereby the application of an exogenous chemical induces gene expression. Modulating of gene expression can also be obtained by a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters include, but are not limited to, the maize ln2-2 promoter, activated by benzene sulfonamide herbicide safeners (De Veylder et al., (1997) Plant Cell Physiol 38:568-77), the maize GST promoter (GST-ll-27, WO93/01294), activated by hydrophobic electrophilic compounds used as pre-emergent herbicides, and the tobacco PR-1 a promoter (Ono et al., (2004) Biosci Biotechnol Biochem 68:803-7) activated by salicylic acid. Promoters which are regulated by antibiotics, such as tetracycline-inducible and tetracycline-repressible promoters (Gatz et al., (1991) Mol Gen Genet 227:229-37; U.S. Patent Nos. 5,814,618 and 5,789,156) can also be used herein.

Translocation to and/or expression in specific plant organelles

[0420] The engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom can include one or more elements for translocation to and/or expression in a specific plant organelle. In some embodiments, a tissue specific promoter can be included in the expression construct. In some embodiments, a tissue localization or organelle localization sequence or signal can be incorporated into the expression constructs. Such promoters and localization signals are described in greater detail elsewhere herein and/or will be appreciated by one of ordinary skill in the art.

Chloroplast targeting

[0421] In some embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom can modify chloroplast genes or to

ensure expression in the chloroplast. Thus in some embodiments, the modified cells contain modified chloroplasts. In some embodiments, chloroplast transformation methods or compartmentalization of the system components to the chloroplast. For instance, the introduction of genetic modifications in the plastid genome can reduce biosafety issues such as gene flow through pollen.

[0422] Methods of chloroplast transformation are known in the art and include particle bombardment, PEG treatment, and microinjection. Additionally, methods involving the translocation of transformation cassettes from the nuclear genome to the plastid can be used as described in WO2010061186.

[0423] In some embodiments, one or more of the polynucleotide modifying system components can be targeted to the plant chloroplast. This can be achieved by incorporating in the expression construct a sequence encoding a chloroplast transit peptide (CTP) or plastid transit peptide, operably linked to the 5' region of the sequence encoding the Cas protein. The CTP is removed in a processing step during translocation into the chloroplast. Chloroplast targeting of expressed proteins is well known to the skilled artisan (see e.g., Protein Transport into Chloroplasts, 2010, Annual Review of Plant Biology,Vol. 61: 157-180). In such embodiments it is also can be desirable to target the guide RNA to the plant chloroplast. Methods and constructs which can be used for translocating guide RNA into the chloroplast by means of a chloroplast localization sequence are described, for instance, in US 20040142476, incorporated herein by reference. Such variations of constructs can be incorporated into the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom to efficiently translocate the genetic modifying agent or system, such as an RNA guided nuclease system.

Introduction of polynucleotides in Algal cells.

[0424] In some embodiments, the modified organism is algae. Modified algae (or other plants such as rape) can be useful in a variety of situations, such as in the production of vegetable oils or biofuels such as alcohols (especially methanol and ethanol) or other products. In some embodiments, such organisms can be engineered to express or overexpress high levels of a useful product. For example, they can be modified to produce oil and/or alcohols for use in the oil or biofuel industries.

[0425] Algae modification using polynucleotide modifying agents has been described in, for example U.S. Pat. No. 8,945,839 and WO 2015086795, which can be adapted to modifying

algae and similar organisms with the polynucleotide modifying agents and systems described herein. In some embodiments, the polynucleotide modifying agent(s) or system thereof can be introduced to the algae using a vector that expresses the polynucleotide modifying agent(s) or system thereof under the control of a constitutive promoter such as Hsp70A-Rbc S2 or Beta2 -tubulin. Some components of the polynucleotide modifying system (such as a guide RNA or other RNAs) can be optionally delivered using a vector containing T7 promoter. In some embodiments, a polynucleotide modifying agent and/or other components of the polynucleotide modifying system mRNA can be expressed and in vitro transcribed guide RNA can be delivered to algal cells. In some embodiments, delivery can be via electroporation. Electroporation protocols are available to the skilled person such as the standard recommended protocol from the GeneArt Chlamydomonas Engineering kit.

[0426] In some embodiments, engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom is/are or contain or can deliver to the algal cells is a split Cas enzyme. Split Cas enzymes used in Algae for targeted genome modification as have been described for Cas9 in WO 2015086795. Use of the Cas split system is suitable for an inducible method of genome targeting and can avoid or mitigate the potential toxic effect of the Cas overexpression within the algae cell. In particular embodiments, said Cas split domains (RuvC and HNH domains in the case of Cas9) can be simultaneously or sequentially introduced into the cell such that said split Cas domain(s) process the target nucleic acid sequence in the algae cell. The reduced size of the split Cas compared to the wild type Cas allows other methods of delivery of the systems to the cells, such as the use of cell penetrating peptides as described herein. Such approaches can be applied and adapted for other genetic modifying agents and systems, including but not limited to other RNA guided nuclease systems. Such approaches can be applied and adapted for the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom.

Introduction of polynucleotides in yeast cells

[0427] In some embodiments, a yeast cell can be modified using the polynucleotide modifying agents and/or systems described herein. Methods for transforming yeast cells which can be used to introduce polynucleotides encoding the systems components are well known to the artisan and are reviewed by Kawai et al., 2010, Bioeng Bugs. 2010 Nov-Dec; 1(6): 395–403). Non-limiting examples include transformation of yeast cells by lithium acetate treatment

(which may further include carrier DNA and PEG treatment), bombardment or by electroporation.

Delivery to the plant cell

[0428] In some embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom are delivered and/or facilitate delivery directly to the plant cell. In particular embodiments, one or more of the polynucleotide modifying agent(s) or components of the system can be prepared outside the plant or plant cell and delivered to the cell. For example, in some embodiments one or more polynucleotide modifying agents, systems, or components thereof are incorporated into one or more of the engineered compositions, systems, formulations, delivery vesicles, etc., such as a cargo, and can be delivered, such as via a vector or vesicle be delivered to a plant cell.

[0429] In some embodiments, the protein polynucleotide modifying agent (e.g., where a CRISPR-Cas system is used, a Cas protein) or system component is prepared in vitro prior to introduction to the plant cell. Proteins can be prepared by various methods known by one of skill in the art and include recombinant production and de novo synthesis. After expression, the protein can be isolated, refolded if needed, purified and optionally treated to remove any purification tags, such as a His-tag. Once crude, partially purified, or more completely purified protein is obtained, the protein may be introduced to the plant cell.

[0430] In some embodiments where a CRISPR-Cas or other RNA guided nuclease system is employed, the Cas or other protein(s) can be mixed with guide RNA(s) targeting the gene(s) of interest to form a pre-assembled ribonucleoprotein. Such complex can optionally be delivered or by an engineered compositions, systems, formulations, delivery vesicles, etc. of the present invention.

[0431] The individual components or pre-assembled ribonucleoprotein, whether incorporated into engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention or not, can be introduced into the plant cell via electroporation, by bombardment with an RNA guided nuclease effector-associated (or other genetic modifying agent effector) gene product coated particles, by chemical transfection or by some other means of transport across a cell membrane. For instance, transfection of a plant protoplast with a pre-assembled CRISPR ribonucleoprotein has been demonstrated to ensure targeted modification of the plant genome (as described by

Woo et al. Nature Biotechnology, 2015; DOI: 10.1038/nbt.3389), which can be adapted for use with the present invention.

[0432] In particular embodiments, the system components are introduced into the plant cells using nanoparticles. The components, either as protein or nucleic acid or in a combination thereof, can be uploaded onto or packaged in nanoparticles and applied to the plants (such as for instance described in WO 2008042156 and US 20130185823). In some embodiments, nanoparticles loaded with or packed with DNA molecule(s) encoding a genetic modifying agent or system, or component(s) thereof can be delivered to the plant cell. A demonstrative and non-limiting example with respect to a CRISPR-Cas system is provided in WO2015089419 and can be applied or adapted for use with the present invention and/or other gene modifying agents and systems.

[0433] In some embodiments, the polynucleotide modifying agent(s) or one or more components of the system to the plant cell is by using cell penetrating peptides (CPP). In some embodiments, the cell penetrating peptide can be linked to a protein polynucleotide modifying agent or other component of a polynucleotide modifying agent or system thereof.

[0434] In some embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. or components thereof of the present invention is/are coupled to one or more CPPs to effectively transport them inside plant protoplasts. See also Ramakrishna. Genome Res. 2014 Jun;24(6):1020-7 which demonstrates such an approach for Cas9 in human cells and can be applied and adapted for use with the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention. In other embodiments, the engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos of the present invention encoded by one or more circular or non-circular DNA molecule(s) which are coupled to one or more CPPs for plant protoplast delivery. The plant protoplasts can then regenerate to produce plant cells and further to plants.

[0435] CPPs are generally described as short peptides of fewer than 35 amino acids either derived from proteins or from chimeric sequences which are capable of transporting biomolecules across cell membrane in a receptor independent manner. CPP can be cationic peptides, peptides having hydrophobic sequences, amphipatic peptides, peptides having proline-rich and anti-microbial sequence, and chimeric or bipartite peptides (Pooga and Langel 2005). CPPs are able to penetrate biological membranes and as such trigger the movement of various biomolecules across cell membranes into the cytoplasm and to improve their

intracellular routing, and hence facilitate interaction of the biomolecule with the target. Examples of CPP include amongst others: Tat, a nuclear transcriptional activator protein required for viral replication by HIV type1, penetratin, Kaposi fibroblast growth factor (FGF) signal peptide sequence, integrin β3 signal peptide sequence; polyarginine peptide Args sequence, Guanine rich-molecular transporters, sweet arrow peptide, etc.

Making genetically modified non-transgenic plants

[0436] In particular embodiments, the systems and methods described herein are used to modify endogenous genes or to modify their expression without the permanent introduction into the genome of the plant of any foreign gene, including those encoding polynucleotide modifying agent(s) or components of a polynucleotide modifying system, so as to avoid the presence of foreign DNA in the genome of the plant. This can be of interest as the regulatory requirements for non-transgenic plants are less rigorous.

[0437] In particular embodiments, this can be achieved by transient expression of the genetically modifying system components. In particular embodiments, one or more of the systems components are expressed on one or more viral vectors, such as those of the present invention, which produce sufficient components of the system to consistently steadily ensure modification of a gene of interest according to a method described herein. In particular embodiments, transient expression of constructs is ensured in plant protoplasts and thus not integrated into the genome. The limited window of expression can be sufficient to allow the system to ensure modification of the target gene(s) as described herein.

[0438] In some embodiments, different components of the system are introduced in the plant cell, protoplast or plant tissue either separately or in mixture, with the aid of particulate delivery molecules such as nanoparticles or CPP molecules, or delivery vesicles of the present invention.

[0439] The engineered compositions, systems, formulations, delivery vesicles, etc. and/or cargos delivered therefrom of the present invention can induce modification of the genome, either by direct activity of a polynucleotide modifying agent (e.g., when a CRISPR-Cas system is employed, a Cas protein) that can be incorporated into the engineered compositions, systems, and delivery vesicles of the present invention, and optionally introduction of template DNA or by modification of genes targeted using the system as described herein. The different strategies described herein above can allow targeted genome editing without requiring the introduction

of the components into the plant genome. Components which are transiently introduced into the plant cell can be, in some embodiments, removed upon crossing.

[0440] Protocols for targeted plant genome editing via CRISPR-Cas are also available based on those disclosed for the CRISPR-Cas9 system in volume 1284 of the series Methods in Molecular Biology pp 239-255 10 February 2015. A detailed procedure to design, construct, and evaluate dual gRNAs for plant codon optimized Cas9 (pcoCas9) mediated genome editing using Arabidopsis thaliana and Nicotiana benthamiana protoplasts s model cellular systems are described. Strategies to apply the CRISPR-Cas9 system to generating targeted genome modifications in whole plants are also discussed. The protocols described in the chapter can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0441] Sugano et al. (Plant Cell Physiol. 2014 Mar;55(3):475-81. doi: 10.1093/pcp/pcu014. Epub 2014 Jan 18) reports the application of CRISPR-Cas9 to targeted mutagenesis in the liverwort Marchantia polymorpha L., which has emerged as a model species for studying land plant evolution. The U6 promoter of M. polymorpha was identified and cloned to express the gRNA. The target sequence of the gRNA was designed to disrupt the gene encoding auxin response factor 1 (ARF1) in M. polymorpha. Using Agrobacterium-mediated transformation, Sugano et al. isolated stable mutants in the gametophyte generation of M. polymorpha. CRISPR-Cas9-based site-directed mutagenesis in vivo was achieved using either the Cauliflower mosaic virus 35S or M. polymorpha EF1α promoter to express Cas9. Isolated mutant individuals showing an auxin-resistant phenotype were not chimeric. Moreover, stable mutants were produced by asexual reproduction of T1 plants. Multiple arf1 alleles were easily established using CRIPSR-Cas9-based targeted mutagenesis. The methods of Sugano et al. can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0442] Lowder et al. (Plant Physiol. 2015 Aug 21. pii: pp.00636.2015) also developed a CRISPR-Cas9 toolbox enables multiplex genome editing and transcriptional regulation of expressed, silenced or non-coding genes in plants. This toolbox provides researchers with a protocol and reagents to quickly and efficiently assemble functional CRISPR-Cas9 T-DNA constructs for monocots and dicots using Golden Gate and Gateway cloning methods. It comes with a full suite of capabilities, including multiplexed gene editing and transcriptional activation or repression of plant endogenous genes. T-DNA based transformation technology

is fundamental to modern plant biotechnology, genetics, molecular biology and physiology. As such, Applicants developed a method for the assembly of Cas (WT, nickase or dCas) and gRNA(s) into a T-DNA destination-vector of interest. The assembly method is based on both Golden Gate assembly and MultiSite Gateway recombination. Three modules are required for assembly. The first module is a Cas entry vector, which contains promoterless Cas or its derivative genes flanked by attL1 and attR5 sites. The second module is a gRNA entry vector which contains entry gRNA expression cassettes flanked by attL5 and attL2 sites. The third module includes attR1-attR2-containing destination T-DNA vectors that provide promoters of choice for Cas expression. The toolbox of Lowder et al. can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0443] Wang et al. (bioRxiv 051342; doi: doi.org/10.1101/051342; Epub. May 12, 2016) demonstrate editing of homoeologous copies of four genes affecting important agronomic traits in hexaploid wheat using a multiplexed gene editing construct with several gRNA-tRNA units under the control of a single promoter. The methods of Wang et al. can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0444] In some embodiments, the plant is a tree. In some ebodiemtns, the plant is a herbaceous systems (see, e.g., Belhaj et al., Plant Methods 9: 39 and Harrison et al., Genes & Development 28: 1859–1872). The engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention can target single nucleotide polymorphisms (SNPs) in trees (see, e.g., Zhou et al., New Phytologist, Volume 208, Issue 2, pages 298–301, October 2015). In the Zhou et al. study, the authors applied a systems in the woody perennial Populus using the 4-coumarate:CoA ligase (4CL) gene family as a case study and achieved 100% mutational efficiency for two 4CL genes targeted, with every transformant examined carrying biallelic modifications. In the Zhou et al., study, the CRISPR-Cas9 system was highly sensitive to single nucleotide polymorphisms (SNPs), as cleavage for a third 4CL gene was abolished due to SNPs in the target sequence. These methods Wang et al. (bioRxiv 051342; doi: doi.org/10.1101/051342; Epub. May 12, 2016) demonstrate editing of homoeologous copies of four genes affecting important agronomic traits in hexaploid wheat using a multiplexed gene editing construct with several gRNA-tRNA units under the control of a single promoter. These techniques and methods can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0445] In particular embodiments, the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention, can be configured for self-cleavage or a cargo such as a polynucleotide modifying agent or system. In these embodiments, the promotor of the genetic modifying agent effector enzyme (e.g., an RNA guided nuclease) and any system component (e.g., gRNA) present can be a constitutive promotor and a second system component (e.g., a second gRNA or other nuclease (such as a restriction endonuclease)) can be introduced in the same transformation cassette, but controlled by an inducible promoter. This second system component (e.g., a second gRNA) can be designated to induce site-specific cleavage in the modifying agent effector gene in order to create a non-functional modifying agent. In some embodiments, the second system component induces cleavage on both ends of the transformation cassette, resulting in the removal of the cassette from the host genome. This configuration can provide a controlled duration of cellular exposure to the modifying agent effector and can further minimize off-target editing or other modification. Furthermore, in the context of a CRISPR-Cas system, cleavage of both ends of a CRISPR/Cas cassette can be used to generate transgene-free T0 plants with bi-allelic mutations (as described for Cas9 e.g. Moore et al., Nucleic Acids Research, 2014; Schaeffer et al., Plant Science, 2015). The methods of Moore et al. can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0446] Kabadi et al. (Nucleic Acids Res. 2014 Oct 29;42(19):e147. doi: 10.1093/nar/gku749. Epub 2014 Aug 13) developed a single lentiviral system to express a Cas9 variant, a reporter gene and up to four sgRNAs from independent RNA polymerase III promoters that are incorporated into the vector by a convenient Golden Gate cloning method. Each sgRNA was efficiently expressed and can mediate multiplex gene editing and sustained transcriptional activation in immortalized and primary human cells. The methods of Kabadi et al. may be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0447] Ling et al. (BMC Plant Biology 2014, 14:327) developed a CRISPR-Cas9 binary vector set based on the pGreen or pCAMBIA backbone, as well as a gRNA This toolkit requires no restriction enzymes besides BsaI to generate final constructs harboring maize-codon optimized Cas9 and one or more gRNAs with high efficiency in as little as one cloning step. The toolkit was validated using maize protoplasts, transgenic maize lines, and transgenic Arabidopsis lines and was shown to exhibit high efficiency and specificity. More importantly,

using this toolkit, targeted mutations of three Arabidopsis genes were detected in transgenic seedlings of the T1 generation. Moreover, the multiple-gene mutations could be inherited by the next generation. (guide RNA) module vector set, as a toolkit for multiplex genome editing in plants. The toolbox of Lin et al. can be applied to engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0448] The methods of Zhou et al. (New Phytologist, Volume 208, Issue 2, pages 298–301, October 2015) may be applied to the present invention as follows. Two 4CL genes, 4CL1 and 4CL2, associated with lignin and flavonoid biosynthesis, respectively are targeted for CRISPR-Cas9 editing. The Populus tremula × alba clone 717-1B4 routinely used for transformation is divergent from the genome-sequenced Populus trichocarpa. Therefore, the 4CL1 and 4CL2 gRNAs designed from the reference genome are interrogated with in-house 717 RNA-Seq data to ensure the absence of SNPs which could limit Cas efficiency. A third gRNA designed for 4CL5, a genome duplicate of 4CL1, is also included. The corresponding 717 sequence harbors one SNP in each allele near/within the PAM, both of which are expected to abolish targeting by the 4CL5-gRNA. All three gRNA target sites are located within the first exon. For 717 transformation, the gRNA is expressed from the Medicago U6.6 promoter, along with a human codon-optimized Cas under control of the CaMV 35S promoter in a binary vector. Transformation with the Cas-only vector can serve as a control. Randomly selected 4CL1 and 4CL2 lines are subjected to amplicon-sequencing. The data is then processed and biallelic mutations are confirmed in all cases. These methods can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0449] The disclosure of International Patent Publications: WO2016/099887, WO2016/025131, WO2016/073433, WO2017/066175, WO2017/100158, WO 2017/105991, WO2017/106414, WO2016/100272, WO2016/100571, WO 2016/100568, WO 2016/100562, WO 2017/019867 can be applied and adapted for use with the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

Detecting modifications in the plant genome- selectable markers

[0450] In particular embodiments, a selectable marker can be included or introduced to allow for identification of modified cells. Selectable markers can be advantageous for many situations, such as when the modification is made to an endogenous target gene of the plant genome. Any suitable method can be used to determine, after the plant, plant part or plant cell

is infected or transfected with the system, whether gene targeting or targeted mutagenesis has occurred at the target site.

[0451] Where the method involves introduction of a transgene, a transformed plant cell, callus, tissue or plant may be identified and isolated by selecting or screening the engineered plant material for the presence of the transgene or for traits encoded by the transgene. Physical and biochemical methods may be used to identify plant or plant cell transformants containing inserted gene constructs or an endogenous DNA modification. These methods include but are not limited to: 1) Southern analysis or PCR amplification for detecting and determining the structure of the recombinant DNA insert or modified endogenous genes; 2) Northern blot, S1 RNase protection, primer-extension or reverse transcriptase-PCR amplification for detecting and examining RNA transcripts of the gene constructs; 3) enzymatic assays for detecting enzyme or ribozyme activity, where such gene products are encoded by the gene construct or expression is affected by the genetic modification; 4) protein gel electrophoresis, Western blot techniques, immunoprecipitation, or enzyme-linked immunoassays, where the gene construct or endogenous gene products are proteins. Additional techniques, such as in situ hybridization, enzyme staining, and immunostaining, also may be used to detect the presence or expression of the recombinant construct or detect a modification of endogenous gene in specific plant organs and tissues. The methods for doing all these assays are well known to those skilled in the art.

[0452] In some embodiments, the expression system encoding the polynucleotide modifying agent and/or system components can be designed to comprise one or more selectable or detectable markers that provide a means to isolate or efficiently select cells that contain and/or have been modified by the system at an early stage and on a large scale.

[0453] In the case of Agrobacterium-mediated transformation, the marker cassette may be adjacent to or between flanking T-DNA borders and contained within a binary vector. In another embodiment, the marker cassette may be outside of the T-DNA. A selectable marker cassette may also be within or adjacent to the same T-DNA borders as the expression cassette or may be somewhere else within a second T-DNA on the binary vector (e.g., a 2 T-DNA system).

[0454] For particle bombardment or with protoplast transformation, the expression system can include one or more isolated linear fragments or may be part of a larger construct that might contain bacterial replication elements, bacterial selectable markers or other detectable

elements. The expression cassette(s) comprising the polynucleotide(s) encoding the polynucleotide modifying agents(s), system component(s), or system can be physically linked to a marker cassette or may be mixed with a second nucleic acid molecule encoding a marker cassette. The marker cassette can include the necessary elements to express a detectable or selectable marker that allows for efficient selection of transformed cells. Such elements will be appreciated by one of ordinary skill in the art.

[0455] The selection procedure for the cells based on the selectable marker will depend on the nature of the marker gene. In particular embodiments, use is made of a selectable marker, i.e., a marker which allows a direct selection of the cells based on the expression of the marker. A selectable marker can confer positive or negative selection and is conditional or non-conditional on the presence of external substrates (Miki et al. 2004, 107(3): 193–232). Most commonly, antibiotic or herbicide resistance genes are used as a marker, whereby selection is be performed by growing the engineered plant material on media containing an inhibitory amount of the antibiotic or herbicide to which the marker gene confers resistance. Examples of such genes are genes that confer resistance to antibiotics, such as hygromycin (hpt) and kanamycin (nptII), and genes that confer resistance to herbicides, such as phosphinothricin (bar) and chlorosulfuron (als),

[0456] Transformed plants and plant cells can also be identified by screening for the activities of a visible marker, typically an enzyme capable of processing a colored substrate (e.g., the β-glucuronidase, luciferase, B or C1 genes). Such selection and screening methodologies are well known to those skilled in the art.

Plant cultures and regeneration

[0457] In particular embodiments, plant cells which have a modified genome and that are produced or obtained by any of the methods described herein, can be cultured to regenerate a whole plant which possesses the transformed or modified genotype and thus the desired phenotype. Conventional regeneration techniques are well known to those skilled in the art. Particular examples of such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, and typically relying on a biocide and/or herbicide marker which has been introduced together with the desired nucleotide sequences. In further particular embodiments, plant regeneration is obtained from cultured protoplasts, plant callus, explants, organs, pollens, embryos or parts thereof (see e.g. Evans et al. (1983), Handbook of Plant Cell Culture, Klee et al (1987) Ann. Rev. of Plant Phys.).

[0458] In particular embodiments, transformed or improved plants as described herein can be self-pollinated to provide seed for homozygous improved plants of the invention (homozygous for the DNA modification) or crossed with non-transgenic plants or different improved plants to provide seed for heterozygous plants. Where a recombinant DNA was introduced into the plant cell, the resulting plant of such a crossing is a plant which is heterozygous for the recombinant DNA molecule. Both such homozygous and heterozygous plants obtained by crossing from the improved plants and comprising the genetic modification (which can be a recombinant DNA) are referred to herein as "progeny". Progeny plants are plants descended from the original transgenic plant and containing the genome modification or recombinant DNA molecule introduced by the methods provided herein. Alternatively, genetically modified plants can be obtained by one of the methods described supra using the Cfp1 enzyme whereby no foreign DNA is incorporated into the genome. Progeny of such plants obtained by further breeding may also contain the genetic modification. Breedings are performed by any breeding methods that are commonly used for different crops (e.g., Allard, Principles of Plant Breeding, John Wiley & Sons, NY, U. of CA, Davis, CA, 50-98 (1960).

Non-human Animals

[0459] The systems and methods may be used to generate modified non-human animals and cells thereof. In an aspect, the invention provides a non-human eukaryotic organism; preferably a multicellular eukaryotic organism, comprising a eukaryotic host cell according to any of the described embodiments. In other aspects, the invention provides a eukaryotic organism; preferably a multicellular eukaryotic organism, comprising a eukaryotic host cell according to any of the described embodiments. The organism in some embodiments of these aspects may be an animal; for example, a mammal. Also, the organism may be an arthropod such as an insect. The present invention may also be extended to other agricultural applications such as, for example, farm and production animals. For example, pigs have many features that make them attractive as biomedical models, especially in regenerative medicine. In particular, pigs with severe combined immunodeficiency (SCID) may provide useful models for regenerative medicine, xenotransplantation (discussed also elsewhere herein), and tumor development and will aid in developing therapies for human SCID patients. Lee et al., (Proc Natl Acad Sci U S A. 2014 May 20;111(20):7260-5) utilized a reporter-guided transcription activator-like effector nuclease (TALEN) system to generated targeted modifications of recombination activating gene (RAG) 2 in somatic cells at high efficiency, including some that

affected both alleles. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified non-human animal or cell thereof.

[0460] The methods of Lee et al., (Proc Natl Acad Sci U S A. 2014 May 20;111(20):7260-5) may be applied to the present invention analogously as follows. Mutated pigs are produced by targeted insertion for example in RAG2 in fetal fibroblast cells followed by SCNT and embryo transfer. Constructs coding for CRISPR Cas and a reporter are electroporated into fetal-derived fibroblast cells. After 48 h, transfected cells expressing the green fluorescent protein are sorted into individual wells of a 96-well plate at an estimated dilution of a single cell per well. Targeted modification of RAG2 are screened by amplifying a genomic DNA fragment flanking any CRISPR Cas cutting sites followed by sequencing the PCR products. After screening and ensuring lack of off-site mutations, cells carrying targeted modification of RAG2 are used for SCNT. The polar body, along with a portion of the adjacent cytoplasm of oocyte, presumably containing the metaphase II plate, are removed, and a donor cell are placed in the perivitelline. The reconstructed embryos are then electrically porated to fuse the donor cell with the oocyte and then chemically activated. The activated embryos are incubated in Porcine Zygote Medium 3 (PZM3) with 0.5 μM Scriptaid (S7817; Sigma-Aldrich) for 14–16 h. Embryos are then washed to remove the Scriptaid and cultured in PZM3 until they were transferred into the oviducts of surrogate pigs. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified non-human animal or cell thereof.

[0461] SNPs of other animals, such as cows can also be modified or generated using one or more polynucleotide modifying agents or systems described herein. Tan et al. (Proc Natl Acad Sci U S A. 2013 Oct 8; 110(41): 16526–16531) expanded the livestock gene editing toolbox to include transcription activator-like (TAL) effector nuclease (TALEN)- and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9- stimulated homology-directed repair (HDR) using plasmid, rAAV, and oligonucleotide templates. Gene specific gRNA sequences were cloned into the Church lab gRNA vector (Addgene ID: 41824) according to their methods (Mali P, et al. (2013) RNA-Guided Human Genome Engineering via Cas9. Science 339(6121):823-826). The Cas9 nuclease was provided either by co-transfection of the hCas9 plasmid (Addgene ID: 41815) or mRNA synthesized from RCIScript-hCas9. This RCIScript-hCas9 was constructed by sub-cloning the XbaI-AgeI fragment from

the hCas9 plasmid (encompassing the hCas9 cDNA) into the RCIScript plasmid. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified non-human animal or cell thereof.

[0462] Heo et al. (Stem Cells Dev. 2015 Feb 1;24(3):393-402. doi: 10.1089/scd.2014.0278. Epub 2014 Nov 3) reported highly efficient gene targeting in the bovine genome using bovine pluripotent cells and clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 nuclease. First, Heo et al. generate induced pluripotent stem cells (iPSCs) from bovine somatic fibroblasts by the ectopic expression of yamanaka factors and GSK3β and MEK inhibitor (2i) treatment. Heo et al. observed that these bovine iPSCs are highly similar to naïve pluripotent stem cells with regard to gene expression and developmental potential in teratomas. Moreover, CRISPR-Cas9 nuclease, which was specific for the bovine NANOG locus, showed highly efficient editing of the bovine genome in bovine iPSCs and embryos. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified non-human animal or cell thereof.

[0463] Igenity® provides a profile analysis of animals, such as cows, to perform and transmit traits of economic traits of economic importance, such as carcass composition, carcass quality, maternal and reproductive traits and average daily gain. The analysis of a comprehensive Igenity® profile begins with the discovery of DNA markers (most often single nucleotide polymorphisms or SNPs). All the markers behind the Igenity® profile were discovered by independent scientists at research institutions, including universities, research organizations, and government entities such as USDA. Markers are then analyzed at Igenity® in validation populations. Igenity® uses multiple resource populations that represent various production environments and biological types, often working with industry partners from the seedstock, cow-calf, feedlot and/or packing segments of the beef industry to collect phenotypes that are not commonly available. Cattle genome databases are widely available, see, e.g., the NAGRP Cattle Genome Coordination Program (www.animalgenome.org/cattle/maps/db.html). Thus, the polynucleotide modifying agent(s) and/or systems described herein can be applied to target bovine SNPs. One of skill in the art may utilize the above protocols for targeting SNPs and apply them to bovine SNPs as described, for example, by Tan et al. or Heo et al.

[0464] Qingjian Zou et al. (Journal of Molecular Cell Biology Advance Access published October 12, 2015) demonstrated increased muscle mass in dogs by targeting the first exon of

the dog Myostatin (MSTN) gene (a negative regulator of skeletal muscle mass). First, the efficiency of the sgRNA was validated, using cotransfection of the sgRNA targeting MSTN with a Cas9 vector into canine embryonic fibroblasts (CEFs). Thereafter, MSTN KO dogs were generated by micro-injecting embryos with normal morphology with a mixture of Cas9 mRNA and MSTN sgRNA and auto-transplantation of the zygotes into the oviduct of the same female dog. The knock-out puppies displayed an obvious muscular phenotype on thighs compared with its wild-type littermate sister. This can also be performed using the polynucleotide agent(s) and/or systems provided herein. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified non-human animal or cell thereof.

Livestock

[0465] Also described herein are modified pigs or cells that can express one or more polynucleotides, genes or alleles of interest. As reported by Kristin M Whitworth and Dr Randall Prather et al. (Nature Biotech 3434 published online 07 December 2015) CD163 (a viral target) was targeted using CRISPR-Cas9 and the offspring of edited pigs were resistant when exposed to PRRSv. One founder male and one founder female, both of whom had mutations in exon 7 of CD163, were bred to produce offspring. The founder male possessed an 11-bp deletion in exon 7 on one allele, which results in a frameshift mutation and missense translation at amino acid 45 in domain 5 and a subsequent premature stop codon at amino acid 64. The other allele had a 2-bp addition in exon 7 and a 377-bp deletion in the preceding intron, which were predicted to result in the expression of the first 49 amino acids of domain 5, followed by a premature stop code at amino acid 85. The sow had a 7 bp addition in one allele that when translated was predicted to express the first 48 amino acids of domain 5, followed by a premature stop codon at amino acid 70. The sow's other allele was unamplifiable. Selected offspring were predicted to be a null animal (CD163–/–), i.e. a CD163 knock out. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified pig that can express a polynucleotide of interest. Thus, also described herein are modified pigs their progeny that also express one or more copies of the gene or allele of interest. This may be for livestock, breeding or modelling purposes (i.e. a porcine model). Semen comprising the modification (e.g. polynucleotide of interest) is also provided.

Other Animals

[0466] Also described herein are other non-human animals that are modified to express one or more polynucleotides, genes or alleles of interest. Suitable polynucleotide modifying agent(s) and/or system thereof described elsewhere herein can be used to generate other non-human animals such as non-human primates, chickens (reviewed in Sid and Schusser et al 2018. Front. Genet. Doi.org/10.3389/fgene.2018.00456) and other avians (e.g. Scott et al. 2010. ILAR J. 51(4):353-361), cattle (Yum et al., 2016. Scientific Reports. 6:27185 and Tait-Burkard et al. 2018. Genome Biology. 19:2014.), sheep and goats (see e.g. Kalds et al., 2019. Front. Genet. Doi.org//10.3389/fgene.2019.00750), horses (see e.g. West and Gill. 2016. J. Equine Vet. Sci. 41:1-6), dogs (see e.g. D. Duan. Nature Biomedical Engineereing. 2018. 2: 795-796), reptiles (see e.g. Rasys et al. 2019. Cell Reports. 28:2288-2292), fish (including but not limited to zebrafish, see e.g. Datsomor et al. 2019. Scientific Reports. 9:7533, Liu et al. 2019. Front. Cell. Dev. Biol. https://doi.org/10.3389/fcell.2019.00013), insects (see e.g. Kotwica-Rolinska et al. 2019. Front. Physiol. https://doi.org/10.3389/fphys.2019.00891; Gantz and Akbari. 2018. Curr. Opin. Insect. Sci. 28:66-72), rabbits (see e.g. Kawano and Honda. 2017. Methods Mol. Biol. 4630:109-120; Liu et al., 2018. Nature Commun. 9:2717; and Liu et al. 2018. Gene. https://doi.org/10.1016/j.gene.2018.01.044), mice (see e.g. Hall et al. 2018. Curr Protoc Cell Biol. 81(1): e57), rats (see e.g. Back et al. 2019. Neuron. 102(1):105-119), amphibians (see e.g. Nakayama et al. 2013. Genesis. 51(12):835-843), nematodes (see e.g. J.B. Lok. 2019. Front. Genet. https://doi.org/10.3389/fgene.2019.00656), molluscs (see e.g. Abe and Kuroda. 2019. Development. 146: dev175976 doi: 10.1242/dev.175976, geckos, shrimp and other crustaceans (see e.g. Gui et al. Genes Genomes Genetics: 6(11): 3757-3764), oysters (Yu et al. 2019; Mar. Biotechnol (NY) 21(3):301-309. doi: 10.1007/s10126-019-09885-y), and sponges (see e.g. Revilla-i-Domingo et al. 2018. Genetics. 210(2)435-443), the teachings of which can be adapted for use with one or more of the modifying agent(s) and/or systems described herein to generate the modified non-human animal or cell thereof.

FORMULATIONS

[0467] Also described herein are pharmaceutical formulations that can contain an amount, effective amount, and/or least effective amount, and/or therapeutically effective amount of one or more compounds, molecules, compositions, vectors, vector systems, cells, or a combination thereof (which are also referred to as the primary active agent or ingredient elsewhere herein) described in greater detail elsewhere herein a pharmaceutically acceptable carrier or excipient.

As used herein, "pharmaceutical formulation" refers to the combination of an active agent, compound, or ingredient with a pharmaceutically acceptable carrier or excipient, making the composition suitable for diagnostic, therapeutic, or preventive use in vitro, in vivo, or ex vivo. As used herein, "pharmaceutically acceptable carrier or excipient" refers to a carrier or excipient that is useful in preparing a pharmaceutical formulation that is generally safe, non-toxic, and is neither biologically or otherwise undesirable, and includes a carrier or excipient that is acceptable for veterinary use as well as human pharmaceutical use. A "pharmaceutically acceptable carrier or excipient" as used in the specification and claims includes both one and more than one such carrier or excipient. When present, the compound can optionally be present in the pharmaceutical formulation as a pharmaceutically acceptable salt. In some embodiments, the pharmaceutical formulation can include, such as an active ingredient, an engineered delivery vesicle generation composition or system, or delivery vesicle generated therefrom, or cell containing the engineered composition, system, or vesicle described in greater detail elsewhere herein. In some embodiments, the pharmaceutical formulation can include, such as an active ingredient, an engineered delivery vesicle generation composition or system, or delivery vesicle generated therefrom, or cell containing the engineered composition, system, or vesicle described in greater detail elsewhere herein. In some embodiments, the pharmaceutical formulation can include, such as an active ingredient one or more modified cells, such as one or more modified cells described in greater detail elsewhere herein.

[0468] In some embodiments, the active ingredient is present as a pharmaceutically acceptable salt of the active ingredient. As used herein, "pharmaceutically acceptable salt" refers to any acid or base addition salt whose counter-ions are non-toxic to the subject to which they are administered in pharmaceutical doses of the salts. Suitable salts include, hydrobromide, iodide, nitrate, bisulfate, phosphate, isonicotinate, lactate, salicylate, acid citrate, tartrate, oleate, tannate, pantothenate, bitartrate, ascorbate, succinate, maleate, gentisinate, fumarate, gluconate, glucaronate, saccharate, formate, benzoate, glutamate, methanesulfonate, ethanesulfonate, benzenesulfonate, p-toluenesulfonate, camphorsulfonate, napthalenesulfonate, propionate, malonate, mandelate, malate, phthalate, and pamoate.

[0469] The pharmaceutical formulations described herein can be administered to a subject in need thereof via any suitable method or route to a subject in need thereof. Suitable administration routes can include, but are not limited to auricular (otic), buccal, conjunctival, cutaneous, dental, electro-osmosis, endocervical, endosinusial, endotracheal, enteral, epidural,

extra-amniotic, extracorporeal, hemodialysis, infiltration, interstitial, intra-abdominal, intra-amniotic, intra-arterial, intra-articular, intrabiliary, intrabronchial, intrabursal, intracardiac, intracartilaginous, intracaudal, intracavernous, intracavitary, intracerebral, intracisternal, intracorneal, intracoronal (dental), intracoronary, intracorporus cavernosum, intradermal, intradiscal, intraductal, intraduodenal, intradural, intraepidermal, intraesophageal, intragastric, intragingival, intraileal, intralesional, intraluminal, intralymphatic, intramedullary, intrameningeal, intramuscular, intraocular, intraovarian, intrapericardial, intraperitoneal, intrapleural, intraprostatic, intrapulmonary, intrasinal, intraspinal, intrasynovial, intratendinous, intratesticular, intrathecal, intrathoracic, intratubular, intratumor, intratympanic, intrauterine, intravascular, intravenous, intravenous bolus, intravenous drip, intraventricular, intravesical, intravitreal, iontophoresis, irrigation, laryngeal, nasal, nasogastric, occlusive dressing technique, ophthalmic, oral, oropharyngeal, other, parenteral, percutaneous, periarticular, peridural, perineural, periodontal, rectal, respiratory (inhalation), retrobulbar, soft tissue, subarachnoid, subconjunctival, subcutaneous, sublingual, submucosal, topical, transdermal, transmucosal, transplacental, transtracheal, transtympanic, ureteral, urethral, and/or vaginal administration, and/or any combination of the above administration routes, which typically depends on the disease to be treated and/or the active ingredient(s).

[0470] Where appropriate, compounds, molecules, compositions, vectors, vector systems, cells, or a combination thereof described in greater detail elsewhere herein can be provided to a subject in need thereof as an ingredient, such as an active ingredient or agent, in a pharmaceutical formulation. As such, also described are pharmaceutical formulations containing one or more of the compounds and salts thereof, or pharmaceutically acceptable salts thereof described herein. Suitable salts include, hydrobromide, iodide, nitrate, bisulfate, phosphate, isonicotinate, lactate, salicylate, acid citrate, tartrate, oleate, tannate, pantothenate, bitartrate, ascorbate, succinate, maleate, gentisinate, fumarate, gluconate, glucaronate, saccharate, formate, benzoate, glutamate, methanesulfonate, ethanesulfonate, benzenesulfonate, p-toluenesulfonate, camphorsulfonate, napthalenesulfonate, propionate, malonate, mandelate, malate, phthalate, and pamoate.

[0471] In some embodiments, the subject in need thereof has or is suspected of having a hematopoietic disease or a symptom thereof. In some embodiments, the subject in need thereof has or is suspected of having, a neurobiological disease or disorder, a brain disease or disorder, a musculoskeletal disease or disorder, a soft tissue disease or disorder, a psychiatric disease or

disorder, a cancer, an autoimmune disease or disorder, a thrombosis disease, a heart disease, a kidney disease, a lung disease, or a blood vessel disease, or a combination thereof. As used herein, "agent" refers to any substance, compound, molecule, and the like, which can be biologically active or otherwise can induce a biological and/or physiological effect on a subject to which it is administered to. As used herein, "active agent" or "active ingredient" refers to a substance, compound, or molecule, which is biologically active or otherwise, induces a biological or physiological effect on a subject to which it is administered to. In other words, "active agent" or "active ingredient" refers to a component or components of a composition to which the whole or part of the effect of the composition is attributed. An agent can be a primary active agent, or in other words, the component(s) of a composition to which the whole or part of the effect of the composition is attributed. An agent can be a secondary agent, or in other words, the component(s) of a composition to which an additional part and/or other effect of the composition is attributed.

Pharmaceutically Acceptable Carriers and Secondary Ingredients and Agents

[0472] The pharmaceutical formulation can include a pharmaceutically acceptable carrier. Suitable pharmaceutically acceptable carriers include, but are not limited to water, salt solutions, alcohols, gum arabic, vegetable oils, benzyl alcohols, polyethylene glycols, gelatin, carbohydrates such as lactose, amylose or starch, magnesium stearate, talc, silicic acid, viscous paraffin, perfume oil, fatty acid esters, hydroxy methylcellulose, and polyvinyl pyrrolidone, which do not deleteriously react with the active composition.

[0473] The pharmaceutical formulations can be sterilized, and if desired, mixed with agents, such as lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, flavoring and/or aromatic substances, and the like which do not deleteriously react with the active compound.

[0474] In some embodiments, the pharmaceutical formulation can also include an effective amount of secondary active agents, including but not limited to, biologic agents or molecules including, but not limited to, e.g. polynucleotides, amino acids, peptides, polypeptides, antibodies, aptamers, ribozymes, hormones, immunomodulators, antipyretics, anxiolytics, antipsychotics, analgesics, antispasmodics, anti-inflammatories, anti-histamines, anti-infectives, chemotherapeutics, and combinations thereof.

Effective Amounts

[0475] In some embodiments, the amount of the primary active agent and/or optional secondary agent can be an effective amount, least effective amount, and/or therapeutically effective amount. As used herein, "effective amount" refers to the amount of the primary and/or optional secondary agent included in the pharmaceutical formulation that achieve one or more therapeutic effects or desired effect. As used herein, "least effective" amount refers to the lowest amount of the primary and/or optional secondary agent that achieves the one or more therapeutic or other desired effects. As used herein, "therapeutically effective amount" refers to the amount of the primary and/or optional secondary agent included in the pharmaceutical formulation that achieves one or more therapeutic effects.

[0476] The effective amount, least effective amount, and/or therapeutically effective amount of the primary and optional secondary active agent described elsewhere herein contained in the pharmaceutical formulation can range from about 0 to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 pg, ng, μg, mg, or g or be any numerical value with any of these ranges.

[0477] In some embodiments, the effective amount, least effective amount, and/or therapeutically effective amount can be an effective concentration, least effective concentration, and/or therapeutically effective concentration, which can each range from about 0 to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 pM, nM, μM, mM, or M or be any numerical value with any of these ranges.

[0478] In other embodiments, the effective amount, least effective amount, and/or therapeutically effective amount of the primary and optional secondary active agent can range from about 0 to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180,

190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000 IU or be any numerical value with any of these ranges.

[0479] In some embodiments, the primary and/or the optional secondary active agent present in the pharmaceutical formulation can range from about 0 to 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.2, 0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3, 0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5, 0.51, 0.52, 0.53, 0.54, 0.55, 0.56, 0.57, 0.58, 0.59, 0.6, 0.61, 0.62, 0.63, 0.64, 0.65, 0.66, 0.67, 0.68, 0.69, 0.7, 0.71, 0.72, 0.73, 0.74, 0.75, 0.76, 0.77, 0.78, 0.79, 0.8, 0.81, 0.82, 0.83, 0.84, 0.85, 0.86, 0.87, 0.88, 0.89, 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.9, to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 % w/w, v/v, or w/v of the pharmaceutical formulation.

[0480] In some embodiments where a cell population is present in the pharmaceutical formulation (e.g., as a primary and/or or secondary active agent), the effective amount of cells can range from about 2 cells to $1X10^{1}$/mL, $1X10^{20}$/mL or more, such as about $1X10^{1}$/mL, $1X10^{2}$/mL, $1X10^{3}$/mL, $1X10^{4}$/mL, $1X10^{5}$/mL, $1X10^{6}$/mL, $1X10^{7}$/mL, $1X10^{8}$/mL, $1X10^{9}$/mL, $1X10^{10}$/mL, $1X10^{11}$/mL, $1X10^{12}$/mL, $1X10^{13}$/mL, $1X10^{14}$/mL, $1X10^{15}$/mL, $1X10^{16}$/mL, $1X10^{17}$/mL, $1X10^{18}$/mL, $1X10^{19}$/mL, to/or about $1X10^{20}$/mL.

[0481] In some embodiments, the amount or effective amount, particularly where an infective particle is being delivered (e.g. a virus particle having the primary or secondary agent as a cargo), the effective amount of virus particles can be expressed as a titer (plaque forming units per unit of volume) or as a MOI (multiplicity of infection). In some embodiments, the effective amount can be $1X10^{1}$ particles per pL, nL, μL, mL, or L to $1X10^{20}$/ particles per pL, nL, μL, mL, or L or more, such as about $1X10^{1}$, $1X10^{2}$, $1X10^{3}$, $1X10^{4}$, $1X10^{5}$, $1X10^{6}$, $1X10^{7}$, $1X10^{8}$, $1X10^{9}$, $1X10^{10}$, $1X10^{11}$, $1X10^{12}$, $1X10^{13}$, $1X10^{14}$, $1X10^{15}$, $1X10^{16}$, $1X10^{17}$, $1X10^{18}$,

$1X10^{19}$, to/or about $1X10^{20}$ particles per pL, nL, μL, mL, or L. In some embodiments, the effective titer can be about $1X10^{1}$ transforming units per pL, nL, μL, mL, or L to $1X10^{20}$/ transforming units per pL, nL, μL, mL, or L or more, such as about $1X10^{1}$, $1X10^{2}$, $1X10^{3}$, $1X10^{4}$, $1X10^{5}$, $1X10^{6}$, $1X10^{7}$, $1X10^{8}$, $1X10^{9}$, $1X10^{10}$, $1X10^{11}$, $1X10^{12}$, $1X10^{13}$, $1X10^{14}$, $1X10^{15}$, $1X10^{16}$, $1X10^{17}$, $1X10^{18}$, $1X10^{19}$, to/or about $1X10^{20}$ transforming units per pL, nL, μL, mL, or L. In some embodiments, the MOI of the pharmaceutical formulation can range from about 0.1 to 10 or more, such as 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9, 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10 or more.

[0482] In some embodiments, the amount or effective amount of the one or more of the active agent(s) described herein contained in the pharmaceutical formulation can range from about 1 pg/kg to about 10 mg/kg based upon the bodyweight of the subject in need thereof or average bodyweight of the specific patient population to which the pharmaceutical formulation can be administered.

[0483] In embodiments where there is a secondary agent contained in the pharmaceutical formulation, the effective amount of the secondary active agent will vary depending on the secondary agent, the primary agent, the administration route, subject age, disease, stage of disease, among other things, which will be one of ordinary skill in the art.

[0484] When optionally present in the pharmaceutical formulation, the secondary active agent can be included in the pharmaceutical formulation or can exist as a stand-alone compound or pharmaceutical formulation that can be administered contemporaneously or sequentially with the compound, derivative thereof, or pharmaceutical formulation thereof.

[0485] In some embodiments, the effective amount of the secondary active agent can range from about 0 to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 % w/w, v/v, or w/v of the total secondary active agent in the pharmaceutical formulation. In additional embodiments, the effective amount of the secondary active agent can range from about 0 to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,

12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9 % w/w, v/v, or w/v of the total pharmaceutical formulation.

Dosage Forms

[0486] In some embodiments, the pharmaceutical formulations described herein can be provided in a dosage form. The dosage form can be administered to a subject in need thereof. The dosage form can be effective generate specific concentration, such as an effective concentration, at a given site in the subject in need thereof. As used herein, "dose," "unit dose," or "dosage" can refer to physically discrete units suitable for use in a subject, each unit containing a predetermined quantity of the primary active agent, and optionally present secondary active ingredient, and/or a pharmaceutical formulation thereof calculated to produce the desired response or responses in association with its administration. In some embodiments, the given site is proximal to the administration site. In some embodiments, the given site is distal to the administration site. In some cases, the dosage form contains a greater amount of one or more of the active ingredients present in the pharmaceutical formulation than the final intended amount needed to reach a specific region or location within the subject to account for loss of the active components such as via first and second pass metabolism.

[0487] The dosage forms can be adapted for administration by any appropriate route. Appropriate routes include, but are not limited to, oral (including buccal or sublingual), rectal, intraocular, inhaled, intranasal, topical (including buccal, sublingual, or transdermal), vaginal, parenteral, subcutaneous, intramuscular, intravenous, internasal, and intradermal. Other appropriate routes are described elsewhere herein. Such formulations can be prepared by any method known in the art.

[0488] Dosage forms adapted for oral administration can discrete dosage units such as capsules, pellets or tablets, powders or granules, solutions, or suspensions in aqueous or non-aqueous liquids; edible foams or whips, or in oil-in-water liquid emulsions or water-in-oil liquid emulsions. In some embodiments, the pharmaceutical formulations adapted for oral administration also include one or more agents which flavor, preserve, color, or help disperse the pharmaceutical formulation. Dosage forms prepared for oral administration can also be in the form of a liquid solution that can be delivered as a foam, spray, or liquid solution. The oral

dosage form can be administered to a subject in need thereof. Where appropriate, the dosage forms described herein can be microencapsulated.

[0489] The dosage form can also be prepared to prolong or sustain the release of any ingredient. In some embodiments, compounds, molecules, compositions, vectors, vector systems, cells, or a combination thereof described herein can be the ingredient whose release is delayed. In some embodiments the primary active agent is the ingredient whose release is delayed. In some embodiments, an optional secondary agent can be the ingredient whose release is delayed. Suitable methods for delaying the release of an ingredient include, but are not limited to, coating or embedding the ingredients in material in polymers, wax, gels, and the like. Delayed release dosage formulations can be prepared as described in standard references such as "Pharmaceutical dosage form tablets," eds. Liberman et. al. (New York, Marcel Dekker, Inc., 1989), "Remington - The science and practice of pharmacy", 20th ed., Lippincott Williams & Wlkins, Baltimore, MD, 2000, and "Pharmaceutical dosage forms and drug delivery systems", 6th Edition, Ansel et al., (Media, PA: Wlliams and Wlkins, 1995). These references provide information on excipients, materials, equipment, and processes for preparing tablets and capsules and delayed release dosage forms of tablets and pellets, capsules, and granules. The delayed release can be anywhere from about an hour to about 3 months or more.

[0490] Examples of suitable coating materials include, but are not limited to, cellulose polymers such as cellulose acetate phthalate, hydroxypropyl cellulose, hydroxypropyl methylcellulose, hydroxypropyl methylcellulose phthalate, and hydroxypropyl methylcellulose acetate succinate; polyvinyl acetate phthalate, acrylic acid polymers and copolymers, and methacrylic resins that are commercially available under the trade name EUDRAGIT® (Roth Pharma, Westerstadt, Germany), zein, shellac, and polysaccharides.

[0491] Coatings may be formed with a different ratio of water-soluble polymer, water insoluble polymers, and/or pH dependent polymers, with or without water insoluble/water soluble non-polymeric excipient, to produce the desired release profile. The coating is either performed on the dosage form (matrix or simple) which includes, but is not limited to, tablets (compressed with or without coated beads), capsules (with or without coated beads), beads, particle compositions, "ingredient as is" formulated as, but not limited to, suspension form or as a sprinkle dosage form.

[0492] Where appropriate, the dosage forms described herein can be a liposome. In these embodiments, primary active ingredient(s), and/or optional secondary active ingredient(s), and/or pharmaceutically acceptable salt thereof where appropriate are incorporated into a liposome. In embodiments where the dosage form is a liposome, the pharmaceutical formulation is thus a liposomal formulation. The liposomal formulation can be administered to a subject in need thereof.

[0493] Dosage forms adapted for topical administration can be formulated as ointments, creams, suspensions, lotions, powders, solutions, pastes, gels, sprays, aerosols, or oils. In some embodiments for treatments of the eye or other external tissues, for example the mouth or the skin, the pharmaceutical formulations are applied as a topical ointment or cream. When formulated in an ointment, a primary active ingredient, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate can be formulated with a paraffinic or water-miscible ointment base. In other embodiments, the primary and/or secondary active ingredient can be formulated in a cream with an oil-in-water cream base or a water-in-oil base. Dosage forms adapted for topical administration in the mouth include lozenges, pastilles, and mouth washes.

[0494] Dosage forms adapted for nasal or inhalation administration include aerosols, solutions, suspension drops, gels, or dry powders. In some embodiments, a primary active ingredient, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate can be in a dosage form adapted for inhalation is in a particle-size-reduced form that is obtained or obtainable by micronization. In some embodiments, the particle size of the size reduced (e.g. micronized) compound or salt or solvate thereof, is defined by a D_{50} value of about 0.5 to about 10 microns as measured by an appropriate method known in the art. Dosage forms adapted for administration by inhalation also include particle dusts or mists. Suitable dosage forms wherein the carrier or excipient is a liquid for administration as a nasal spray or drops include aqueous or oil solutions/suspensions of an active (primary and/or secondary) ingredient, which may be generated by various types of metered dose pressurized aerosols, nebulizers, or insufflators. The nasal/inhalation formulations can be administered to a subject in need thereof.

[0495] In some embodiments, the dosage forms are aerosol formulations suitable for administration by inhalation. In some of these embodiments, the aerosol formulation contains a solution or fine suspension of a primary active ingredient, secondary active ingredient, and/or

pharmaceutically acceptable salt thereof where appropriate and a pharmaceutically acceptable aqueous or non-aqueous solvent. Aerosol formulations can be presented in single or multi-dose quantities in sterile form in a sealed container. For some of these embodiments, the sealed container is a single dose or multi-dose nasal or an aerosol dispenser fitted with a metering valve (e.g. metered dose inhaler), which is intended for disposal once the contents of the container have been exhausted.

[0496] Where the aerosol dosage form is contained in an aerosol dispenser, the dispenser contains a suitable propellant under pressure, such as compressed air, carbon dioxide, or an organic propellant, including but not limited to a hydrofluorocarbon. The aerosol formulation dosage forms in other embodiments are contained in a pump-atomizer. The pressurized aerosol formulation can also contain a solution or a suspension of a primary active ingredient, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof. In further embodiments, the aerosol formulation also contains co-solvents and/or modifiers incorporated to improve, for example, the stability and/or taste and/or fine particle mass characteristics (amount and/or profile) of the formulation. Administration of the aerosol formulation can be once daily or several times daily, for example 2, 3, 4, or 8 times daily, in which 1, 2, 3 or more doses are delivered each time. The aerosol formulations can be administered to a subject in need thereof.

[0497] For some dosage forms suitable and/or adapted for inhaled administration, the pharmaceutical formulation is a dry powder inhalable-formulations. In addition to a primary active agent, optional secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate, such a dosage form can contain a powder base such as lactose, glucose, trehalose, manitol, and/or starch. In some of these embodiments, a primary active agent, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate is in a particle-size reduced form. In further embodiments, a performance modifier, such as L-leucine or another amino acid, cellobiose octaacetate, and/or metals salts of stearic acid, such as magnesium or calcium stearate. In some embodiments, the aerosol formulations are arranged so that each metered dose of aerosol contains a predetermined amount of an active ingredient, such as the one or more of the compositions, compounds, vector(s), molecules, cells, and combinations thereof described herein.

[0498] Dosage forms adapted for vaginal administration can be presented as pessaries, tampons, creams, gels, pastes, foams, or spray formulations. Dosage forms adapted for rectal

administration include suppositories or enemas. The vaginal formulations can be administered to a subject in need thereof.

[0499] Dosage forms adapted for parenteral administration and/or adapted for injection can include aqueous and/or non-aqueous sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, solutes that render the composition isotonic with the blood of the subject, and aqueous and non-aqueous sterile suspensions, which can include suspending agents and thickening agents. The dosage forms adapted for parenteral administration can be presented in a single-unit dose or multi-unit dose containers, including but not limited to sealed ampoules or vials. The doses can be lyophilized and re-suspended in a sterile carrier to reconstitute the dose prior to administration. Extemporaneous injection solutions and suspensions can be prepared in some embodiments, from sterile powders, granules, and tablets. The parenteral formulations can be administered to a subject in need thereof.

[0500] For some embodiments, the dosage form contains a predetermined amount of a primary active agent, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate per unit dose. In an embodiment, the predetermined amount of primary active agent, secondary active ingredient, and/or pharmaceutically acceptable salt thereof where appropriate can be an effective amount, a least effect amount, and/or a therapeutically effective amount. In other embodiments, the predetermined amount of a primary active agent, secondary active agent, and/or pharmaceutically acceptable salt thereof where appropriate, can be an appropriate fraction of the effective amount of the active ingredient.

Co-Therapies and Combination Therapies

[0501] In some embodiments, the pharmaceutical formulation(s) described herein can be part of a combination treatment or combination therapy. The combination treatment can include the pharmaceutical formulation described herein and an additional treatment modality. The additional treatment modality can be a chemotherapeutic, a biological therapeutic, surgery, radiation, diet modulation, environmental modulation, a physical activity modulation, and combinations thereof.

[0502] In some embodiments, the co-therapy or combination therapy can additionally include but not limited to, polynucleotides, amino acids, peptides, polypeptides, antibodies, aptamers, ribozymes, hormones, immunomodulators, antipyretics, anxiolytics, antipsychotics,

analgesics, antispasmodics, anti-inflammatories, anti-histamines, anti-infectives, chemotherapeutics, and combinations thereof.

Administration of the Pharmaceutical Formulations

[0503] The pharmaceutical formulations or dosage forms thereof described herein can be administered one or more times hourly, daily, monthly, or yearly (e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more times hourly, daily, monthly, or yearly). In some embodiments, the pharmaceutical formulations or dosage forms thereof described herein can be administered continuously over a period of time ranging from minutes to hours to days. Devices and dosages forms are known in the art and described herein that are effective to provide continuous administration of the pharmaceutical formulations described herein. In some embodiments, the first one or a few initial amount(s) administered can be a higher dose than subsequent doses. This is typically referred to in the art as a loading dose or doses and a maintenance dose, respectively. In some embodiments, the pharmaceutical formulations can be administered such that the doses over time are tapered (increased or decreased) overtime so as to wean a subject gradually off of a pharmaceutical formulation or gradually introduce a subject to the pharmaceutical formulation.

[0504] As previously discussed, the pharmaceutical formulation can contain a predetermined amount of a primary active agent, secondary active agent, and/or pharmaceutically acceptable salt thereof where appropriate. In some of these embodiments, the predetermined amount can be an appropriate fraction of the effective amount of the active ingredient. Such unit doses may therefore be administered once or more than once a day, month, or year (e.g. 1, 2, 3, 4, 5, 6, or more times per day, month, or year). Such pharmaceutical formulations may be prepared by any of the methods well known in the art.

[0505] Where co-therapies or multiple pharmaceutical formulations are to be delivered to a subject, the different therapies or formulations can be administered sequentially or simultaneously. Sequential administration is administration where an appreciable amount of time occurs between administrations, such as more than about 15, 20, 30, 45, 60 minutes or more. The time between administrations in sequential administration can be on the order of hours, days, months, or even years, depending on the active agent present in each administration. Simultaneous administration refers to administration of two or more formulations at the same time or substantially at the same time (e.g. within seconds or just a

few minutes apart), where the intent is that the formulations be administered together at the same time.

KITS

[0506] Any of the engineered compositions, formulations, systems particles, vesicles, cells, of the present invention or a combination thereof can be presented as a combination kit. As used herein, the terms "combination kit" or "kit of parts" refers to the compounds, compositions, formulations, particles, cells and any additional components that are used to package, sell, market, deliver, and/or administer the combination of elements or a single element, such as the active ingredient, contained therein. Such additional components include, but are not limited to, packaging, syringes, blister packages, bottles, and the like. When one or more of the engineered compositions, formulations, systems particles, vesicles, cells, of the present invention or a combination thereof (e.g., agents) contained in the kit are administered simultaneously, the combination kit can contain the active agents in a single formulation, such as a pharmaceutical formulation, (e.g., a tablet, capsule, vial, etc.) or in separate formulations. When the engineered compositions, formulations, systems particles, vesicles, cells, of the present invention or a combination thereof and/or kit components are not administered simultaneously, the combination kit can contain each agent or other component in separate pharmaceutical formulations. The separate kit components can be contained in a single package or in separate packages within the kit.

[0507] In some embodiments, the combination kit also includes instructions printed on or otherwise contained in a tangible medium of expression. The instructions can provide information regarding the content of the engineered compositions, formulations, systems particles, vesicles, cells, of the present invention described herein or a combination thereof contained therein, safety information regarding the content of the engineered compositions, formulations (e.g. pharmaceutical formulations), systems particles, vesicles, cells, of the present invention described herein or a combination thereof contained therein, information regarding the dosages, indications for use, and/or recommended treatment regimen(s) for engineered compositions, formulations, systems particles, vesicles, cells, of the present invention contained therein. In some embodiments, the instructions can provide directions for administering the engineered compositions, formulations, systems particles, vesicles, cells, of the present invention or a combination thereof to a subject. In some embodiments, the subject is a subject in need of a treatment or therapy that can be delivered or otherwise provided by the

engineered compositions, formulations, systems particles, vesicles, cells, of the present invention.

METHODS OF MAKING AND LOADING ENGINEERED DELIVERY VESICLES

[0508] Engineered delivery vesicles can be made by a producer cell. A producer cell can also be considered a bioreactor or bioreactor cell. One or more of the engineered compositions, systems, and/or formulations thereof can be delivered to a cell, whereby the cell expresses one or more of the compositions, systems, and/or formulations thereof and generates delivery vesicles. Suitable methods of culturing, expanding, and expressing vectors and vector systems, such as those of the present invention so as to express one or more polynucleotides of the system and/or generate vesicles, such as viral particles, are described in greater detail elsewhere herein and in the art.

[0509] In some embodiments, the cells also package one or more cargos within the delivery vesicles. In some embodiments, the cargo(s) is/are included in the one or more engineered compositions, systems, and/or formulations thereof and are optionally expressed and/or packaged in the delivery vesicles during vesicle generation. In other embodiments, the cargo(s) are exogenous to the engineered compositions, systems, or formulations thereof. In some embodiments, the cargos are endogenous or otherwise present into the producer cell and are encapsulated or otherwise incorporated into the delivery vesicles as is described in greater detail elsewhere herein.

[0510] In some embodiments, the producer cell, prior to making delivery particles, can uptake exogenous molecules, compounds, and the like (e.g., small molecule therapeutics, imaging agents, etc.). When delivery vesicles are formed by the producer cell, these exogenous molecules, compounds, and the like are incorporated into the vesicles. In some embodiments, a separate expression system is provided to the producer cell that makes one or more gene products. Such gene products can be produced by the cell and incorporated into delivery vesicles as they are being formed by the producer cell so as to load the vesicles with a cargo.

[0511] Unloaded and Cargo-loaded delivery vesicles of the present invention can be exposed to cells (e.g. *in vitro*, *ex vivo*, or *in vivo*) where the delivery vesicles delivers the cargo to the target cell, for example, by transduction. Delivery vesicles can be optionally concentrated prior to exposure to target cells. As used in this context herein “transduction” refers to the process by which foreign nucleic acids and/or proteins are introduced to a cell (prokaryote or eukaryote) by a viral or pseudo viral particle.

[0512] In some embodiments, the producer cells can be delivered to a subject so as to deliver a vesicle or cargo therein to the subject.

Vesicle Production

[0513] In some embodiments, one or more engineered delivery vesicle generation system viral vectors and/or system thereof can be delivered to a suitable cell line for production of virus particles containing the cargo polynucleotide or other payload to be delivered to a host cell.

Delivery

[0514] Any suitable delivery for delivering polynucleotides and vectors can be used such as physical delivery methods and other non-vector delivery methods. Physical methods for delivery of the polynucleotides and vectors of the present invention include, but are not limited to, microinjection, electroporation, hydrodynamic delivery, transfection, transduction, and biolistics. Non-vector based delivery can include, but is not limited to, delivery via lipid particles (e.g. lipid nanoparticles, liposomes, stable nucleic-acid-lipid-particles (SNAPs)) lipoplexes, sugar-based particles, cell penetrating peptides, DNA Nanoclews and cages, nanoparticles (metal, silica, polymer, inorganic, composites, etc.), streptolysin O (SLO), iTOP, exosomes, spherical nucleic acids, self-assembling nanoparticles, and supercharged proteins.

Physical Delivery

[0515] In some embodiments, the cargos may be introduced to cells by physical delivery methods. Examples of physical methods include microinjection, electroporation, and hydrodynamic delivery. Both nucleic acid and proteins may be delivered using such methods.

Microinjection

[0516] Microinjection of the cargo directly to cells can achieve high efficiency, e.g., above 90% or about 100%. In some embodiments, microinjection may be performed using a microscope and a needle (e.g., with 0.5–5.0 μm in diameter) to pierce a cell membrane and deliver the cargo directly to a target site within the cell. Microinjection may be used for *in vitro* and *ex vivo* delivery.

[0517] Plasmids or other vectors can be microinjected. In some cases, microinjection may be used i) to deliver DNA directly to a cell nucleus, and/or ii) to deliver mRNA (e.g., in vitro transcribed) to a cell nucleus or cytoplasm.

[0518] Microinjection may also be used to generate genetically modified animals. For example, gene editing cargos may be injected into zygotes to allow for efficient germline

modification. Such approach can yield normal embryos and full-term mouse pups harboring the desired modification(s). Microinjection can also be used to provide transiently up- or down-regulate a specific gene within the genome of a cell.

Electroporation

[0519] In some embodiments, the cargos and/or delivery vehicles may be delivered by electroporation. Electroporation may use pulsed high-voltage electrical currents to transiently open nanometer-sized pores within the cellular membrane of cells suspended in buffer, allowing for components with hydrodynamic diameters of tens of nanometers to flow into the cell. In some cases, electroporation may be used on various cell types and efficiently transfer cargo into cells. Electroporation may be used for *in vitro* and *ex vivo* delivery.

[0520] Electroporation may also be used to deliver the cargo to into the nuclei of mammalian cells by applying specific voltage and reagents, e.g., by nucleofection. Such approaches include those described in Wu Y, et al. (2015). Cell Res 25:67–79; Ye L, et al. (2014). Proc Natl Acad Sci USA 111:9591–6; Choi PS, Meyerson M. (2014). Nat Commun 5:3728; Wang J, Quake SR. (2014). Proc Natl Acad Sci 111:13157–62. Electroporation may also be used to deliver the cargo in vivo, e.g., with methods described in Zuckermann M, et al. (2015). Nat Commun 6:7391.

Hydrodynamic Delivery

[0521] Hydrodynamic delivery may also be used for delivering the cargos, e.g., for *in vivo* delivery. In some examples, hydrodynamic delivery may be performed by rapidly pushing a large volume (8–10% body weight) solution containing the gene editing cargo into the bloodstream of a subject (e.g., an animal or human), e.g., for mice, via the tail vein. As blood is incompressible, the large bolus of liquid may result in an increase in hydrodynamic pressure that temporarily enhances permeability into endothelial and parenchymal cells, allowing for cargo not normally capable of crossing a cellular membrane to pass into cells. This approach may be used for delivering naked DNA plasmids and proteins. The delivered cargos may be enriched in liver, kidney, lung, muscle, and/or heart.

Transfection

[0522] The vectors, nucleic acids, and/or polypeptides, may be introduced to cells by transfection methods for introducing nucleic acids into cells. Examples of transfection methods include calcium phosphate-mediated transfection, cationic transfection, liposome transfection,

dendrimer transfection, heat shock transfection, magnetofection, lipofection, impalefection, optical transfection, proprietary agent-enhanced uptake of nucleic acid.

Transduction

[0523] The vectors, nucleic acids and/or polypeptides, can be introduced to cells by transduction by a viral or pseudoviral particle. Methods of packaging the cargos in viral particles can be accomplished using any suitable viral vector or vector systems. Such viral vector and vector systems are described in greater detail elsewhere herein. As used in this context herein "transduction" refers to the process by which foreign nucleic acids and/or proteins are introduced to a cell (prokaryote or eukaryote) by a viral or pseudo viral particle. After packaging in a viral particle or pseudo viral particle, the viral particles can be exposed to cells (e.g. in vitro, ex vivo, or in vivo) where the viral or pseudoviral particle infects the cell and delivers the cargo to the cell via transduction. Viral and pseudoviral particles can be optionally concentrated prior to exposure to target cells. In some embodiments, the virus titer of a composition containing viral and/or pseudoviral particles can be obtained and a specific titer be used to transduce cells.

Biolistics

[0524] The vectors nucleic acids and/or polypeptides, can be introduced to cells using a biolistic method or technique. The term of art "biolistic", as used herein refers to the delivery of nucleic acids to cells by high-speed particle bombardment. In some embodiments, the cargo(s) can be attached, associated with, or otherwise coupled to particles, which than can be delivered to the cell via a gene-gun (see e.g., Liang et al. 2018. Nat. Protocol. 13:413-430; Svitashev et al. 2016. Nat. Comm. 7:13274; Ortega-Escalante et al., 2019. Plant. J. 97:661-672). In some embodiments, the particles can be gold, tungsten, palladium, rhodium, platinum, or iridium particles.

Non-Vector Delivery Vehicles

[0525] The delivery vehicles may comprise non-viral vehicles. In general, methods and vehicles capable of delivering nucleic acids and/or proteins may be used for delivering the systems compositions herein. Examples of non-viral vehicles include lipid nanoparticles, cell-penetrating peptides (CPPs), DNA nanoclews, metal nanoparticles, streptolysin O, multifunctional envelope-type nanodevices (MENDs), lipid-coated mesoporous silica particles, and other inorganic nanoparticles.

Lipid Particles

[0526] The delivery vehicles may comprise lipid particles, e.g., lipid nanoparticles (LNPs) and liposomes. Lipofection is described in e.g., U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Felgner, International Patent Publication Nos. WO 91/17424 and WO 91/16024. The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known to one of skill in the art (see, e.g., Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); 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, and 4,946,787).

Lipid nanoparticles (LNPs)

[0527] LNPs may encapsulate nucleic acids within cationic lipid particles (e.g., liposomes), and may be delivered to cells with relative ease. In some examples, lipid nanoparticles do not contain any viral components, which helps minimize safety and immunogenicity concerns. Lipid particles may be used for *in vitro*, *ex vivo*, and *in vivo* deliveries. Lipid particles may be used for various scales of cell populations.

[0528] In some examples. LNPs may be used for delivering DNA molecules (e.g., those comprising coding sequences of Cas and/or gRNA) and/or RNA molecules (e.g., mRNA of Cas, gRNAs). In certain cases, LNPs may be use for delivering RNP complexes of Cas/gRNA.

[0529] Components in LNPs may comprise cationic lipids 1,2- dilineoyl-3-dimethylammonium-propane (DLinDAP), 1,2-dilinoleyloxy-3-N,N- dimethylaminopropane (DLinDMA), 1,2-dilinoleyloxyketo-N,N-dimethyl-3-aminopropane (DLinK-DMA), 1,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLinKC2-DMA), (3- o-[2"-(methoxypolyethyleneglycol 2000) succinoyl]-1,2-dimyristoyl-sn-glycol (PEG-S-DMG), R-3-[(ro-methoxy-poly(ethylene glycol)2000) carbamoyl]-1,2-dimyristyloxlpropyl-3-amine (PEG-C-DOMG, and any combination thereof. Preparation of LNPs and encapsulation may be adapted from Rosin et al, Molecular Therapy, vol. 19, no. 12, pages 1286-2200, Dec. 2011).

[0530] In some embodiments, an LNP delivery vehicle can be used to deliver a virus particle containing a cargo thereof. In some embodiments, the virus particle(s) can be adsorbed

to the lipid particle, such as through electrostatic interactions, and/or can be attached to the liposomes via a linker.

[0531] In some embodiments, the LNP contains a nucleic acid, wherein the charge ratio of nucleic acid backbone phosphates to cationic lipid nitrogen atoms is about 1: 1.5 – 7 or about 1:4.

[0532] In some embodiments, the LNP also includes a shielding compound, which is removable from the lipid composition under *in vivo* conditions. In some embodiments, the shielding compound is a biologically inert compound. In some embodiments, the shielding compound does not carry any charge on its surface or on the molecule as such. In some embodiments, the shielding compounds are polyethylenglycoles (PEGs), hydroxyethylglucose (HEG) based polymers, polyhydroxyethyl starch (polyHES) and polypropylene. In some embodiments, the PEG, HEG, polyHES, and a polypropylene weight between about 500 to 10,000 Da or between about 2000 to 5000 Da. In some embodiments, the shielding compound is PEG2000 or PEG5000.

[0533] In some embodiments, the LNP can include one or more helper lipids. In some embodiments, the helper lipid can be a phosphor lipid or a steroid. In some embodiments, the helper lipid is between about 20 mol % to 80 mol % of the total lipid content of the composition. In some embodiments, the helper lipid component is between about 35 mol % to 65 mol % of the total lipid content of the LNP. In some embodiments, the LNP includes lipids at 50 mol% and the helper lipid at 50 mol% of the total lipid content of the LNP.

[0534] Other non-limiting, exemplary LNP delivery vehicles are described in U.S. Patent Publication Nos. US 20160174546, US 20140301951, US 20150105538, US 20150250725, Wang et al., J. Control Release, 2017 Jan 31. pii: S0168-3659(17)30038-X. doi: 10.1016/j.jconrel.2017.01.037. [Epub ahead of print]; Altınoğlu et al., Biomater Sci., 4(12):1773-80, Nov. 15, 2016; Wang et al., PNAS, 113(11):2868-73 March 15, 2016; Wang et al., PloS One, 10(11): e0141860. doi: 10.1371/journal.pone.0141860. eCollection 2015, Nov. 3, 2015; Takeda et al., Neural Regen Res. 10(5):689-90, May 2015; Wang et al., Adv. Healthc Mater., 3(9):1398-403, Sep. 2014; and Wang et al., Agnew Chem Int Ed Engl., 53(11):2893-8, Mar. 10, 2014; James E. Dahlman and Carmen Barnes et al. Nature Nanotechnology (2014) published online 11 May 2014, doi:10.1038/nnano.2014.84; Coelho et al., N Engl J Med 2013; 369:819-29; Aleku *et al.*, *Cancer Res.*, 68(23): 9788-98 (Dec. 1, 2008), Strumberg *et al.*, *Int. J. Clin. Pharmacol. Ther.*, 50(1): 76-8 (Jan. 2012), Schultheis *et al.*, *J.*

Clin. Oncol., 32(36): 4141-48 (Dec. 20, 2014), and Fehring *et al.*, *Mol. Ther.*, 22(4): 811-20 (Apr. 22, 2014); Novobrantseva, Molecular Therapy–Nucleic Acids (2012) 1, e4; doi:10.1038/mtna.2011.3; WO2012135025; US 20140348900; US 20140328759; US 20140308304; WO 2005/105152; WO 2006/069782; WO 2007/121947; US 2015/082080; US 20120251618; 7,982,027; 7,799,565; 8,058,069; 8,283,333; 7,901,708; 7,745,651; 7,803,397; 8,101,741; 8,188,263; 7,915,399; 8,236,943 and 7,838,658 and European Pat. Nos 1766035; 1519714; 1781593 and 1664316.

Liposomes

[0535] In some embodiments, a lipid particle may be liposome. Liposomes are spherical vesicle structures composed of a uni- or multilamellar lipid bilayer surrounding internal aqueous compartments and a relatively impermeable outer lipophilic phospholipid bilayer. In some embodiments, liposomes are biocompatible, nontoxic, can deliver both hydrophilic and lipophilic drug molecules, protect their cargo from degradation by plasma enzymes, and transport their load across biological membranes and the blood brain barrier (BBB).

[0536] Liposomes can be made from several different types of lipids, e.g., phospholipids. A liposome may comprise natural phospholipids and lipids such as l,2-distearoryl-sn-glycero-3 -phosphatidyl choline (DSPC), sphingomyelin, egg phosphatidylcholines, monosialoganglioside, or any combination thereof.

[0537] Several other additives may be added to liposomes in order to modify their structure and properties. For instance, liposomes may further comprise cholesterol, sphingomyelin, and/or l,2-dioleoyl-sn-glycero-3- phosphoethanolamine (DOPE), e.g., to increase stability and/or to prevent the leakage of the liposomal inner cargo.

[0538] In some embodiments, a liposome delivery vehicle can be used to deliver an engineered delivery vesicle generation system or component thereof and/or delivery vesicle to a cell. In some embodiments, the virus particle(s) can be adsorbed to the liposome, such as through electrostatic interactions, and/or can be attached to the liposomes via a linker.

[0539] In some embodiments, the liposome can be a Trojan Horse liposome (also known in the art as Molecular Trojan Horses), see e.g. http://cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long, the teachings of which can be applied and/or adapted to generated and/or deliver the CRISPR-Cas systems described herein.

[0540] Other non-limiting, exemplary liposomes can be those as set forth in Wang et al., ACS Synthetic Biology, 1, 403-07 (2012); Wang et al., PNAS, 113(11) 2868-2873 (2016);

Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679; WO 2008/042973; US Pat. No. 8,071,082; WO 2014/186366; 20160257951; US20160129120; US 20160244761; 20120251618; WO2013/093648; Lipofectin (a combination of DOTMA and DOPE), Lipofectase, LIPOFECTAMINE.RTM. (e.g., LIPOFECTAMINE.RTM. 2000, LIPOFECTAMINE.RTM. 3000, LIPOFECTAMINE.RTM. RNAiMAX, LIPOFECTAMINE.RTM. LTX), SAINT-RED (Synvolux Therapeutics, Groningen Netherlands), DOPE, Cytofectin (Gilead Sciences, Foster City, Calif.), and Eufectins (JBL, San Luis Obispo, Calif.).

Stable nucleic-acid-lipid particles (SNALPs)

[0541] In some embodiments, the lipid particles may be stable nucleic acid lipid particles (SNALPs). SNALPs may comprise an ionizable lipid (DLinDMA) (e.g., cationic at low pH), a neutral helper lipid, cholesterol, a diffusible polyethylene glycol (PEG)-lipid, or any combination thereof. In some examples, SNALPs may comprise synthetic cholesterol, dipalmitoylphosphatidylcholine, 3-N-[(w-methoxy polyethylene glycol)2000)carbamoyl]-l,2-dimyrestyloxypropylamine, and cationic 1,2-dilinoleyloxy-3-N,Ndimethylaminopropane. In some examples, SNALPs may comprise synthetic cholesterol, 1,2-distearoyl-sn-glycero-3-phosphocholine, PEG- cDMA, and 1,2-dilinoleyloxy-3-(N;N-dimethyl)aminopropane (DLinDMAo).

[0542] Other non-limiting, exemplary SNALPs that can be used to deliver an engineered delivery vesicle generation system or component thereof and/or delivery vesicle can be any such SNALPs as described in Morrissey et al., Nature Biotechnology, Vol. 23, No. 8, August 2005, Zimmerman et al., Nature Letters, Vol. 441, 4 May 2006; Geisbert et al., Lancet 2010; 375: 1896-905; Judge, J. Clin. Invest. 119:661-673 (2009); and Semple et al., Nature Niotechnology, Volume 28 Number 2 February 2010, pp. 172-177.

Other Lipids

[0543] The lipid particles may also comprise one or more other types of lipids, e.g., cationic lipids, such as amino lipid 2,2-dilinoleyl-4-dimethylaminoethyl-[l,3]- dioxolane (DLin-KC2-DMA), DLin-KC2-DMA4, C12- 200 and colipids disteroylphosphatidyl choline, cholesterol, and PEG-DMG.

[0544] In some embodiments, the delivery vehicle can be or include a lipidoid, such as any of those set forth in, for example, US 20110293703.

[0545] In some embodiments, the delivery vehicle can be or include an amino lipid, such as any of those set forth in, for example, Jayaraman, Angew. Chem. Int. Ed. 2012, 51, 8529 – 8533.

[0546] In some embodiments, the delivery vehicle can be or include a lipid envelope, such as any of those set forth in, for example, Korman et al., 2011. Nat. Biotech. 29:154-157.

Lipoplexes/polyplexes

[0547] In some embodiments, the delivery vehicles comprise lipoplexes and/or polyplexes. Lipoplexes may bind to negatively charged cell membrane and induce endocytosis into the cells. Examples of lipoplexes may be complexes comprising lipid(s) and non-lipid components. Examples of lipoplexes and polyplexes include FuGENE-6 reagent, a non-liposomal solution containing lipids and other components, zwitterionic amino lipids (ZALs), Ca2þ (e.g., forming DNA/Ca^{2+} microcomplexes), polyethenimine (PEI) (e.g., branched PEI), and poly(L-lysine) (PLL).

Sugar-Based Particles

[0548] In some embodiments, the delivery vehicle can be a sugar-based particle. In some embodiments, the sugar-based particles can be or include GalNAc, such as any of those described in WO2014118272; US 20020150626; Nair, JK et al., 2014, Journal of the American Chemical Society 136 (49), 16958-16961; Østergaard et al., Bioconjugate Chem., 2015, 26 (8), pp 1451–1455;

Cell Penetrating Peptides

[0549] In some embodiments, the delivery vehicles to deliver an engineered delivery vesicle generation system or component thereof and/or delivery vesicle to a cell can comprise cell penetrating peptides (CPPs). CPPs are short peptides that facilitate cellular uptake of various molecular cargo (e.g., from nanosized particles to small chemical molecules and large fragments of DNA).

[0550] CPPs may be of different sizes, amino acid sequences, and charges. In some examples, CPPs can translocate the plasma membrane and facilitate the delivery of various molecular cargoes to the cytoplasm or an organelle. CPPs may be introduced into cells via different mechanisms, e.g., direct penetration in the membrane, endocytosis-mediated entry, and translocation through the formation of a transitory structure.

[0551] CPPs may have an amino acid composition that either contains a high relative abundance of positively charged amino acids such as lysine or arginine or has sequences that

contain an alternating pattern of polar/charged amino acids and non-polar, hydrophobic amino acids. These two types of structures are referred to as polycationic or amphipathic, respectively. A third class of CPPs are the hydrophobic peptides, containing only apolar residues, with low net charge or have hydrophobic amino acid groups that are crucial for cellular uptake. Another type of CPPs is the trans-activating transcriptional activator (Tat) from Human Immunodeficiency Virus 1 (HIV-1). Examples of CPPs include to Penetratin, Tat (48-60), Transportan, and (R-AhX-R4) (Ahx refers to aminohexanoyl), Kaposi fibroblast growth factor (FGF) signal peptide sequence, integrin β3 signal peptide sequence, polyarginine peptide Args sequence, Guanine rich-molecular transporters, and sweet arrow peptide. Examples of CPPs and related applications also include those described in US Patent 8,372,951.

[0552] CPPs can be used for *in vitro* and *ex vivo* work quite readily, and extensive optimization for each cargo and cell type is usually required. In some examples, CPPs may be covalently attached to the Cas protein directly, which is then complexed with the gRNA and delivered to cells. In some examples, separate delivery of CPP–Cas and CPP–gRNA to multiple cells may be performed. CPP may also be used to delivery RNPs.

[0553] CPPs may be used to deliver the compositions and systems to plants. In some examples, CPPs may be used to deliver the components to plant protoplasts, which are then regenerated to plant cells and further to plants.

DNA Nanoclews

[0554] In some embodiments, the delivery vehicles comprise DNA nanoclews. A DNA nanoclew refers to a sphere-like structure of DNA (e.g., with a shape of a ball of yarn). The nanoclew may be synthesized by rolling circle amplification with palindromic sequences that aide in the self-assembly of the structure. The sphere may then be loaded with a payload. An example of DNA nanoclew is described in Sun W et al, J Am Chem Soc. 2014 Oct 22;136(42):14722-5; and Sun W et al, Angew Chem Int Ed Engl. 2015 Oct 5;54(41):12029-33. DNA nanoclew may have a palindromic sequences to be partially complementary to the gRNA within the Cas:gRNA ribonucleoprotein complex. A DNA nanoclew may be coated, e.g., coated with PEI to induce endosomal escape.

Metal Nanoparticles

[0555] In some embodiments, the delivery vehicles comprise gold nanoparticles (also referred to AuNPs or colloidal gold). Gold nanoparticles may form complex with cargos, e.g., an engineered delivery vesicle generation system or component thereof and/or delivery vesicle.

Gold nanoparticles may be coated, e.g., coated in a silicate and an endosomal disruptive polymer, PAsp(DET). Examples of gold nanoparticles include AuraSense Therapeutics' Spherical Nucleic Acid (SNA™) constructs, and those described in Mout R, et al. (2017). ACS Nano 11:2452–8; Lee K, et al. (2017). Nat Biomed Eng 1:889–901. Other metal nanoparticles can also be complexed with cargo(s). Such metal particles include tungsten, palladium, rhodium, platinum, and iridium particles. Other non-limiting, exemplary metal nanoparticles are described in US 20100129793.

iTOP

[0556] In some embodiments, the delivery vehicles comprise iTOP. iTOP refers to a combination of small molecules drives the highly efficient intracellular delivery of native proteins, independent of any transduction peptide. iTOP may be used for induced transduction by osmocytosis and propanebetaine, using NaCl-mediated hyperosmolality together with a transduction compound (propanebetaine) to trigger macropinocytotic uptake into cells of extracellular macromolecules. Examples of iTOP methods and reagents include those described in D'Astolfo DS, Pagliero RJ, Pras A, et al. (2015). Cell 161:674–690.

Polymer-based Particles

[0557] In some embodiments, the delivery vehicles may comprise polymer-based particles (e.g., nanoparticles). In some embodiments, the polymer-based particles may mimic a viral mechanism of membrane fusion. The polymer-based particles may be a synthetic copy of Influenza virus machinery and form transfection complexes with various types of nucleic acids ((siRNA, miRNA, plasmid DNA or shRNA, mRNA) that cells take up via the endocytosis pathway, a process that involves the formation of an acidic compartment. The low pH in late endosomes acts as a chemical switch that renders the particle surface hydrophobic and facilitates membrane crossing. Once in the cytosol, the particle releases its payload for cellular action. This Active Endosome Escape technology is safe and maximizes transfection efficiency as it is using a natural uptake pathway. In some embodiments, the polymer-based particles may comprise alkylated and carboxyalkylated branched polyethylenimine. In some examples, the polymer-based particles are VIROMER, e.g., VIROMER RNAi, VIROMER RED, VIROMER mRNA, VIROMER CRISPR. Example methods of delivering the systems and compositions herein include those described in Bawage SS et al., Synthetic mRNA expressed Cas13a mitigates RNA virus infections, www.biorxiv.org/content/10.1101/370460v1.full doi: doi.org/10.1101/370460, Viromer® RED, a powerful tool for transfection of keratinocytes.

doi: 10.13140/RG.2.2.16993.61281, Viromer® Transfection - Factbook 2018: technology, product overview, users' data., doi:10.13140/RG.2.2.23912.16642. Other exemplary and non-limiting polymeric particles are described in US 20170079916, US 20160367686, US 20110212179, US 20130302401, 6,007,845, 5,855,913, 5,985,309, 5,543,158, WO2012135025, US 20130252281, US 20130245107, US 20130244279; US 20050019923, 20080267903.

Streptolysin O (SLO)

[0558] The delivery vehicles may be streptolysin O (SLO). SLO is a toxin produced by Group A streptococci that works by creating pores in mammalian cell membranes. SLO may act in a reversible manner, which allows for the delivery of proteins (e.g., up to 100 kDa) to the cytosol of cells without compromising overall viability. Examples of SLO include those described in Sierig G, et al. (2003). Infect Immun 71:446–55; Walev I, et al. (2001). Proc Natl Acad Sci U S A 98:3185–90; Teng KW, et al. (2017). Elife 6:e25460.

Multifunctional Envelope-Type Nanodevice (MEND)

[0559] The delivery vehicles may comprise multifunctional envelope-type nanodevice (MENDs). MENDs may comprise condensed plasmid DNA, a PLL core, and a lipid film shell. A MEND may further comprise cell-penetrating peptide (e.g., stearyl octaarginine). The cell penetrating peptide may be in the lipid shell. The lipid envelope may be modified with one or more functional components, e.g., one or more of: polyethylene glycol (e.g., to increase vascular circulation time), ligands for targeting of specific tissues/cells, additional cell-penetrating peptides (e.g., for greater cellular delivery), lipids to enhance endosomal escape, and nuclear delivery tags. In some examples, the MEND may be a tetra-lamellar MEND (T-MEND), which may target the cellular nucleus and mitochondria. In certain examples, a MEND may be a PEG-peptide-DOPE-conjugated MEND (PPD-MEND), which may target bladder cancer cells. Examples of MENDs include those described in Kogure K, et al. (2004). J Control Release 98:317–23; Nakamura T, et al. (2012). Acc Chem Res 45:1113–21.

Lipid-coated mesoporous silica particles

[0560] The delivery vehicles may comprise lipid-coated mesoporous silica particles. Lipid-coated mesoporous silica particles may comprise a mesoporous silica nanoparticle core and a lipid membrane shell. The silica core may have a large internal surface area, leading to high cargo loading capacities. In some embodiments, pore sizes, pore chemistry, and overall particle sizes may be modified for loading different types of cargos. The lipid coating of the particle

may also be modified to maximize cargo loading, increase circulation times, and provide precise targeting and cargo release. Examples of lipid-coated mesoporous silica particles include those described in Du X, et al. (2014). Biomaterials 35:5580–90; Durfee PN, et al. (2016). ACS Nano 10:8325–45.

Inorganic nanoparticles

[0561] The delivery vehicles may comprise inorganic nanoparticles. Examples of inorganic nanoparticles include carbon nanotubes (CNTs) (e.g., as described in Bates K and Kostarelos K. (2013). Adv Drug Deliv Rev 65:2023–33.), bare mesoporous silica nanoparticles (MSNPs) (e.g., as described in Luo GF, et al. (2014). Sci Rep 4:6064), and dense silica nanoparticles (SiNPs) (as described in Luo D and Saltzman WM. (2000). Nat Biotechnol 18:893–5).

Exosomes

[0562] The delivery vehicles may comprise exosomes. Exosomes include membrane bound extracellular vesicles, which can be used to contain and delivery various types of biomolecules, such as proteins, carbohydrates, lipids, and nucleic acids, and complexes thereof (e.g., RNPs). Examples of exosomes include those described in Schroeder A, et al., J Intern Med. 2010 Jan;267(1):9-21; El-Andaloussi S, et al., Nat Protoc. 2012 Dec;7(12):2112-26; Uno Y, et al., Hum Gene Ther. 2011 Jun;22(6):711-9; Zou W, et al., Hum Gene Ther. 2011 Apr;22(4):465-75.

[0563] In some examples, the exosome may form a complex (e.g., by binding directly or indirectly) to one or more components of the cargo. In certain examples, a molecule of an exosome may be fused with first adapter protein and a component of the cargo may be fused with a second adapter protein. The first and the second adapter protein may specifically bind each other, thus associating the cargo with the exosome. Examples of such exosomes include those described in Ye Y, et al., Biomater Sci. 2020 Apr 28. doi: 10.1039/d0bm00427h.

[0564] Other non-limiting, exemplary exosomes include any of those set forth in Alvarez-Erviti et al. 2011, Nat Biotechnol 29: 341; [1401] El-Andaloussi et al. (Nature Protocols 7:2112–2126(2012); and Wahlgren et al. (Nucleic Acids Research, 2012, Vol. 40, No. 17 e130).

Spherical Nucleic Acids (SNAs)

[0565] In some embodiments, the delivery vehicle can be a SNA. SNAs are three dimensional nanostructures that can be composed of densely functionalized and highly oriented

nucleic acids that can be covalently attached to the surface of spherical nanoparticle cores. The core of the spherical nucleic acid can impart the conjugate with specific chemical and physical properties, and it can act as a scaffold for assembling and orienting the oligonucleotides into a dense spherical arrangement that gives rise to many of their functional properties, distinguishing them from all other forms of matter. In some embodiments, the core is a crosslinked polymer. Non-limiting, exemplary SNAs can be any of those set forth in Cutler et al., J. Am. Chem. Soc. 2011 133:9254-9257, Hao et al., Small. 2011 7:3158-3162, Zhang et al., ACS Nano. 2011 5:6962-6970, Cutler et al., J. Am. Chem. Soc. 2012 134:1376-1391, Young et al., Nano Lett. 2012 12:3867-71, Zheng et al., Proc. Natl. Acad. Sci. USA. 2012 109:11975-80, Mirkin, Nanomedicine 2012 7:635-638 Zhang et al., J. Am. Chem. Soc. 2012 134:16488-1691, Weintraub, Nature 2013 495:S14-S16, Choi et al., Proc. Natl. Acad. Sci. USA. 2013 110(19):7625-7630, Jensen et al., Sci. Transl. Med. 5, 209ra152 (2013) and Mirkin, et al., and Small, 10:186-192.

Self-Assembling Nanoparticles

[0566] In some embodiments, the delivery vehicle is a self-assembling nanoparticle. The self-assembling nanoparticles can contain one or more polymers. The self-assembling nanoparticles can be PEGylated. Self-assembling nanoparticles are known in the art. Non-limiting, exemplary self-assembling nanoparticles can any as set forth in Schiffelers et al., Nucleic Acids Research, 2004, Vol. 32, No. 19, Bartlett et al. (PNAS, September 25, 2007, vol. 104, no. 39; Davis et al., Nature, Vol 464, 15 April 2010.

Supercharged Proteins

[0567] In some embodiments, the delivery vehicle can be a supercharged protein. As used herein "Supercharged proteins" are a class of engineered or naturally occurring proteins with unusually high positive or negative net theoretical charge. Non-limiting, exemplary supercharged proteins can be any of those set forth in Lawrence et al., 2007, Journal of the American Chemical Society 129, 10110–10112.

Targeted Delivery

[0568] In some embodiments, delivery of an engineered delivery vesicle generation system or component thereof can be targeted such that only or substantially delivery is only to target cells. In this way, only certain cells can become producer cells despite delivery being otherwise nonspecific. In some embodiments, the delivery vehicle can allow for targeted delivery to a specific cell, tissue, organ, or system. In such embodiments, the delivery vehicle can include

one or more targeting moieties that can direct targeted delivery of the cargo(s). In an embodiment, the delivery vehicle comprises a targeting moiety, such as active targeting of a lipid entity of the invention, e.g., lipid particle or nanoparticle or liposome or lipid bilayer of the invention comprising a targeting moiety for active targeting.

[0569] With regard to targeting moieties, mention is made of Deshpande et al, "Current trends in the use of liposomes for tumor targeting," Nanomedicine (Lond). 8(9), doi:10.2217/nnm.13.118 (2013), and the documents it cites, all of which are incorporated herein by reference and the teachings of which can be applied and/or adapted for targeted delivery of one or more CRISPR-Cas molecules described herein. Mention is also made of International Patent Publication No. WO 2016/027264, and the documents it cites, all of which are incorporated herein by reference, the teachings of which can be applied and/or adapted for targeted delivery of one or more CRISPR-Cas molecules described herein. And mention is made of Lorenzer et al, "Going beyond the liver: Progress and challenges of targeted delivery of siRNA therapeutics," Journal of Controlled Release, 203: 1–15 (2015), and the documents it cites, all of which are incorporated herein by reference, the teachings of which can be applied and/or adapted for targeted delivery of one or more CRISPR-Cas molecules described herein.

[0570] An actively targeting lipid particle or nanoparticle or liposome or lipid bilayer delivery system (generally as to embodiments of the invention, "lipid entity of the invention" delivery systems) are prepared by conjugating targeting moieties, including small molecule ligands, peptides and monoclonal antibodies, on the lipid or liposomal surface; for example, certain receptors, such as folate and transferrin (Tf) receptors (TfR), are overexpressed on many cancer cells and have been used to make liposomes tumor cell specific. Liposomes that accumulate in the tumor microenvironment can be subsequently endocytosed into the cells by interacting with specific cell surface receptors. To efficiently target liposomes to cells, such as cancer cells, it is useful that the targeting moiety have an affinity for a cell surface receptor and to link the targeting moiety in sufficient quantities to have optimum affinity for the cell surface receptors; and determining these embodiments are within the ambit of the skilled artisan. In the field of active targeting, there are a number of cell-, e.g., tumor-, specific targeting ligands.

[0571] Also, as to active targeting, with regard to targeting cell surface receptors such as cancer cell surface receptors, targeting ligands on liposomes can provide attachment of liposomes to cells, e.g., vascular cells, via a noninternalizing epitope; and, this can increase the extracellular concentration of that which is being delivered, thereby increasing the amount

delivered to the target cells. A strategy to target cell surface receptors, such as cell surface receptors on cancer cells, such as overexpressed cell surface receptors on cancer cells, is to use receptor-specific ligands or antibodies. Many cancer cell types display upregulation of tumor-specific receptors. For example, TfRs and folate receptors (FRs) are greatly overexpressed by many tumor cell types in response to their increased metabolic demand. Folic acid can be used as a targeting ligand for specialized delivery owing to its ease of conjugation to nanocarriers, its high affinity for FRs and the relatively low frequency of FRs, in normal tissues as compared with their overexpression in activated macrophages and cancer cells, e.g., certain ovarian, breast, lung, colon, kidney and brain tumors. Overexpression of FR on macrophages is an indication of inflammatory diseases, such as psoriasis, Crohn's disease, rheumatoid arthritis and atherosclerosis; accordingly, folate-mediated targeting of the invention can also be used for studying, addressing or treating inflammatory disorders, as well as cancers. Folate-linked lipid particles or nanoparticles or liposomes or lipid bylayers of the invention ("lipid entity of the invention") deliver their cargo intracellularly through receptor-mediated endocytosis. Intracellular trafficking can be directed to acidic compartments that facilitate cargo release, and, most importantly, release of the cargo can be altered or delayed until it reaches the cytoplasm or vicinity of target organelles. Delivery of cargo using a lipid entity of the invention having a targeting moiety, such as a folate-linked lipid entity of the invention, can be superior to nontargeted lipid entity of the invention. The attachment of folate directly to the lipid head groups may not be favorable for intracellular delivery of folate-conjugated lipid entity of the invention, since they may not bind as efficiently to cells as folate attached to the lipid entity of the invention surface by a spacer, which may can enter cancer cells more efficiently. A lipid entity of the invention coupled to folate can be used for the delivery of complexes of lipid, e.g., liposome, e.g., anionic liposome and virus or capsid or envelope or virus outer protein, such as those herein discussed such as adenovirous or AAV. Tf is a monomeric serum glycoprotein of approximately 80 KDa involved in the transport of iron throughout the body. Tf binds to the TfR and translocates into cells via receptor-mediated endocytosis. The expression of TfR is can be higher in certain cells, such as tumor cells (as compared with normal cells and is associated with the increased iron demand in rapidly proliferating cancer cells. Accordingly, the invention comprehends a TfR-targeted lipid entity of the invention, e.g., as to liver cells, liver cancer, breast cells such as breast cancer cells, colon such as colon cancer cells, ovarian

cells such as ovarian cancer cells, head, neck and lung cells, such as head, neck and non-small-cell lung cancer cells, cells of the mouth such as oral tumor cells.

[0572] Also, as to active targeting, a lipid entity of the invention can be multifunctional, i.e., employ more than one targeting moiety such as CPP, along with Tf; a bifunctional system; e.g., a combination of Tf and poly-L-arginine which can provide transport across the endothelium of the blood–brain barrier. EGFR, is a tyrosine kinase receptor belonging to the ErbB family of receptors that mediates cell growth, differentiation and repair in cells, especially non-cancerous cells, but EGF is overexpressed in certain cells such as many solid tumors, including colorectal, non-small-cell lung cancer, squamous cell carcinoma of the ovary, kidney, head, pancreas, neck and prostate, and especially breast cancer. The invention comprehends EGFR-targeted monoclonal antibody(ies) linked to a lipid entity of the invention. HER-2 is often overexpressed in patients with breast cancer, and is also associated with lung, bladder, prostate, brain and stomach cancers. HER-2, encoded by the ERBB2 gene. The invention comprehends a HER-2-targeting lipid entity of the invention, e.g., an anti-HER-2-antibody(or binding fragment thereof)-lipid entity of the invention, a HER-2-targeting-PEGylated lipid entity of the invention (e.g., having an anti-HER-2-antibody or binding fragment thereof), a HER-2-targeting-maleimide-PEG polymer- lipid entity of the invention (e.g., having an anti-HER-2-antibody or binding fragment thereof). Upon cellular association, the receptor-antibody complex can be internalized by formation of an endosome for delivery to the cytoplasm.

[0573] With respect to receptor-mediated targeting, the skilled artisan takes into consideration ligand/target affinity and the quantity of receptors on the cell surface, and that PEGylation can act as a barrier against interaction with receptors. The use of antibody-lipid entity of the invention targeting can be advantageous. Multivalent presentation of targeting moieties can also increase the uptake and signaling properties of antibody fragments. In practice of the invention, the skilled person takes into account ligand density (e.g., high ligand densities on a lipid entity of the invention may be advantageous for increased binding to target cells). Preventing early by macrophages can be addressed with a sterically stabilized lipid entity of the invention and linking ligands to the terminus of molecules such as PEG, which is anchored in the lipid entity of the invention (e.g., lipid particle or nanoparticle or liposome or lipid bilayer). The microenvironment of a cell mass such as a tumor microenvironment can be targeted; for instance, it may be advantageous to target cell mass vasculature, such as the tumor

vasculature microenvironment. Thus, the invention comprehends targeting VEGF. VEGF and its receptors are well-known proangiogenic molecules and are well-characterized targets for antiangiogenic therapy. Many small-molecule inhibitors of receptor tyrosine kinases, such as VEGFRs or basic FGFRs, have been developed as anticancer agents and the invention comprehends coupling any one or more of these peptides to a lipid entity of the invention, e.g., phage IVO peptide(s) (e.g., via or with a PEG terminus), tumor-homing peptide APRPG such as APRPG-PEG-modified. VCAM, the vascular endothelium plays a key role in the pathogenesis of inflammation, thrombosis and atherosclerosis. CAMs are involved in inflammatory disorders, including cancer, and are a logical target, E- and P-selectins, VCAM-1 and ICAMs. Can be used to target a lipid entity of the invention., e.g., with PEGylation.

[0574] Matrix metalloproteases (MMPs) belong to the family of zinc-dependent endopeptidases. They are involved in tissue remodeling, tumor invasiveness, resistance to apoptosis and metastasis. There are four MMP inhibitors called TIMP1–4, which determine the balance between tumor growth inhibition and metastasis; a protein involved in the angiogenesis of tumor vessels is MT1-MMP, expressed on newly formed vessels and tumor tissues. The proteolytic activity of MT1-MMP cleaves proteins, such as fibronectin, elastin, collagen and laminin, at the plasma membrane and activates soluble MMPs, such as MMP-2, which degrades the matrix. An antibody or fragment thereof such as a Fab′ fragment can be used in the practice of the invention such as for an antihuman MT1-MMP monoclonal antibody linked to a lipid entity of the invention, e.g., via a spacer such as a PEG spacer. αβ-integrins or integrins are a group of transmembrane glycoprotein receptors that mediate attachment between a cell and its surrounding tissues or extracellular matrix.

[0575] Integrins contain two distinct chains (heterodimers) called α- and β-subunits. The tumor tissue-specific expression of integrin receptors can be been utilized for targeted delivery in the invention, e.g., whereby the targeting moiety can be an RGD peptide such as a cyclic RGD.

[0576] Aptamers are ssDNA or RNA oligonucleotides that impart high affinity and specific recognition of the target molecules by electrostatic interactions, hydrogen bonding and hydrophobic interactions as opposed to the Watson–Crick base pairing, which is typical for the bonding interactions of oligonucleotides. Aptamers as a targeting moiety can have advantages over antibodies: aptamers can demonstrate higher target antigen recognition as compared with antibodies; aptamers can be more stable and smaller in size as compared with antibodies;

aptamers can be easily synthesized and chemically modified for molecular conjugation; and aptamers can be changed in sequence for improved selectivity and can be developed to recognize poorly immunogenic targets. Such moieties as a sgc8 aptamer can be used as a targeting moiety (e.g., via covalent linking to the lipid entity of the invention, e.g., via a spacer, such as a PEG spacer).

[0577] Also, as to active targeting, the invention also comprehends intracellular delivery. Since liposomes follow the endocytic pathway, they are entrapped in the endosomes (pH 6.5–6) and subsequently fuse with lysosomes (pH <5), where they undergo degradation that results in a lower therapeutic potential. The low endosomal pH can be taken advantage of to escape degradation. Fusogenic lipids or peptides, which destabilize the endosomal membrane after the conformational transition/activation at a lowered pH. Amines are protonated at an acidic pH and cause endosomal swelling and rupture by a buffer effect Unsaturated dioleoylphosphatidylethanolamine (DOPE) readily adopts an inverted hexagonal shape at a low pH, which causes fusion of liposomes to the endosomal membrane. This process destabilizes a lipid entity containing DOPE and releases the cargo into the cytoplasm; fusogenic lipid GALA, cholesteryl-GALA and PEG-GALA may show a highly efficient endosomal release; a pore-forming protein listeriolysin O may provide an endosomal escape mechanism; and, histidine-rich peptides have the ability to fuse with the endosomal membrane, resulting in pore formation, and can buffer the proton pump causing membrane lysis.

[0578] The invention comprehends a lipid entity of the invention modified with CPP(s), for intracellular delivery that may proceed via energy dependent macropinocytosis followed by endosomal escape. The invention further comprehends organelle-specific targeting. A lipid entity of the invention surface-functionalized with the triphenylphosphonium (TPP) moiety or a lipid entity of the invention with a lipophilic cation, rhodamine 123 can be effective in delivery of cargo to mitochondria. DOPE/sphingomyelin/stearyl-octa-arginine can delivers cargos to the mitochondrial interior via membrane fusion. A lipid entity of the invention surface modified with a lysosomotropic ligand, octadecyl rhodamine B can deliver cargo to lysosomes. Ceramides are useful in inducing lysosomal membrane permeabilization; the invention comprehends intracellular delivery of a lipid entity of the invention having a ceramide. The invention further comprehends a lipid entity of the invention targeting the nucleus, e.g., via a DNA-intercalating moiety. The invention also comprehends multifunctional liposomes for targeting, i.e., attaching more than one functional group to the surface of the lipid entity of the

invention, for instance to enhances accumulation in a desired site and/or promotes organelle-specific delivery and/or target a particular type of cell and/or respond to the local stimuli such as temperature (e.g., elevated), pH (e.g., decreased), respond to externally applied stimuli such as a magnetic field, light, energy, heat or ultrasound and/or promote intracellular delivery of the cargo. All of these are considered actively targeting moieties.

[0579] Thus, in an embodiment of the engineered delivery system, of the present invention the targeting moiety comprises a receptor ligand, such as, for example, hyaluronic acid for CD44 receptor, galactose for hepatocytes, or antibody or fragment thereof such as a binding antibody fragment against a desired surface receptor, and as to each of a targeting moiety comprising a receptor ligand, or an antibody or fragment thereof such as a binding fragment thereof, such as against a desired surface receptor, there is an embodiment of the invention wherein the delivery system comprises a targeting moiety comprising a receptor ligand, or an antibody or fragment thereof such as a binding fragment thereof, such as against a desired surface receptor, or hyaluronic acid for CD44 receptor, galactose for hepatocytes (*see, e.g.*, Surace et al, "Lipoplexes targeting the CD44 hyaluronic acid receptor for efficient transfection of breast cancer cells," J. Mol Pharm 6(4):1062-73; doi: 10.1021/mp800215d (2009); Sonoke et al, "Galactose-modified cationic liposomes as a liver-targeting delivery system for small interfering RNA," Biol Pharm Bull. 34(8):1338-42 (2011); Torchilin, "Antibody-modified liposomes for cancer chemotherapy," Expert Opin. Drug Deliv. 5 (9), 1003-1025 (2008); Manjappa et al, "Antibody derivatization and conjugation strategies: application in preparation of stealth immunoliposome to target chemotherapeutics to tumor," J. Control. Release 150 (1), 2-22 (2011); Sofou S "Antibody-targeted liposomes in cancer therapy and imaging," Expert Opin. Drug Deliv. 5 (2): 189-204 (2008); Gao J et al, "Antibody-targeted immunoliposomes for cancer treatment," Mini. Rev. Med. Chem. 13(14): 2026-2035 (2013); Molavi et al, "Anti-CD30 antibody conjugated liposomal doxorubicin with significantly improved therapeutic efficacy against anaplastic large cell lymphoma," Biomaterials 34(34):8718-25 (2013), each of which and the documents cited therein are hereby incorporated herein by reference), the teachings of which can be applied and/or adapted for targeted delivery of one or more CRISPR-Cas molecules and/or other cargo described herein with the engineered systems of the present invention.

[0580] Other exemplary targeting moieties are described elsewhere herein, such as epitope tags and the like.

Responsive Delivery

[0581] In some embodiments, the delivery vehicle can allow for responsive delivery of the engineered delivery vesicle generation system so as to provide control over delivery vesicle production, for example, and/or delivery of a cargo. Responsive delivery, as used in this context herein, refers to delivery of cargo(s) by the delivery vehicle in response to an external stimuli. Examples of suitable stimuli include, without limitation, an energy (light, heat, cold, and the like), a chemical stimuli (e.g. chemical composition, etc.), and a biologic or physiologic stimuli (e.g. environmental pH, osmolarity, salinity, biologic molecule, etc.). In some embodiments, the targeting moiety can be responsive to an external stimuli and facilitate responsive delivery. In other embodiments, responsiveness is determined by a non-targeting moiety component of the delivery vehicle.

[0582] The delivery vehicle can be stimuli-sensitive, e.g., sensitive to an externally applied stimuli, such as magnetic fields, ultrasound or light; and pH-triggering can also be used, e.g., a labile linkage can be used between a hydrophilic moiety such as PEG and a hydrophobic moiety such as a lipid entity of the invention, which is cleaved only upon exposure to the relatively acidic conditions characteristic of the a particular environment or microenvironment such as an endocytic vacuole or the acidotic tumor mass. pH-sensitive copolymers can also be incorporated in embodiments of the invention can provide shielding; diortho esters, vinyl esters, cysteine-cleavable lipopolymers, double esters and hydrazones are a few examples of pH-sensitive bonds that are quite stable at pH 7.5, but are hydrolyzed relatively rapidly at pH 6 and below, e.g., a terminally alkylated copolymer of N-isopropylacrylamide and methacrylic acid that copolymer facilitates destabilization of a lipid entity of the invention and release in compartments with decreased pH value; or, the invention comprehends ionic polymers for generation of a pH-responsive lipid entity of the invention (e.g., poly(methacrylic acid), poly(diethylaminoethyl methacrylate), poly(acrylamide) and poly(acrylic acid)).

[0583] Temperature-triggered delivery is also within the ambit of the invention. Many pathological areas, such as inflamed tissues and tumors, show a distinctive hyperthermia compared with normal tissues. Utilizing this hyperthermia is an attractive strategy in cancer therapy since hyperthermia is associated with increased tumor permeability and enhanced uptake. This technique involves local heating of the site to increase microvascular pore size and blood flow, which, in turn, can result in an increased extravasation of embodiments of the invention. Temperature-sensitive lipid entity of the invention can be prepared from

thermosensitive lipids or polymers with a low critical solution temperature. Above the low critical solution temperature (e.g., at site such as tumor site or inflamed tissue site), the polymer precipitates, disrupting the liposomes to release. Lipids with a specific gel-to-liquid phase transition temperature are used to prepare these lipid entities of the invention; and a lipid for a thermosensitive embodiment can be dipalmitoylphosphatidylcholine. Thermosensitive polymers can also facilitate destabilization followed by release, and a useful thermosensitive polymer is poly (N-isopropylacrylamide). Another temperature triggered system can employ lysolipid temperature-sensitive liposomes.

[0584] The invention also comprehends redox-triggered delivery. The difference in redox potential between normal and inflamed or tumor tissues, and between the intra- and extra-cellular environments has been exploited for delivery, e.g., GSH is a reducing agent abundant in cells, especially in the cytosol, mitochondria and nucleus. The GSH concentrations in blood and extracellular matrix are just one out of 100 to one out of 1000 of the intracellular concentration, respectively. This high redox potential difference caused by GSH, cysteine and other reducing agents can break the reducible bonds, destabilize a lipid entity of the invention and result in release of payload. The disulfide bond can be used as the cleavable/reversible linker in a lipid entity of the invention, because it causes sensitivity to redox owing to the disulfideto-thiol reduction reaction; a lipid entity of the invention can be made reduction sensitive by using two (e.g., two forms of a disulfide-conjugated multifunctional lipid as cleavage of the disulfide bond (e.g., via tris(2-carboxyethyl)phosphine, dithiothreitol, L-cysteine or GSH), can cause removal of the hydrophilic head group of the conjugate and alter the membrane organization leading to release of payload. Calcein release from reduction-sensitive lipid entity of the invention containing a disulfide conjugate can be more useful than a reduction-insensitive embodiment.

[0585] Enzymes can also be used as a trigger to release payload. Enzymes, including MMPs (e.g. MMP2), phospholipase A2, alkaline phosphatase, transglutaminase or phosphatidylinositol-specific phospholipase C, have been found to be overexpressed in certain tissues, e.g., tumor tissues. In the presence of these enzymes, specially engineered enzyme-sensitive lipid entity of the invention can be disrupted and release the payload. an MMP2-cleavable octapeptide (Gly-Pro-Leu-Gly-Ile-Ala-Gly-Gln (SEQ ID NO:48)) can be incorporated into a linker, and can have antibody targeting, e.g., antibody 2C5.

[0586] The invention also comprehends light-or energy-triggered delivery, e.g., the lipid entity of the invention can be light-sensitive, such that light or energy can facilitate structural and conformational changes, which lead to direct interaction of the lipid entity of the invention with the target cells via membrane fusion, photo-isomerism, photofragmentation or photopolymerization; such a moiety therefor can be benzoporphyrin photosensitizer. Ultrasound can be a form of energy to trigger delivery; a lipid entity of the invention with a small quantity of particular gas, including air or perfluorated hydrocarbon can be triggered to release with ultrasound, e.g., low-frequency ultrasound (LFUS). Magnetic delivery: A lipid entity of the invention can be magnetized by incorporation of magnetites, such as Fe3O4 or γ-Fe2O3, e.g., those that are less than 10 nm in size. Targeted delivery can be then by exposure to a magnetic field.

Vesicle Production

[0587] Suitable host cells for virus production from viral vectors and systems thereof described herein are known in the art and are commercially available. They are also referred to herein and in the art as producer cells. For example, suitable host cells include HEK 293 cells and its variants (HEK 293T and HEK 293TN cells). In some embodiments, the suitable host cell for virus production from viral vectors and systems thereof described herein can stably express one or more genes involved in packaging (e.g. pol, gag, and/or VSV-G) and/or other supporting genes. Other cells for producer cells are described elsewhere herein.

[0588] It is possible to propagate and isolate amounts of retroviral vector particles (e.g., to prepare suitable titers of retroviral vector particles) for the subsequent transduction of, for example, a site of interest (such as adult brain tissue) using producer/packaging cell lines. Producer cell lines are usually better for large-scale production of vector particles. Transient transfection has several advantages over the cell packaging method. In this regard, transient transfection avoids the longer time required to produce stable vector producing cell lines and is used if the vector genome or the retroviral packaging components are toxic to cells. If the vector genome encodes toxic genes or genes that interfere with host cell replication, such as cell cycle inhibitors or apoptosis inducing genes, it may be difficult to produce stable vector producing cell lines, but transient transfection to produce the vector prior to cell death. Cell lines using transient infection that produce vector titer levels that are comparable to levels obtained from stable vector cell lines (Pear et al., 1993, PNAS 90: 8392-8396) have also been developed.

[0589] In some embodiments, the retroviral particles are lentiviral particles, as described elsewhere herein.

[0590] Populations of cells are described above. However, cells may be of any suitable cell type. The cells are generally mammalian cells but may be, for example, insect cells.

[0591] Suitable cells may further contain the elements necessary for the production of the infectious recombinant virus which are not present in the RNA genome. Typically, such packaging cells contain one or more producing plasmids that are capable of expressing viral structural proteins (such as codon optimized gag-pol and env) but these do not contain a packaging signal. The term "packaging signal", also referred to as "packaging sequence" is used in reference to the non-coding, cis-acting sequence required for encapsidation of retroviral RNA strands during formation of the viral particle. In HIV-1, this sequence was mapped to loci extending from the upstream main donor site (SD) up to at least the gag initiation codon.

[0592] Packaging cell lines suitable for use with the above-described vector constructs (see also WO 92/05266) can be used to create cell lines for the production of retroviral vector particles. It has been found that single packaging cell lines, comprising a provirus in which the packaging signal has been eliminated, originate rapid production by recombination of unwanted replication-competent viruses. To improve safety, a second generation of cell lines was produced in which the 3 'LTR of the provirus was deleted. In such cells, two recombinations would be required to produce a wild-type virus. A further improvement involves the introduction of the gag-pol genes and the env gene into separate constructs designated the third generation of packaging cell lines. These constructs are introduced in sequence to prevent recombination during transfection.

[0593] In some embodiments, the packaging cell lines are second generation packaging cell lines. In some embodiments, the packaging cell lines are third generation packaging cell lines.

[0594] In these separate constructions, third generation cell lines, further reduction of codon modification recombination can be performed. This technique, based on the redundancy of the genetic code, is intended to reduce the homology between the separate constructs, for example, between the overlapping regions in the open reading frames of gag-pol and env.

[0595] Packaging cell lines are useful for providing the gene products necessary to encapsidate and provide the membrane protein for the production of vector particles with a high titer. The packaging cell may be an in vitro cultured cell such as a tissue culture cell line.

Suitable cell lines include, but are not limited to, mammalian cells, such as murine fibroblast derived cell lines or human cell lines. Preferably, the packaging cell line is a primate or human cell line, such as for example: HEK293, 293T, 293FT, TE671, or HT1080. In specific embodiments, suitable cell lines comprise 293T or 293FT cells.

[0596] Alternatively, the packaging cell may be a cell derived from the subject being treated, such as a monocyte, macrophage, blood cell or fibroblast. The cell may be isolated from an individual and the packaging components and the vector be administered ex vivo followed by re-administration of the autologous packaging cells. It is highly desirable to use high titer virus preparations in experimental and practical applications.

[0597] Different cells differ in their use of particular codons. This codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. Expression can be increased by altering the codons in the sequence so that they are adjusted to match the relative abundance of the corresponding tRNAs. Likewise, it is possible to decrease expression by deliberate selection of codons whose corresponding tRNAs are known to be rare in the particular cell type. In this way, an additional degree of translation control is available.

[0598] Many viruses, including HIV and other lentiviruses, utilize a large number of rare codons and the increased expression of the packaging components in mammalian producing cells can be performed by modifying them to correspond to commonly used mammalian codons. Codon usage tables for mammalian cells as well as for various other organisms are known in the art. This so-called codon optimization has several other advantages. Because of changes in their sequences, the nucleotide sequences encoding the packaging components of the viral particles required for assembling viral particles in producer cells/packaging cells have the RNA instability sequences removed. Simultaneously, the amino acid coding sequence for the packaging components is conserved so that the viral components encoded by the sequences remain the same or at least sufficiently similar so that the function of the packaging components is not compromised. Codon optimization also circumvents the need for Rev/RRE for export, yielding optimized Rev-independent sequences. Codon optimization also reduces homologous recombination between different constructs within the vector system (e.g. between overlapping regions in open reading frames gag-pol and env). The overall effect of codon optimization is therefore a notable increase in viral titer and improved safety.

[0599] In some embodiments, after delivery of one or more viral vectors to the suitable host cells for or virus production from viral vectors and systems thereof, the cells are incubated

for an appropriate length of time to allow for viral gene expression from the vectors, packaging of the polynucleotide to be delivered (e.g. a cargo polynucleotide), and virus particle assembly, and secretion of mature virus particles into the culture media. Various other methods and techniques are generally known to those of ordinary skill in the art.

[0600] Mature virus particles can be collected from the culture media by a suitable method. In some embodiments, this can involve centrifugation to concentrate the virus. The titer of the composition containing the collected virus particles can be obtained using a suitable method. Such methods can include transducing a suitable cell line (e.g. NIH 3T3 cells) and determining transduction efficiency, infectivity in that cell line by a suitable method. Suitable methods include PCR-based methods, flow cytometry, and antibiotic selection-based methods. Various other methods and techniques are generally known to those of ordinary skill in the art. The concentration of virus particle can be adjusted as needed. In some embodiments, the resulting composition containing virus particles can contain 1 $X10^{1}$ -1 X 10^{20} particles/mL.

[0601] Lentiviruses may be prepared from any lentiviral vector or vector system described herein. In one example embodiment, after cloning pCasES10 (which contains a lentiviral transfer plasmid backbone), HEK293FT at low passage (p=5) can be seeded in a T-75 flask to 50% confluence the day before transfection in DMEM with 10% fetal bovine serum and without antibiotics. After 20 hours, the media can be changed to OptiMEM (serum-free) media and transfection of the lentiviral vectors can done 4 hours later. Cells can be transfected with 10 µg of lentiviral transfer plasmid (pCasES10) and the appropriate packaging plasmids (e.g., 5 µg of pMD2.G (VSV-g pseudotype), and 7.5ug of psPAX2 (gag/pol/rev/tat)). Transfection can be carried out in 4mL OptiMEM with a cationic lipid delivery agent (50uL Lipofectamine 2000 and 100ul Plus reagent). After 6 hours, the media can be changed to antibiotic-free DMEM with 10% fetal bovine serum. These methods can use serum during cell culture, but serum-free methods are preferred.

[0602] Following transfection and allowing the producing cells (also referred to as packaging cells) to package and produce virus particles with packaged cargo, the lentiviral particles can be purified. In an exemplary embodiment, virus-containing supernatants can be harvested after 48 hours. Collected virus-containing supernatants can first be cleared of debris and filtered through a 0.45um low protein binding (PVDF) filter. They can then be spun in an ultracentrifuge for 2 hours at 24,000 rpm. The resulting virus-containing pellets can be

resuspended in 50ul of DMEM overnight at 4 degrees C. They can be then aliquoted and used immediately or immediately frozen at -80 degrees C for storage.

METHODS OF USING THE ENGINEERED NUCLEIC ACIDS AND SYSTEMS THEREOF

Overview

[0603] Also provided herein are methods of diagnosing, prognosing, treating, and/or preventing a disease, state, or condition in or of a subject using the engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention. In some embodiments, the compositions, systems, and/or formulations described herein can be provided directly to a cell and/or nucleic acids and modify one or more cells and/or nucleic acids. In some embodiments, the compositions, systems, and/or formulations thereof of the present disclosure can deliver or otherwise provide one or more cargos that can and modify one or more cells and/or nucleic acids. Generally, the methods of diagnosing, prognosing, treating, and/or preventing a disease, state, or condition in or of a subject can include modifying a polynucleotide in a subject or cell thereof using a engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention and/or include detecting a diseased or healthy polynucleotide in a subject or cell thereof using a engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention system or component thereof. In some embodiments, the method of treatment or prevention can include using a system engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention or component thereof to modify a polynucleotide of an infectious organism (e.g., bacterial or virus) within a subject or cell thereof. In some embodiments, the method of treatment or prevention can include using a system engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention or component thereof to modify a polynucleotide of an infectious organism or symbiotic organism within a subject. The systems engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention and components thereof can be used to develop models of diseases, states, or conditions. The systems engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention and components thereof can be used to detect a

disease state or correction thereof, such as by a method of treatment or prevention described herein. The systems engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby of the present invention and components thereof can be used to screen and select cells that can be used, for example, as treatments or preventions described herein. The engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby and components thereof can be used to develop biologically active agents that can be used to modify one or more biologic functions or activities in a subject or a cell thereof.

Polynucleotide Modification

[0604] Also described herein are methods of modifying one or more nucleic acid components of cells, using the compositions, systems, and/or formulations described herein, such as the engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby. Nucleic acid components that can be modified include, without limitation, DNA and RNA, such as genomic DNA, nuclear RNA, and cytoplasmic DNA and RNA molecules. Modification includes insertions, deletions, additions, substitutions and combinations thereof. Such modification can be in the context of a genetic therapy for prevention or treatment of a disease, disorder, or condition of a human, non-human animal, plant, single celled eukaryote or bacteria. Such modification can also be in the context of a genetic enhancement, improvement, or other non-therapeutic benefit in the case of non-human animals, plants, single-celled eukaryotes, and prokaryotes the modification.

[0605] In some embodiments, polynucleotide modification can include the introduction, deletion, or substitution of 1-75 nucleotides at one or more positions, such as at one or more or each target sequence, of a polynucleotide of said cell(s). The modification can include the introduction, deletion, or substitution of 1, 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at one or more positions, such as at one or more or each target sequence(s), of a polynucleotide. The modification can include the introduction, deletion, or substitution of 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at one or more positions in a polynucleotide, such as at one or more or each target sequence(s). The modification can include the introduction, deletion, or substitution of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,

23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at one or more positions of a polynucleotide, such as at one or more or each target sequence(s). The modification can include the introduction, deletion, or substitution of 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, or 75 nucleotides at one or more positions of a polynucleotide, such as at one or more or each target sequence(s). The modification can include the introduction, deletion, or substitution of 40, 45, 50, 75, 100, 200, 300, 400 or 500 nucleotides at one or more positions of a polynucleotide, such as at one or more or each target sequence(s). The modification can include the introduction, deletion, or substitution of 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000, 5100, 5200, 5300, 5400, 5500, 5600, 5700, 5800, 5900, 6000, 6100, 6200, 6300, 6400, 6500, 6600, 6700, 6800, 6900, 7000, 7100, 7200, 7300, 7400, 7500, 7600, 7700, 7800, 7900, 8000, 8100, 8200, 8300, 8400, 8500, 8600, 8700, 8800, 8900, 9000, 9100, 9200, 9300, 9400, 9500, 9600, 9700, 9800, or 9900 to 10000 or more nucleotides at one or more positions of a polynucleotide, such as at one or more or each target sequence(s).

[0606] In some embodiments, an engineered composition or system of the present invention is capable of modifying a polynucleotide.

[0607] In some embodiments engineered composition or system of the present invention is an engineered delivery system and is configured to and capable of delivering a cargo, where the cargo is capable of modifying a polynucleotide. In some embodiments, the cargo is an RNA guided-nuclease system or component thereof. In some embodiments, the RNA guided nuclease system is a CRISPR-Cas system. Other polynucleotide systems that can be delivered as cargo by the engineered delivery system of the present invention are discussed elsewhere herein in greater detail.

[0608] In some embodiments, the method of polynucleotide modification also includes delivery of compositions or systems of the present invention or components there of that promote and/or inhibit the repair process of Non-Homologous End-Joining (NHEJ). In some embodiments, the method of polynucleotide modification includes delivery of compositions and systems in addition to those of the present invention that promote and/or inhibit Non-Homologous End-Joining (NHEJ). In some embodiments, the method of polynucleotide modification also includes delivery of compositions or systems of the present invention or

components there of that promote and/or inhibit the repair process of homology directed repair (HDR) In some embodiments, the method of polynucleotide modification includes delivery of compositions and systems in addition to those of the present invention that promote and/or inhibit HDR. See e.g., Vitor et al., Front. Mol. Biosci., 21 February 2020; Arras et al. PLoS One. 2016; 11(9): e0163049; Maruyama et al., Nat Biotechnol. 2015 May; 33(5): 538–542; Vartak et al. FEBS Journal 282 (2015) 4289–4294; Liu et al., Front. Genet., 07 January 2019 | https://doi.org/10.3389/fgene.2018.00691; Yang et al., Int. J. Mol. Sci. 2020, 21, 6461; doi:10.3390/ijms21186461; Nambiar et al., Nature Communications volume 10, Article number: 3395 (2019); Fu et al., Nucleic Acids Research, Volume 49, Issue 2, 25 January 2021, Pages 969–985, https://doi.org/10.1093/nar/gkaa1251; Ryu et al., BMB Reports 2019; 52(8): 475-481; Zhang et al., Genome Biology volume 18, Article number: 35 (2017); Riesenberg et al., Nat Commun. 2018; 9: 2164; Yu et al., Cell Stem Cell. 2015 Feb 5; 16(2): 142–147; Zhang et al., 2020. eLife 2020;9:e56008 DOI: 10.7554/eLife.56008; and Devkota, S. BMB Reports 2018; 51(9): 437-443 https://doi.org/10.5483/BMBRep.2018.51.9.187.

[0609] In general, the method can include delivering an engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby and/or component thereof to a subject or cell thereof, or to an infectious or symbiotic organism by a suitable delivery technique and/or composition or formulation. Once administered the components and/or cargos delivered by the compositions and/or system can function as described elsewhere herein to elicit a nucleic acid modification event or otherwise modify the cell to which it is delivered. In some embodiments, the nucleic acid modification event can occur at the genomic, epigenomic, and/or transcriptomic level. DNA and/or RNA cleavage, gene activation, and/or gene deactivation can occur as a result of action by the delivered composition, system, component thereof and/or cargo(s). Additional features, uses, and advantages are described in greater detail below. On the basis of this concept, several variations are appropriate to elicit a genomic locus event, including DNA cleavage, gene activation, or gene deactivation. Using the provided compositions, systems, and formulations, the person skilled in the art can advantageously and specifically target single or multiple loci with the same or different functional domains to elicit one or more genomic locus events. In addition to treating and/or preventing a disease in a subject, the compositions can be applied in a wide variety of methods for screening in libraries in cells and functional modeling in vivo (e.g., gene activation of lincRNA and identification of function; gain-of-function modeling; loss-of-

function modeling; the use the compositions of the invention to establish cell lines and transgenic animals for optimization and screening purposes).

Genetic Therapies

[0610] The engineered systems and components thereof described elsewhere herein can be used to treat and/or prevent a disease, such as a genetic and/or epigenetic disease, in a subject. It will be appreciated that the subject can be any subject suffering from a disease with a genetic or epigenetic component, infectious agent, and others that can be prevented or treated (either directly or indirectly) with a genetic or epigenetic modification. Modification can be provided by any composition or system capable of modifying DNA and/or RNA and/or expression and/or abundance thereof.

[0611] The engineered compositions and/or systems and/or components thereof described elsewhere herein can be used to treat and/or prevent infectious diseases in a subject, such as bacterial infections, viral infections, fungal infections, parasite infections, and combinations thereof. The engineered compositions and/or systems and/or components thereof described elsewhere herein can be used to modify the composition or profile of a microbiome in a subject, which can in turn modify the health status of the subject. The engineered compositions and/or systems and/or components thereof described herein can be used to modify cells *ex vivo*, which can then be administered to the subject whereby the modified cells can treat or prevent a disease or symptom thereof. Cell therapies are described in greater detail elsewhere herein. The engineered compositions and/or systems and/or components thereof described herein can be used to treat mitochondrial diseases, where the mitochondrial disease etiology involves a mutation in the mitochondrial DNA.

[0612] Also described herein are methods of inducing one or more polynucleotide modifications in a eukaryotic or prokaryotic cell or component thereof (e.g., a mitochondria) of a subject, infectious organism, and/or organism of the microbiome of the subject so as to treat or prevent a disease or symptom thereof. The method can generally include delivery of an engineered compositions, system, component thereof and/or one or more cargos to a eukaryotic cell, prokaryotic cell, population thereof or component thereof so as to modify the cell to which it is delivered.

[0613] Suitable subjects include those where it is desirable to treat or prevent a disease, disorder, and/or infection by a pathogen. In some embodiments, the subject is a human or a cell or cell population thereof. In some embodiments, the subject is a non-human animal or a

cell or cell population thereof. In some embodiments, the subject is a plant (including algae) or a cell or cell population thereof. In some embodiments, the subject is a single-cell eukaryote (e.g., yeast). In some embodiments, the subject is prokaryote where it is desirable to treat or prevent a disease, infection, condition or other disorder.

[0614] In some embodiments, a method of gene therapy includes administering an engineered system that is capable of or is capable of delivering a system capable of modifying a target polynucleotide to a subject in need thereof or cell thereof and manipulation of a target sequence within a coding, non-coding, or regulatory element of a genomic locus in a target sequence in a subject need thereof by the delivered engineered system capable of polynucleotide modification, where the condition or disease is susceptible to treatment or inhibition by manipulation of the target sequence by a delivered engineered system capable of polynucleotide modification.

[0615] The engineered composition, system, formulation thereof, or particle or vesicle produced therefrom of the present disclosure can be used in method of *ex vivo* or *in vivo* gene therapy and/or be used in an *in vitro*, *ex vivo* or *in vivo* gene therapy. The engineered composition, system, formulation thereof, or particle or vesicle produced therefrom of the present disclosure can be used in the manufacture of a medicament for *in vitro*, *ex vivo*, or *in vivo* gene therapy.

[0616] In some embodiments the method is or includes an individualized or personalized treatment of a disease or disorder. In some embodiments, aa method of individualized or personalized treatment of a genetic disease in a subject in need of such treatment can include: (a) introducing one or more mutations *ex vivo* in a tissue, organ, or a cell line, or *in vivo* in a transgenic non-human mammal or plant, where such introduction includes delivering to cell(s) of the tissue, organ, cell, plant, or non-human animal a composition comprising the engineered compositions or system thereof or component(s) thereof of the present invention, where the specific mutations or precise sequence substitutions are or have been correlated to the genetic disease of the subject; (b) testing treatment(s) for the genetic disease on the cell(s) to which the vector has been delivered that have the specific mutation(s) or precise sequence substitution(s) correlated to the genetic disease; and (c) treating the subject based on results from the testing of treatment(s) of step (b).

[0617] In some embodiments, a method of treating and/or preventing a genetic disease can include administering an engineered composition, system, formulation thereof, or particle or

vesicle produced therefrom of the present disclosure to a subject, where the engineered composition, system, formulation thereof, or particle or vesicle produced therefrom of the present disclosure and/or cargo delivered therefrom is capable of modifying one or more copies of one or more genes associated with the genetic disease or a disease with a genetic and/or epigenetic aspect in one or more cells of the subject. In some embodiments, modifying one or more copies of one or more genes associated with a genetic disease or a disease with a genetic and/or epigenetic aspect in the subject can eliminate a genetic disease or a symptom thereof in the subject. In some embodiments, modifying one or more copies of one or more genes associated with a genetic disease or a disease with a genetic and/or epigenetic aspect in the subject can decrease the severity of a genetic disease or a symptom thereof in the subject. In some embodiments, the engineered composition, system, formulation thereof, or particle or vesicle produced therefrom of the present disclosure and/or cargo delivered therefrom can modify one or more genes or polynucleotides associated with one or more diseases, including genetic diseases and/or those having a genetic aspect and/or epigenetic aspect, including but not limited to, any one or more set forth in **Tables 3 and 4** and those set forth at mitomap.org and described elsewhere herein. It will be appreciated that those diseases and associated genes listed herein are non-exhaustive and non-limiting. Further, it will be appreciated some genes play roles in the development of multiple diseases, thus treatment by correcting a defective or missing gene can provide therapy for one or more diseases or symptoms thereof.

[0618] In some embodiments, the subject treated with a gene therapy is a human. In some embodiments, the subject is a non-human animal. Exemplary, non-limiting, diseases and/or genes that can be the target of a gene therapy using the engineered compositions, systems and formulations thereof of the present invention are set forth in **Tables 3 and 4** as well as are set forth at mitomap.org, such as in the case of a mitochondrial disease identified as being caused by or otherwise attributed to a mtDNA mutation.

[0619] Further non-limiting examples of disease-associated genes and polynucleotides and disease specific information is available from McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), available on the World Wide Web.

Table 3. Exemplary Genetic and Other Diseases and Associated Genes

Disease Name	Primary Tissues or System Affected	Additional Tissues/Systems Affected	Genes

Achondroplasia	Bone and Muscle		fibroblast growth factor receptor 3 (FGFR3)
Achromatopsia	eye		CNGA3, CNGB3, GNAT2, PDE6C, PDE6H, ACHM2, ACHM3,
Acute Renal Injury	kidney		NFkappaB, AATF, p85alpha, FAS, Apoptosis cascade elements (e.g. FASR, Caspase 2, 3, 4, 6, 7, 8, 9, 10, AKT, TNF alpha, IGF1, IGF1R, RIPK1), p53
Age Related Macular Degeneration	eye		Abcr; CCL2; CC2; CP (ceruloplasmin); Timp3; cathepsinD; VLDLR, CCR2
AIDS	Immune System		KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1
Albinism (including oculocutaneous albinism (types 1-7) and ocular albinism)	Skin, hair, eyes,		TYR, OCA2, TYRP1, and SLC45A2, SLC24A5 and C10orf11
Alkaptonuria	Metabolism of amino acids	Tissues/organs where homogentisic acid accumulates, particularly cartilage (joints), heart valves, kidneys	HGD
alpha-1 antitrypsin deficiency (AATD or A1AD)	Lung	Liver, skin, vascular system, kidneys, GI	SERPINA1, those set forth in WO2017165862, PiZ allele
ALS	CNS		SOD1; ALS2; ALS3; ALS5; ALS7;STEX; FUS; TARDBP; VEGF (VEGF-a; VEGF-b; VEGF-c); DPP6; NEFH, PTGS1, SLC1A2, TNFRSF10B, PRPH, HSP90AA1, CRIA2, IFNG, AMPA2 S100B, FGF2, AOX1, CS, TXN, RAPHJ1, MAP3K5, NBEAL1, GPX1, ICA1L, RAC1, MAPT, ITPR2, ALS2CR4, GLS, ALS2CR8, CNTFR, ALS2CR11, FOLH1, FAM117B, P4HB, CNTF, SQSTM1, STRADB, NAIP, NLR, YWHAQ, SLC33A1, TRAK2, SCA1, NIF3L1, NIF3, PARD3B, COX8A, CDK15, HECW1, HECT, C2, WW 15, NOS1, MET, SOD2, HSPB1, NEFL, CTSB, ANG, HSPA8, RNase A, VAPB, VAMP, SNCA, alpha HGF, CAT, ACTB, NEFM, TH, BCL2, FAS, CASP3, CLU, SMN1, G6PD, BAX, HSF1, RNF19A, JUN, ALS2CR12, HSPA5, MAPK14, APEX1, TXNRD1, NOS2, TIMP1, CASP9, XIAP, GLG1, EPO, VEGFA, ELN, GDNF, NFE2L2, SLC6A3, HSPA4, APOE, PSMB8, DCTN2, TIMP3, KIFAP3, SLC1A1, SMN2, CCNC, STUB1, ALS2,

			PRDX6, SYP, CABIN1, CASP1, GART, CDK5, ATXN3, RTN4, C1QB, VEGFC, HTT, PARK7, XDH, GFAP, MAP2, CYCS, FCGR3B, CCS, UBL5, MMP9m SLC18A3, TRPM7, HSPB2, AKT1, DEERL1, CCL2, NGRN, GSR, TPPP3, APAF1, BTBD10, GLUD1, CXCR4, S:C1A3, FLT1, PON1, AR, LIF, ERBB3, :GA:S1, CD44, TP53, TLR3, GRIA1, GAPDH, AMPA, GRIK1, DES, CHAT, FLT4, CHMP2B, BAG1, CHRNA4, GSS, BAK1, KDR, GSTP1, OGG1, IL6
Alzheimer's Disease	Brain		E1; CHIP; UCH; UBB; Tau; LRP; PICALM; CLU; PS1; SORL1; CR1; VLDLR; UBA1; UBA3; CHIP28; AQP1; UCHL1; UCHL3; APP, AAA, CVAP, AD1, APOE, AD2, DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3, ALAS2, ABCA1, BIN1, BDNF, BTNL8, C1ORF49, CDH4, CHRNB2, CKLFSF2, CLEC4E, CR1L, CSF3R, CST3, CYP2C, DAPK1, ESR1, FCAR, FCGR3B, FFA2, FGA, GAB2, GALP, GAPDHS, GMPB, HP, HTR7, IDE, IF127, IFI6, IFIT2, IL1RN, IL-1RA, IL8RA, IL8RB, JAG1, KCNJ15, LRP6, MAPT, MARK4, MPHOSPH1, MTHFR, NBN, NCSTN, NIACR2, NMNAT3, NTM, ORM1, P2RY13, PBEF1, PCK1, PICALM, PLAU, PLXNC1, PRNP, PSEN1, PSEN2, PTPRA, RALGPS2, RGSL2, SELENBP1, SLC25A37, SORL1, Mitoferrin-1, TF, TFAM, TNF, TNFRSF10C, UBE1C
Amyloidosis			APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB
Amyloid neuropathy			TTR, PALB
Anemia	Blood		CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT
Angelman Syndrome	Nervous system, brain		UBE3A
Attention Deficit Hyperactivity Disorder (ADHD)	Brain		PTCHD1
Autoimmune lymphoproliferative syndrome	Immune system		TNFRSF6, APT1, FAS, CD95, ALPS1A
Autism, Autism spectrum disorders (ASDs), including Asperger's and a general diagnostic category called	Brain		PTCHD1; Mecp2; BZRAP1; MDGA2; Sema5A; Neurexin 1; GLO1, RTT, PPMX, MRX16, RX79, NLGN3, NLGN4, KIAA1260, AUTSX2, FMR1, FMR2; FXR1; FXR2;

Pervasive Developmental Disorders (PDDs)			MGLUR5, ATP10C, CDH10, GRM6, MGLUR6, CDH9, CNTN4, NLGN2, CNTNAP2, SEMA5A, DHCR7, NLGN4X, NLGN4Y, DPP6, NLGN5, EN2, NRCAM, MDGA2, NRXN1, FMR2, AFF2, FOXP2, OR4M2, OXTR, FXR1, FXR2, PAH, GABRA1, PTEN, GABRA5, PTPRZ1, GABRB3, GABRG1, HIRIP3, SEZ6L2, HOXA1, SHANK3, IL6, SHBZRAP1, LAMB1, SLC6A4, SERT, MAPK3, TAS2R1, MAZ, TSC1, MDGA2, TSC2, MECP2, UBE3A, WNT2, see also 20110023145
autosomal dominant polycystic kidney disease (ADPKD) - (includes diseases such as von Hippel-Lindau disease and tubreous sclerosis complex disease)	kidney	liver	PKD1, PKD2
Autosomal Recessive Polycystic Kidney Disease (ARPKD)	kidney	liver	PKDH1
Ataxia-Telangiectasia (a.k.a Louis Bar syndrome)	Nervous system, immune system	various	ATM
B-Cell Non-Hodgkin Lymphoma			BCL7A, BCL7
Bardet-Biedl syndrome	Eye, musculoskeletal system, kidney, reproductive organs	Liver, ear, gastrointestinal system, brain	ARL6, BBS1, BBS2, BBS4, BBS5, BBS7, BBS9, BBS10, BBS12, CEP290, INPP5E, LZTFL1, MKKS, MKS1, SDCCAG8, TRIM32, TTC8
Bare Lymphocyte Syndrome	blood		TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5
Bartter's Syndrome (types I, II, III, IVA and B, and V)	kidney		SLC12A1 (type I), KCNJ1 (type II), CLCNKB (type III), BSND (type IV A), or both the CLCNKA CLCNKB genes (type IV B), CASR (type V).
Becker muscular dystrophy	Muscle		DMD, BMD, MYF6
Best Disease (Vitelliform Macular Dystrophy type 2)	eye		VMD2
Bleeding Disorders	blood		TBXA2R, P2RX1, P2X1
Blue Cone Monochromacy	eye		OPN1LW, OPN1MW, and LCR
Breast Cancer	Breast tissue		BRCA1, BRCA2, COX-2
Bruton's Disease (aka X-linked Agammglobulinemia)	Immune system, specifically B cells		BTK
Cancers (e.g., lymphoma, chronic lymphocytic leukemia (CLL), B cell acute lymphocytic leukemia (B-ALL), acute lymphoblastic leukemia, acute myeloid leukemia, non-Hodgkin's lymphoma (NHL), diffuse large cell lymphoma (DLCL), multiple myeloma, renal cell carcinoma (RCC), neuroblastoma, colorectal	Various		FAS, BID, CTLA4, PDCD1, CBLB, PTPN6, TRAC, TRBC, those described in WO2015048577

cancer, breast cancer, ovarian cancer, melanoma, sarcoma, prostate cancer, lung cancer, esophageal cancer, hepatocellular carcinoma, pancreatic cancer, astrocytoma, mesothelioma, head and neck cancer, and medulloblastoma			
Cardiovascular Diseases	heart	Vascular system	IL1B, XDH, TP53, PTGS, MB, IL4, ANGPT1, ABCGu8, CTSK, PTGIR, KCNJ11, INS, CRP, PDGFRB, CCNA2, PDGFB, KCNJ5, KCNN3, CAPN10, ADRA2B, ABCG5, PRDX2, CPAN5, PARP14, MEX3C, ACE, RNF, IL6, TNF, STN, SERPINE1, ALB, ADIPOQ, APOB, APOE, LEP, MTHFR, APOA1, EDN1, NPPB, NOS3, PPARG, PLAT, PTGS2, CETP, AGTR1, HMGCR, IGF1, SELE, REN, PPARA, PON1, KNG1, CCL2, LPL, VWF, F2, ICAM1, TGFB, NPPA, IL10, EPO, SOD1, VCAM1, IFNG, LPA, MPO, ESR1, MAPK, HP, F3, CST3, COG2, MMP9, SERPINC1, F8, HMOX1, APOC3, IL8, PROL1, CBS, NOS2, TLR4, SELP, ABCA1, AGT, LDLR, GPT, VEGFA, NR3C2, IL18, NOS1, NR3C1, FGB, HGF, IL1A, AKT1, LIPC, HSPD1, MAPK14, SPP1, ITGB3, CAT, UTS2, THBD, F10, CP, TNFRSF11B, EGFR, MMP2, PLG, NPY, RHOD, MAPK8, MYC, FN1, CMA1, PLAU, GNB3, ADRB2, SOD2, F5, VDR, ALOX5, HLA-DRB1, PARP1, CD40LG, PON2, AGER, IRS1, PTGS1, ECE1, F7, IRMN, EPHX2, IGFBP1, MAPK10, FAS, ABCB1, JUN, IGFBP3, CD14, PDE5A, AGTR2, CD40, LCAT, CCR5, MMP1, TIMP1, ADM, DYT10, STAT3, MMP3, ELN, USF1, CFH, HSPA4, MMP12, MME, F2R, SELL, CTSB, ANXA5, ADRB1, CYBA, FGA, GGT1, LIPG, HIF1A, CXCR4, PROC, SCARB1, CD79A, PLTP, ADD1, FGG, SAA1, KCNH2, DPP4, NPR1, VTN, KIAA0101, FOS, TLR2, PPIG, IL1R1, AR, CYP1A1, SERPINA1, MTR, RBP4, APOA4, CDKN2A, FGF2, EDNRB, ITGA2, VLA-2, CABIN1, SHBG, HMGB1, HSP90B2P, CYP3A4, GJA1, CAV1, ESR2, LTA, GDF15, BDNF, CYP2D6, NGF, SP1, TGIF1, SRC, EGF, PIK3CG, HLA-A, KCNQ1, CNR1, FBN1, CHKA, BEST1, CTNNB1, IL2, CD36, PRKAB1, TPO,

			ALDH7A1, CX3CR1, TH, F9, CH1, TF, HFE, IL17A, PTEN, GSTM1, DMD, GATA4, F13A1, TTR, FABP4, PON3, APOC1, INSR, TNFRSF1B, HTR2A, CSF3, CYP2C9, TXN, CYP11B2, PTH, CSF2, KDR, PLA2G2A, THBS1, GCG, RHOA, ALDH2, TCF7L2, NFE2L2, NOTCH1, UGT1A1, IFNA1, PPARD, SIRT1, GNHR1, PAPPA, ARR3, NPPC, AHSP, PTK2, IL13, MTOR, ITGB2, GSTT1, IL6ST, CPB2, CYP1A2, HNF4A, SLC64A, PLA2G6, TNFSF11, SLC8A1, F2RL1, AKR1A1, ALDH9A1, BGLAP, MTTP, MTRR, SULT1A3, RAGE, C4B, P2RY12, RNLS, CREB1, POMC, RAC1, LMNA, CD59, SCM5A, CYP1B1, MIF, MMP13, TIMP2, CYP19A1, CUP21A2, PTPN22, MYH14, MBL2, SELPLG, AOC3, CTSL1, PCNA, IGF2, ITGB1, CAST, CXCL12, IGHE, KCNE1, TFRC, COL1A1, COL1A2, IL2RB, PLA2G10, ANGPT2, PROCR, NOX4, HAMP, PTPN11, SLCA1, IL2RA, CCL5, IRF1, CF:AR, CA:CA, EIF4E, GSTP1, JAK2, CYP3A5, HSPG2, CCL3, MYD88, VIP, SOAT1, ADRBK1, NR4A2, MMP8, NPR2, GCH1, EPRS, PPARGC1A, F12, PECAM1, CCL4, CERPINA34, CASR, FABP2, TTF2, PROS1, CTF1, SGCB, YME1L1, CAMP, ZC3H12A, AKR1B1, MMP7, AHR, CSF1, HDAC9, CTGF, KCNMA1, UGT1A, PRKCA, COMT, S100B, EGR1, PRL, IL15, DRD4, CAMK2G, SLC22A2, CCL11, PGF, THPO, GP6, TACR1, NTS, HNF1A, SST, KCDN1, LOC646627, TBXAS1, CUP2J2, TBXA2R, ADH1C, ALOX12, AHSG, BHMT, GJA4, SLC25A4, ACLY, ALOX5AP, NUMA1, CYP27B1, CYSLTR2, SOD3, LTC4S, UCN, GHRL, APOC2, CLEC4A, KBTBD10, TNC, TYMS, SHC1, LRP1, SOCS3, ADH1B, KLK3, HSD11B1, VKORC1, SERPINB2, TNS1, RNF19A, EPOR, ITGAM, PITX2, MAPK7, FCGR3A, LEEPR, ENG, GPX1, GOT2, HRH1, NR112, CRH, HTR1A, VDAC1, HPSE, SFTPD, TAP2, RMF123, PTK2Bm NTRK2, IL6R, ACHE, GLP1R, GHR, GSR, NQO1, NR5A1, GJB2, SLC9A1, MAOA, PCSK9, FCGR2A, SERPINF1, EDN3, UCP2, TFAP2A,

			C4BPA, SERPINF2, TYMP, ALPP, CXCR2, SLC3A3, ABCG2, ADA, JAK3, HSPA1A, FASN, FGF1, F11, ATP7A, CR1, GFPA, ROCK1, MECP2, MYLK, BCHE, LIPE, ADORA1, WRN, CXCR3, CD81, SMAD7, LAMC2, MAP3K5, CHGA, IAPP, RHO, ENPP1, PTHLH, NRG1, VEGFC, ENPEP, CEBPB, NAGLU,. F2RL3, CX3CL1, BDKRB1, ADAMTS13, ELANE, ENPP2, CISH, GAST, MYOC, ATP1A2, NF1, GJB1, MEF2A, VCL, BMPR2, TUBB, CDC42, KRT18, HSF1, MYB, PRKAA2, ROCK2, TFP1, PRKG1, BMP2, CTNND1, CTH, CTSS, VAV2, NPY2R, IGFBP2, CD28, GSTA1, PPIA, APOH, S100A8, IL11, ALOX15, FBLN1, NR1H3, SCD, GIP, CHGB, PRKCB, SRD5A1,HSD11B2, CALCRL, GALNT2, ANGPTL4, KCNN4, PIK3C2A, HBEGF, CYP7A1, HLA-DRB5, BNIP3, GCKR, S100A12, PADI4, HSPA14, CXCR1, H19, KRTAP19-3, IDDM2, RAC2, YRY1, CLOCK, NGFR, DBH, CHRNA4, CACNA1C, PRKAG2, CHAT, PTGDS, NR1H2, TEK, VEGFB, MEF2C, MAPKAPK2, TNFRSF11A, HSPA9, CYSLTR1, MAT1A, OPRL1, IMPA1, CLCN2, DLD, PSMA6, PSMB8, CHI3L1, ALDH1B1, PARP2,STAR, LBP, ABCC6, RGS2, EFNB2, GJB6, APOA2, AMPD1, DYSF, FDFT1,EMD2, CCR6, GJB3, IL1RL1, ENTPD1, BBS4, CELSR2, F11R, RAPGEF3, HYAL1, ZNF259, ATOX1, ATF6, KHK, SAT1, GGH, TIMP4, SLC4A4, PDE2A, PDE3B, FADS1, FADS2, TMSB4X, TXNIP, LIMS1, RHOB, LY96, FOXO1, PNPLA2,TRH, GJC1, S:C17A5, FTO, GJD2, PRSC1, CASP12, GPBAR1, PXK, IL33, TRIB1, PBX4, NUPR1, 15-SEP, CILP2, TERC, GGT2, MTCO1, UOX, AVP, ANGPLT3
Cataract	eye		CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1,

			GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1
CDKL-5 Deficiencies or Mediated Diseases	Brain, CNS		CDKL5
Charcot-Marie-Tooth (CMT) disease (Types 1, 2, 3, 4,)	Nervous system	Muscles (dystrophy)	PMP22 (CMT1A and E), MPZ (CMT1B), LITAF (CMT1C), EGR2 (CMT1D), NEFL (CMT1F), GJB1 (CMT1X), MFN2 (CMT2A), KIF1B (CMT2A2B), RAB7A (CMT2B), TRPV4 (CMT2C), GARS (CMT2D), NEFL (CMT2E), GAPD1 (CMT2K), HSPB8 (CMT2L), DYNC1H1, CMT2O), LRSAM1 (CMT2P), IGHMBP2 (CMT2S), MORC2 (CMT2Z), GDAP1 (CMT4A), MTMR2 or SBF2/MTMR13 (CMT4B), SH3TC2 (CMT4C), NDRG1 (CMT4D), PRX (CMT4F), FIG4 (CMT4J), NT-3
Chédiak-Higashi Syndrome	Immune system	Skin, hair, eyes, neurons	LYST
Choroidermia			CHM, REP1,
Chorioretinal atrophy	eye		PRDM13, RGR, TEAD1
Chronic Granulomatous Disease	Immune system		CYBA, CYBB, NCF1, NCF2, NCF4
Chronic Mucocutaneous Candidiasis	Immune system		AIRE, CARD9, CLEC7A IL12B, IL12B1, IL1F, IL17RA, IL17RC, RORC, STAT1, STAT3, TRAF31P2
Cirrhosis	liver		KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988
Colon cancer (Familial adenomatous polyposis (FAP) and hereditary nonpolyposis colon cancer (HNPCC))	Gastrointestinal		FAP: APC HNPCC: MSH2, MLH1, PMS2, SH6, PMS1
Combined Immunodeficiency	Immune System		IL2RG, SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228
Cone(-rod) dystrophy	eye		AIPL1, CRX, GUA1A, GUCY2D, PITPM3, PROM1, PRPH2, RIMS1, SEMA4A, ABCA4, ADAM9, ATF6, C21ORF2, C8ORF37, CACNA2D4, CDHR1, CERKL, CNGA3, CNGB3, CNNM4, CNAT2, IFT81, KCNV2, PDE6C, PDE6H, POC1B, RAX2, RDH5, RPGRIP1, TTLL5, RetCG1, GUCY2E
Congenital Stationary Night Blindness	eye		CABP4, CACNA1F, CACNA2D4, GNAT1, CPR179, GRK1, GRM6, LRIT3, NYX, PDE6B, RDH5, RHO, RLBP1, RPE65, SAG, SLC24A1, TRPM1,
Congenital Fructose Intolerance	Metabolism		ALDOB
Cori's Disease (Glycogen Storage Disease Type III)	Various-wherever glycogen accumulates, particularly		AGL

	liver, heart, skeletal muscle		
Corneal clouding and dystrophy	eye		APOA1, TGFBI, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD
Cornea plana congenital			KERA, CNA2
Cri du chat Syndrome, also known as 5p syndrome and cat cry syndrome			Deletions involving only band 5p15.2 to the entire short arm of chromosome 5, e.g. CTNND2, TERT,
Cystic Fibrosis (CF)	Lungs and respiratory system	Pancreas, liver, digestive system, reproductive system, exocrine, glands,	CTFR, ABCC7, CF, MRP7, SCNN1A, those described in WO2015157070
Diabetic nephropathy	kidney		Gremlin, 12/15- lipoxygenase, TIM44,
Dent Disease (Types 1 and 2)	Kidney		Type 1: CLCN5, Type 2: ORCL
Dentatorubro-Pallidoluysian Atrophy (DRPLA) (aka Haw River and Naito-Oyanagi Disease)	CNS, brain, muscle		Atrophin-1 and Atn1
Down Syndrome	various		Chromosome 21 trisomy
Drug Addiction	Brain		Prkce; Drd2; Drd4; ABAT; GRIA2;Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1
Duane syndrome (Types 1, 2, and 3, including subgroups A, B and C). Other names for this condition include: Duane's Retraction Syndrome (or DR syndrome), Eye Retraction Syndrome, Retraction Syndrome, Congenital retraction syndrome and Stilling-Turk-Duane Syndrome	eye		CHN1, indels on chromosomes 4 and 8
Duchenne muscular dystrophy (DMD)	muscle	Cardiovascular, respiratory	DMD, BMD, dystrophin gene, intron flanking exon 51 of DMD gene, exon 51 mutations in DMD gene, see also WO2013163628 and US Pat. Pub. 20130145487
Edward's Syndrome (Trisomy 18)			Complete or partial trisomy of chromosome 18
Ehlers-Danlos Syndrome (Types I-VI)	Various depending on type: including musculoskeletal, eye, vasculature, immune, and skin		COL5A1, COL5A2, COL1A1, COL3A1, TNXB, PLOD1, COL1A2, FKBP14 and ADAMTS2
Emery-Dreifuss muscular dystrophy	muscle		LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A
Enhanced S-Cone Syndrome	eye		NR2E3, NRL
Fabry's Disease	Various – including skin,		GLA

	eyes, and gastrointestinal system, kidney, heart, brain, nervous system		
Facioscapulohumeral muscular dystrophy	muscles		FSHMD1A, FSHD1A, FRG1,
Factor H and Factor H-like 1	blood		HF1, CFH, HUS
Factor V Leiden thrombophilia and Factor V deficiency	blood		Factor V (F5)
Factor V and Factor VII deficiency	blood		MCFD2
Factor VII deficiency	blood		F7
Factor X deficiency	blood		F10
Factor XI deficiency	blood		F11
Factor XII deficiency	blood		F12, HAF
Factor XIIIA deficiency	blood		F13A1, F13A
Factor XIIIB deficiency	blood		F13B
Familial Hypercholestereolemia	Cardiovascular system		APOB, LDLR, PCSK9
Familial Mediterranean Fever (FMF) also called recurrent polyserositis or familial paroxysmal polyserositis	Various-organs/tissues with serous or synovial membranes, skin, joints	Heart, kidney, brain/CNS, reproductive organs	MEFV
Fanconi Anemia	Various – blood (anemia), immune system, cognitive, kidneys, eyes, musculoskeletal		FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCC, FANCG, RAD51, BRCA1, BRCA2, BRIP1, BACH1, FANCJ, FANCB, FANCD1, FANCD2, FANCD, FAD, FANCE, FACE, FANCF, FANCI, ERCC4, FANCL, FANCM, PALB2, RAD51C, SLX4, UBE2T, FANCB, XRCC9, PHF9, KIAA1596
Fanconi Syndrome Types I (Childhood onset) and II (Adult Onset)	kidneys		FRTS1, GATM
Fragile X syndrome and related disorders	brain		FMR1, FMR2; FXR1; FXR2; mGLUR5
Fragile XE Mental Retardation (aka Martin Bell syndrome)	Brain, nervous system		FMR1
Friedreich Ataxia (FRDA)	Brain, nervous system	heart	FXN/X25
Fuchs endothelial corneal dystrophy	Eye		TCF4; COL8A2
Galactosemia	Carbohydrate metabolism disorder	Various-where galactose accumulates – liver, brain, eyes	GALT, GALK1, and GALE
Gastrointestinal Epithelial Cancer, GI cancer			CISH
Gaucher Disease (Types 1, 2, and 3, as well as other unusual forms that may not fit into these types)	Fat metabolism disorder	Various-liver, spleen, blood, CNS, skeletal system	GBA

Griscelli syndrome			
Glaucoma	eye		MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A, those described in WO2015153780
Glomerulo sclerosis	kidney		CC chemokine ligand 2
Glycogen Storage Diseases Types I-VI -See also Cori's Disease, Pompe's Disease, McArdle's disease, Hers Disease, and Von Gierke's disease	Metabolism Diseases		SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM, see also Cori's Disease, Pompe's Disease, McArdle's disease, Hers Disease, and Von Gierke's disease
RBC Glycolytic enzyme deficiency	blood		any mutations in a gene for an enzyme in the glycolysis pathway including mutations in genes for hexokinases I and II, glucokinase, phosphoglucose isomerase, phosphofructokinase, aldolase Bm triosephosphate isomerease, glyceraldehydee-3-phosphate dehydrogenase, phosphoglycerokinase, phosphoglycerate mutase, enolase I, pyruvate kinase
Hartnup's disease	Malabsorption disease	Various- brain, gastrointestinal, skin,	SLC6A19
Hearing Loss	ear		NOX3, Hes5, BDNF,
Hemochromatosis (HH)	Iron absorption regulation disease	Various- wherever iron accumulates, liver, heart, pancreas, joints, pituitary gland	HFE and H63D
Hemophagocytic lymphohistiocytosis disorders	blood		PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3
Hemorrhagic disorders	blood		PI, ATT, F5
Hers disease (Glycogen storage disease Type VI)	liver	muscle	PYGL
Hereditary angioedema (HAE)			kalikrein B1
Hereditary Hemorrhagic Telangiectasia (Osler-Weber-Rendu Syndrome)	Skin and mucous membranes		ACVRL1, ENG and SMAD4
Hereditary Spherocytosis	blood		NK1, EPB42, SLC4A1, SPTA1, and SPTB
Hereditary Persistence of Fetal Hemoglobin	blood		HBG1, HBG2, BCL11A, promoter region of HBG 1 and/or 2 (in the CCAAT box)
Hemophilia (hemophilia A (Classic) a B (aka Christmas disease) and C)	blood		A: FVIII, F8C, HEMA B: FVIX, HEMB, FIX C: F9, F11
Hepatic adenoma	liver		TCF1, HNF1A, MODY3
Hepatic failure, early onset, and neurologic disorder	liver		SCOD1, SCO1
Hepatic lipase deficiency	liver		LIPC

Hepatoblastoma, cancer and carcinomas	liver		CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5
Hermansky-Pudlak syndrome	Skin, eyes, blood, lung, kidneys, intestine		HPS1, HPS3, HPS4, HPS5, HPS6, HPS7, DTNBP1, BLOC1, BLOC1S2, BLOC3
HIV susceptibility or infection	Immune system		IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCKR5 (CCR5), those in WO2015148670A1
Holoprosencephaly (HPE) (Alobar, Semilobar, and Lobar)	brain		ACVRL1, ENG, SMAD4
Homocystinuria	Metabolic disease	Various- connective tissue, muscles, CNS, cardiovascular system	CBS, MTHFR, MTR, MTRR, and MMADHC
HPV			HPV16 and HPV18 E6/E7
HSV1, HSV2, and related keratitis	eye		HSV1 genes (immediate early and late HSV-1 genes (UL1, 1.5, 5, 6, 8, 9, 12, 15, 16, 18, 19, 22, 23, 26, 26.5, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 42, 48, 49.5, 50, 52, 54, S6, RL2, RS1, those described in WO2015153789, WO2015153791
Hunter's Syndrome (aka Mucopolysaccharidosis type II)	Lysosomal storage disease	Various- liver, spleen, eye, joint, heart, brain, skeletal	IDS
Huntington's disease (HD) and HD-like disorders	Brain, nervous system		HD, HTT, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17, PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; and TGM2, and those described in WO2013130824, WO2015089354
Hurler's Syndrome (aka mucopolysaccharidosis type I H, MPS IH)	Lysosomal storage disease	Various- liver, spleen, eye, joint, heart, brain, skeletal	IDUA, α-L-iduronidase
Hurler-Scheie syndrome (aka mucopolysaccharidosis type I H-S, MPS I H-S)	Lysosomal storage disease	Various- liver, spleen, eye, joint, heart, brain, skeletal	IDUA, α-L-iduronidase
hyaluronidase deficiency (aka MPS IX)	Soft and connective tissues		HYAL1
Hyper IgM syndrome	Immune system		CD40L
Hyper- tension caused renal damage	kidney		Mineral corticoid receptor
Immunodeficiencies	Immune System		CD3E, CD3G, AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TACI

Inborn errors of metabolism: including urea cycle disorders, organic acidemias), fatty acid oxidation defects, amino acidopathies, carbohydrate disorders, mitochondrial disorders	Metabolism diseases, liver	Various organs and cells	See also: Carbohydrate metabolism disorders (e.g. galactosemia), Amino acid Metabolism disorders (e.g. phenylketonuria), Fatty acid metabolism (e.g. MCAD deficiency), Urea Cycle disorders (e.g. Citrullinemia), Organic acidemias (e.g. Maple Syrup Urine disease), Mitochondrial disorders (e.g. MELAS), peroxisomal disorders (e.g. Zellweger syndrome)
Inflammation	Various		IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-17b; IL-17c; IL-17d; IL-17f); II-23; Cx3cr1; ptpn22; TNFa; NOD2/CARD15 for IBD; IL-6; IL-12 (IL-12a; IL-12b); CTLA4; Cx3cl1
Inflammatory Bowel Diseases (e.g. Ulcerative Colitis and Chron's Disease)	Gastrointestinal	Joints, skin	NOD2, IRGM, LRRK2, ATG5, ATG16L1, IRGM, GATM, ECM1, CDH1, LAMB1, HNF4A, GNA12, IL10, CARD9/15. CCR6, IL2RA, MST1, TNFSF15, REL, STAT3, IL23R, IL12B, FUT2
Interstitial renal fibrosis	kidney		TGF-β type II receptor
Job's Syndrome (aka Hyper IgE Syndrome)	Immune System		STAT3, DOCK8
Juvenile Retinoschisis	eye		RS1, XLRS1
Kabuki Syndrome 1			MLL4, KMT2D
Kennedy Disease (aka Spinobulbar Muscular Atrophy)	Muscles, brain, nervous system		SBMA/SMAX1/AR
Klinefelter syndrome	Various-particularly those involved in development of male characteristics		Extra X chromosome in males
Lafora Disease	Brain, CNS		EMP2A and EMP2B
Leber Congenital Amaurosis	eye		CRB1, RP12, CORD2, CRD, CRX, IMPDH1, OTX2, AIPL1, CABP4, CCT2, CEP290, CLUAP1, CRB1, CRX, DTHD1, GDF6, GUCY2D, IFT140, IQCB1, KCNJ13, LCA5, LRAT, NMNAT1, PRPH2, RD3, RDH12, RPE65, RP20, RPGRIP1, SPATA7, TULP1, LCA1, LCA4, GUC2D, CORD6, LCA3,
Lesch-Nyhan Syndrome	Metabolism disease	Various – joints, cognitive, brain, nervous system	HPRT1
Leukocyte deficiencies and disorders	blood		ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH, CLE, EIF2B4
Leukemia	Blood		TAL1, TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL,

			ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B, AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN
Limb-girdle muscular dystrophy diseases	muscle		LGMD
Lowe syndrome	brain, eyes, kidneys		OCRL
Lupus glomerulo- nephritis	kidney		MAPK1
Machado-Joseph's Disease (also known as Spinocerebellar ataxia Type 3)	Brain, CNS, muscle		ATX3
Macular degeneration	eye		ABC4, CBC1, CHM1, APOE, C1QTNF5, C2, C3, CCL2, CCR2, CD36, CFB, CFH, CFHR1, CFHR3, CNGB3, CP, CRP, CST3, CTSD, CX3CR1, ELOVL4, ERCC6, FBLN5, FBLN6, FSCN2, HMCN1, HTRA1, IL6, IL8, PLEKHA1, PROM1, PRPH2, RPGR, SERPING1, TCOF1, TIMP3, TLR3
Macular Dystrophy	eye		BEST1, C1QTNF5, CTNNA1, EFEMP1, ELOVL4, FSCN2, GUCA1B, HMCN1, IMPG1, OTX2, PRDM13, PROM1, PRPH2, RP1L1, TIMP3, ABCA4, CFH, DRAM2, IMG1, MFSD8, ADMD, STGD2, STGD3, RDS, RP7, PRPH, AVMD, AOFMD, VMD2
Malattia Leventinesse	eye		EFEMP1, FBLN3
Maple Syrup Urine Disease	Metabolism disease		BCKDHA, BCKDHB, and DBT
Marfan syndrome	Connective tissue	Musculoskeletal	FBN1
Maroteaux-Lamy Syndrome (aka MPS VI)	Musculoskeletal system, nervous system	Liver, spleen	ARSB
McArdle's Disease (Glycogen Storage Disease Type V)	Glycogen storage disease	muscle	PYGM
Medullary cystic kidney disease	kidney		UMOD, HNFJ, FJHN, MCKD2, ADMCKD2
Metachromatic leukodystrophy	Lysosomal storage disease	Nervous system	ARSA

Methylmalonic acidemia (MMA)	Metabolism disease		MMAA, MMAB, MUT, MMACHC, MMADHC, LMBRD1
Morquio Syndrome (aka MPS IV A and B)	Connective tissue, skin, bone, eyes	heart	GALNS
Mucopolysaccharidosis diseases (Types I H/S, I H, II, III A B and C, I S, IVA and B, IX, VII, and VI)	Lysosomal storage disease – affects various organs/tissues		See also Hurler/Scheie syndrome, Hurler disease, Sanfillipo syndrome, Scheie syndrome, Morquio syndrome, hyaluronidase deficiency, Sly syndrome, and Maroteaux-Lamy syndrome
Muscular Atrophy	muscle		VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1
Muscular dystrophy	muscle		FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPN1, SELN, RSMD1, PLEC1, PLTN, EBS1
Myotonic dystrophy (Type 1 and Type 2)	Muscles	Eyes, heart, endocrine	CNBP (Type 2) and DMPK (Type 1)
Neoplasia			PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4; Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF; HIF1a; HIF3a; Met; HRG; Bcl2; PPAR alpha; PPAR gamma; WT1 (Wilms Tumor); FGF Receptor Family members (5 members: 1, 2, 3, 4, 5); CDKN2a; APC; RB (retinoblastoma); MEN1; VHL; BRCA1; BRCA2; AR (Androgen Receptor); TSG101; IGF; IGF Receptor; Igf1 (4 variants); Igf2 (3 variants); Igf 1 Receptor; Igf 2 Receptor; Bax; Bcl2; caspases family (9 members: 1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc
Neurofibromatosis (NF) (NF1, formerly Recklinghausen's NF, and NF2)	brain, spinal cord, nerves, and skin		NF1, NF2
Niemann-Pick Lipidosis (Types A, B, and C)	Lysosomal Storage Disease	Various- where sphingomyelin accumulates,	Types A and B: SMPD1; Type C: NPC1 or NPC2

		particularly spleen, liver, blood, CNS	
Noonan Syndrome	Various – musculoskeletal, heart, eyes, reproductive organs, blood		PTPN11, SOS1, RAF1 and KRAS
Norrie Disease or X-linked Familial Exudative Vitreoretinopathy	eye		NDP
North Carolina Macular Dystrophy	eye		MCDR1
Osteogenesis imperfecta (OI) (Types I, II, III, IV, V, VI, VII)	bones, musculoskeletal		COL1A1, COL1A2, CRTAP, P3H
Osteopetrosis	bones		LRP5, BMND1, LRP7, LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1
Patau's Syndrome (Trisomy 13)	Brain, heart, skeletal system		Additional copy of chromosome 13
Parkinson's disease (PD)	Brain, nervous system		SNCA (PARK1), UCHL1 (PARK 5), and LRRK2 (PARK8), (PARK3), PARK2, PARK4, PARK7 (PARK7), PINK1 (PARK6); x-Synuclein, DJ-1, Parkin, NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, NCAP, PRKN, PDJ, DBH, NDUFV2
Pattern Dystrophy of the RPE	eye		RDS/peripherin
Phenylketonuria (PKU)	Metabolism disorder	Various due to build-up of phenylalanine, phenyl ketones in tissues and CNS	PAH, PKU1, QDPR, DHPR, PTS
Polycystic kidney and hepatic disease	Kidney, liver		FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63
Pompe's Disease	Glycogen storage disease	Various – heart, liver, spleen	GAA
Porphyria (actually refers to a group of different diseases all having a specific heme production process abnormality)	Various-wherever heme precursors accumulate		ALAD, ALAS2, CPOX, FECH, HMBS, PPOX, UROD, or UROS
posterior polymorphous corneal dystrophy	eyes		TCF4; COL8A2
Primary Hyperoxaluria (e.g. type 1)	Various – eyes, heart, kidneys, skeletal system		LDHA (lactate dehydrogenase A) and hydroxyacid oxidase 1 (HAO1)
Primary Open Angle Glaucoma (POAG)	eyes		MYOC
Primary sclerosing cholangitis	Liver, gallbladder		TCF4; COL8A2
Progeria (also called Hutchinson-Gilford progeria syndrome)	All		LMNA
Prader-Willi Syndrome	Musculoskeletal system, brain,		Deletion of region of short arm of chromosome 15, including UBE3A

	reproductive and endocrine system		
Prostate Cancer	prostate		HOXB13, MSMB, GPRC6A, TP53
Pyruvate Dehydrogenase Deficiency	Brain, nervous system		PDHA1
Kidney/Renal carcinoma	kidney		RLIP76, VEGF
Rett Syndrome	Brain		MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-Synuclein, DJ-1
Retinitis pigmentosa (RP)	eye		ADIPOR1, ABCA4, AGBL5, ARHGEF18, ARL2BP, ARL3, ARL6, BEST1, BBS1, BBS2, C2ORF71, C8ORF37, CA4, CERKL, CLRN1, CNGA1, CMGB1, CRB1, CRX, CYP4V2, DHDDS, DHX38, EMC1, EYS, FAM161A, FSCN2, GPR125, GUCA1B, HK1, HPRPF3, HGSNAT, IDH3B, IMPDH1, IMPG2, IFT140, IFT172, KLHL7, KIAA1549, KIZ, LRAT, MAK, MERTK, MVK, NEK2, NUROD1, NR2E3, NRL, OFD1, PDE6A, PDE6B, PDE6G, POMGNT1, PRCD, PROM1, PRPF3, PRPF4, PRPF6, PRPF8, PRPF31, PRPH2, RPB3, RDH12, REEP6, RP39, RGR, RHO, RLBP1, ROM1, RP1, RP1L1, RPY, RP2, RP9, RPE65, RPGR, SAMD11, SAG, SEMA4A, SLC7A14, SNRNP200, SPP2, SPATA7, TRNT1, TOPORS, TTC8, TULP1, USH2A, ZFN408, ZNF513, see also 20120204282
Scheie syndrome (also known as mucopolysaccharidosis type I S(MPS I-S))	Various- liver, spleen, eye, joint, heart, brain, skeletal		IDUA, α-L-iduronidase
Schizophrenia	Brain		Neuregulin1 (Nrg1); Erb4 (receptor for Neuregulin); Complexin1 (Cplx1); Tph1 Tryptophan hydroxylase; Tph2 Tryptophan hydroxylase 2; Neurexin 1; GSK3; GSK3a; GSK3b; 5-HTT (Slc6a4); COMT; DRD (Drd1a); SLC6A3; DAOA; DTNBP1; Dao (Dao1); TCF4; COL8A2
Secretase Related Disorders	Various		APH-1 (alpha and beta); PSEN1; NCSTN; PEN-2; Nos1, Parp1, Nat1, Nat2, CTSB, APP, APH1B, PSEN2, PSENEN, BACE1, ITM2B, CTSD, NOTCH1, TNF, INS, DYT10, ADAM17, APOE, ACE, STN, TP53, IL6, NGFR, IL1B, ACHE, CTNNB1, IGF1, IFNG, NRG1, CASP3, MAPK1, CDH1, APBB1, HMGCR, CREB1,

			PTGS2, HES1, CAT, TGFB1, ENO2, ERBB4, TRAPPC10, MAOB, NGF, MMP12, JAG1, CD40LG, PPARG, FGF2, LRP1, NOTCH4, MAPK8, PREP, NOTCH3, PRNP, CTSG, EGF, REN, CD44, SELP, GHR, ADCYAP1, INSR, GFAP, MMP3, MAPK10, SP1, MYC, CTSE, PPARA, JUN, TIMP1, IL5, IL1A, MMP9, HTR4, HSPG2, KRAS, CYCS, SMG1, IL1R1, PROK1, MAPK3, NTRK1, IL13, MME, TKT, CXCR2, CHRM1, ATXN1,PAWR, NOTCJ2, M6PR, CYP46A1, CSNK1D, MAPK14, PRG2, PRKCA, L1 CAM, CD40, NR1I2, JAG2, CTNND1, CMA1, SORT1, DLK1, THEM4, JUP, CD46, CCL11, CAV3, RNASE3, HSPA8, CASP9, CYP3A4, CCR3, TFAP2A, SCP2, CDK4, JOF1A, TCF7L2, B3GALTL, MDM2, RELA, CASP7, IDE, FANP4, CASK, ADCYAP1R1, ATF4, PDGFA, C21ORF33, SCG5, RMF123, NKFB1, ERBB2, CAV1, MMP7, TGFA, RXRA, STX1A, PSMC4, P2RY2, TNFRSF21, DLG1, NUMBL, SPN, PLSCR1, UBQLN2, UBQLN1, PCSK7, SPON1, SILV, QPCT, HESS, GCC1
Selective IgA Deficiency	Immune system		Type 1: MSH5; Type 2: TNFRSF13B
Severe Combined Immunodeficiency (SCID) and SCID-X1, and ADA-SCID	Immune system		JAK3, JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4, those identified in US Pat. App. Pub. 20110225664, 20110091441, 20100229252, 20090271881 and 20090222937;
Sickle cell disease	blood		HBB, BCL11A, BCL11Ae, cis-regulatory elements of the B-globin locus, HBG 1/2 promoter, HBG distal CCAAT box region between -92 and -130 of the HBG Transcription Start Site, those described in WO2015148863, WO 2013/126794, US Pat. Pub. 20110182867
Sly Syndrome (aka MPS VII)			GUSB
Spinocerebellar Ataxias (SCA types 1, 2, 3, 6, 7, 8, 12 and 17)			ATXN1, ATXN2, ATX3
Sorsby Fundus Dystrophy	eye		TIMP3
Stargardt disease	eye		ABCR, ELOVL4, ABCA4, PROM1
Tay-Sachs Disease	Lysosomal Storage disease	Various – CNS, brain, eye	HEX-A
Thalassemia (Alpha, Beta, Delta)	blood		HBA1, HBA2 (Alpha), HBB (Beta), HBB and HBD (delta), LCRB, BCL11A, BCL11Ae, cis-regulatory elements of the B-globin locus, HBG

			1/2 promoter, those described in WO2015148860, US Pat. Pub. 20110182867, 2015/148860
Thymic Aplasia (DiGeorge Syndrome;22q11.2 deletion syndrome)	Immune system, thymus		deletion of 30 to 40 genes in the middle of chromosome 22 at a location known as 22q11.2, including TBX1, DGCR8
Transthyretin amyloidosis (ATTR)	liver		TTR (transthyretin)
trimethylaminuria	Metabolism disease		FMO3
Trinucleotide Repeat Disorders (generally)	Various		HTT; SBMA/SMAX1/AR; FXN/X25 ATX3; ATXN1; ATXN2; DMPK; Atrophin-1 and Atn1 (DRPLA Dx); CBP (Creb-BP - global instability); VLDLR; Atxn7; Atxn10; FEN1, TNRC6A, PABPN1, JPH3, MED15, ATXN1, ATXN3, TBP, CACNA1A, ATXN80S, PPP2R2B, ATXN7, TNRC6B, TNRC6C, CELF3, MAB21L1, MSH2, TMEM185A, SIX5, CNPY3, RAXE, GNB2, RPL14, ATXN8, ISR, TTR, EP400, GIGYF2, OGG1, STC1, CNDP1, C10ORF2, MAML3, DKC1, PAXIP1, CASK, MAPT, SP1, POLG, AFF2, THBS1, TP53, ESR1, CGGBP1, ABT1, KLK3, PRNP, JUN, KCNN3, BAX, FRAXA, KBTBD10, MBNL1, RAD51, NCOA3, ERDA1, TSC1, COMP, GGLC, RRAD, MSH3, DRD2, CD44, CTCF, CCND1, CLSPN, MEF2A, PTPRU, GAPDH, TRIM22, WT1, AHR, GPX1, TPMT, NDP, ARX, TYR, EGR1, UNG, NUMBL, FABP2, EN2, CRYGC, SRP14, CRYGB, PDCD1,HOXA1, ATXN2L, PMS2, GLA, CBL, FTH1, IL12RB2, OTX2, HOXA5, POLG2, DLX2, AHRR, MANF, RMEM158, see also 20110016540
Turner's Syndrome (XO)	Various – reproductive organs, and sex characteristics, vasculature		Monosomy X
Tuberous Sclerosis	CNS, heart, kidneys		TSC1, TSC2
Usher syndrome (Types I, II, and III)	Ears, eyes		ABHD12, CDH23, CIB2, CLRN1, DFNB31, GPR98, HARS, MYO7A, PCDH15, USH1C, USH1G, USH2A, USH11A, those described in WO2015134812A1
Velocardiofacial syndrome (aka 22q11.2 deletion syndrome, DiGeorge syndrome, conotruncal	Various – skeletal, heart,		Many genes are deleted, COM, TBX1, and other are associated with symptoms

anomaly face syndrome (CTAF), autosomal dominant Opitz G/BB syndrome or Cayler cardiofacial syndrome)	kidney, immune system, brain		
Von Gierke's Disease (Glycogen Storage Disease type I)	Glycogen Storage disease	Various – liver, kidney	G6PC and SLC37A4
Von Hippel-Lindau Syndrome	Various – cell growth regulation disorder	CNS, Kidney, Eye, visceral organs	VHL
Von Willebrand Disease (Types I, II and III)	blood		VWF
Wilson Disease	Various - Copper Storage Disease	Liver, brains, eyes, other tissues where copper builds up	ATP7B
Wiskott-Aldrich Syndrome	Immune System		WAS
Xeroderma Pigmentosum	Skin	Nervous system	POLH
XXX Syndrome	Endocrine, brain		X chromosome trisomy

Table 4. Exemplary Genes controlling Cellular Functions

CELLULAR FUNCTION	GENES
PI3K/AKT Signaling	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2; PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1; AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2; PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2; ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3; PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7; YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A; CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1; PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2; TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK; HSP90AA1; RPS6KB1
ERK/MAPK Signaling	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1; PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A; PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN; EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ; PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1; PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1; CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK

Glucocorticoid Receptor Signaling	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1; MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I; PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2; MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1; MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13; RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1; PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3; MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP; CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1; ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1; STAT1; IL6; HSP90AA1
Axonal Guidance Signaling	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12; IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2; ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2; CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11; PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1; GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B; AKT3; PRKCA
Ephrin Receptor Signaling Actin Cytoskeleton Signaling	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2; DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14; CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4; AKT3; SGK
	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1; PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6; ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8; F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7; PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1; MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3; ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL; BRAF; VAV3; SGK

Huntington's Disease Signaling	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2; MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2; PIK3CA; HDAC5; CREB1; PRKCI; HSPA5; REST; GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1; GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2; HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A; HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1; PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX; ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
Apoptosis Signaling	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1; BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB; CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8; BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF; RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2; CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2; BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK; CASP3; BIRC3; PARP1
B Cell Receptor Signaling	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11; AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A; MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1; MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9; EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1; NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN; GSK3B; ATF4; AKT3; VAV3; RPS6KB1
Leukocyte Extravasation Signaling	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA; RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11; MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12; PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB; MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK; MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2; CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK; CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
Integrin Signaling	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A; TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2; CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8; CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA; SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1; TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2; CRKL; BRAF; GSK3B; AKT3

Acute Phase Response Signaling	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11; AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14; PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1; TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN; AKT3; IL1R1; IL6
PTEN Signaling	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11; MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA; CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1; MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR; RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2; AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1; NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2; GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
p53 Signaling	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A; BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2; PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1; PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9; CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A; HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1; SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN; SNAI2; GSK3B; BAX; AKT3
Aryl Hydrocarbon Receptor Signaling	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1; NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1; SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1; MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1; SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF; CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1; CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1; HSP90AA1
Xenobiotic Metabolism Signaling	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1; NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A; PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1; ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD; GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL; NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1; CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1; NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1; HSP90AA1

SAPK/JNK Signaling	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1; GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA; FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1; GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2; PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1; CRKL; BRAF; SGK
PPAr/RXR Signaling	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN; RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2; ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8; IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A; NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1; TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1; ADIPOQ
NF-KB Signaling	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6; TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2; MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2; KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF; INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1; PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10; GSK3B; AKT3; TNFAIP3; IL1R1
Neuregulin Signaling	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1; MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI; CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS; PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2; ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3; EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL; AKT3; PRKCA; HSP90AA1; RPS6KB1
Wnt & Beta catenin Signaling	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO; AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A; WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK; LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1; PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1; GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B; AKT3; SOX2
Insulin Receptor Signaling	PTEN; INS; EIF4E; PTPN1; PRKCZ; MAPK1; TSC1; PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3; MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1; SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1; GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK; RPS6KB1

IL-6 Signaling	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3; MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3; MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
Hepatic Cholestasis	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA; RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8; PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1; TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8; CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4; JUN; IL1R1; PRKCA; IL6
IGF-1 Signaling	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2; PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8; IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A; YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3; FOXO1; SRF; CTGF; RPS6KB1
NRF2-mediated Oxidative Stress Response	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1; NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL; NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP; MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1; GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
Hepatic Fibrosis/Hepatic Stellate Cell Activation	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF; SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9; IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8; PDGFRA; NFKB1; TGFBR1; SMAD4; VEGFA; BAX; IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
PPAR Signaling	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB; NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3; NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2; PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA; MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
Fc Epsilon RI Signaling	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD; MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3; VAV3; PRKCA

G-Protein Coupled Receptor Signaling	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB; PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB; PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1; IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK; PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3; PRKCA
Inositol Phosphate Metabolism	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6; MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3; MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1; MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
PDGF Signaling	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC; PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2; PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC; JUN; CRKL; PRKCA; SRF; STAT1; SPHK2
VEGF Signaling	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF; AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3; BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN; RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN; VEGFA; AKT3; FOXO1; PRKCA
Natural Killer Cell Signaling	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11; KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6; PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1; PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA
Cell Cycle: G1/S Checkpoint Regulation	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC; ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11; HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1; E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1; GSK3B; RBL1; HDAC6
T Cell Receptor Signaling	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS; NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN; MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10; JUN; VAV3
Death Receptor Signaling	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD; FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8; DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB; CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3; BIRC3

FGF Signaling	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11; AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8; MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1; AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4; AKT3; PRKCA; HGF
GM-CSF Signaling	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A; STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3; ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3; STAT1
Amyotrophic Lateral Sclerosis Signaling	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2; PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1; PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1; APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
JAK/Stat Signaling	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B; PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A; PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3; STAT1
Nicotinate and Nicotinamide Metabolism	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1; PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1; PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2; MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
Chemokine Signaling	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ; CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13; RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1; MAP2K2; MAP2K1; JUN; CCL2; PRKCA
IL-2 Signaling	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS; STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2; JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
Synaptic Long Term Depression	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS; PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3; KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA; YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
Estrogen Receptor Signaling	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2; SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1; HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP; MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
Protein Ubiquitination Pathway	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4; CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7; USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8; USP1; VHL; HSP90AA1; BIRC3

IL-10 Signaling	TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF; IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1; JUN; IL1R1; IL6
VDR/RXR Activation	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1; NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD; RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1; LRP5; CEBPB; FOXO1; PRKCA
TGF-beta Signaling	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1; FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2; SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2; MAP2K1; TGFBR1; SMAD4; JUN; SMAD5
Toll-like Receptor Signaling	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK; NFKB1; TLR2; JUN
p38 MAPK Signaling	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS; CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2; MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1; SRF; STAT1
Neurotrophin/TRK Signaling	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A; RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; CDC42; JUN; ATF4
FXR/RXR Activation	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8; APOB; MAPK10; PPARG; MTTP; MAPK9; PPARGC1A; TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
Synaptic Long Term Potentiation	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1; PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS; PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1; ATF4; PRKCA
Calcium Signaling	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1; CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11; HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4; HDAC6
EGF Signaling	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1; STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
Hypoxia Signaling in the Cardiovascular System	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT; HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM; VEGFA; JUN; ATF4; VHL; HSP90AA1
LPS/IL-1 Mediated Inhibition of RXR Function	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1; MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2; TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1

LXR/RXR Activation	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA; NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1; SREBF1; IL1R1; CCL2; IL6; MMP9
Amyloid Processing	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2; CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1; PSEN1; CSNK1A1; GSK3B; AKT3; APP
IL-4 Signaling	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1; PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1; FRAP1; AKT3; RPS6KB1
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC; CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A; PRKDC; ATM; SFN; CDKN2A
Nitric Oxide Signaling in the Cardiovascular System	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3; CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1; VEGFA; AKT3; HSP90AA1
Purine Metabolism	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4; PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C; NT5E; POLD1; NME1
cAMP-mediated Signaling	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3; SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
Mitochondrial Dysfunction Notch Signaling	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9; PARK7; PSEN1; PARK2; APP; CASP3 HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2; PSEN1; NOTCH3; NOTCH1; DLL4
Endoplasmic Reticulum Stress Pathway Pyrimidine Metabolism	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4; EIF2AK3; CASP3 NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B; NT5E; POLD1; NME1
Parkinson's Signaling	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7; PARK2; CASP3
Cardiac & Beta Adrenergic Signaling	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC; PPP2R5C
Glycolysis/Gluconeogenesis	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
Interferon Signaling	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3
Sonic Hedgehog Signaling	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B
Glycerophospholipid Metabolism	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2
Phospholipid Degradation	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
Tryptophan Metabolism	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
Lysine Degradation	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
Nucleotide Excision Repair Pathway	ERCC5; ERCC4; XPA; XPC; ERCC1
Starch and Sucrose Metabolism	UCHL1; HK2; GCK; GPI; HK1
Aminosugars Metabolism	NQO1; HK2; GCK; HK1

Arachidonic Acid Metabolism	PRDX6; GRN; YWHAZ; CYP1B1
Circadian Rhythm Signaling	CSNK1E; CREB1; ATF4; NR1D1
Coagulation System	BDKRB1; F2R; SERPINE1; F3
Dopamine Receptor Signaling	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
Glutathione Metabolism	IDH2; GSTP1; ANPEP; IDH1
Glycerolipid Metabolism	ALDH1A1; GPAM; SPHK1; SPHK2
Linoleic Acid Metabolism	PRDX6; GRN; YWHAZ; CYP1B1
Methionine Metabolism	DNMT1; DNMT3B; AHCY; DNMT3A
Pyruvate Metabolism	GLO1; ALDH1A1; PKM2; LDHA
Arginine and Proline Metabolism	ALDH1A1; NOS3; NOS2A
Eicosanoid Signaling	PRDX6; GRN; YWHAZ
Fructose and Mannose Metabolism	HK2; GCK; HK1
Galactose Metabolism	HK2; GCK; HK1
Stilbene, Coumarine and Lignin Biosynthesis	PRDX6; PRDX1; TYR
Antigen Presentation Pathway	CALR; B2M
Biosynthesis of Steroids	NQO1; DHCR7
Butanoate Metabolism	ALDH1A1; NLGN1
Citrate Cycle	IDH2; IDH1
Fatty Acid Metabolism	ALDH1A1; CYP1B1
Glycerophospholipid Metabolism	PRDX6; CHKA
Histidine Metabolism	PRMT5; ALDH1A1
Inositol Metabolism	ERO1L; APEX1
Metabolism of Xenobiotics by Cytochrome p450	GSTP1; CYP1B1
Methane Metabolism	PRDX6; PRDX1
Phenylalanine Metabolism	PRDX6; PRDX1
Propanoate Metabolism	ALDH1A1; LDHA
Selenoamino Acid Metabolism	PRMT5; AHCY
Sphingolipid Metabolism	SPHK1; SPHK2
Aminophosphonate Metabolism	PRMT5
Androgen and Estrogen Metabolism	PRMT5
Ascorbate and Aldarate Metabolism	ALDH1A1
Bile Acid Biosynthesis	ALDH1A1
Cysteine Metabolism	LDHA
Fatty Acid Biosynthesis	FASN
Glutamate Receptor Signaling	GNB2L1

NRF2-mediated Oxidative Stress Response	PRDX1
Pentose Phosphate Pathway	GPI
Pentose and Glucuronate Interconversions	UCHL1
Retinol Metabolism	ALDH1A1
Riboflavin Metabolism	TYR
Tyrosine Metabolism	PRMT5, TYR
Ubiquinone Biosynthesis	PRMT5
Valine, Leucine and Isoleucine Degradation	ALDH1A1
Glycine, Serine and Threonine Metabolism	CHKA
Lysine Degradation	ALDH1A1
Pain/Taste	TRPM5; TRPA1
Pain	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2; Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca; Prkacb; Prkar1a; Prkar2a
Mitochondrial Function	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2
Developmental Neurology	BMP-4; Chordin (Chrd); Noggin (Nog); WNT (Wnt2; Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b; Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); beta-catenin; Dkk-1; Frizzled related proteins; Otx-2; Gbx2; FGF-8; Reelin; Dab1; unc-86 (Pou4f1 or Brn3a); Numb; Reln

Treating Diseases of the Circulatory System

[0620] In some embodiments, the gene therapy is used to treat and/or prevent a circulatory system disease or symptom thereof in a human subject. Exemplary diseases of the circulatory system suitable for gene therapy are identified in, for example, **Tables 3 and 4** and at mitomap.org, such as in the case of those attributed to mitochondrial DNA mutations.

[0621] In some embodiments delivery of an engineered compositions or system described herein of the present invention can be delivered to blood cells, for example, via the engineered delivery vesicle generation systems, delivery vesicles produced therefrom, and/or cargos delivered thereby, plasma exosomes (see e.g., Wahlgren et al., Nucleic Acids Research, 2012, Vol. 40, No. 17 e130), other lentiviral particles based on systems generally known in the art (see e.g., Drakopoulou, Stem Cells International, Volume 2011, Article ID 987980, 10 pages, doi:10.4061/2011/987980). In some embodiments, the circulatory disease can be treated by correcting hematopoietic stem cells as demonstrated by various exemplary approaches using CRISPR-Cas systems. See e.g., Negre et al., Hum Gene Ther. 2016 Feb 1; 27(2): 148–165; Cavazzana M, et al. Blood 2014;124:4797; Cavazzana-Calvo, Nature 467, 318–322 ((2010)

doi:10.1038/nature09328; Nienhuis, "Development of Gene Therapy for Thalassemia, Cold Spring Harbor Perspectives in Medicine, doi: 10.1101/cshperspect.a011833 (2012); Xie et al., Genome Research gr.173427.114 (2014) http://www.genome.org/cgi/doi/10.1101/gr.173427.114 (Cold Spring Harbor Laboratory Press; and Watts et al, "Hematopoietic Stem Cell Expansion and Gene Therapy" Cytotherapy 13(10):1164–1171. doi:10.3109/14653249.2011.620748 (2011). In some embodiments, iPSCs can be modified using an engineered compositions or system of the present invention to correct a disease polynucleotide associated with a circulatory disease. In this regard, the teachings of Xu et al. (Sci Rep. 2015 Jul 9;5:12065. doi: 10.1038/srep12065) and Song et al. (Stem Cells Dev. 2015 May 1;24(9):1053-65. doi: 10.1089/scd.2014.0347. Epub 2015 Feb 5) with respect to modifying iPSCs can be adapted for use in view of the description herein for use with the engineered compositions systems of the present invention. Administration of the engineered compositions or systems or formulations thereof described herein of the present invention configured for gene therapy for a circulatory disease can be directly to the blood, such as via intravenous administration.

[0622] The term "Hematopoietic Stem Cell" or "HSC" refers broadly those cells considered to be an HSC, e.g., blood cells that give rise to all the other blood cells and are derived from mesoderm; located in the red bone marrow, which is contained in the core of most bones. HSCs of the invention include cells having a phenotype of hematopoietic stem cells, identified by small size, lack of lineage (lin) markers, and markers that belong to the cluster of differentiation series, like: CD34, CD38, CD90, CD133, CD105, CD45, and also c-kit, - the receptor for stem cell factor. Hematopoietic stem cells are negative for the markers that are used for detection of lineage commitment, and are, thus, called Lin-; and, during their purification by FACS, a number of up to 14 different mature blood-lineage markers, e.g., CD13 & CD33 for myeloid, CD71 for erythroid, CD19 for B cells, CD61 for megakaryocytic, etc. for humans; and, B220 (murine CD45) for B cells, Mac-1 (CD11b/CD18) for monocytes, Gr-1 for Granulocytes, Ter119 for erythroid cells, Il7Ra, CD3, CD4, CD5, CD8 for T cells, etc. Mouse HSC markers: CD34lo/-, SCA-1+, Thy1.1+/lo, CD38+, C-kit+, lin-, and Human HSC markers: CD34+, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, and lin-. HSCs are identified by markers. Hence in embodiments discussed herein, the HSCs can be CD34+ cells. HSCs can also be hematopoietic stem cells that are CD34-/CD38-. Stem cells that may lack c-

kit on the cell surface that are considered in the art as HSCs are within the ambit of the invention, as well as CD133+ cells likewise considered HSCs in the art.

[0623] In some embodiments, the treatment or prevention for treating a circulatory system or blood disease can include modifying a human cord blood cell with any modification described herein. In some embodiments, the treatment or prevention for treating a circulatory system or blood disease can include modifying a granulocyte colony-stimulating factor-mobilized peripheral blood cell (mPB) with any modification described herein. In some embodiments, the human cord blood cell or mPB can be CD34+. In some embodiments, the cord blood cell(s) or mPB cell(s) modified can be autologous. In some embodiments, the cord blood cell(s) or mPB cell(s) can be allogenic. In addition to the modification of the disease gene(s), allogenic cells can be further modified using the composition, system, described herein to reduce the immunogenicity of the cells when delivered to the recipient. Such techniques are described elsewhere herein and e.g. Cartier, "MINI-SYMPOSIUM: X-Linked Adrenoleukodystrophypa, Hematopoietic Stem Cell Transplantation and Hematopoietic Stem Cell Gene Therapy in X-Linked Adrenoleukodystrophy," Brain Pathology 20 (2010) 857–862, which can be adapted for use with the composition, system, herein. The modified cord blood cell(s) or mPB cell(s) can be optionally expanded in vitro. The modified cord blood cell(s) or mPB cell(s) can be derived to a subject in need thereof using any suitable delivery technique.

[0624] In some embodiments, the engineered compositions, systems, or components thereof or cargos thereof are codon optimized for expression in a mammalian HSC, iPSC, such as a human HSC or iPSC.

Treating Diseases of the Brain and Nervous System

[0625] In some embodiments, the gene therapy is used to treat and/or prevent a brain and/or nervous system disease or symptom thereof in a human subject. Exemplary diseases of the brain and/or nervous system suitable for gene therapy are identified in, for example, **Tables 3 and 4** and at mitomap.org, such as in the case of those attributed to mitochondrial DNA mutations. The engineered compositions, systems, vesicles, and/or formulations thereof of the present invention can be delivered to the brain or other part of the central or peripheral nervous system by any suitable administration route, such as intracerebroventicaluar injection, spinal canal injection, or delivery to the subarachnoid space. In some embodiments, the engineered compositions, systems, vesicles, and/or formulations thereof are configured for trans-blood brain barrier delivery. Such configurations include, without limitations, encapsulation of the

engineered compositions, systems, vesicles, and/or components thereof of the present invention into liposomes and/or conjugating them to mole molecular Trojan horses. Molecular Trojan horses have been demonstrated to be effective for trans blood brain barrier delivery and can be adapted for use with the engineered compositions, systems, vesicles, and the like of the present invention. See e.g., Xia et al., Mol Pharm. 2009 May-Jun;6(3):747-51. doi: 10.1021/mp800194). In some embodiments, the combined use a receptor-specific antibody, such as a monoclonal antibody (mAb) is conjugated to the engineered composition(s), system(s), vesicle(s), and/or the like of the present invention and avidin-biotin technology (see e.g., *Id.*). In some embodiments, engineered vesicles, such as virus or virus like particles, of the present invention can be used for delivery of one or more cargos to the central nervous system. In some embodiments, other artificial virus particles can be generated for central nervous system and/or brain delivery. See e.g., Zhang et al. (Mol Ther. 2003 Jan;7(1):11-8.)), the teachings of which can be adapted for use with the engineered compositions and systems of the present invention.

Treating Diseases of the Auditory System

[0626] In some embodiments the engineered compositions, systems, vesicles, and the like of the present invention can be used to treat a hearing disease or hearing loss in one or both ears of a human subject. Deafness is often caused by lost or damaged hair cells that cannot relay signals to auditory neurons. In such cases, cochlear implants may be used to respond to sound and transmit electrical signals to the nerve cells. But these neurons often degenerate and retract from the cochlea as fewer growth factors are released by impaired hair cells.

[0627] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention are delivered to one or both ears for treating or preventing hearing disease or loss by any suitable method or technique. Suitable methods and techniques include, but are not limited to, those set forth in U.S. Publication No. 20120328580, which describes injection of a pharmaceutical composition into the ear (e.g., auricular administration), such as into the luminae of the cochlea (e.g., the Scala media, Sc vestibulae, and Sc tympani), e.g., using a syringe, e.g., a single-dose syringe. For example, one or more of the engineered compositions, systems, vesicles, and the like of the present invention described herein can be administered by intratympanic injection (e.g., into the middle ear), and/or injections into the outer, middle, and/or inner ear; administration in situ, via a catheter or pump (see e.g., McKenna et al., (U.S. Publication No. 2006/0030837) and Jacobsen et al., (U.S. Pat. No.

7,206,639); administration in combination with a mechanical device such as a cochlear implant or a hearing aid, which is worn in the outer ear (see e.g., U.S. Publication No. 2007/0093878, which provides an exemplary cochlear implant suitable for delivery of the engineered compositions, systems, vesicles, and the like of the present invention described herein to the ear). Such methods are routinely used in the art, for example, for the administration of steroids and antibiotics into human ears. Injection can be, for example, through the round window of the ear or through the cochlear capsule. Other inner ear administration methods are known in the art (see, e.g., Salt and Plontke, Drug Discovery Today, 10:1299-1306, 2005) and can be used for delivery of the engineered compositions, systems, vesicles, and the like of the present invention. In some embodiments, a catheter or pump can be positioned, e.g., in the ear (e.g., the outer, middle, and/or inner ear) of a patient during a surgical procedure to deliver the engineered compositions, systems, vesicles, and the like of the present invention. In some embodiments, a catheter or pump can be positioned, e.g., in the ear (e.g., the outer, middle, and/or inner ear) of a patient without the need for a surgical procedure to deliver the engineered compositions, systems, vesicles, and the like of the present invention.

[0628] The engineered compositions, systems, vesicles, and the like of the present invention can be delivered to the ear by direct application of pharmaceutical composition to the outer ear, with compositions modified from US Published application, 20110142917. In some embodiments the pharmaceutical composition is applied to the ear canal. Delivery to the ear can also be referred to as aural or otic delivery.

[0629] The engineered compositions, systems, vesicles, and the like of the present invention can be delivered to ear via a transfection to the inner ear through the intact round window by a novel proteidic delivery technology which may be applied to the nucleic acid-targeting system of the present invention (see e.g., Qi et al., Gene Therapy (2013), 1-9). In some embodiments, the dosage for administration of a nucleic acid to an ear can be about 40 μl of a 10mM nucleic acid formulation. In some embodiments the nucleic acid is DNA, RNA, or both.

[0630] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention can be combined with a cochlear implant therapy (see e.g., Rejali et al. Hear Res. 2007 Jun;228(1-2):180-7). In some embodiments, the combination of cochlear implant therapy with gene therapy provided by the engineered compositions, systems, vesicles, and the like of the present invention results in an increase in the survival and/or preservation

of spiral ganglion neurons. In some embodiments, the gene therapy provided by engineered compositions, systems, vesicles, and the like of the present invention includes delivery of a functional brain derived neurotrophic factor, which has been demonstrated to enhance spiral ganglion survival in experimentally defended ears. In some embodiments, cells containing and/or expressing an engineered composition, system, vesicle, and the like of the present invention configured to deliver or generate a functional brain derived neurotrophic factor in the cells, are coated on the cochlear implant used in the combination therapy. In some embodiments, the cells are fibroblasts.

[0631] In some embodiments, the system set forth in Mukherjea et al. (Antioxidants & Redox Signaling, Volume 13, Number 5, 2010) can be adapted for transtympanic administration of the engineered compositions, systems, vesicles, and the like of the present invention to the ear. In some embodiments, the dosage for human administration is about 2 mg to about 4 mg of one or more components of the engineered compositions, systems, vesicles, and the like of the present invention.

[0632] In some embodiments, the system set forth in [Jung et al. (Molecular Therapy, vol. 21 no. 4, 834–841 Apr. 2013) can be adapted for vestibular epithelial delivery of the engineered compositions, systems, vesicles, and the like of the present invention to the ear. In some embodiments, a dosage for human administration is about 1 to about 30 mg of one or more components of the engineered compositions, systems, vesicles, and the like of the present invention.

Treating Diseases in Non-Dividing Cells

[0633] In some embodiments, the gene or transcript to be corrected is in a non-dividing cell. Exemplary non-dividing cells are muscle cells or neurons. Non-dividing (especially non-dividing, fully differentiated) cell types present issues for gene targeting or genome engineering, for example because homologous recombination (HR) is generally suppressed in the G1 cell-cycle phase. However, while studying the mechanisms by which cells control normal DNA repair systems, Durocher discovered a previously unknown switch that keeps HR "off" in non-dividing cells and devised a strategy to toggle this switch back on. Orthwein et al. (Daniel Durocher's lab at the Mount Sinai Hospital in Ottawa, Canada) recently reported (Nature 16142, published online 9 Dec 2015) have shown that the suppression of HR can be lifted and gene targeting successfully concluded in both kidney (293T) and osteosarcoma (U2OS) cells. Tumor suppressors, BRCA1, PALB2 and BRAC2 are known to promote DNA

DSB repair by HR. They found that formation of a complex of BRCA1 with PALB2 - BRAC2 is governed by a ubiquitin site on PALB2, such that action on the site by an E3 ubiquitin ligase. This E3 ubiquitin ligase is composed of KEAP1 (a PALB2 -interacting protein) in complex with cullin-3 (CUL3)–RBX1. PALB2 ubiquitylation suppresses its interaction with BRCA1 and is counteracted by the deubiquitylase USP11, which is itself under cell cycle control. Restoration of the BRCA1–PALB2 interaction combined with the activation of DNA-end resection is sufficient to induce homologous recombination in G1, as measured by a number of methods including a CRISPR–Cas9-based gene-targeting assay directed at USP11 or KEAP1 (expressed from a pX459 vector). However, when the BRCA1–PALB2 interaction was restored in resection-competent G1 cells using either KEAP1 depletion or expression of the PALB2-KR mutant, a robust increase in gene-targeting events was detected. These teachings can be adapted for and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention.

[0634] Thus, reactivation of HR in cells, especially non-dividing, fully differentiated cell types is preferred, in some embodiments. In some embodiments, promotion of the BRCA1–PALB2 interaction is preferred in some embodiments. In some embodiments, the target cell for gene treatment is a non-dividing cell. In some embodiments, the target cell for treatment is a neuron or muscle cell. In some embodiments, the target cell is targeted in vivo. In some embodiments, the cell is in G1 and HR is suppressed. In some embodiments, use of KEAP1 depletion, for example inhibition of expression of KEAP1 activity, is preferred. KEAP1 depletion can be achieved through siRNA, for example as shown in Orthwein et al. In some embodiments, expression of the PALB2-KR mutant (lacking all eight Lys residues in the BRCA1-interaction domain) is used, either in combination with KEAP1 depletion or alone. PALB2-KR interacts with BRCA1 irrespective of cell cycle position. Thus, promotion or restoration of the BRCA1-PALB2 interaction, especially in G1 cells, is preferred in some embodiments, especially where the target cells are non-dividing, or where removal and return (ex vivo gene targeting) is problematic, for example neuron or muscle cells. In some embodiments, a BRCA1–PALB2 complex is delivered to the G1 cell. In some embodiments, PALB2 deubiquitylation is promoted, for example, by increased expression of the deubiquitylase USP11. In some embodiments a composition or system is configured and/or co-delivered that promotes or up-regulates expression and/or activity of the deubiquitylase USP11.

Treating Diseases of the Ocular System

[0635] In some embodiments, the disease to be treated is a disease or disorder that affects the eyes. Thus, in some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention is delivered to one or both eyes. The engineered compositions, systems, vesicles, and the like of the present invention can be used to correct ocular defects that arise from several genetic mutations further described in Genetic Diseases of the Eye, Second Edition, edited by Elias I. Traboulsi, Oxford University Press, 2012.

[0636] In some embodiments, the eye disorder is or includes glaucoma. In some embodiments, the eye disorder is or includes a retinal degenerative disease. In some embodiments, the retinal degenerative disease is selected from Stargardt disease, Bardet-Biedl Syndrome, Best disease, Blue Cone Monochromacy, Choroidermia, Cone-rod dystrophy, Congenital Stationary Night Blindness, Enhanced S-Cone Syndrome, Juvenile X-Linked Retinoschisis, Leber Congenital Amaurosis, Malattia Leventinesse, Norrie Disease or X-linked Familial Exudative Vitreoretinopathy, Pattern Dystrophy, Sorsby Dystrophy, Usher Syndrome, Retinitis Pigmentosa, Achromatopsia or Macular dystrophies or degeneration, Retinitis Pigmentosa, Achromatopsia, and age related macular degeneration. In some embodiments, the retinal degenerative disease is Leber Congenital Amaurosis (LCA) or Retinitis Pigmentosa. Other exemplary eye diseases and disorders are described in greater detail elsewhere herein.

[0637] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention is/are delivered to the eye, optionally via intravitreal injection or subretinal injection. Intraocular injections may be performed with the aid of an operating microscope. For subretinal and intravitreal injections, eyes can be prolapsed by gentle digital pressure and fundi visualized using a contact lens system consisting of a drop of a coupling medium solution on the cornea covered with a glass microscope slide coverslip. In some embodiments, subretinal injections are performed as follows: the tip of a 10-mm 34-gauge needle, mounted on a 5-µl Hamilton syringe is advanced under direct visualization through the superior equatorial sclera tangentially towards the posterior pole until the aperture of the needle is visible in the subretinal space. Then, about 2 µl of formulation containing one or more engineered compositions, systems, vesicles, and the like of the present invention can be injected to produce a superior bullous retinal detachment, thus confirming subretinal vector administration. This approach creates a self-sealing sclerotomy allowing the vector suspension to be retained in the subretinal space until it is absorbed by the RPE, usually within 48 h of the

procedure. This procedure can be repeated in the inferior hemisphere to produce an inferior retinal detachment. This technique results in the exposure of approximately 70% of neurosensory retina and RPE to the formulation containing the engineered compositions, systems, vesicles, and the like of the present invention.

[0638] In some embodiments, intravitreal injections are performed as follows: the needle tip is advanced through the sclera about 1 mm posterior to the corneoscleral limbus and about 2 µl of a formulation containing one or more engineered compositions, systems, vesicles, and the like of the present invention is injected into the vitreous cavity. In some embodiments, intracameral injections can be performed as follows: the needle tip is advanced through a corneoscleral limbal paracentesis, directed towards the central cornea, and about 2 µl of a formulation containing one or more engineered compositions, systems, vesicles, and the like of the present invention is injected. In some embodiments, intracameral injections are performed as follows: the needle tip is advanced through a corneoscleral limbal paracentesis, directed towards the central cornea, and about 2 µl of a formulation containing one or more engineered compositions, systems, vesicles, and the like of the present invention is injected. When viral delivery particles are used, such as engineered viral particles or delivery vesicles of the present invention, they can be injected at titers ranging from $1.0–1.4 \times 10^{10}$ to$1.0–1.4 \times 10^{9}$ transducing units (TU)/ml or any value or range of values therein.

[0639] In some embodiments, the engineered system of the present invention configured for gene therapy or delivery thereof to the eye is a retroviral or lentiviral-based system. In some embodiments, the engineered lentiviral vector based system of the present invention is or is based upon an equine infectious anemia virus (EIAV) vector. Exemplary EIAV vectors for eye delivery are described in Balagaan, J Gene Med 2006; 8: 275 – 285, Published online 21 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jgm.845; Binley et al., HUMAN GENE THERAPY 23:980–991 (September 2012), which can be adapted for use with the engineered compositions, systems, vesicles, and the like of the present invention. In some embodiments, the dosage can be 1.1 x 10^{5} transducing units per eye (TU/eye) in a total volume of 100 µl.

[0640] Other viral vectors can also be used for delivery to the eye, such as AAV vectors, such as those described in Campochiaro et al., Human Gene Therapy 17:167-176 (February 2006), Millington-Ward et al. (Molecular Therapy, vol. 19 no. 4, 642–649 apr. 2011; Dalkara et al. (Sci Transl Med 5, 189ra76 (2013)), which can be adapted for use with the engineered

compositions, systems, vesicles, and the like of the present invention. In some embodiments, the dose can range from about 10^6 to $10^{9.5}$ particle units or be any value or range thereof therein. In the context of the Millington-Ward AAV vectors, a dose of about 2 x 10^{11} to about 6 x 10^{13} virus particles can be administered. In the context of Dalkara vectors, a dose for human administration can be about 1 x 10^{15} to about 1 x 10^{16} vg/ml.

[0641] In some embodiments, the sd-rxRNA® system of RXi Pharmaceuticals are used/and or adapted for engineered compositions, systems, vesicles, and the like of the present invention to the eye. In this system, a single intravitreal administration of 3 μg of sd-rxRNA results in sequence-specific reduction of PPIB mRNA levels for 14 days. The sd-rxRNA® system can be applied to the engineered compositions, systems, vesicles, and the like of the present invention. In some embodiments, a dose for human administration is about 3 to 20 mg of engineered compositions, systems, vesicles, and the like of the present invention.

[0642] In some embodiments, the methods of US Patent Publication No. 20130183282, which is directed to methods of cleaving a target sequence from the human rhodopsin gene, are modified and adapted for use with the engineered compositions, systems, vesicles, and the like of the present invention for delivery to the eye.

[0643] In some embodiments, the methods of US Patent Publication No. 20130202678, which is directed to treating retinopathies and sight-threatening ophthalmologic disorders relating to delivering a Puf-A gene to the sub-retinal or intravitreal space in the eye, are modified and/or adapted for use with the engineered compositions, systems, vesicles, and the like of the present invention. In particular, desirable targets are zgc:193933, prdm1a, spata2, tex10, rbb4, ddx3, zp2.2, Blimp-1 and HtrA2, all of which may be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention.

[0644] Wu (Cell Stem Cell,13:659–62, 2013) designed a guide RNA that led Cas9 to a single base pair mutation that causes cataracts in mice, where it induced DNA cleavage. Then using either the other wild-type allele or oligos given to the zygotes repair mechanisms corrected the sequence of the broken allele and corrected the cataract-causing genetic defect in mutant mouse. This approach can be adapted to and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention.

[0645] US Patent Publication No. 20120159653, describes use of zinc finger nucleases to genetically modify cells, animals and proteins associated with macular degeneration (MD), the

teachings of which can be applied to and/or adapted for the engineered compositions, systems, vesicles, and the like of the present invention.

[0646] One aspect of US Patent Publication No. 20120159653 relates to editing of any chromosomal sequences that encode proteins associated with MD, which may be applied to the engineered compositions, systems, vesicles, and the like of the present invention.

Treating Diseases of the Musculoskeletal and Cardiovascular System

[0647] The engineered compositions, systems, vesicles, and the like of the present invention can be used to provide and/or deliver a gene therapy treat and/or prevent a muscle disease and associated circulatory or cardiovascular disease or disorder. The engineered compositions, systems, vesicles, and the like of the present invention can be to the heart and/or other portion of the cardiovascular system. In some embodiments, a myocardium tropic adeno-associated virus (AAVM) or myocardium tropic AAV, or myocardium tropic lentiviral virus is used for delivery. In some embodiments, an engineered delivery vesicle generation system of the present invention is used for delivery of a gene modifying agent or system. In some embodiments, the engineered compositions and systems and the like of the present invention are delivered by an AAVM41 virus or AAVM41 based virus system, which can preferentially deliver a cargo to the heart (see e.g., Lin-Yanga et al., PNAS, March 10, 2009, vol. 106, no. 10). Administration may be systemic or local. In some embodiments, dosage for systemic administration is about 1-10 x 10^{14} vector genomes (vg). See also e.g., Eulalio et al. (2012) Nature 492: 376 and Somasuntharam et al. (2013) Biomaterials 34: 7790, the teachings of which can be adapted for and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention.

[0648] US Patent Publication No. 20110023139 describes use of zinc finger nucleases to genetically modify cells, animals and proteins associated with cardiovascular disease, the teachings of which can be adapted for and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention. Cardiovascular diseases generally include high blood pressure, heart attacks, heart failure, and stroke and TIA. Any chromosomal sequence involved in cardiovascular disease or the protein encoded by any chromosomal sequence involved in cardiovascular disease may be utilized in a treatment, such as a gene therapy, provided by the engineered compositions, systems, vesicles, and the like of the present invention. For example, in some embodiments, such sequences can be included as a target for a gene therapy provided or delivered by the engineered compositions, systems, vesicles, and

the like of the present invention. The cardiovascular-related proteins targeted for treatment are typically selected based on an experimental association of the cardiovascular-related protein to the development of cardiovascular disease. For example, the production rate or circulating concentration of a cardiovascular-related protein may be elevated or depressed in a population having a cardiovascular disorder relative to a population lacking the cardiovascular disorder. Differences in protein levels may be assessed using proteomic techniques including but not limited to Western blot, immunohistochemical staining, enzyme linked immunosorbent assay (ELISA), and mass spectrometry. In some embodiments, the cardiovascular-related proteins targeted for treatment can be identified by obtaining gene expression profiles of the genes encoding the proteins using genomic techniques including but not limited to DNA microarray analysis, serial analysis of gene expression (SAGE), and quantitative real-time polymerase chain reaction (Q-PCR). Exemplary non-limiting chromosomal sequences can be found in **Table 3**.

[0649] The engineered compositions, systems, vesicles, and the like of the present invention can be used for treating diseases of the muscular system. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention are delivered to muscle(s) and can provide and/or deliver a treatment thereto.

[0650] In some embodiments, the muscle disease to be treated is a muscle dystrophy such as DMD. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention, is or can deliver a system capable of DNA and/or RNA modification to provide a corrective therapy. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention can provide and/or deliver a system capable of exon skipping to achieve correction of a diseased gene. As used herein, the term “exon skipping” refers to the modification of pre-mRNA splicing by the targeting of splice donor and/or acceptor sites within a pre-mRNA with one or more complementary antisense oligonucleotide(s) (AONs). By blocking access of a spliceosome to one or more splice donor or acceptor site, an AON may prevent a splicing reaction thereby causing the deletion of one or more exons from a fully-processed mRNA. Exon skipping may be achieved in the nucleus during the maturation process of pre-mRNAs. In some examples, exon skipping may include the masking of key sequences involved in the splicing of targeted exons by using engineered compositions, systems, vesicles, and the like of the present invention described herein capable of RNA modification and/or delivering a system capable of

RNA modification. In some embodiments, exon skipping can be achieved in dystrophin mRNA. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention or system delivered by the engineered compositions, systems, vesicles, and the like of the present invention can induce exon skipping at exon 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 45, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, or any combination thereof of the dystrophin mRNA. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention or system delivered by the engineered compositions, systems, vesicles, and the like of the present invention can induce exon skipping at exon 43, 44, 50, 51, 52, 55, or any combination thereof of the dystrophin mRNA. Mutations in these exons, can also be corrected using non-exon skipping polynucleotide modification compositions, systems, and methods, such as by the engineered compositions, systems, vesicles, and the like of the present invention or such systems capable of polynucleotide modification that are delivered by the engineered compositions, systems, vesicles, and the like of the present invention.

[0651] In some embodiments, for treatment of a muscle disease, the method of Bortolanza et al. Molecular Therapy vol. 19 no. 11, 2055–2064 Nov. 2011) can be applied and/or adapted to engineered compositions, systems, vesicles, and the like of the present invention. A dosage for administration via injection to humans can be about 2×10^{15} or 2×10^{16} vg of vector.

[0652] In some embodiments, the method of Dumonceaux et al. (Molecular Therapy vol. 18 no. 5, 881–887 May 2010) can be applied and/or adapted to the engineered compositions, systems, vesicles, and the like of the present invention. A dosage for administration via injection into humans, for example, can be about 10^{14} to about 10^{15} vg of vector.

[0653] In some embodiments, the method of Kinouchi et al. (Gene Therapy (2008) 15, 1126–1130) can be applied and/or adapted to the engineered compositions, systems, vesicles, and the like of the present invention. A dosage for administration via injection to a human via injection into muscle, for example, can be about 500 to 1000 ml of a 40 μM formulation.

[0654] In some embodiments, the method of Hagstrom et al. (Molecular Therapy Vol. 10, No. 2, August 2004) can be adapted for and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention. A dosage for administration via injection to a human via injection into the great saphenous vein can be about 15 to about 50 mg.

Treating Diseases of the Hepatic and Renal System

[0655] The engineered compositions, systems, vesicles, and the like of the present invention can be used to treat a disease of the kidney or liver. Thus, in some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention are delivered to the liver and/or kidney.

[0656] In some embodiments, the engineered compositions, systems, vesicles, and/or the like of the present invention can be delivered using strategies to induce cellular uptake of the therapeutic nucleic acids or other components of the engineered compositions, systems etc. of the present invention include physical force or vector systems such as viral-, lipid- or complex-based delivery, or nanocarriers.

[0657] Csaba Révész and Péter Hamar (2011) Delivery Methods to Target RNAs in the Kidney, Gene Therapy Applications, Prof. Chunsheng Kang (Ed.), ISBN: 978-953-307-541-9, InTech, Available from: www.intechopen.com/books/gene-therapy-applications/delivery-methods-to-target-rnas-inthe-kidney) describes a variety of methods of delivery of nucleic acids to kidney cells. Such techniques and methods can be adapted and applied to the engineered compositions, systems, vesicles, and the like of the present invention for delivery to kidney cells. Delivery methods to the kidney can include those set forth in Yuan et al. (Am J Physiol Renal Physiol 295: F605–F617, 2008). The method of Yuang et al. can be adapted and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention. A dose of about 1-2 g of the engineered compositions, systems, vesicles, and/or the like of the present invention with cholesterol can be used for subcutaneous injection for delivery to the kidneys. In some embodiments, the method of Molitoris et al. (J Am Soc Nephrol 20: 1754–1764, 2009) can be adapted to the engineered compositions, systems, vesicles, and the like of the present invention and a cumulative dose of 12- 20 mg/kg for delivery to the proximal tubule cells of the kidneys in a human can be used. In some embodiments, the methods of Thompson et al. (Nucleic Acid Therapeutics, Volume 22, Number 4, 2012) can be adapted to the engineered compositions, systems, vesicles, and the like of the present invention of the present invention and a dose of up to 25 mg/kg can be delivered via i.v. administration. In some embodiments, the method of Shimizu et al. (J Am Soc Nephrol 21: 622–633, 2010) can be adapted to the engineered compositions, systems, vesicles, and the like of the present invention and a dose of about of 10-20 μmol engineered

compositions, systems, vesicles, and the like of the present invention optionally complexed with nanocarriers in about 1-2 liters of a physiologic fluid for i.p. administration can be used.

[0658] In addition to delivery vesicles of the present invention, other various delivery vehicles can be used to deliver the engineered compositions, systems, vesicles, and the like of the present invention to the kidney such as viral, hydrodynamic, lipid, polymer nanoparticles, aptamers and various combinations thereof (see e.g., Larson et al., Surgery, (Aug 2007), Vol. 142, No. 2, pp. (262-269); Hamar et al., Proc Natl Acad Sci, (Oct 2004), Vol. 101, No. 41, pp. (14883-14888); Zheng et al., Am J Pathol, (Oct 2008), Vol. 173, No. 4, pp. (973–980); Feng et al., Transplantation, (May 2009), Vol. 87, No. 9, pp. (1283–1289); Q. Zhang et al., PloS ONE, (Jul 2010), Vol. 5, No. 7, e11709, pp. (1-13); Kushibikia et al., J Controlled Release, (Jul 2005), Vol. 105, No. 3, pp. (318-331); Wang et al., Gene Therapy, (Jul 2006), Vol. 13, No. 14, pp. (1097-1103); Kobayashi et al., Journal of Pharmacology and Experimental Therapeutics, (Feb 2004), Vol. 308, No. 2, pp. (688-693); Wolfrum et al., Nature Biotechnology, (Sep 2007), Vol. 25, No. 10, pp. (1149-1157); Molitoris et al., J Am Soc Nephrol, (Aug 2009), Vol. 20, No. 8 pp. (1754-1764); Mikhaylova et al., Cancer Gene Therapy, (Mar 2011), Vol. 16, No. 3, pp. (217-226); Y. Zhang et al., J Am Soc Nephrol, (Apr 2006), Vol. 17, No. 4, pp. (1090–1101); Singhal et al., Cancer Res, (May 2009), Vol. 69, No. 10, pp. (4244-4251); Malek et al., Toxicology and Applied Pharmacology, (Apr 2009), Vol. 236, No. 1, pp. (97-108); Shimizu et al., J Am Soc Nephrology, (Apr 2010), Vol. 21, No. 4, pp. (622-633); Jiang et al., Molecular Pharmaceutics, (May-Jun 2009), Vol. 6, No. 3, pp. (727-737); Cao et al, J Controlled Release, (Jun 2010), Vol. 144, No. 2, pp. (203-212); Ninichuk et al., Am J Pathol, (Mar 2008), Vol. 172, No. 3, pp. (628-637); Purschke et al., Proc Natl Acad Sci, (Mar 2006), Vol. 103, No. 13, pp. (5173-5178).

[0659] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention are delivered to liver cells. In some embodiments, the liver cell is a hepatocyte. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention can be delivered by via viral vectors, such as AAV (and in particular AAV2/6) vectors. In some embodiments, delivery of a cargo to a liver cell can be by an engineered delivery composition, system, or vesicle of the present invention. These can be administered to a subject by intravenous injection.

[0660] In some embodiments the target for the liver gene therapy, whether in vitro or in vivo, is the albumin gene. This is a so-called ‘safe harbor” as albumin is expressed at very high

levels and so some reduction in the production of albumin following successful gene editing is tolerated. It is also preferred as the high levels of expression seen from the albumin promoter/enhancer allows for useful levels of correct or transgene production (from the inserted donor template) to be achieved even if only a small fraction of hepatocytes are edited. See sites identified by Wechsler et al. (reported at the 57th Annual Meeting and Exposition of the American Society of Hematology - abstract available online at https://ash.confex.com/ash/2015/webprogram/Paper86495.html and presented on 6th December 2015) which can be adapted for use with the engineered compositions, systems, vesicles, and the like of the present invention.

[0661] Exemplary liver and kidney diseases that can be treated and/or prevented are described elsewhere herein, such as in **Tables 3 and 4**.

Treating Diseases Epithelial and Respiratory System

[0662] In some embodiments, the disease treated or prevented by the engineered compositions, systems, vesicles, and the like of the present invention or a cargo delivered by the engineered compositions, systems, vesicles, and the like of the present invention can be a lung or epithelial disease. The engineered compositions, systems, vesicles, and the like of the present invention described herein can be used for treating epithelial and/or lung diseases. Delivery can be to one or both lungs.

[0663] In some embodiments, as viral vector, including but not limited to any of those of the present invention, can be used to deliver the engineered compositions, systems, vesicles, and the like of the present invention to the lungs. In some embodiments, the viral vector is an engineered delivery composition, system, delivery vesicle, and the like of the present invention. In some embodiments, delivery of the engineered compositions, systems, vesicles, and the like of the present invention can be by, based on, or otherwise include an AAV viral vector system. In some embodiments, the AAV is an AAV-1, AAV-2, AAV-5, AAV-6, and/or AAV-9 for delivery to the lungs. (see, e.g., Li et al., Molecular Therapy, vol. 17 no. 12, 2067-2077 Dec 2009). In some embodiments, the MOI delivered by a viral particle or vesicle of the present invention can range from 1×10^3 to 4×10^5 vector genomes/cell. In some embodiments, the delivery vector can be, be based on, or otherwise include an RSV vector as in Zamora et al. (Am J Respir Crit Care Med Vol 183. pp 531–538, 2011. The method of Zamora et al. can be applied to the engineered compositions, systems, vesicles, and the like of the present invention of the present invention and an aerosolized engineered compositions, systems, vesicles, and

the like of the present invention. , A dosage for administration, such as to a human, of the engineered compositions, systems, vesicles, and the like of the present invention can be about 0.6 mg/kg.

[0664] Subjects treated for a lung disease can, for example, receive pharmaceutically effective amount of aerosolized AAV vector system per lung endobronchially delivered while spontaneously breathing. As such, aerosolized delivery is preferred for AAV and other viral or vesicle delivery in general, including those engineered delivery vesicles of the present invention. An adenovirus or an AAV particle may be used for delivery. Suitable gene constructs, each operably linked to one or more regulatory sequences, may be cloned into the delivery vector, such as any of those of the present invention. In this instance, the following constructs are provided as examples: Cbh or EF1a promoter for a Cas (Cas (e.g. Cas9 and/or Cas12)), U6 or H1 promoter for guide RNA). A preferred arrangement is to use a CFTRdelta508 targeting guide, a repair template for deltaF508 mutation and a codon optimized Cas (e.g. Cas9 and/or Cas12) enzyme, with optionally one or more nuclear localization signal or sequence(s) (NLS(s)), e.g., two (2) NLSs.

Treating Diseases of the Skin

[0665] The engineered compositions, systems, vesicles, and the like of the present invention can be used for the treatment of skin diseases and/or delivery of a treatment for one or more skin diseases. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention are capable of delivering a cargo, such as an engineered guided nuclease system, (e.g., a CRISPR-Cas system described herein), to the skin.

[0666] In some embodiments, delivery to the skin (intradermal delivery) of the engineered compositions, systems, vesicles, and the like of the present invention and/or cargo thereof can be via one or more microneedles or microneedle containing device. For example, in some embodiments the device and methods of Hickerson et al. (Molecular Therapy—Nucleic Acids (2013) 2, e129) can be used and/or adapted to deliver the engineered compositions, systems, vesicles, and the like of the present invention. A dosage for administration to the skin of a subject, such as a human, via this method can be to 300 μl of 0.1 mg/ml of the engineered compositions, systems, vesicles, and the like of the present invention.

[0667] In some embodiments, the methods and techniques of Leachman et al. (Molecular Therapy, vol. 18 no. 2, 442–446 Feb. 2010) can be used and/or adapted for delivery of the

engineered compositions, systems, vesicles, and the like of the present invention to the skin of a subject, such as a human.

[0668] In some embodiments, the methods and techniques of Zheng et al. (PNAS, July 24, 2012, vol. 109, no. 30, 11975–11980) can be used and/or adapted for nanoparticle delivery of a engineered compositions, systems, vesicles, and the like of the present invention to the skin. A dosage for administration to the skin of a subject, such as a human can be about 25 nM and can be administered in a single application.

Treating Cancer

[0669] The engineered compositions, systems, vesicles, and the like of the present invention can be used for the treatment of cancer and/or be used to deliver a cargo for treatment of cancer. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention are delivered to one or more cancer cells and/or deliver a cargo to one or more cancer cells. Also, as is described elsewhere herein the engineered compositions, systems, vesicles, and the like of the present invention can be used to modify an immune cell, such as a CAR or CAR T cell, which can then in turn be used to treat and/or prevent cancer. This is also described in WO2015161276, the disclosure of which is hereby incorporated by reference and described in greater detail elsewhere herein.

[0670] Target genes suitable for the treatment or prophylaxis of cancer can include those set forth in **Tables 3 and 4** and those identified at mitoMap.org (for those mitochondrial mutations that can contribute to or cause a cancer). In some embodiments, target genes for cancer treatment and prevention also include those described in WO2015048577, which is hereby incorporated by reference and can be adapted for and/or applied to the engineered compositions, systems, vesicles, and the like of the present invention.

Genetic Therapies for Infectious diseases

[0671] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargo delivered therefrom can be used to diagnose, prognose, treat, and/or prevent an infectious disease caused by a microorganism, such as bacteria, virus, fungi, parasites, or combinations thereof.

[0672] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be capable of targeting specific microorganism within a mixed population. Exemplary methods of such techniques are described in e.g. Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. 2014.

Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. mBio 5:e00928-13; Citorik RJ, Mimee M, Lu TK. 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat Biotechnol 32:1141–1145, the teachings of which can be adapted for use with the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos thereof.

[0673] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be capable of targeting pathogenic and/or drug-resistant microorganisms, such as bacteria, virus, parasites, and fungi. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be capable of targeting and modifying one or more polynucleotides in a pathogenic microorganism such that the microorganism is less virulent, killed, inhibited, or is otherwise rendered incapable of causing disease and/or infecting and/or replicating in a host cell.

[0674] In some embodiments, the pathogenic bacteria that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom herein include, but are not limited to, those of the genus Actinomyces (e.g. A. israelii), Bacillus (e.g., B. anthracis, B. cereus), Bactereoides (e.g., B. fragilis), Bartonella (B. henselae, B. quintana), Bordetella (B. pertussis), Borrelia (e.g., B. burgdorferi, B. garinii, B. afzelii, and B. recurreentis), Brucella (e.g., B. abortus, B. canis, B. melitensis, and B. suis), Campylobacter (e.g., C. jejuni), Chlamydia (e.g., C. pneumoniae and C. trachomatis), Chlamydophila (e.g., C. psittaci), Clostridium (e.g., C. botulinum, C. difficile, C. perfringens. C. tetani), Corynebacterium (e.g., C. diptheriae), Enterococcus (e.g. E. Faecalis, E. faecium), Ehrlichia (E. canis and E. chaffensis) Escherichia (e.g. E. coli), Francisella (e.g. F. tularensis), Haemophilus (e.g. H. influenzae), Helicobacter (H. pylori), Klebsiella (E.g. K. pneumoniae), Legionella (e.g. L. pneumophila), Leptospira (e.g. L. interrogans, L. santarosai, L. weilii, L. noguchii), Listereia (e.g. L. monocytogeenes), Mycobacterium (e.g. M. leprae, M. tuberculosis, M. ulcerans), Mycoplasma (M. pneumoniae), Neisseria (N. gonorrhoeae and N. menigitidis), Nocardia (e.g. N. asteeroides), Pseudomonas (P. aeruginosa), Rickettsia (R. rickettsia), Salmonella (S. typhi and S. typhimurium), Shigella (S. sonnei and S. dysenteriae), Staphylococcus (S. aureus, S. epidermidis, and S. saprophyticus), Streeptococcus (S. agalactiaee, S. pneumoniae, S. pyogenes), Treponema (T. pallidum), Ureeaplasma (e.g. U.

urealyticum), Vibrio (e.g. V. cholerae), Yersinia (e.g. Y pestis, Y, enteerocolitica, and Y, pseudotuberculosis).

[0675] In some embodiments, the pathogenic virus that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom include, but are not limited to, a double-stranded DNA virus, a partly double-stranded DNA virus, a single-stranded DNA virus, a positive single-stranded RNA virus, a negative single-stranded RNA virus, or a double stranded RNA virus. In some embodiments, the pathogenic virus can be from the family Adenoviridae (e.g. Adenovirus), Herpeesviridae (e.g. Herpes simplex, type 1, Herpes simplex, type 2, Varicella-zoster virus, Epstein–Barr virus, Human cytomegalovirus, Human herpesvirus, type 8), Papillomaviridae (e.g. Human papillomavirus), Polyomaviridae (e.g. BK virus, JC virus), Poxviridae (e.g. smallpox), Hepadnaviridae (e.g. Hepatitis B), Parvoviridae (e.g. Parvovirus B19), Astroviridae (e.g. Human astrovirus), Caliciviridae (e.g. Norwalk virus), Picornaviridae (e.g. coxsackievirus, hepatitis A virus, poliovirus, rhinovirus), Coronaviridae (e.g. Severe acute respiratory syndrome-related coronavirus, strains: Severe acute respiratory syndrome virus, Severe acute respiratory syndrome coronavirus 2 (COVID-19)), Flaviviridae (e.g. Hepatitis C virus, yellow fever virus, dengue virus, West Nile virus,TBE virus), Togaviridae (e.g. Rubella virus), Hepeviridae (e.g. Hepatitis E virus), Retroviridae (Human immunodeficiency virus (HIV)), Orthomyxoviridae (e.g. Influenza virus), Arenaviridae (e.g. Lassa virus), Bunyaviridae (e.g. Crimean-Congo hemorrhagic fever virus, Hantaan virus), Filoviridae (e.g. Ebola virus and Marburg virus), Paramyxoviridae (e.g. Measles virus, Mumps virus, Parainfluenza virus, Respiratory syncytial virus), Rhabdoviridae (Rabies virus), Hepatits D virus, Reoviridae (e.g. Rotavirus, Orbivirus, Coltivirus, Banna virus).

[0676] In some embodiments, the pathogenic fungi that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom include, but are not limited to, those of the genus Candida (e.g. C. albicans), Aspergillus (e.g. A. fumigatus, A. flavus, A. clavatus), Cryptococcus (e.g. C. neoformans, C. gattii), Histoplasma (H. capsulatum), Pneumocystis (e.g. P. jiroveecii), Stachybotrys (e.g. S. chartarum).

[0677] In some embodiments, the pathogenic parasites that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom include, but are not limited to, protozoa, helminths, and

ectoparasites. In some embodiments, the pathogenic protozoa that can be targeted and/or modified by the CRISPR-Cas system(s) and/or component(s) thereof described herein include, but are not limited to, those from the groups Sarcodina (e.g. ameba such as Entamoeba), Mastigophora (e.g. flagellates such as Giardia and Leishmania), Cilophora (e.g. ciliates such as Balantidum), and sporozoa (e.g. plasmodium and cryptosporidium). In some embodiments, the pathogenic helminths that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom include, but are not limited to, flatworms (platyhelminths), thorny-headed worms (acanthoceephalins), and roundworms (nematodes). In some embodiments, the pathogenic ectoparasites that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom include, but are not limited to, ticks, fleas, lice, and mites.

[0678] In some embodiments, the pathogenic parasite that can be targeted and/or modified by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom include, but are not limited to, Acanthamoeba spp., Balamuthia mandrillaris, Babesiosis spp. (e.g. Babesia B. divergens, B. bigemina, B. equi, B. microfti, B. duncani), Balantidiasis spp. (e.g. Balantidium coli), Blastocystis spp., Cryptosporidium spp., Cyclosporiasis spp. (e.g. Cyclospora cayetanensis), Dientamoebiasis spp. (e.g. Dientamoeba fragilis), Amoebiasis spp. (e.g. Entamoeba histolytica), Giardiasis spp. (e.g. Giardia lamblia), Isosporiasis spp. (e.g. Isospora belli), Leishmania spp., Naegleria spp. (e.g. Naegleria fowleri), Plasmodium spp. (e.g. Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale curtisi, Plasmodium ovale wallikeri, Plasmodium malariae, Plasmodium knowlesi), Rhinosporidiosis spp. (e.g. Rhinosporidium seeberi), Sarcocystosis spp. (e.g. Sarcocystis bovihominis, Sarcocystis suihominis), Toxoplasma spp. (e.g. Toxoplasma gondii), Trichomonas spp. (e.g. Trichomonas vaginalis), Trypanosoma spp. (e.g. Trypanosoma brucei), Trypanosoma spp. (e.g. Trypanosoma cruzi), Tapeworm (e.g. Cestoda, Taenia multiceps, Taenia saginata, Taenia solium), Diphyllobothrium latum spp., Echinococcus spp. (e.g. Echinococcus granulosus, Echinococcus multilocularis, E. vogeli, E. oligarthrus), Hymenolepis spp. (e.g. Hymenolepis nana, Hymenolepis diminuta), Bertiella spp. (e.g. Bertiella mucronata, Bertiella studeri), Spirometra (e.g. Spirometra erinaceieuropaei), Clonorchis spp. (e.g. Clonorchis sinensis; Clonorchis viverrini), Dicrocoelium spp. (e.g. Dicrocoelium dendriticum), Fasciola spp. (e.g. Fasciola hepatica, Fasciola gigantica), Fasciolopsis spp. (e.g. Fasciolopsis buski),

Metagonimus spp. (e.g. Metagonimus yokogawai), Metorchis spp. (e.g. Metorchis conjunctus), Opisthorchis spp. (e.g. Opisthorchis viverrini, Opisthorchis felineus), Clonorchis spp. (e.g. Clonorchis sinensis), Paragonimus spp. (e.g. Paragonimus westermani; Paragonimus africanus; Paragonimus caliensis; Paragonimus kellicotti; Paragonimus skrjabini; Paragonimus uterobilateralis), Schistosoma sp., Schistosoma spp. (e.g. Schistosoma mansoni, Schistosoma haematobium, Schistosoma japonicum, Schistosoma mekongi, and Schistosoma intercalatum), Echinostoma spp. (e.g. E. echinatum), Trichobilharzia spp. (e.g. Trichobilharzia regent), Ancylostoma spp. (e.g. Ancylostoma duodenale), Necator spp. (e.g. Necator americanus), Angiostrongylus spp., Anisakis spp., Ascaris spp. (e.g. Ascaris lumbricoides), Baylisascaris spp. (e.g. Baylisascaris procyonis), Brugia spp. (e.g. Brugia malayi, Brugia timori), Dioctophyme spp. (e.g. Dioctophyme renale), Dracunculus spp. (e.g. Dracunculus medinensis), Enterobius spp. (e.g. Enterobius vermicularis, Enterobius gregorii), Gnathostoma spp. (e.g. Gnathostoma spinigerum, Gnathostoma hispidum), Halicephalobus spp. (e.g. Halicephalobus gingivalis), Loa loa spp. (e.g. Loa loa filaria), Mansonella spp. (e.g. Mansonella streptocerca), Onchocerca spp. (e.g. Onchocerca volvulus), Strongyloides spp. (e.g. Strongyloides stercoralis), Thelazia spp. (e.g. Thelazia californiensis, Thelazia callipaeda), Toxocara spp. (e.g. Toxocara canis, Toxocara cati, Toxascaris leonine), Trichinella spp. (e.g. Trichinella spiralis, Trichinella britovi, Trichinella nelsoni, Trichinella nativa), Trichuris spp. (e.g. Trichuris trichiura, Trichuris vulpis), Wuchereria spp. (e.g. Wuchereria bancrofti), Dermatobia spp. (e.g. Dermatobia hominis), Tunga spp. (e.g. Tunga penetrans), Cochliomyia spp. (e.g. Cochliomyia hominivorax), Linguatula spp. (e.g. Linguatula serrata), Archiacanthocephala sp., Moniliformis sp. (e.g. Moniliformis moniliformis), Pediculus spp. (e.g. Pediculus humanus capitis, Pediculus humanus humanus), Pthirus spp. (e.g. Pthirus pubis), Arachnida spp. (e.g. Trombiculidae, Ixodidae, Argaside), Siphonaptera spp (e.g. Siphonaptera: Pulicinae), Cimicidae spp. (e.g. Cimex lectularius and Cimex hemipterus), Diptera spp., Demodex spp. (e.g. Demodex folliculorum/brevis/canis), Sarcoptes spp. (e.g. Sarcoptes scabiei), Dermanyssus spp. (e.g. Dermanyssus gallinae), Ornithonyssus spp. (e.g. Ornithonyssus sylviarum, Ornithonyssus bursa, Ornithonyssus bacoti), Laelaps spp. (e.g. Laelaps echidnina), Liponyssoides spp. (e.g. Liponyssoides sanguineus).

[0679] In some embodiments the gene targets can be any of those as set forth in Table 1 of Strich and Chertow. 2019. J. Clin. Microbio. 57:4 e01307-18, which is incorporated herein as if expressed in its entirety herein.

[0680] In some embodiments, the method can include delivering a engineered compositions, systems, vesicles, and the like of the present invention and/or cargos contained by the engineered compositions, systems, vesicle, and the like of the present invention to a pathogenic organism described herein, allowing the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom to specifically bind and modify one or more targets in the pathogenic organism, whereby the modification kills, inhibits, reduces the pathogenicity of the pathogenic organism, or otherwise renders the pathogenic organism non-pathogenic. In some embodiments, delivery of the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom occurs in vivo (i.e. in the subject being treated). In some embodiments occurs by an intermediary, such as microorganism or phage that is non-pathogenic to the subject but is capable of transferring polynucleotides and/or infecting the pathogenic microorganism. In some embodiments, the intermediary microorganism can be an engineered bacteria, virus, or phage that contains the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom. The method can include administering an intermediary microorganism containing the engineered compositions, systems, vesicles, and the like of the present invention to the subject to be treated. The intermediary microorganism can then produce the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom or transfer the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom to the pathogenic organism. In embodiments, where the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom is transferred to the pathogenic microorganism, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom is then produced in the pathogenic microorganism and modifies the pathogenic microorganism such that it is less virulent, killed, inhibited, or is otherwise rendered incapable of causing disease and/or infecting and/or replicating in a host or cell thereof.

[0681] In some embodiments, where the pathogenic microorganism inserts its genetic material into the host cell's genome (e.g., a virus), the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be designed such that it modifies the host cell's genome such that the viral DNA or cDNA cannot be replicated by the host cell's machinery into a functional virus. In some embodiments, where

the pathogenic microorganism inserts its genetic material into the host cell's genome (e.g., a virus), the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be designed such that it modifies the host cell's genome such that the viral DNA or cDNA is deleted from the host cell's genome.

[0682] It will be appreciated that inhibiting or killing the pathogenic microorganism, the disease and/or condition that its infection causes in the subject can be treated or prevented. Thus, also provided herein are methods of treating and/or preventing one or more diseases or symptoms thereof caused by any one or more pathogenic microorganisms, such as any of those described herein.

Genetic Treatment of Mitochondrial Diseases

[0683] Some of the most challenging mitochondrial disorders arise from mutations in mitochondrial DNA (mtDNA), a high copy number genome that is maternally inherited. In some embodiments, mtDNA mutations can be modified using the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom. In some embodiments, the mitochondrial disease that can be diagnosed, prognosed, treated, and/or prevented can be MELAS (mitochondrial myopathy encephalopathy, and lactic acidosis and stroke-like episodes), CPEO/PEO (chronic progressive external ophthalmoplegia syndrome/progressive external ophthalmoplegia), KSS (Kearns-Sayre syndrome), MIDD (maternally inherited diabetes and deafness), MERRF (myoclonic epilepsy associated with ragged red fibers), NIDDM (noninsulin-dependent diabetes mellitus), LHON (Leber hereditary optic neuropathy), LS (Leigh Syndrome) an aminoglycoside induced hearing disorder, NARP (neuropathy, ataxia, and pigmentary retinopathy), Extrapyramidal disorder with akinesia-rigidity, psychosis and SNHL, Nonsyndromic hearing loss a cardiomyopathy, an encephalomyopathy, Pearson's syndrome, a disease identified as being caused or attributed to a mtDNA mutation set forth at mitomap.org, or a combination thereof.

[0684] In some embodiments, the mtDNA of a subject can be modified in vivo or ex vivo. In some embodiments, where the mtDNA is modified ex vivo, after modification the cells containing the modified mitochondria can be administered back to the subject. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom is capable of correcting an mtDNA mutation such as any one or more of those that can be found at mitomap.org.

[0685] In some embodiments, at least one of the one or more mtDNA mutations is selected from the group consisting of: A3243G, C3256T, T3271C, G1019A, A1304T, A15533G, C1494T, C4467A, T1658C, G12315A, A3421G, A8344G, T8356C, G8363A, A13042T, T3200C, G3242A, A3252G, T3264C, G3316A, T3394C, T14577C, A4833G, G3460A, G9804A, G11778A, G14459A, A14484G, G15257A, T8993C, T8993G, G10197A, G13513A, T1095C, C1494T, A1555G, G1541A, C1634T, A3260G, A4269G, T7587C, A8296G, A8348G, G8363A, T9957C, T9997C, G12192A, C12297T, A14484G, G15059A, duplication of CCCCCTCCCC-tandem repeats at positions 305-314 and/or 956-965, deletion at positions from 8,469-13,447, 4,308-14,874, and/or 4,398-14,822, 961ins/delC, the mitochondrial common deletion (e.g. mtDNA 4,977 bp deletion), and combinations thereof.

[0686] In some embodiments, the mitochondrial mutation can be any mutation as set forth in or as identified by use of one or more bioinformatic tools available at Mitomap available at mitomap.org. Such tools include, but are not limited to, "Variant Search, aka Market Finder", Find Sequences for Any Haplogroup, aka "Sequence Finder", "Variant Info", "POLG Pathogenicity Prediction Server", "MITOMASTER", "Allele Search", "Sequence and Variant Downloads", "Data Downloads". MitoMap contains reports of mutations in mtDNA that can be associated with disease and maintains a database of reported mitochondrial DNA Base Substitution Diseases: rRNA/tRNA mutations.

[0687] In some embodiments, the method includes delivering an engineered composition, system, vesicle, and the like of the present invention and/or cargo delivered therefrom to a cell, and more specifically one or more mitochondria in a cell, allowing the engineered composition, system, vesicle, and the like of the present invention and/or cargo delivered therefrom to modify one or more target polynucleotides in the cell, and more specifically one or more mitochondria in the cell. The target polynucleotides can correspond to a mutation in the mtDNA, such as any one or more of those described herein. In some embodiments, the modification can alter a function of the mitochondria such that the mitochondria functions normally or at least is/are less dysfunctional as compared to an unmodified mitochondria. Modification can occur in vivo or ex vivo. Where modification is performed ex vivo, cells containing modified mitochondria can be administered to a subject in need thereof in an autologous or allogenic manner.

Genetic Microbiome Modification

[0688] Microbiomes play important roles in health and disease. For example, the gut microbiome can play a role in health by controlling digestion, preventing growth of pathogenic microorganisms and have been suggested to influence mood and emotion. Imbalanced microbiomes can promote disease and are suggested to contribute to weight gain, unregulated blood sugar, high cholesterol, cancer, and other disorders. A healthy microbiome has a series of joint characteristics that can be distinguished from non-healthy individuals, thus detection and identification of the disease-associated microbiome can be used to diagnose and detect disease in an individual. The engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be used to screen the microbiome cell population and be used to identify a disease associated microbiome. Cell screening methods utilizing engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom are described elsewhere herein and can be applied to screening a microbiome, such as a gut, skin, vagina, and/or oral microbiome, of a subject.

[0689] In some embodiments, the microbe population of a microbiome in a subject can be modified using the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be used to identify and select one or more cell types in the microbiome and remove them from the microbiome population. Exemplary methods of selecting cells using the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom are described elsewhere herein. In this way the make-up or microorganism profile of the microbiome can be altered. In some embodiments, the alteration causes a change from a diseased microbiome composition to a healthy microbiome composition. In this way the ratio of one type or species of microorganism to another can be modified, such as going from a diseased ratio to a healthy ratio. In some embodiments, the cells selected are pathogenic microorganisms.

[0690] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be used to modify a polynucleotide in a microorganism of a microbiome in a subject. In some embodiments, the microorganism is a pathogenic microorganism. In some embodiments, the microorganism is a

commensal and non-pathogenic microorganism. Methods of modifying polynucleotides in a cell in the subject are described elsewhere herein and can be applied to these embodiments.

Treatment and Prevention of Plant Pathogens and Disease

[0691] In particular embodiments, present invention can be used as a therapy for virus removal in plant systems as it is able to cleave both viral DNA and RNA. Previous studies in human systems have demonstrated the success of utilizing CRISPR in targeting the single strand RNA virus, hepatitis C (A. Price, et al., Proc. Natl. Acad. Sci, 2015) as well as the double stranded DNA virus, hepatitis B (V. Ramanan, et al., Sci. Rep, 2015). These methods may also be adapted for using the systems in plants.

[0692] The engineered compositions, systems, formulations, cargos, and delivery vesicles of the present invention can be configured to induce mutations, to analyze the genome of sources of resistance genes, and in varieties having desired characteristics or traits employ the present invention to induce the rise of resistance genes, with more precision than previous mutagenic agents and hence accelerate and improve plant breeding programs. Further, the engineered compositions, systems, formulations, cargos, and delivery vesicles of the present invention can be used to generate plants with one or more disease resistant genes or alleles. Thus, the engineered plants of the present invention can be used to identify pathogen and disease resistance genes as well as be modified so as to be resistant to one or more pathogens and/or diseases.

[0693] In some embodiments, the modified plants are modified so as to be resistant to Fusarium oxysporum f. sp. Lycopersici, F. oxysporum f. dianthii Puccinia graminis f. sp. tritici attacks only wheat. In some embodiments, the plant modified is modified so that it has non-host resistance (e.g., the host and pathogen are incompatible) to one or more pathogens. In some embodiments, the plant is modified such that it has horizontal resistance (e.g., partial resistance against all races of a pathogen). Such modification can include modification of and/or introduction of one or more genes. In some embodiments, the modified plant is modified such that it has vertical resistance (e.g., complete resistance to some races of a pathogen but not to other races). Such modification can include modification of and/or introduction of one or more genes, but typically less than the number of modifications/insertions employed for horizontal resistance.

[0694]

Non-Disease Genetic Enhancements

Generation of Engineered Cells and Organisms

[0695] The engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be used to generate engineered cells, tissues, and organisms. The modified cells, tissues, and organisms can have one or more modifications (e.g., insertions, deletions, substitutions, etc.) in one or more polynucleotides (e.g., DNA and/or RNA). Modifications can be transient (e.g., made directly to RNA) or stable (e.g., made to genomic DNA). The cells can be modified by delivering a polynucleotide modifying agent or system (e.g., an RNA guided nuclease or other genetic modifying system) described in greater detail elsewhere herein or a component thereof into a cell by a suitable delivery mechanism, such as the engineered compositions, systems, vesicles, etc. of the present invention. In other words, in some embodiments, the polynucleotide modifying agent or system is a cargo that can be delivered by one or more of the engineered compositions, systems, vesicles, or formulations thereof of the present invention. Suitable delivery methods and techniques include but are not limited to, transfection via a vector, transduction with viral particles, electroporation, endocytic methods, and others, which are described elsewhere herein and will be appreciated by those of ordinary skill in the art in view of this disclosure.

[0696] The modified cells can be further optionally cultured and/or expanded in vitro or ex vivo using any suitable cell culture techniques or conditions, which unless specified otherwise herein, will be appreciated by one of ordinary skill in the art in view of this disclosure. In some embodiments, the cells can be modified, optionally cultured and/or expanded, and administered to a subject in need thereof. In some embodiments, cells can be isolated from a subject, subsequently modified and optionally cultured and/or expanded, and administered back to the subject. Such administration can be referred to as autologous administration. In some embodiments, cells can be isolated from a first subject, subsequently modified, optionally cultured and/or expanded, and administered to a second subject, where the first subject and the second subject are different. Such administration can be referred to as non-autologous administration.

[0697] Specific methods of generating modified cells are described in greater detail elsewhere herein.

Cargo Delivery

[0698] As described elsewhere herein, the engineered retroviral delivery vesicle generation systems and retroviral delivery vesicle systems described herein can deliver a cargo to a cell, such as a recipient cell or a target cell or cell in need thereof. In some embodiments, a method can include delivering or administering an engineered retroviral delivery vesicle containing one or more cargos to a subject in need thereof, including to one or more cells. In some embodiments, the method includes expressing one or more engineered retroviral delivery systems of the present invention in one or more cells and delivery the cell and/or vesicles produced therefrom to a subject in need thereof. This is further described elsewhere herein

Cell Selection

[0699] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom (such as RNA guided nuclease systems) can be used in a method to screen and/or select cells. In some embodiments, an RNA guided nuclease system (e.g., a CRISPR-Cas system)-based screening/selection method can be used to identify diseased cells in a cell population. In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom form, produce, or are a component of an RNA guided nuclease system (such as a CRISPR-Cas system). It will be appreciated that other configurations and/or cargos of the engineered compositions, systems, vesicles, and the like of the present invention can be used in an appropriate manner for cell selection. For example, by modifying a cell to produce a selectable tag or biomarker, which can then be positively or negatively selected for so as to select for or against a cell state or cell type. In some embodiments, selection of the cells results in a modification in the cells such that the selected cells die. In this way, diseased cells can be identified, and removed from the healthy cell population. In some embodiments, the diseased cells can be a cancer cell, pre-cancerous cell, a virus or other pathogenic organism infected cells, or otherwise abnormal cell. In some embodiments, the modification can impart another detectable change in the cells to be selected (e.g., a functional change and/or genomic barcode) that facilitates selection of the desired cells. In some embodiments a negative selection scheme can be used to obtain a desired cell population. In these embodiments, the cells to be selected against are modified, thus can be removed from the cell population based on their death or identification or sorting based the detectable change imparted on the cells. Thus, in these embodiments, the remaining cells after selection are the desired cell population.

[0700] In some embodiments, a method of selecting one or more cell(s) containing a polynucleotide modification can include: introducing one or more the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom into the cell(s), wherein the the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom vectors or vector systems contains and/or is capable of expressing one or more of: an RNA guided nuclease effector (e.g. a Cas or IscB), a suitable guide sequence optionally linked to a tracr mate sequence, a tracr sequence, and optionally an editing template. In some embodiments, that which is being expressed is within and expressedo by the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom, the editing template contains the one or more mutations that abolish RNA guided nuclease effector cleavage; allowing homologous recombination of the editing template with the target polynucleotide in the cell(s) to be selected; allowing a CRISPR or other RNA guided nuclease complex to bind to a target polynucleotide to effect cleavage of the target polynucleotide within said gene, wherein the CRISPR or other RNA guided nuclease complex that contains the RNA guided nuclease effector complexed with (1) the guide sequence that is hybridized to the target sequence within the target polynucleotide, and optionally (2) the tracr mate sequence that is hybridized to the tracr sequence, where binding of the RNA guided nuclease complex to the target polynucleotide induces cell death or imparts some other detectable change to the cell, thereby allowing one or more cell(s) in which one or more mutations have been introduced to be selected. In a preferred embodiment, the RNA guided nuclease effector is a Cas (e.g., Cas9 or Cas12) or an IscB. In some embodiments, the cell to be selected can be a eukaryotic cell. In some embodiments, the cell to be selected may be a prokaryotic cell. Selection of specific cells via the methods herein can be performed without requiring a selection marker or a two-step process that may include a counter-selection system.

Therapeutic Agent Development

[0701] The engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be used to develop biologically active agents, such as small molecule therapeutics, RNA guided nuclease systems (e.g., CRISPR-Cas and/or IscB systems), and the like. As used herein, “active agent” or “active ingredient” refers to a substance, compound, or molecule, which is biologically active or otherwise, induces a biological or physiological effect on a subject to which it is administered to. In other words,

"active agent" or "active ingredient" refers to a component or components of a composition to which the whole or part of the effect of the composition is attributed. As used herein, "agent" refers to any substance, compound, molecule, and the like, which can be biologically active or otherwise can induce a biological and/or physiological effect on a subject to which it is administered to. An agent can be a primary active agent, or in other words, the component(s) of a composition to which the whole or part of the effect of the composition is attributed. An agent can be a secondary agent, or in other words, the component(s) of a composition to which an additional part and/or other effect of the composition is attributed. Thus, described herein are methods for developing a biologically active agent that modulates a cell function and/or signaling event associated with a disease and/or disease gene. In some embodiments, the method comprises (a) contacting a test compound with a diseased cell and/or a cell containing a disease gene cell; and (b) detecting a change in a readout that is indicative of a reduction or an augmentation of a cell signaling event or other cell functionality associated with said disease or disease gene, thereby developing said biologically active agent that modulates said cell signaling event or other functionality associated with said disease gene. In some embodiments, the diseased cell is a model cell described elsewhere herein. In some embodiments, the diseased cell is a diseased cell isolated from a subject in need of treatment. In some embodiments, the test compound is a small molecule agent. In some embodiments, test compound is a small molecule agent. In some embodiments, the test compound is a biologic molecule agent.

[0702] In some embodiments, the method involves developing a therapeutic based on the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom. In some embodiments, the cargo is an RNA guided nuclease system. In particular embodiments, the therapeutic comprises an RNA guided nuclease effector (e.g., a Cas or an IscB) and/or a guide RNA capable of hybridizing to a target sequence of interest. In particular embodiments, the therapeutic is an RNA guided nuclease vector or vector system that can contain a) a first regulatory element operably linked to a nucleotide sequence encoding the RNA guided nuclease effector protein(s); and b) a second regulatory element operably linked to one or more nucleotide sequences encoding one or more nucleic acid molecules comprising a guide RNA comprising a guide sequence, a direct repeat sequence; wherein components (a) and (b) are located on same or different vectors. In particular embodiments, the biologically active agent is a composition comprising a delivery system operably configured to deliver an RNA guided nuclease system or components thereof, and/or or one or

more polynucleotide sequences, vectors, or vector systems containing or encoding said components into a cell and capable of forming a RNA guided nuclease complex, and where the RNA guided nuclease complex is operable in the cell. In some embodiments, the RNA guided nuclease complex can include a Cas or IscB effector protein(s) as described herein, guide RNA comprising the guide sequence, and one or more direct repeat sequences. In any such compositions, the delivery system can be an engineered delivery system of the present invention, a yeast system, a lipofection system, a microinjection system, a biolistic system, virosomes, liposomes, immunoliposomes, polycations, lipid:nucleic acid conjugates or artificial virions, or any other system as described herein. In particular embodiments, the delivery is via a particle, a nanoparticle, a lipid or a cell penetrating peptide (CPP).

[0703] Also described herein are methods for developing or designing an RNA guided nuclease system, optionally a CRISPR-Cas or IscB system based therapy or therapeutic, comprising (a) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest gRNA target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, and optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more gRNA recognizing one or more of said (sub)selected target sites.

[0704] In some embodiments, the method for developing or designing a gRNA for use in a RNA guided nuclease system, optionally a CRISPR-Cas or IscB system based therapy or therapeutic, can include (a) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest gRNA target sites, wherein a gRNA directed against said target sites recognizes a minimal number of

off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more gRNA recognizing one or more of said (sub)selected target sites.

[0705] In some embodiments, the method for developing or designing an RNA guided nuclease system, optionally a CRISPR-Cas or IscB system based therapy or therapeutic in a population, can include (a) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest gRNA target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more gRNA recognizing one or more of said (sub)selected target sites.

[0706] In some embodiments the method for developing or designing a gRNA for use in a RNA guided nuclease system, optionally a CRISPR-Cas or IscB system based therapy or therapeutic in a population, can include (a) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, and from said selected target sites subselecting target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, or (b) selecting for a (therapeutic) locus of interest gRNA target sites, wherein said target sites have minimal sequence variation across a population, or selecting for a (therapeutic) locus of interest gRNA target sites, wherein a gRNA directed against said target sites recognizes a minimal number of off-target sites across said population, and optionally estimating the number of (sub)selected target sites needed to treat or otherwise modulate or manipulate a population, optionally validating one or more of the (sub)selected target sites for an individual subject, optionally designing one or more gRNA recognizing one or more of said (sub)selected target sites.

[0707] In some embodiments, the method for developing or designing an RNA guided nuclease system, such as a CRISPR-Cas or IscB system based therapy or therapeutic, optionally in a population; or for developing or designing a gRNA for use in an RNA guided nuclease system, optionally a CRISPR-Cas system or IscB based therapy or therapeutic, optionally in a population, can include: selecting a set of target sequences for one or more loci in a target population, wherein the target sequences do not contain variants occurring above a threshold allele frequency in the target population (i.e. platinum target sequences); removing from said selected (platinum) target sequences any target sequences having high frequency off-target candidates (relative to other (platinum) targets in the set) to define a final target sequence set; preparing one or more, such as a set of RNA guided nuclease systems based on the final target sequence set, optionally wherein a number of RNA guided nuclease systems prepared is based (at least in part) on the size of a target population.

[0708] In certain embodiments, off-target candidates/off-targets, PAM restrictiveness, target cleavage efficiency, or effector protein specificity is identified or determined using a sequencing-based double-strand break (DSB) detection assay, such as described herein elsewhere. In certain embodiments, off-target candidates/off-targets are identified or determined using a sequencing-based double-strand break (DSB) detection assay, such as described herein elsewhere. In certain embodiments, off-targets, or off target candidates have at least 1, preferably 1-3, mismatches or (distal) PAM mismatches, such as 1 or more, such as 1, 2, 3, or more (distal) PAM mismatches. In certain embodiments, sequencing-based DSB detection assay comprises labeling a site of a DSB with an adapter comprising a primer binding site, labeling a site of a DSB with a barcode or unique molecular identifier, or combination thereof, as described herein elsewhere.

[0709] It will be understood that the guide sequence of the gRNA is 100% complementary to the target site, i.e., does not comprise any mismatch with the target site. It will be further understood that "recognition" of an (off-)target site by a gRNA presupposes CRISPR-Cas system or other RNA guided nuclease system functionality, i.e. an (off-)target site is only recognized by a gRNA if binding of the gRNA to the (off-)target site leads to CRISPR-Cas system activity or other RNA guided nuclease system (such as induction of single or double strand DNA cleavage, transcriptional modulation, etc.).

[0710] In certain embodiments, the target sites having minimal sequence variation across a population are characterized by absence of sequence variation in at least 99%, preferably at

least 99.9%, more preferably at least 99.99% of the population. In certain embodiments, optimizing target location comprises selecting target sequences or loci having an absence of sequence variation in at least 99%, %, preferably at least 99.9%, more preferably at least 99.99% of a population. These targets are referred to herein elsewhere also as "platinum targets". In certain embodiments, said population comprises at least 1000 individuals, such as at least 5000 individuals, such as at least 10000 individuals, such as at least 50000 individuals.

[0711] In certain embodiments, the off-target sites are characterized by at least one mismatch between the off-target site and the gRNA. In certain embodiments, the off-target sites are characterized by at most five, preferably at most four, more preferably at most three mismatches between the off-target site and the gRNA. In certain embodiments, the off-target sites are characterized by at least one mismatch between the off-target site and the gRNA and by at most five, preferably at most four, more preferably at most three mismatches between the off-target site and the gRNA.

[0712] In certain embodiments, said minimal number of off-target sites across said population is determined for high-frequency haplotypes in said population. In certain embodiments, said minimal number of off-target sites across said population is determined for high-frequency haplotypes of the off-target site locus in said population. In certain embodiments, said minimal number of off-target sites across said population is determined for high-frequency haplotypes of the target site locus in said population. In certain embodiments, the high-frequency haplotypes are characterized by occurrence in at least 0.1% of the population.

[0713] In certain embodiments, the number of (sub)selected target sites needed to treat a population is estimated based on based low frequency sequence variation, such as low frequency sequence variation captured in large scale sequencing datasets. In certain embodiments, the number of (sub)selected target sites needed to treat a population of a given size is estimated.

[0714] In certain embodiments, the method further include obtaining genome sequencing data of a subject to be treated; and treating the subject with a CRISPR-Cas system or other RNA guided nuclease system selected from the set of CRISPR-Cas or other RNA guided nucleases systems, wherein the CRISPR-Cas or other guided RNA nuclease system selected is based (at least in part) on the genome sequencing data of the individual. In certain

embodiments, the ((sub)selected) target is validated by genome sequencing, preferably whole genome sequencing.

[0715] In certain embodiments, target sequences or loci as described herein are (further) selected based on optimization of one or more parameters, such as PAM type (natural or modified), PAM nucleotide content, PAM length, target sequence length, PAM restrictiveness, target cleavage efficiency, and target sequence position within a gene, a locus or other genomic region. Methods of optimization are discussed in greater detail elsewhere herein.

[0716] In certain embodiments, target sequences or loci as described herein are (further) selected based on optimization of one or more of target loci location, target length, target specificity, and PAM characteristics. As used herein, PAM characteristics may comprise for instance PAM sequence, PAM length, and/or PAM GC contents. In certain embodiments, optimizing PAM characteristics comprises optimizing nucleotide content of a PAM. In certain embodiments, optimizing nucleotide content of PAM is selecting a PAM with a motif that maximizes abundance in the one or more target loci, minimizes mutation frequency, or both. Minimizing mutation frequency can for instance be achieved by selecting PAM sequences devoid of or having low or minimal CpG.

[0717] In certain embodiments, the effector protein for each RNA guided nuclease system in the set of RNA guided nuclease systems is selected based on optimization of one or more parameters selected from the group consisting of; effector protein size, ability of effector protein to access regions of high chromatin accessibility, degree of uniform enzyme activity across genomic targets, epigenetic tolerance, mismatch/budge tolerance, effector protein specificity, effector protein stability or half-life, effector protein immunogenicity or toxicity. Methods of optimization are discussed in greater detail elsewhere herein.

Gene Drives

[0718] In some embodiments, the engineered compositions, systems, vesicles, and the like of the present invention and/or cargos delivered therefrom can be used, to provide and/or generate RNA-guided gene drives, for example in systems analogous to gene drives described in PCT Patent Publication WO 2015/105928. Systems of this kind may for example provide methods for altering eukaryotic germline cells, by introducing into the germline cell a nucleic acid sequence encoding an RNA-guided DNA nuclease and one or more guide RNAs. The guide RNAs may be designed to be complementary to one or more target locations on genomic DNA of the germline cell. The nucleic acid sequence encoding the RNA guided DNA nuclease

and the nucleic acid sequence encoding the guide RNAs may be provided on constructs between flanking sequences, with promoters arranged such that the germline cell may express the RNA guided DNA nuclease and the guide RNAs, together with any desired cargo-encoding sequences that are also situated between the flanking sequences. The flanking sequences will typically include a sequence which is identical to a corresponding sequence on a selected target chromosome, so that the flanking sequences work with the components encoded by the construct to facilitate insertion of the foreign nucleic acid construct sequences into genomic DNA at a target cut site by mechanisms such as homologous recombination, to render the germline cell homozygous for the foreign nucleic acid sequence. In this way, gene-drive systems are capable of introgressing desired cargo genes throughout a breeding population (Gantz et al., 2015, Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi, PNAS 2015, published ahead of print November 23, 2015, doi:10.1073/pnas.1521077112; Esvelt et al., 2014, Concerning RNA-guided gene drives for the alteration of wild populations eLife 2014;3:e03401). In select embodiments, target sequences may be selected which have few potential off-target sites in a genome. Targeting multiple sites within a target locus, using multiple guide RNAs, may increase the cutting frequency and hinder the evolution of drive resistant alleles. Truncated guide RNAs may reduce off-target cutting. Paired nickases may be used instead of a single nuclease, to further increase specificity. Gene drive constructs may include cargo sequences encoding transcriptional regulators, for example to activate homologous recombination genes and/or repress non-homologous end-joining. Target sites may be chosen within an essential gene, so that non-homologous end-joining events may cause lethality rather than creating a drive-resistant allele. The gene drive constructs can be engineered to function in a range of hosts at a range of temperatures (Cho et al. 2013, Rapid and Tunable Control of Protein Stability in Caenorhabditis elegans Using a Small Molecule, PLoS ONE 8(8): e72393. doi:10.1371/journal.pone.0072393).

METHODS OF USING ENGINEERED CELLS AND ORGANSIMS

[0719] ...

Cell and Tissue Based Therapies

[0720] The compositions, systems, and formulations described elsewhere herein can be used in the context of Cell and Tissue based therapies, including but not limited to Adoptive Cell Transfer, cell transfer and replacement therapies, and regenerative medicine. For example,

the compositions and systems described herein can be used to deliver a cargo to one or more cells, which can subsequently modify said cells to provide cells with a desired functionality(ies) and/or characteristic(s). The modified cells can then be transferred to a subject in need thereof to prevent and/or treat a disease or condition or a symptom thereof. It will be appreciated that the term of art "Adoptive Cell Transfer" is generally used in connection with the modification and/or transfer of immune cells or the use of cells for an immune-based functionality. It will be appreciated that the term cell transfer and cell replacement therapy can refer to both immune and non-immune cell based therapies. The term "regenerative medicine or therapy" generally refers to cell and tissue based therapies aimed at regenerating and/or replacing missing, dead, diseased, non-functional, or otherwise ineffective cells, tissue, organs, or systems to reestablish function, structure and/or prevent or treat a disease, condition, or a symptom thereof that results from the missing, dead, diseased, non-functional, or otherwise ineffective cells, tissue, organs, or systems. Methods of modifying cells using the compositions, formulations, and systems provided herein are described in greater detail elsewhere herein. Generally, cells can be modified for cell based therapies and/or regenerative medicine *in vitro*, *in vivo*, and/or *ex vivo*.

Adoptive Cell Transfer

[0721] The compositions, systems, and components thereof described herein can be used to modify cells for an adoptive cell therapy. As used herein, "ACT", "adoptive cell therapy" and "adoptive cell transfer" are used interchangeably herein and refer. to the transfer of cells to a patient with the goal of transferring the functionality and characteristics into the new host by engraftment of the cells (see, e.g., Mettananda et al., Editing an α-globin enhancer in primary human hematopoietic stem cells as a treatment for β-thalassemia, Nat Commun. 2017 Sep 4;8(1):424). As used herein, the term "engraft" or "engraftment" refers to the process of cell incorporation into a tissue of interest *in vivo* through contact with existing cells of the tissue.

[0722] Cells used in ACT can be autologous or allogenic. The use autologous cells can be advantageous in some embodiments as it helps the recipient by minimizing graft-versus-host disease (GVDH) and associated issues. As described further herein, allogenic cells can be modified (e.g., such as edited) to reduce alloreactivity and prevent GVHD. Thus, use of allogenic cells allows for cells to be obtained from healthy donors and prepared for use in patients as opposed to preparing autologous cells from a patient after diagnosis. Further, allogenic cells can be prepared in advance and ready as "off the shelf" cells that can be

delivered to a recipient, allowing for more immediate treatment than if the therapy was autologous cell-based. Formulations containing cells for ACT have a concentration of cells that can range from, in some embodiments, about $1x10^3$ to $1x10^4$, $1x10^5$, $1x10^6$, $1x10^7$,$1x10^8$,$1x10^9$, $1x10^{10}$, $1x10^{11}$, $1x10^{12}$, $1x10^{13}$, $1x10^{14}$, $1x10^{15}$ cells/mL or more.

[0723] Adoptive cell therapy (ACT) refers, in some embodiments, to the transfer of cells, most commonly immune-derived cells, back into the same patient or into a new recipient host with the goal of transferring the immunologic functionality and characteristics into the new host. In some embodiments, a method of ACT includes delivery of immune system cells, such as T cells, specific for selected antigens, such as tumor associated antigens or tumor specific neoantigens (see, e.g., Maus et al., 2014, Adoptive Immunotherapy for Cancer or Viruses, Annual Review of Immunology, Vol. 32: 189-225; Rosenberg and Restifo, 2015, Adoptive cell transfer as personalized immunotherapy for human cancer, Science Vol. 348 no. 6230 pp. 62-68; Restifo et al., 2015, Adoptive immunotherapy for cancer: harnessing the T cell response. Nat. Rev. Immunol. 12(4): 269-281; and Jenson and Riddell, 2014, Design and implementation of adoptive therapy with chimeric antigen receptor-modified T cells. Immunol Rev. 257(1): 127–144; and Rajasagi et al., 2014, Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. Blood. 2014 Jul 17;124(3):453-62).

[0724] The adoptive transfer of autologous tumor infiltrating lymphocytes (TIL) (Zacharakis et al., (2018) Nat Med. 2018 Jun;24(6):724-730; Besser et al., (2010) Clin. Cancer Res 16 (9) 2646–55; Dudley et al., (2002) Science 298 (5594): 850–4; and Dudley et al., (2005) Journal of Clinical Oncology 23 (10): 2346–57.) or genetically re-directed peripheral blood mononuclear cells (Johnson et al., (2009) Blood 114 (3): 535–46; and Morgan et al., (2006) Science 314(5796) 126-9) has been used to successfully treat patients with advanced solid tumors, including melanoma, metastatic breast cancer and colorectal carcinoma, as well as patients with CD19-expressing hematologic malignancies (Kalos et al., (2011) Science Translational Medicine 3 (95): 95ra73).

[0725] In certain embodiments, allogenic cells immune cells are transferred (see e.g., Ren et al., (2017) Clin Cancer Res 23 (9) 2255-2266).

[0726] In certain embodiments, an antigen (such as a tumor antigen) to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) may be selected from a group consisting of: MR1 (see, e.g., Crowther, et al., 2020, Genome-wide CRISPR–Cas9 screening reveals ubiquitous T cell cancer

targeting via the monomorphic MHC class I-related protein MR1, Nature Immunology volume 21, pages178–185), B cell maturation antigen (BCMA) (see, e.g., Friedman et al., Effective Targeting of Multiple BCMA-Expressing Hematological Malignancies by Anti-BCMA CAR T Cells, Hum Gene Ther. 2018 Mar 8; Berdeja JG, et al. Durable clinical responses in heavily pretreated patients with relapsed/refractory multiple myeloma: updated results from a multicenter study of bb2121 anti-Bcma CAR T cell therapy. Blood. 2017;130:740; and Mouhieddine and Ghobrial, Immunotherapy in Multiple Myeloma: The Era of CAR T Cell Therapy, Hematologist, May-June 2018, Volume 15, issue 3); PSA (prostate-specific antigen); prostate-specific membrane antigen (PSMA); PSCA (Prostate stem cell antigen); Tyrosine-protein kinase transmembrane receptor ROR1; fibroblast activation protein (FAP); Tumor-associated glycoprotein 72 (TAG72); Carcinoembryonic antigen (CEA); Epithelial cell adhesion molecule (EPCAM); Mesothelin; Human Epidermal growth factor Receptor 2 (ERBB2 (Her2/neu)); Prostase; Prostatic acid phosphatase (PAP); elongation factor 2 mutant (ELF2M); Insulin-like growth factor 1 receptor (IGF-1R); gplOO; BCR-ABL (breakpoint cluster region-Abelson); tyrosinase; New York esophageal squamous cell carcinoma 1 (NY-ESO-1); κ-light chain, LAGE (L antigen); MAGE (melanoma antigen); Melanoma-associated antigen 1 (MAGE-A1); MAGE A3; MAGE A6; legumain; Human papillomavirus (HPV) E6; HPV E7; prostein; survivin; PCTA1 (Galectin 8); Melan-A/MART-1; Ras mutant; TRP-1 (tyrosinase related protein 1, or gp75); Tyrosinase-related Protein 2 (TRP2); TRP-2/INT2 (TRP-2/intron 2); RAGE (renal antigen); receptor for advanced glycation end products 1 (RAGE1); Renal ubiquitous 1, 2 (RU1, RU2); intestinal carboxyl esterase (iCE); Heat shock protein 70-2 (HSP70-2) mutant; thyroid stimulating hormone receptor (TSHR); CD123; CD171; CD19; CD20; CD22; CD26; CD30; CD33; CD44v7/8 (cluster of differentiation 44, exons 7/8); CD53; CD92; CD100; CD148; CD150; CD200; CD261; CD262; CD362; CS-1 (CD2 subset 1, CRACC, SLAMF7, CD319, and 19A24); C-type lectin-like molecule-1 (CLL-1); ganglioside GD3 (aNeu5Ac(2-8)aNeu5Ac(2-3)bDGalp(1-4)bDGlcp(1-1)Cer); Tn antigen (Tn Ag); Fms-Like Tyrosine Kinase 3 (FLT3); CD38; CD138; CD44v6; B7H3 (CD276); KIT (CD117); Interleukin-13 receptor subunit alpha-2 (IL-13Ra2); Interleukin 11 receptor alpha (IL-11Ra); prostate stem cell antigen (PSCA); Protease Serine 21 (PRSS21); vascular endothelial growth factor receptor 2 (VEGFR2); Lewis(Y) antigen; CD24; Platelet-derived growth factor receptor beta (PDGFR-beta); stage-specific embryonic antigen-4 (SSEA-4); Mucin 1, cell surface associated (MUC1); mucin 16 (MUC16); epidermal growth factor

receptor (EGFR); epidermal growth factor receptor variant III (EGFRvIII); neural cell adhesion molecule (NCAM); carbonic anhydrase IX (CAIX); Proteasome (Prosome, Macropain) Subunit, Beta Type, 9 (LMP2); ephrin type-A receptor 2 (EphA2); Ephrin B2; Fucosyl GM1; sialyl Lewis adhesion molecule (sLe); ganglioside GM3 (aNeu5Ac(2-3)bDGalp(1-4)bDGlcp(1-1)Cer); TGS5; high molecular weight-melanoma-associated antigen (HMWMAA); o-acetyl-GD2 ganglioside (OAcGD2); Folate receptor alpha; Folate receptor beta; tumor endothelial marker 1 (TEM1/CD248); tumor endothelial marker 7-related (TEM7R); claudin 6 (CLDN6); G protein-coupled receptor class C group 5, member D (GPRC5D); chromosome X open reading frame 61 (CXORF61); CD97; CD179a; anaplastic lymphoma kinase (ALK); Polysialic acid; placenta-specific 1 (PLAC1); hexasaccharide portion of globoH glycoceramide (GloboH); mammary gland differentiation antigen (NY-BR-1); uroplakin 2 (UPK2); Hepatitis A virus cellular receptor 1 (HAVCR1); adrenoceptor beta 3 (ADRB3); pannexin 3 (PANX3); G protein-coupled receptor 20 (GPR20); lymphocyte antigen 6 complex, locus K 9 (LY6K); Olfactory receptor 51E2 (OR51E2); TCR Gamma Alternate Reading Frame Protein (TARP); Wilms tumor protein (WT1); ETS translocation-variant gene 6, located on chromosome 12p (ETV6-AML); sperm protein 17 (SPA17); X Antigen Family, Member 1A (XAGE1); angiopoietin-binding cell surface receptor 2 (Tie 2); CT (cancer/testis (antigen)); melanoma cancer testis antigen-1 (MAD-CT-1); melanoma cancer testis antigen-2 (MAD-CT-2); Fos-related antigen 1; p53; p53 mutant; human Telomerase reverse transcriptase (hTERT); sarcoma translocation breakpoints; melanoma inhibitor of apoptosis (ML-IAP); ERG (transmembrane protease, serine 2 (TMPRSS2) ETS fusion gene); N-Acetyl glucosaminyl-transferase V (NA17); paired box protein Pax-3 (PAX3); Androgen receptor; Cyclin B1; Cyclin D1; v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog (MYCN); Ras Homolog Family Member C (RhoC); Cytochrome P450 1B1 (CYP1B1); CCCTC-Binding Factor (Zinc Finger Protein)-Like (BORIS); Squamous Cell Carcinoma Antigen Recognized By T Cells-1 or 3 (SART1, SART3); Paired box protein Pax-5 (PAX5); proacrosin binding protein sp32 (OY-TES1); lymphocyte-specific protein tyrosine kinase (LCK); A kinase anchor protein 4 (AKAP-4); synovial sarcoma, X breakpoint-1, -2, -3 or -4 (SSX1, SSX2, SSX3, SSX4); CD79a; CD79b; CD72; Leukocyte-associated immunoglobulin-like receptor 1 (LAIR1); Fc fragment of IgA receptor (FCAR); Leukocyte immunoglobulin-like receptor subfamily A member 2 (LILRA2); CD300 molecule-like family member f (CD300LF); C-type lectin domain family 12 member A (CLEC12A); bone marrow

stromal cell antigen 2 (BST2); EGF-like module-containing mucin-like hormone receptor-like 2 (EMR2); lymphocyte antigen 75 (LY75); Glypican-3 (GPC3); Fc receptor-like 5 (FCRL5); mouse double minute 2 homolog (MDM2); livin; alphafetoprotein (AFP); transmembrane activator and CAML Interactor (TACI); B-cell activating factor receptor (BAFF-R); V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS); immunoglobulin lambda-like polypeptide 1 (IGLL1); 707-AP (707 alanine proline); ART-4 (adenocarcinoma antigen recognized by T4 cells); BAGE (B antigen; b-catenin/m, b-catenin/mutated); CAMEL (CTL-recognized antigen on melanoma); CAP1 (carcinoembryonic antigen peptide 1); CASP-8 (caspase-8); CDC27m (cell-division cycle 27 mutated); CDK4/m (cycline-dependent kinase 4 mutated); Cyp-B (cyclophilin B); DAM (differentiation antigen melanoma); EGP-2 (epithelial glycoprotein 2); EGP-40 (epithelial glycoprotein 40); Erbb2, 3, 4 (erythroblastic leukemia viral oncogene homolog-2, -3, 4); FBP (folate binding protein); , fAchR (Fetal acetylcholine receptor); G250 (glycoprotein 250); GAGE (G antigen); GnT-V (N-acetylglucosaminyltransferase V); HAGE (helicose antigen); ULA-A (human leukocyte antigen-A); HST2 (human signet ring tumor 2); KIAA0205; KDR (kinase insert domain receptor); LDLR/FUT (low density lipid receptor/GDP L-fucose: b-D-galactosidase 2-a-L fucosyltransferase); L1CAM (L1 cell adhesion molecule); MC1R (melanocortin 1 receptor); Myosin/m (myosin mutated); MUM-1, -2, -3 (melanoma ubiquitous mutated 1, 2, 3); NA88-A (NA cDNA clone of patient M88); KG2D (Natural killer group 2, member D) ligands; oncofetal antigen (h5T4); p190 minor bcr-abl (protein of 190KD bcr-abl); Pml/RARa (promyelocytic leukaemia/retinoic acid receptor a); PRAME (preferentially expressed antigen of melanoma); SAGE (sarcoma antigen); TEL/AML1 (translocation Ets-family leukemia/acute myeloid leukemia 1); TPI/m (triosephosphate isomerase mutated); CD70; and any combination thereof.

[0727] In certain embodiments, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a tumor-specific antigen (TSA).

[0728] In certain embodiments, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a neoantigen.

[0729] In certain embodiments, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a tumor-associated antigen (TAA).

[0730] In certain embodiments, an antigen to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is a universal tumor antigen. In some embodiments, the universal tumor antigen is selected from the group of: a human telomerase reverse transcriptase (hTERT), survivin, mouse double minute 2 homolog (MDM2), cytochrome P450 1B 1 (CYP1B), HER2/neu, Wilms' tumor gene 1 (WT1), livin, alphafetoprotein (AFP), carcinoembryonic antigen (CEA), mucin 16 (MUC16), MUC1, prostate-specific membrane antigen (PSMA), p53, cyclin (Dl), and any combinations thereof.

[0731] In certain embodiments, an antigen (such as a tumor antigen) to be targeted in adoptive cell therapy (such as particularly CAR or TCR T-cell therapy) of a disease (such as particularly of tumor or cancer) is selected from a group of: CD19, BCMA, CD70, CLL-1, MAGE A3, MAGE A6, HPV E6, HPV E7, WT1, CD22, CD171, ROR1, MUC16, and SSX2. In some embodiments, the antigen is CD19. For example, CD19 can be targeted in hematologic malignancies, such as in lymphomas, more particularly in B-cell lymphomas, such as without limitation in diffuse large B-cell lymphoma, primary mediastinal b-cell lymphoma, transformed follicular lymphoma, marginal zone lymphoma, mantle cell lymphoma, acute lymphoblastic leukemia including adult and pediatric ALL, non-Hodgkin lymphoma, indolent non-Hodgkin lymphoma, or chronic lymphocytic leukemia. For example, BCMA can be targeted in multiple myeloma or plasma cell leukemia (see, e.g., 2018 American Association for Cancer Research (AACR) Annual meeting Poster: Allogeneic Chimeric Antigen Receptor T Cells Targeting B Cell Maturation Antigen). For example, CLL1 can be targeted in acute myeloid leukemia. For example, MAGE A3, MAGE A6, SSX2, and/or KRAS can be targeted in solid tumors. For example, HPV E6 and/or HPV E7 may be targeted in cervical cancer or head and neck cancer. For example, WT1 can be targeted in acute myeloid leukemia (AML), myelodysplastic syndromes (MDS), chronic myeloid leukemia (CML), non-small cell lung cancer, breast, pancreatic, ovarian or colorectal cancers, or mesothelioma. For example, CD22 can be targeted in B cell malignancies, including non-Hodgkin lymphoma, diffuse large B-cell lymphoma, or acute lymphoblastic leukemia. For example, CD171 can be targeted in neuroblastoma, glioblastoma, or lung, pancreatic, or ovarian cancers. For example, ROR1 may

be targeted in ROR1+ malignancies, including non-small cell lung cancer, triple negative breast cancer, pancreatic cancer, prostate cancer, ALL, chronic lymphocytic leukemia, or mantle cell lymphoma. For example, MUC16 can be targeted in MUC16ecto+ epithelial ovarian, fallopian tube or primary peritoneal cancer. For example, CD70 can be targeted in both hematologic malignancies as well as in solid cancers such as renal cell carcinoma (RCC), gliomas (e.g., GBM), and head and neck cancers (HNSCC). CD70 is expressed in both hematologic malignancies as well as in solid cancers, while its expression in normal tissues is restricted to a subset of lymphoid cell types (see, e.g., 2018 American Association for Cancer Research (AACR) Annual meeting Poster: Allogeneic CRISPR Engineered Anti-CD70 CAR-T Cells Demonstrate Potent Preclinical Activity Against Both Solid and Hematological Cancer Cells).

[0732] Various strategies can be employed to genetically modify T cells by altering the specificity of the T cell receptor (TCR), for example, by introducing new TCR α and β chains with selected peptide specificity (see e.g., U.S. Patent No. 8,697,854; PCT Patent Publications: WO2003020763, WO2004033685, WO2004044004, WO2005114215, WO2006000830, WO2008038002, WO2008039818, WO2004074322, WO2005113595, WO2006125962, WO2013166321, WO2013039889, WO2014018863, WO2014083173; U.S. Patent No. 8,088,379).

[0733] As an alternative to, or addition to, TCR modifications, chimeric antigen receptors (CARs) can be used in order to generate immunoresponsive cells, such as T cells, specific for selected targets, such as malignant cells, with a wide variety of receptor chimera constructs having been described (see e.g., U.S. Patent Nos. 5,843,728; 5,851,828; 5,912,170; 6,004,811; 6,284,240; 6,392,013; 6,410,014; 6,753,162; 8,211,422; and PCT Publication WO9215322).

[0734] In general, CARs are composed of an extracellular domain, a transmembrane domain, and an intracellular domain, wherein the extracellular domain comprises an antigen-binding domain that is specific for a predetermined target. While the antigen-binding domain of a CAR is often an antibody or antibody fragment (e.g., a single chain variable fragment, scFv), the binding domain is not particularly limited so long as it results in specific recognition of a target. For example, in some embodiments, the antigen-binding domain is composed of a receptor, such that the CAR is capable of binding to the ligand of the receptor. In other embodiments, the antigen-binding domain is composed of a ligand, such that the CAR is capable of binding the endogenous receptor of that ligand.

[0735] The antigen-binding domain of a CAR is generally separated from the transmembrane domain by a hinge or spacer. The spacer is also not particularly limited, and it is designed to provide the CAR with flexibility. For example, in some embodiments a spacer domain may comprise a portion of a human Fc domain, including a portion of the CH3 domain, or the hinge region of any immunoglobulin, such as IgA, IgD, IgE, IgG, or IgM, or variants thereof. Furthermore, the hinge region can be modified so as to prevent off-target binding by FcRs or other potential interfering objects. For example, the hinge can include an IgG4 Fc domain with or without a S228P, L235E, and/or N297Q mutation (according to Kabat numbering) in order to decrease binding to FcRs. Additional spacers/hinges include, but are not limited to, CD4, CD8, and CD28 hinge regions.

[0736] The transmembrane domain of a CAR can be derived either from a natural or from a synthetic source. Where the source is natural, the domain may be derived from any membrane bound or transmembrane protein. In some embodiments, transmembrane regions are derived from CD8, CD28, CD3, CD45, CD4, CD5, CDS, CD9, CD 16, CD22, CD33, CD37, CD64, CD80, CD86, CD 134, CD137, CD 154, TCR. In some embodiments, the transmembrane domain is synthetic, in which case it can include predominantly hydrophobic residues such as leucine and valine. In some embodiments, a triplet of phenylalanine, tryptophan and valine will be found at each end of a synthetic transmembrane domain. Optionally, a short oligo- or polypeptide linker, preferably between 2 and 10 amino acids in length can form the linkage between the transmembrane domain and the cytoplasmic signaling domain of the CAR. In some embodiments, the linker is a glycine-serine doublet, which can be a particularly suitable linker.

[0737] Alternative CAR constructs can be characterized as belonging to successive generations. First-generation CARs typically are composed of a single-chain variable fragment of an antibody specific for an antigen, for example including a VL linked to a VH of a specific antibody, linked by a flexible linker, for example by a CD8α hinge domain and a CD8α transmembrane domain, to the transmembrane and intracellular signaling domains of either CD3ζ or FcRγ (scFv-CD3ζ or scFv-FcRγ; see e.g., U.S. Patent No. 7,741,465; U.S. Patent No. 5,912,172; U.S. Patent No. 5,906,936). Second-generation CARs incorporate the intracellular domains of one or more costimulatory molecules, such as CD28, OX40 (CD134), or 4-1BB (CD137) within the endodomain (for example scFv-CD28/OX40/4-1BB-CD3ζ; see e.g., U.S. Patent Nos. 8,911,993; 8,916,381; 8,975,071; 9,101,584; 9,102,760; 9,102,761). Third-

generation CARs include a combination of costimulatory endodomains, such as a CD3ζ-chain, CD97, GDI la-CD18, CD2, ICOS, CD27, CD154, CDS, OX40, 4-1BB, CD2, CD7, LIGHT, LFA-1, NKG2C, B7-H3, CD30, CD40, PD-1, or CD28 signaling domains (for example scFv-CD28-4-1BB-CD3ζ or scFv-CD28-OX40-CD3ζ; see e.g., U.S. Patent No. 8,906,682; U.S. Patent No. 8,399,645; U.S. Pat. No. 5,686,281; PCT Publication No. WO2014134165; PCT Publication No. WO2012079000). In certain embodiments, the primary signaling domain comprises a functional signaling domain of a protein selected from the group of CD3 zeta, CD3 gamma, CD3 delta, CD3 epsilon, common FcR gamma (FCERIG), FcR beta (Fc Epsilon R1b), CD79a, CD79b, Fc gamma RIIa, DAP10, and/or DAP12. In some embodiments, the primary signaling domain includes a functional signaling domain of CD3ζ or FcRγ. In certain embodiments, the one or more costimulatory signaling domains comprise a functional signaling domain of a protein selected, each independently, from the group of: CD27, CD28, 4-1BB (CD137), OX40, CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3 and/or a ligand that specifically binds with CD83, CDS, ICAM-1, GITR, BAFFR, HVEM (LIGHTR), SLAMF7, NKp80 (KLRF1), CD160, CD19, CD4, CD8 alpha, CD8 beta, IL2R beta, IL2R gamma, IL7R alpha, ITGA4, VLA1, CD49a, ITGA4, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD11d, ITGAE, CD103, ITGAL, CD11a, LFA-1, ITGAM, CD11b, ITGAX, CD11c, ITGB1, CD29, ITGB2, CD18, ITGB7, TNFR2, TRANCE/RANKL, DNAM1 (CD226), SLAMF4 (CD244, 2B4), CD84, CD96 (Tactile), CEACAM1, CRTAM, Ly9 (CD229), CD160 (BY55), PSGL1, CD100 (SEMA4D), CD69, SLAMF6 (NTB-A, Lyl08), SLAM (SLAMF1, CD150, IPO-3), BLAME (SLAMF8), SELPLG (CD162), LTBR, LAT, GADS, SLP-76, PAG/Cbp, NKp44, NKp30, NKp46, and/or NKG2D. In certain embodiments, the one or more costimulatory signaling domains include a functional signaling domain of a protein selected, each independently, from the group of: 4-1BB, CD27, and CD28. In certain embodiments, a chimeric antigen receptor can have the design as described in U.S. Patent No. 7,446,190 and can be composed of an intracellular domain of CD3ζ chain (such as amino acid residues 52-163 of the human CD3 zeta chain, as shown in SEQ ID NO: 14 of US 7,446,190) and a signaling region from CD28 and an antigen-binding element (or portion or domain; such as scFv). The CD28 portion, when between the zeta chain portion and the antigen-binding element, can suitably include the transmembrane and signaling domains of CD28 (such as amino acid residues 114-220 of SEQ ID NO: 10, full sequence shown in SEQ ID NO: 6 of US 7,446,190; these can include the

following portion of CD28 as set forth in Genbank identifier NM_006139 (sequence version 1, 2 or 3): IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPLFPGPSKPFWVLVVVGGVLACYSLLVT VAFIIFWVRSKRSRLLHSDYMNMTPRRPGPTRKHYQPYAPPRDFAAYRS)) (SEQ ID NO:49). In some embodiments, when the zeta sequence lies between the CD28 sequence and the antigen-binding element, the intracellular domain of CD28 can be used alone (such as amino sequence set forth in SEQ ID NO: 9 of US 7,446,190). Hence, certain embodiments employ a CAR comprising (a) a zeta chain portion including the intracellular domain of human CD3ζ chain, (b) a costimulatory signaling region, and (c) an antigen-binding element (or portion or domain), where the costimulatory signaling region includes the amino acid sequence encoded by SEQ ID NO: 6 of US 7,446,190.

[0738] In other embodiments, co-stimulation is achieved by expressing CARs in antigen-specific T cells, which can be chosen so as to be activated and expanded following engagement of their native αβTCR by, for example, antigen on professional antigen-presenting cells with attendant co-stimulation. In some embodiments, additional engineered receptors can be included on the immunoresponsive cells and can be capable of improving targeting of a T-cell attack and/or minimizing side effects.

[0739] In a non-limiting example, Kochenderfer et al., (2009) J Immunother. 32 (7): 689-702 described anti-CD19 chimeric antigen receptors (CAR). FMC63-28Z CAR contained a single chain variable region moiety (scFv) recognizing CD19 derived from the FMC63 mouse hybridoma (described in Nicholson et al., (1997) Molecular Immunology 34: 1157–1165), a portion of the human CD28 molecule, and the intracellular component of the human TCR-ζ molecule. FMC63-CD828BBZ CAR contained the FMC63 scFv, the hinge and transmembrane regions of the CD8 molecule, the cytoplasmic portions of CD28 and 4-1BB, and the cytoplasmic component of the TCR-ζ molecule. The exact sequence of the CD28 molecule included in the FMC63-28Z CAR corresponded to Genbank identifier NM_006139; the sequence included all amino acids starting with the amino acid sequence IEVMYPPPY (SEQ ID NO:50) and continuing all the way to the carboxy-terminus of the protein. To encode the anti-CD19 scFv component of the vector, the authors designed a DNA sequence which was based on a portion of a previously published CAR (Cooper et al., (2003) Blood 101: 1637–1644). This sequence encoded the following components in frame from the 5' end to the 3' end: an XhoI site, the human granulocyte-macrophage colony-stimulating factor (GM-CSF)

receptor α-chain signal sequence, the FMC63 light chain variable region (as in Nicholson et al., supra), a linker peptide (as in Cooper et al., supra), the FMC63 heavy chain variable region (as in Nicholson et al., supra), and a NotI site. A plasmid encoding this sequence was digested with XhoI and NotI. To form the MSGV-FMC63-28Z retroviral vector, the XhoI and NotI-digested fragment encoding the FMC63 scFv was ligated into a second XhoI and NotI-digested fragment that encoded the MSGV retroviral backbone (as in Hughes et al., (2005) Human Gene Therapy 16: 457–472) as well as part of the extracellular portion of human CD28, the entire transmembrane and cytoplasmic portion of human CD28, and the cytoplasmic portion of the human TCR-ζ molecule (as in Maher et al., 2002) Nature Biotechnology 20: 70–75). The FMC63-28Z CAR is included in the KTE-C19 (axicabtagene ciloleucel) anti-CD19 CAR-T therapy product in development by Kite Pharma, Inc. for the treatment of inter alia patients with relapsed/refractory aggressive B-cell non-Hodgkin lymphoma (NHL). Accordingly, in certain embodiments, cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may express the FMC63-28Z CAR as described by Kochenderfer et al. (supra). Hence, in certain embodiments, cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may comprise a CAR comprising an extracellular antigen-binding element (or portion or domain; such as scFv) that specifically binds to an antigen, an intracellular signaling domain comprising an intracellular domain of a CD3ζ chain, and a costimulatory signaling region comprising a signaling domain of CD28. Preferably, the CD28 amino acid sequence is as set forth in Genbank identifier NM_006139 (sequence version 1, 2 or 3) starting with the amino acid sequence IEVMYPPPY (SEQ ID NO:50) and continuing all the way to the carboxy-terminus of the protein. In some embodiments, the CD28 has a sequence that is 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 percent identical to IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPLFPGPSKPFWVLVVVGGVLACYSLLVT VAFIIFWVRSKRSRLLHSDYMNMTPRRPGPTRKHYQPYAPPRDFAAYRS (SEQ ID NO:49)). Preferably, the antigen is CD19, more preferably the antigen-binding element is an anti-CD19 scFv, even more preferably the anti-CD19 scFv as described by Kochenderfer et al. (supra).

[0740] Additional anti-CD19 CARs that can be included in some embodiments are further described in WO2015187528. More particularly Example 1 and Table 1 of WO2015187528, incorporated by reference herein, demonstrate the generation of anti-CD19 CARs based on a

fully human anti-CD19 monoclonal antibody (47G4, as described in US20100104509) and murine anti-CD19 monoclonal antibody (as described in Nicholson et al. and explained above). Various combinations of a signal sequence (human CD8-alpha or GM-CSF receptor), extracellular and transmembrane regions (human CD8-alpha) and intracellular T-cell signaling domains (CD28-CD3ζ; 4-1BB-CD3ζ; CD27-CD3ζ; CD28-CD27-CD3ζ, 4-1BB-CD27-CD3ζ; CD27-4-1BB-CD3ζ; CD28-CD27-FcεRI gamma chain; or CD28-FcεRI gamma chain) were disclosed. Hence, in certain embodiments, cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, include a CAR having an extracellular antigen-binding element that specifically binds to an antigen, an extracellular and transmembrane region as set forth in Table 1 of WO2015187528 and an intracellular T-cell signaling domain as set forth in Table 1 of WO2015187528. In some embodiments, the antigen is CD19. In some embodiments, the antigen-binding element is an anti-CD19 scFv. In some embodiments, the antigen-binding element is the mouse or human anti-CD19 scFv as described in Example 1 of WO2015187528. In certain embodiments, the CAR includes or is composed only of an amino acid sequence of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, or SEQ ID NO: 13 as set forth in Table 1 of WO2015187528.

[0741] By means of an example and without limitation, chimeric antigen receptor that recognizes the CD70 antigen is described in WO2012058460A2 (see also, Park et al., CD70 as a target for chimeric antigen receptor T cells in head and neck squamous cell carcinoma, Oral Oncol. 2018 Mar;78:145-150; and Jin et al., CD70, a novel target of CAR T-cell therapy for gliomas, Neuro Oncol. 2018 Jan 10;20(1):55-65). CD70 is expressed by diffuse large B-cell and follicular lymphoma and also by the malignant cells of Hodgkins lymphoma, Waldenstrom's macroglobulinemia and multiple myeloma, and by HTLV-1- and EBV-associated malignancies. (Agathanggelou et al. Am.J.Pathol. 1995;147: 1152-1160; Hunter et al., Blood 2004; 104:4881. 26; Lens et al., J Immunol. 2005;174:6212-6219; Baba et al., J Virol. 2008;82:3843-3852.) In addition, CD70 is expressed by non-hematological malignancies such as renal cell carcinoma and glioblastoma. (Junker et al., J Urol. 2005;173:2150-2153; Chahlavi et al., Cancer Res 2005;65:5428-5438) Physiologically, CD70 expression is transient and restricted to a subset of highly activated T, B, and dendritic cells.

[0742] Other suitable and non-limiting chimeric antigen receptors that recognize BCMA are set forth in e.g., US20160046724A1; WO2016014789A2; WO2017211900A1; WO2015158671A1; US20180085444A1; WO2018028647A1; US20170283504A1; and WO2013154760A1, and can be included in one or more embodiments herein.

[0743] In certain embodiments, the immune cell, in addition to a CAR or exogenous TCR as described herein, further includes a chimeric inhibitory receptor (inhibitory CAR) that specifically binds to a second target antigen and is capable of inducing an inhibitory or immunosuppressive or repressive signal to the cell upon recognition of the second target antigen. In certain embodiments, the chimeric inhibitory receptor includes an extracellular antigen-binding element (or portion thereof or domain thereof) configured to specifically bind to a target antigen, a transmembrane domain, and an intracellular immunosuppressive or repressive signaling domain. In certain embodiments, the second target antigen is an antigen that is not expressed on the surface of a cancer cell or infected cell or the expression of which is downregulated on a cancer cell or an infected cell. In certain embodiments, the second target antigen is an MHC-class I molecule. In certain embodiments, the intracellular signaling domain comprises a functional signaling portion of an immune checkpoint molecule, such as for example PD-1 or CTLA4. The inclusion of such inhibitory CAR reduces the chance of the engineered immune cells attacking non-target (e.g., non-cancer) tissues, which can be advantageous.

[0744] In some embodiments, T-cells expressing CARs are further modified to reduce or eliminate expression of endogenous TCRs in order to reduce off-target effects. Reduction or elimination of endogenous TCRs can reduce off-target effects and increase the effectiveness of the T cells (see e.g., U.S. 9,181,527). T cells stably lacking expression of a functional TCR can be produced using a variety of approaches. T cells internalize, sort, and degrade the entire T cell receptor as a complex, with a half-life of about 10 hours in resting T cells and 3 hours in stimulated T cells (von Essen, M. et al. 2004. J. Immunol. 173:384-393). Proper functioning of the TCR complex relies on the proper stoichiometric ratio of the proteins that compose the TCR complex. TCR function also relies on two functioning TCR zeta proteins with ITAM motifs. The activation of the TCR upon engagement of its MHC-peptide ligand requires the engagement of several TCRs on the same T cell, which all must signal properly. Thus, if a TCR complex is destabilized with proteins that do not associate properly or cannot signal optimally, the T cell will not become activated sufficiently to begin a cellular response.

[0745] Accordingly, in some embodiments, TCR expression is eliminated using RNA interference (e.g., shRNA, siRNA, miRNA, etc.), CRISPR-Cas systems, and/or other methods that target the nucleic acids encoding specific TCRs (e.g., TCR-α and TCR-β) and/or CD3 chains in primary T cells. By blocking expression of one or more of these proteins, the T cell will no longer produce one or more of the key components of the TCR complex, thereby destabilizing the TCR complex and preventing cell surface expression of a functional TCR.

[0746] In some instances, the CAR also includes a switch mechanism for controlling expression and/or activation of the CAR. For example, a CAR can have an extracellular, transmembrane, and intracellular domain, in which the extracellular domain can have a target-specific binding element that includes a label, binding domain, or tag that is specific for a molecule other than the target antigen that is expressed on, in, or by a target cell. In such embodiments, the specificity of the CAR is provided by a second construct that includes a target antigen binding domain (e.g., an scFv or a bispecific antibody that is specific for both the target antigen and the label or tag on the CAR) and a domain that is recognized by or binds to the label, binding domain, or tag on the CAR. See, e.g., WO 2013/044225, WO 2016/000304, WO 2015/057834, WO 2015/057852, WO 2016/070061, US 9,233,125, US 2016/0129109. In this way, a T-cell that expresses the CAR can be administered to a subject, but the CAR cannot bind its target antigen until the second composition comprising an antigen-specific binding domain is administered.

[0747] Other switch mechanisms include, without limitation, CARs that require multimerization in order to activate their signaling function (see, e.g., US 2015/0368342, US 2016/0175359, US 2015/0368360) and/or an exogenous signal, such as a small molecule drug (US 2016/0166613, Yung et al., Science, 2015), in order to elicit a T-cell response. Some CARs may also include a "suicide switch" to induce cell death of the CAR T-cells following treatment (see e.g., Buddee et al., PLoS One, 2013 Dec 17;8(12):e82742) or to downregulate expression of the CAR following binding to the target antigen (see e.g., WO 2016/011210).

[0748] Other techniques can be used to modify target immunoresponsive cells, such as protoplast fusion, lipofection, transfection or electroporation. A wide variety of vectors may be used, such as retroviral vectors, lentiviral vectors, adenoviral vectors, adeno-associated viral vectors, plasmids or transposons, such as a Sleeping Beauty transposon (see e.g., U.S. Patent Nos. 6,489,458; 7,148,203; 7,160,682; 7,985,739; 8,227,432), can be used to introduce CARs, for example using second generation antigen-specific CARs signaling through CD3ζ and either

CD28 or CD137. Viral vectors can include vectors based on HIV, SV40, EBV, HSV or BPV. Other vectors and systems suitable modifying target immunoresponsive cells are also described in greater detail elsewhere herein.

[0749] Cells that are suitable for modification can include T cells, Natural Killer (NK) cells, cytotoxic T lymphocytes (CTL), regulatory T cells, human embryonic stem cells, tumor-infiltrating lymphocytes (TIL) and/or pluripotent stem cells from which lymphoid cells may be differentiated. T cells expressing a desired CAR may for example be selected through co-culture with γ-irradiated activating and propagating cells (AaPC), which co-express the cancer antigen and co-stimulatory molecules. The engineered CAR T-cells can be expanded in culture. For example, engineered CAR T-cells can be expanded by co-culture on AaPC in presence of soluble factors, such as IL-2 and IL-21. In some embodiments, expansion is carried out such that memory CAR+ T cells are generated. Successful generation of memory CAR+ T cells or other types of CAR T-cells can be confirmed by various assays, such as non-enzymatic digital array and/or multi-panel flow cytometry. In this way, CAR T-cells can be provided that have specific cytotoxic activity against antigen-bearing tumors. This can optionally be in conjunction with production of desired chemokines such as interferon-γ. CAR T-cells of this kind can be used in various applications, including but not limited to, animal models, for example to treat tumor xenografts.

[0750] In certain embodiments, ACT includes co-transferring CD4+ Th1 cells and CD8+ CTLs to induce a synergistic antitumour response (see, e.g., Li et al., Adoptive cell therapy with CD4+ T helper 1 cells and CD8+ cytotoxic T cells enhances complete rejection of an established tumour, leading to generation of endogenous memory responses to non-targeted tumour epitopes. Clin Transl Immunology. 2017 Oct; 6(10): e160).

[0751] In certain embodiments, Th17 cells are transferred to a subject in need thereof. Th17 cells have been reported to directly eradicate melanoma tumors in mice to a greater extent than Th1 cells (Muranski P, et al., Tumor-specific Th17-polarized cells eradicate large established melanoma. Blood. 2008 Jul 15; 112(2):362-73; and Martin-Orozco N, et al., T helper 17 cells promote cytotoxic T cell activation in tumor immunity. Immunity. 2009 Nov 20; 31(5):787-98). Those studies involved an adoptive T cell transfer (ACT) therapy approach, which takes advantage of CD4+ T cells that express a TCR recognizing tyrosinase tumor antigen. Exploitation of the TCR leads to rapid expansion of Th17 populations to large numbers *ex vivo* for reinfusion into the autologous tumor-bearing hosts. In some embodiments, the Th17 cells

used in ACT described herein can rely on the same approach as described by Muranski et al. Martin-Orozco et al.

[0752] Unlike T-cell receptors (TCRs) that are MHC restricted, CARs can potentially bind any cell surface-expressed antigen and can thus be more universally used to treat patients (see Irving et al., Engineering Chimeric Antigen Receptor T-Cells for Racing in Solid Tumors: Don't Forget the Fuel, Front. Immunol., 03 April 2017, doi.org/10.3389/fimmu.2017.00267). In certain embodiments, in the absence of endogenous T-cell infiltrate (e.g., due to aberrant antigen processing and presentation), which precludes the use of TIL therapy and immune checkpoint blockade, the transfer of CAR T-cells may be used to treat patients (see, e.g., Hinrichs CS, Rosenberg SA. Exploiting the curative potential of adoptive T-cell therapy for cancer. Immunol Rev (2014) 257(1):56–71. doi:10.1111/ imr.12132).

[0753] Unlike T-cell receptors (TCRs) that are MHC restricted, CARs can potentially bind any cell surface-expressed antigen and can thus be more universally used to treat patients (see e.g., Irving et al., Engineering Chimeric Antigen Receptor T-Cells for Racing in Solid Tumors: Don't Forget the Fuel, Front. Immunol., 03 April 2017, doi.org/10.3389/fimmu.2017.00267). In certain embodiments, in the absence of endogenous T-cell infiltrate (e.g., due to aberrant antigen processing and presentation), which precludes the use of TIL therapy and immune checkpoint blockade, the transfer of CAR T-cells can be used to treat patients (see e.g., Hinrichs CS, Rosenberg SA. Exploiting the curative potential of adoptive T-cell therapy for cancer. Immunol Rev (2014) 257(1):56–71. doi:10.1111/ imr.12132).

[0754] Approaches such as the foregoing can be adapted to provide methods of treating and/or increasing survival of a subject having a disease, such as a neoplasia, for example by administering an effective amount of an immunoresponsive cell having an antigen recognizing receptor that binds a selected antigen, wherein the binding activates the immunoresponsive cell, thereby treating or preventing the disease (such as a neoplasia, a pathogen infection, an autoimmune disorder, or an allogeneic transplant reaction).

[0755] In certain embodiments, the treatment is administered after lymphodepleting pretreatment in the form of chemotherapy (typically a combination of cyclophosphamide and fludarabine) or radiation therapy. Not being bound by a theory, lymphodepleting pretreatment may eliminate or reduce immune suppressor cells like Tregs and MDSCs (which may attenuate the activity of the transferred cells by out competing for the necessary cytokines) allowing the TILs to persist.

[0756] In some embodiments, the ACT is administrated into patients undergoing an immunosuppressive treatment (e.g., glucocorticoid treatment). The cells or population thereof can be modified such that they are resistant to at least one immunosuppressive agent due to the inactivation of a gene encoding a receptor for such immunosuppressive agent. In certain embodiments, the immunosuppressive treatment provides for the selection and expansion of the immunoresponsive T cells within the patient.

[0757] In certain embodiments, the ACT is administered before a primary treatment (e.g., surgery or radiation therapy) to shrink a tumor before the primary treatment. In another embodiment, the ACT is administered after primary treatment to remove any remaining cancer cells. In some embodiments, the ACT is administered only before a primary treatment, only after a primary treatment, or is administered both before and after a primary treatment.

[0758] In certain embodiments, immunometabolic barriers can be targeted therapeutically prior to and/or during ACT to enhance responses to ACT or CAR T-cell therapy and to support endogenous immunity (see, e.g., Irving et al., Engineering Chimeric Antigen Receptor T-Cells for Racing in Solid Tumors: Don't Forget the Fuel, Front. Immunol., 03 April 2017, doi.org/10.3389/fimmu.2017.00267).

[0759] The cells can be administered to a subject in need thereof in an ACT method by any suitable and/or convenient manner method to one or more regions or places on or in the subject. Such methods include, without limitation, by aerosol inhalation, injection, ingestion, transfusion, implantation or transplantation. The cells or population of cells may be administered to a patient subcutaneously, intradermally, intratumorally, intranodally, intramedullary, intramuscularly, intrathecally, by intravenous or intralymphatic injection, or intraperitoneally. In some embodiments, the engineered CARs or other engineered immune cells are delivered or administered into a cavity formed by the resection of tumor tissue (i.e., intracavity delivery) or directly into a I prior to resection (i.e., intratumoral delivery). In one embodiment, the cell for ACT as described herein are administered by intravenous injection. In another embodiment, administration of the cells for ACT is parenteral. The administration can be an intravenous administration. The administration can be directly done by injection within a tumor.

[0760] In some embodiments, of 10^4- 10^9 cells per kg body weight, inclusive all integer values and ranges therein, can be administered for ACT. In some embodiments, 10^5- 10^6 cells per kg body weight, inclusive all integer values and ranges therein, can be administered for

ACT. In some embodiments, dosing in CAR T-cell ACT therapies can include administration of 10^6 to 10^9 cells/kg bodyweight, with or without a course of lymphodepletion, for example with cyclophosphamide.

[0761] The cells or population of cells can be administrated in one or more doses. In another embodiment, the effective amount of cells are administrated as a single dose. In another embodiment, the effective amount of cells are administrated as more than one dose over a period time. Timing of administration is within the judgment of managing physician and depends on the clinical condition of the patient. The cells or population of cells may be obtained from any source, such as a blood bank or a donor. While individual needs vary, determination of optimal ranges of effective amounts of a given cell type for a particular disease or conditions are within the skill of one in the art. An effective amount means an amount which provides a therapeutic or prophylactic benefit. The dosage administrated will be dependent upon the age, health and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment and the nature of the effect desired.

[0762] To guard against possible adverse reactions, engineered immunoresponsive cells can include a transgenic safety switch, in the form of a transgene that renders the cells vulnerable to exposure to a specific signal. For example, the herpes simplex viral thymidine kinase (TK) gene may be used in this way, for example by introduction into allogeneic T lymphocytes used as donor lymphocyte infusions following stem cell transplantation (Greco, et al., Improving the safety of cell therapy with the TK-suicide gene. Front. Pharmacol. 2015; 6: 95). In such cells, administration of a nucleoside prodrug such as ganciclovir or acyclovir causes cell death. Alternative safety switch constructs include inducible caspase 9, for example triggered by administration of a small-molecule dimerizer that brings together two nonfunctional icasp9 molecules to form the active enzyme. A wide variety of alternative approaches to implementing cellular proliferation controls have been described (see U.S. Patent Publication No. 20130071414; PCT Patent Publication WO2011146862; PCT Patent Publication WO2014011987; PCT Patent Publication WO2013040371; Zhou et al. BLOOD, 2014, 123/25:3895 – 3905; Di Stasi et al., The New England Journal of Medicine 2011; 365:1673-1683; Sadelain M, The New England Journal of Medicine 2011; 365:1735-173; Ramos et al., Stem Cells 28(6):1107-15 (2010)).

[0763] In some embodiments, genome editing is used to tailor immunoresponsive cells to alternative implementations, for example providing edited CAR T-cells (see e.g., Poirot et al.,

2015, Multiplex genome edited T-cell manufacturing platform for "off-the-shelf" adoptive T-cell immunotherapies, Cancer Res 75 (18): 3853; Ren et al., 2017, Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition, Clin Cancer Res. 2017 May 1;23(9):2255-2266. doi: 10.1158/1078-0432.CCR-16-1300. Epub 2016 Nov 4; Qasim et al., 2017, Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells, Sci Transl Med. 2017 Jan 25;9(374); Legut, et al., 2018, CRISPR-mediated TCR replacement generates superior anticancer transgenic T cells. Blood, 131(3), 311-322; and Georgiadis et al., Long Terminal Repeat CRISPR-CAR-Coupled "Universal" T Cells Mediate Potent Anti-leukemic Effects, Molecular Therapy, In Press, Corrected Proof, Available online 6 March 2018). Cells can be edited using any CRISPR system and method of use thereof, including, but not limited to those described elsewhere herein.

[0764] In some embodiments, cells are edited or otherwise modified *ex vivo* and transferred to a subject in need thereof. Immunoresponsive cells, CAR T cells or any cells used for adoptive cell transfer can be edited or otherwise modified using another gene modification technique. Editing or other suitable gene modification technique can be performed, for example, to insert or knock-in an exogenous gene, such as an exogenous gene encoding a CAR or a TCR, at a preselected locus in a cell (e.g. TRAC locus); to eliminate potential alloreactive T-cell receptors (TCR) or to prevent inappropriate pairing between endogenous and exogenous TCR chains, such as to knock-out or knock-down expression of an endogenous TCR in a cell; to disrupt the target of a chemotherapeutic agent in a cell; to block an immune checkpoint, such as to knock-out or knock-down expression of an immune checkpoint protein or receptor in a cell; to knock-out or knock-down expression of other gene or genes in a cell, the reduced expression or lack of expression of which can enhance the efficacy of adoptive therapies using the cell; to knock-out or knock-down expression of an endogenous gene in a cell, said endogenous gene encoding an antigen targeted by an exogenous CAR or TCR; to knock-out or knock-down expression of one or more MHC constituent proteins in a cell; to activate a T cell; to modulate cells such that the cells are resistant to exhaustion or dysfunction; and/or increase the differentiation and/or proliferation of functionally exhausted or dysfunctional CD8+ T-cells (see e.g., PCT Patent Publications: WO2013176915, WO2014059173, WO2014172606, WO2014184744, and WO2014191128).

[0765] In certain embodiments, modification (e.g., via gene editing or other gene modification technique) results in inactivation of a gene. The phrase "inactivating a gene"

means that the gene of interest is not expressed in a functional protein form. In a particular embodiment, the system specifically catalyzes cleavage in one targeted gene thereby inactivating said targeted gene. The nucleic acid strand breaks caused are commonly repaired through the distinct mechanisms of homologous recombination or non-homologous end joining (NHEJ). However, NHEJ is an imperfect repair process that often results in changes to the DNA sequence at the site of the cleavage. Repair via non-homologous end joining (NHEJ) often results in small insertions or deletions (Indel) and can be used for the creation of specific gene knockouts. Cells in which a cleavage induced mutagenesis event has occurred can be identified and/or selected by well-known methods in the art. In certain embodiments, homology directed repair (HDR) is used to concurrently inactivate a gene (e.g., TRAC) and insert an endogenous TCR or CAR into the inactivated locus. Hence, in certain embodiments, editing of cells intended for adoptive cell therapies, including but not limited to immunoresponsive cells such as T cells, can be performed to insert or knock-in an exogenous gene, such as an exogenous gene encoding a CAR or a TCR, at a preselected locus in a cell. Conventionally, nucleic acid molecules encoding CARs or TCRs are transfected or transduced to cells using randomly integrating vectors, which, depending on the site of integration, may lead to clonal expansion, oncogenic transformation, variegated transgene expression and/or transcriptional silencing of the transgene. Directing of transgene(s) to a specific locus in a cell can minimize or avoid such risks and advantageously provide for uniform expression of the transgene(s) by the cells. Without limitation, suitable 'safe harbor' loci for directed transgene integration include CCR5 or AAVS1. Homology-directed repair (HDR) strategies are known and described elsewhere in this specification allowing to insert transgenes into desired loci (e.g., TRAC locus).

[0766] Further suitable loci for insertion of transgenes, in particular CAR or exogenous TCR transgenes, include without limitation loci comprising genes coding for constituents of endogenous T-cell receptor, such as T-cell receptor alpha locus (TRA) or T-cell receptor beta locus (TRB), for example T-cell receptor alpha constant (TRAC) locus, T-cell receptor beta constant 1 (TRBC1) locus or T-cell receptor beta constant 2 (TRBC1) locus. Advantageously, insertion of a transgene into such locus can simultaneously achieve expression of the transgene, potentially controlled by the endogenous promoter, and knock-out expression of the endogenous TCR. This approach has been exemplified in Eyquem et al., (2017) Nature 543: 113-117, wherein the authors used CRISPR/Cas9 gene editing to knock-in a DNA molecule

encoding a CD19-specific CAR into the TRAC locus downstream of the endogenous promoter; the CAR-T cells obtained by CRISPR were significantly superior in terms of reduced tonic CAR signaling and exhaustion.

[0767] T cell receptors (TCR) are cell surface receptors that participate in the activation of T cells in response to the presentation of antigen. The TCR is generally made from two chains, α and β, which assemble to form a heterodimer and associates with the CD3-transducing subunits to form the T cell receptor complex present on the cell surface. Each α and β chain of the TCR consists of an immunoglobulin-like N-terminal variable (V) and constant (C) region, a hydrophobic transmembrane domain, and a short cytoplasmic region. As for immunoglobulin molecules, the variable region of the α and β chains are generated by V(D)J recombination, creating a large diversity of antigen specificities within the population of T cells. However, in contrast to immunoglobulins that recognize intact antigen, T cells are activated by processed peptide fragments in association with an MHC molecule, introducing an extra dimension to antigen recognition by T cells, known as MHC restriction. Recognition of MHC disparities between the donor and recipient through the T cell receptor leads to T cell proliferation and the potential development of graft versus host disease (GVHD). The inactivation of TCRα or TCRβ can result in the elimination of the TCR from the surface of T cells preventing recognition of alloantigen and thus GVHD. However, TCR disruption generally results in the elimination of the CD3 signaling component and alters the means of further T cell expansion.

[0768] Hence, in certain embodiments, editing of cells, particularly cells intended for adoptive cell therapies, more particularly immunoresponsive cells such as T cells, may be performed to knock-out or knock-down expression of an endogenous TCR in a cell. For example, NHEJ-based or HDR-based gene editing approaches can be employed to disrupt the endogenous TCR alpha and/or beta chain genes. For example, gene editing system or systems, such as CRISPR/Cas system or systems, can be designed to target a sequence found within the TCR beta chain conserved between the beta 1 and beta 2 constant region genes (TRBC1 and TRBC2) and/or to target the constant region of the TCR alpha chain (TRAC) gene.

[0769] Allogeneic cells are rapidly rejected by the host immune system. It has been demonstrated that, allogeneic leukocytes present in non-irradiated blood products will persist for no more than 5 to 6 days (Boni, Muranski et al. 2008 Blood 1;112(12):4746-54). Thus, to prevent rejection of allogeneic cells, the host's immune system usually has to be suppressed to some extent. However, in the case of adoptive cell transfer the use of immunosuppressive drugs

also have a detrimental effect on the introduced therapeutic T cells. Therefore, to effectively use an adoptive immunotherapy approach in these conditions, the introduced cells would need to be resistant to the immunosuppressive treatment. Thus, in some embodiments, the immune cells (such as T-cells) for ACT are modified to make them resistant to an immunosuppressive agent, preferably by inactivating at least one gene encoding a target for an immunosuppressive agent. An immunosuppressive agent is an agent that suppresses immune function by one of several mechanisms of action. An immunosuppressive agent can be, but is not limited to a calcineurin inhibitor, a target of rapamycin, an interleukin-2 receptor α-chain blocker, an inhibitor of inosine monophosphate dehydrogenase, an inhibitor of dihydrofolic acid reductase, a corticosteroid or an immunosuppressive antimetabolite. The present invention allows conferring immunosuppressive resistance to T cells for immunotherapy by inactivating the target of the immunosuppressive agent in T cells. As non-limiting examples, targets for an immunosuppressive agent can be a receptor for an immunosuppressive agent such as: CD52, glucocorticoid receptor (GR), a FKBP family gene member and a cyclophilin family gene member.

[0770] In certain embodiments, immune cells (including but not limited to T-cells) for ACT are modified (via gene editing or other suitable gene modification technique) to block an immune checkpoint, such as to knock-out or knock-down expression of an immune checkpoint protein or receptor in a cell. Immune checkpoints are inhibitory pathways that slow down or stop immune reactions and prevent excessive tissue damage from uncontrolled activity of immune cells. In certain embodiments, the immune checkpoint targeted (or modified) is the programmed death-1 (PD-1 or CD279) gene (PDCD1). In other embodiments, the immune checkpoint targeted (or modified) is cytotoxic T-lymphocyte-associated antigen (CTLA-4). In additional embodiments, the immune checkpoint targeted is another member of the CD28 and CTLA4 Ig superfamily such as BTLA, LAG3, ICOS, PDL1 or KIR. In some embodiments, the immune checkpoint targeted (or modified) is a member of the TNFR superfamily such as CD40, OX40, CD137, GITR, CD27 or TIM-3. Additional immune checkpoints that can be modified include, but are not limited to, Src homology 2 domain-containing protein tyrosine phosphatase 1 (SHP-1) (Watson HA, et al., SHP-1: the next checkpoint target for cancer immunotherapy? Biochem Soc Trans. 2016 Apr 15;44(2):356-62) and T cell immunoreceptor with Ig and ITIM domains (TIGIT/Vstm3/WUCAM/VSIG9) and VISTA (Le Mercier I, et al.,

(2015) Beyond CTLA-4 and PD-1, the generation Z of negative checkpoint regulators. Front. Immunol. 6:418).

[0771] WO2014172606 relates to the use of MT1 and/or MT2 inhibitors to increase proliferation and/or activity of exhausted CD8+ T-cells and to decrease CD8+ T-cell exhaustion (e.g., decrease functionally exhausted or unresponsive CD8+ immune cells). In certain embodiments, metallothioneins are targeted (or modified) in the engineered T cells for ACT described herein.

[0772] In certain embodiments, one or more loci and/or genes involved in the expression of an immune checkpoint protein is modified in the engineered immune cell described herein. Such targets may include, but are not limited to CTLA4, PPP2CA, PPP2CB, PTPN6, PTPN22, PDCD1, ICOS (CD278), PDL1, KIR, LAG3, HAVCR2, BTLA, CD160, TIGIT, CD96, CRTAM, LAIR1, SIGLEC7, SIGLEC9, CD244 (2B4), TNFRSF10B, TNFRSF10A, CASP8, CASP10, CASP3, CASP6, CASP7, FADD, FAS, TGFBRII, TGFRBRI, SMAD2, SMAD3, SMAD4, SMAD10, SKI, SKIL, TGIF1, IL10RA, IL10RB, HMOX2, IL6R, IL6ST, EIF2AK4, CSK, PAG1, SIT1, FOXP3, PRDM1, BATF, VISTA, GUCY1A2, GUCY1A3, GUCY1B2, GUCY1B3, MT1, MT2, CD40, OX40, CD137, GITR, CD27, SHP-1, TIM-3, CEACAM-1, CEACAM-3, CEACAM-5, or any combination thereof. In preferred embodiments, the gene locus involved in the expression of PD-1 or CTLA-4 genes is targeted. In other preferred embodiments, combinations of genes are targeted, such as but not limited to PD-1 and TIGIT. Exemplary and non-limiting gene disruptions are found in e.g., WO2016196388, WO2015142675, and Ren et al., (2017) Clin Cancer Res 23 (9) 2255-2266.

[0773] In certain embodiments, cells for ACT are modified to express a CAR, where expression and/or function of methylcytosine dioxygenase genes (TET1, TET2 and/or TET3) in the cells has been reduced or eliminated, (such as the composition or system herein) (see e.g., WO201704916).

[0774] In certain embodiments, cells for ACT are modified such that they have one or more endogenous genes (or expression thereof) knocked down or knocked out, where each encodes an antigen targeted by an exogenous CAR or TCR, thereby reducing the likelihood of targeting of the engineered cells. In certain embodiments, the targeted antigen can be one or more antigen selected from the group consisting of CD38, CD138, CS-1, CD33, CD26, CD30, CD53, CD92, CD100, CD148, CD150, CD200, CD261, CD262, CD362, human telomerase reverse transcriptase (hTERT), survivin, mouse double minute 2 homolog (MDM2), cytochrome P450

1B1 (CYP1B), HER2/neu, Wilms' tumor gene 1 (WT1), livin, alphafetoprotein (AFP), carcinoembryonic antigen (CEA), mucin 16 (MUC16), MUC1, prostate-specific membrane antigen (PSMA), p53, cyclin (D1), B cell maturation antigen (BCMA), transmembrane activator and CAML Interactor (TACI), and/or B-cell activating factor receptor (BAFF-R) (see e.g., WO2016011210 and WO2017011804).

[0775] In some embodiments, cells for ACT are modified such that the gene encoding on e or more MHC constituent proteins or the expression thereof is knocked-out or knocked-down. Exemplary MHC constituent protein encoding genes that can be modified in this way include, without limitation, one or more HLA proteins and/or beta-2 microglobulin (B2M), in a cell, whereby rejection of non-autologous (e.g., allogeneic) cells by the recipient's immune system can be reduced or avoided. In preferred embodiments, one or more HLA class I proteins, such as HLA-A, B and/or C, and/or B2M may be knocked-out or knocked-down. See e.g., Ren et al., (2017) Clin Cancer Res 23 (9) 2255-2266.

[0776] In some embodiments, at least two genes are modified in the engineered cells for ACT. Non-limiting examples of pairs of genes that can be modified engineered cells for ACT include, but are not limited to, PD1 and TCRα, PD1 and TCRβ, CTLA-4 and TCRα, CTLA-4 and TCRβ, LAG3 and TCRα, LAG3 and TCRβ, Tim3 and TCRα, Tim3 and TCRβ, BTLA and TCRα, BTLA and TCRβ, BY55 and TCRα, BY55 and TCRβ, TIGIT and TCRα, TIGIT and TCRβ, B7H5 and TCRα, B7H5 and TCRβ, LAIR1 and TCRα, LAIR1 and TCRβ, SIGLEC10 and TCRα, SIGLEC10 and TCRβ, 2B4 and TCRα, 2B4 and TCRβ, B2M and TCRα, B2M and TCRβ.

[0777] In certain embodiments, multiple genes may be modified in the engineered cells for ACT. Such modifications can be accomplished by serial or multiplexed modifications to the cells. In certain embodiments, a cell may be multiplied edited (multiplex genome editing) as taught herein to (1) knock-out or knock-down expression of an endogenous TCR (for example, TRBC1, TRBC2 and/or TRAC), (2) knock-out or knock-down expression of an immune checkpoint protein or receptor (for example PD1, PD-L1 and/or CTLA4); and (3) knock-out or knock-down expression of one or more MHC constituent proteins (for example, HLA-A, B and/or C, and/or B2M, preferably B2M).

[0778] In some embodiments, T cells used in ACT can be activated and expanded before and/or after modification. Such activation and/or expansion can be achieved generally by using methods such as those described, for example, in U.S. Patents 6,352,694; 6,534,055; 6,905,680;

5,858,358; 6,887,466; 6,905,681; 7,144,575; 7,232,566; 7,175,843; 5,883,223; 6,905,874; 6,797,514; 6,867,041; and 7,572,631. T cells can be expanded *in vitro*, *ex vivo*, or *in vivo*.

[0779] T cells can be obtained from a number of sources, including, but not limited to, peripheral blood mononuclear cells (PBMC), bone marrow, lymph node tissue, spleen tissue, and tumors. Immune cells can be obtained using any method known in the art. In one embodiment, allogenic T cells may be obtained from healthy subjects. In one embodiment T cells that have infiltrated a tumor are isolated. T cells may be removed during surgery. T cells may be isolated after removal of tumor tissue by biopsy. T cells may be isolated by any means known in the art. In one embodiment, T cells are obtained by apheresis. In one embodiment, the method includes obtaining a bulk population of T cells from a tumor sample by any suitable method known in the art. For example, a bulk population of T cells can be obtained from a tumor sample by dissociating the tumor sample into a cell suspension from which specific cell populations can be selected. Suitable methods of obtaining a bulk population of T cells may include, but are not limited to, any one or more of mechanically dissociating (e.g., mincing) the tumor, enzymatically dissociating (e.g., digesting) the tumor, and aspiration (e.g., as with a needle). The bulk population of T cells obtained from a tumor sample can include any suitable type of T cell. In some embodiments, the bulk population of T cells obtained from a tumor sample includes tumor infiltrating lymphocytes (TILs). The tumor sample may be obtained from any mammal. Unless stated otherwise, as used herein, the term "mammal" refers to any mammal including, but not limited to, mammals of the order Logomorpha, such as rabbits; the order Carnivora, including Felines (cats) and Canines (dogs); the order Artiodactyla, including Bovines (cows) and Swines (pigs); or of the order Perssodactyla, including Equines (horses). The mammals may be non-human primates, e.g., of the order Primates, Ceboids, or Simoids (monkeys) or of the order Anthropoids (humans and apes). In some embodiments, the mammal may be a mammal of the order Rodentia, such as mice and hamsters. Preferably, the mammal is a non-human primate or a human. An especially preferred mammal is the human.

[0780] In certain embodiments, T cells can be obtained from a unit of blood collected from a subject using any number of techniques known to the skilled artisan, such as Ficoll separation. In one preferred embodiment, cells from the circulating blood of an individual are obtained by apheresis or leukapheresis. The apheresis product typically contains lymphocytes, including T cells, monocytes, granulocytes, B cells, other nucleated white blood cells, red blood cells, and platelets. In one embodiment, the cells collected by apheresis may be washed to remove the

plasma fraction and to place the cells in an appropriate buffer or media for subsequent processing steps. In one embodiment of the invention, the cells are washed with phosphate buffered saline (PBS). In an alternative embodiment, the wash solution lacks calcium and may lack magnesium or may lack many if not all divalent cations. Initial activation steps in the absence of calcium lead to magnified activation. As those of ordinary skill in the art would readily appreciate a washing step can be accomplished by methods known to those in the art, such as by using a semi-automated "flow-through" centrifuge (for example, the Cobe 2991 cell processor) according to the manufacturer's instructions. After washing, the cells can be resuspended in a variety of biocompatible buffers, such as, for example, Ca-free, Mg-free PBS. Alternatively, the undesirable components of the apheresis sample may be removed and the cells directly resuspended in culture media.

[0781] In some embodiments, T cells are isolated from peripheral blood lymphocytes by lysing the red blood cells and depleting the monocytes, for example, by centrifugation through a PERCOLL™ gradient. A specific subpopulation of T cells, such as CD28+, CD4+, CDC, CD45RA+, and CD45RO+ T cells, can be further isolated by positive or negative selection techniques. For example, in one exemplary embodiment, T cells are isolated by incubation with anti-CD3/anti-CD28 (i.e., 3×28)-conjugated beads, such as DYNABEADS® M-450 CD3/CD28 T, or XCYTE DYNABEADS™ for a time period sufficient for positive selection of the desired T cells. In one embodiment, the time period is about 30 minutes. In a further embodiment, the time period ranges from 30 minutes to 36 hours or longer and all integer values there between. In a further embodiment, the time period is at least 1, 2, 3, 4, 5, or 6 hours. In some embodiments, the time period is 10 to 24 hours. In one preferred embodiment, the incubation time period is 24 hours. For isolation of T cells from patients with leukemia, use of longer incubation times, such as 24 hours, can increase cell yield. Longer incubation times may be used to isolate T cells in any situation where there are few T cells as compared to other cell types, such in isolating tumor infiltrating lymphocytes (TIL) from tumor tissue or from immunocompromised individuals. Further, use of longer incubation times can increase the efficiency of capture of CD8+ T cells.

[0782] Enrichment of a T cell population by negative selection can be accomplished with a combination of antibodies directed to surface markers unique to the negatively selected cells. A preferred method is cell sorting and/or selection via negative magnetic immunoadherence or flow cytometry that uses a cocktail of monoclonal antibodies directed to cell surface markers

present on the cells negatively selected. For example, to enrich for CD4+ cells by negative selection, a monoclonal antibody cocktail typically includes antibodies to CD14, CD20, CD11b, CD16, HLA-DR, and CD8.

[0783] Further, monocyte populations (e.g., CD14+ cells) may be depleted from blood preparations by a variety of methodologies, including anti-CD14 coated beads or columns, or utilization of the phagocytotic activity of these cells to facilitate removal. Accordingly, in one embodiment, the invention uses paramagnetic particles of a size sufficient to be engulfed by phagocytotic monocytes. In certain embodiments, the paramagnetic particles are commercially available beads, for example, those produced by Life Technologies under the trade name Dynabeads™. In one embodiment, other non-specific cells are removed by coating the paramagnetic particles with "irrelevant" proteins (e.g., serum proteins or antibodies). Irrelevant proteins and antibodies refers to molecules (e.g., proteins and antibodies or fragments thereof) that do not specifically target and/or bind the T cells to be isolated. In certain embodiments, the irrelevant beads include beads coated with sheep anti-mouse antibodies, goat anti-mouse antibodies, and human serum albumin.

[0784] In some embodiments, such depletion of monocytes is performed by preincubating T cells isolated from whole blood, apheresed peripheral blood, or tumors with one or more varieties of irrelevant or non-antibody coupled paramagnetic particles at any amount that allows for removal of monocytes (approximately a 20:1 bead:cell ratio) for about 30 minutes to 2 hours at 22 to 37 degrees C., followed by magnetic removal of cells which have attached to or engulfed the paramagnetic particles. Such separation can be performed using standard methods available in the art. For example, any magnetic separation methodology may be used including a variety of which are commercially available, (e.g., DYNAL® Magnetic Particle Concentrator (DYNAL MPC®)). Assurance of requisite depletion can be monitored by a variety of methodologies known to those of ordinary skill in the art, including flow cytometric analysis of CD14 positive cells, before and after depletion.

[0785] For isolation of a desired population of cells by positive or negative selection, the concentration of cells and surface (e.g., particles such as beads) can be varied. In certain embodiments, the volume in which beads and cells are mixed together is significantly decreased which increases the concentration of the cells, which can ensure and increased and/or maximum contact of cells and beads. For example, in one embodiment, a concentration of 2 billion cells/ml is used. In one embodiment, a concentration of 1 billion cells/ml is used. In a

further embodiment, greater than 100 million cells/ml is used. In a further embodiment, a concentration of cells of 10, 15, 20, 25, 30, 35, 40, 45, or 50 million cells/ml is used. In yet another embodiment, a concentration of cells from 75, 80, 85, 90, 95, or 100 million cells/ml is used. In further embodiments, concentrations of 125 or 150 million cells/ml can be used. Using high concentrations can result in increased cell yield, cell activation, and cell expansion. Further, use of high cell concentrations can allow for more efficient capture of cells that may, for example, weakly express target antigens of interest, such as CD28-negative T cells, or from samples where there are many tumor cells present (e.g., leukemic blood, tumor tissue, etc.). Such populations of cells can have therapeutic value and in some embodiments be desirable to obtain. For example, using high concentration of cells allows more efficient selection of CD8+ T cells that normally have weaker CD28 expression.

[0786] In some embodiments, it can be desirable to use lower concentrations of cells. By significantly diluting the mixture of T (or other) cells and surface (e.g., particles such as beads), interactions between the particles and cells is minimized. This can select for cells that express high amounts of desired antigens to be bound to the particles. For example, CD4+ T cells express higher levels of CD28 and are more efficiently captured than CD8+ T cells in dilute concentrations. In one embodiment, the concentration of cells used is 5×10^6/ml. In other embodiments, the concentration used can be from about 1×105/ml to 1×106/ml, and any integer value in between.

[0787] T cells can also be frozen. Wishing not to be bound by theory, the freeze and subsequent thaw step provides a more uniform product by removing granulocytes and to some extent monocytes in the cell population. After a washing step to remove plasma and platelets, the cells may be suspended in a freezing solution. While many freezing solutions and parameters are known in the art and will be useful in this context, one method involves using PBS containing 20% DMSO and 8% human serum albumin, or other suitable cell freezing media, the cells then are frozen to −80° C at a rate of 1° per minute and stored in the vapor phase of a liquid nitrogen storage tank. Other methods of controlled freezing may be used as well as uncontrolled freezing immediately at −20° C. or in liquid nitrogen.

[0788] T cells for use in ACT can be antigen-specific T cells. For example, tumor-specific T cells can be used. In certain embodiments, antigen-specific T cells can be isolated from a subject in need thereof, such as a patient afflicted with a cancer or an infectious disease. In some embodiments, neoepitopes are determined for a subject and T cells specific to these

antigens are isolated. Antigen-specific cells for use in expansion can also be generated in vitro using any number of methods known in the art, for example, as described in U.S. Patent Publication No. US 2004/0224402 or in U.S. Pat. No. 6,040,177. Antigen-specific cells for use ACT can also be generated using any number of methods known in the art, for example, as described in Current Protocols in Immunology, or Current Protocols in Cell Biology, both published by John Wiley & Sons, Inc., Boston, Mass.

[0789] In some embodiments, the antigen-specific cells are sorted or otherwise positively selected (e.g., via magnetic selection) prior to and/or following one, two, or more rounds of expansion. Sorting and/or otherwise positively selecting antigen-specific cells can be carried out using peptide-MHC tetramers (Altman, et al., Science. 1996 Oct. 4; 274(5284):94-6), and/or via use of an adaptable tetramer technology (Andersen et al., 2012 Nat Protoc. 7:891-902). Peptide-MHC tetramers can be generated using techniques known in the art and can be made with any MHC molecule of interest and any antigen of interest as described herein. Specific epitopes to be used in this context can be identified using numerous assays known in the art. For example, the ability of a polypeptide to bind to MHC class I can be evaluated indirectly by monitoring the ability to promote incorporation of 125I labeled β2-microglobulin (β2m) into MHC class I/β2m/peptide heterotrimeric complexes (see e.g., Parker et al., J. Immunol. 152:163, 1994).

[0790] In some embodiments, cells are directly labeled with an epitope-specific reagent for isolation by flow cytometry followed by characterization of phenotype and TCRs. In some embodiments, T cells are isolated by contacting with T cell specific antibodies. Sorting of antigen-specific T cells, or generally any cells described herein (including those for ACT and others), can be carried out using any of a variety of commercially available cell sorters, including, but not limited to, MoFlo sorter (DakoCytomation, Fort Collins, Colo.), FACSAria™, FACSArray™, FACSVantage™, BD™ LSR II, and FACSCalibur™ (BD Biosciences, San Jose, Calif.).

[0791] In some embodiments, the cells that also express CD3 are selected. Selection of CD3 expression cells can occur via any suitable method or technique. In some embodiments, selection of CD3 expressing cells is carried out using flow cytometry by any suitable flow cytometry method known in the art. In some embodiments, the flow cytometry method uses suitable antibodies and stains to select cells. Suitable antibodies are those that specifically recognizes and binds to the particular biomarker being selected. For example, the specific

selection of CD3, CD8, TIM-3, LAG-3, 4-1BB, and/or PD-1 can be carried out using anti-CD3, anti-CD8, anti-TIM-3, anti-LAG-3, anti-4-lBB, or anti-PD-1 antibodies, respectively. It will be appreciated that this approach can be extrapolated to other biomarkers of interest. The antibody(ies) can be conjugated to a bead (e.g., a magnetic bead) or to a fluorochrome. In some embodiments, the flow cytometry is fluorescence-activated cell sorting (FACS). TCRs expressed on T cells can, in some embodiments, be selected based on reactivity to autologous tumors. Additionally, T cells that are reactive to tumors can be selected for based on markers using the methods described in International Patent Application Publication Nos. WO2014133567 and WO2014133568, herein incorporated by reference in their entirety. In some embodiments, activated T cells can be selected for based on surface expression of CD107a.

[0792] In one embodiment of the invention, the method further comprises expanding the numbers of T cells in the enriched cell population. Such methods are described in U.S. Patent No. 8,637,307 and is herein incorporated by reference in its entirety. The numbers of T cells may be increased at least about 3-fold (or 4-, 5-, 6-, 7-, 8-, or 9-fold), more preferably at least about 10-fold (or 20-, 30-, 40-, 50-, 60-, 70-, 80-, or 90-fold), more preferably at least about 100-fold, more preferably at least about 1,000-fold, or most preferably at least about 100,000-fold. The numbers of T cells may be expanded using any suitable method known in the art. Exemplary methods of expanding the numbers of cells are described in patent publication No. WO 2003057171, U.S. Patent No. 8,034,334, and U.S. Patent Application Publication No. 2012/0244133, each of which is incorporated herein by reference.

[0793] *Ex vivo* T cell expansion can be performed by isolation of T cells and subsequent stimulation or activation followed by further expansion. In some embodiments, the T cells are stimulated or activated by a single agent. In some embodiments, T cells are stimulated or activated with two agents, one that induces a primary signal and a second that is a co-stimulatory signal. Ligands useful for stimulating a single signal or stimulating a primary signal and an accessory molecule that stimulates a second signal may be used in soluble form. Ligands can be attached to the surface of a cell, to an Engineered Multivalent Signaling Platform (EMSP), or immobilized on a surface. In some embodiments, both primary and secondary agents are co-immobilized on a surface, for example a bead or a cell. In one embodiment, the molecule providing the primary activation signal may be a CD3 ligand, and the co-stimulatory molecule may be a CD28 ligand or 4-1BB ligand.

[0794] In certain embodiments, T cells comprising a CAR or an exogenous TCR, can be manufactured as described in WO2015120096, other methods described herein, or those known in the art. In some embodiments, such a method can include obtaining and/or enriching a population of lymphocytes obtained from a donor subject; stimulating the population of lymphocytes with one or more T-cell stimulating agents to produce a population of activated T cells, where the stimulation is performed in a closed system using serum-free culture medium; transducing the population of activated T cells with a viral vector or viral vector system including a nucleic acid molecule which encodes the CAR or TCR, using one or more cycle transduction to produce a population of transduced T cells, where the transduction is performed in a closed system using serum-free culture medium; and expanding the population of transduced T cells for a predetermined time to produce a population of engineered T cells, where the expansion is performed in a closed system using serum-free culture medium.

[0795] The predetermined time for expanding the population of transduced T cells can be 1, 2, 3, 4, 5, 6, or more days. In some embodiments, the predetermined time is about 3 days. The time from enriching the population of lymphocytes to producing the engineered T cells can be about 1, 2, 3, 4, 5, 6, or more days. In some embodiments, the time from enriching the population of lymphocytes to producing the engineered T cells is about 6 days. The closed system can be a closed bag system. Further provided is a population of T cells comprising a CAR or an exogenous TCR obtainable or obtained by said method, and a pharmaceutical composition comprising such cells.

[0796] In certain embodiments, T cell maturation or differentiation in vitro may be delayed or inhibited by the method as described in WO2017070395. In some embodiments the method of delaying or inhibiting can include contacting one or more T cells from a subject in need of a T cell therapy with an AKT inhibitor (such as, e.g., one or a combination of two or more AKT inhibitors disclosed in claim 8 of WO2017070395) and at least one of exogenous Interleukin-7 (IL-7) and exogenous Interleukin-15 (IL-15), wherein the resulting T cells exhibit delayed maturation or differentiation, and/or wherein the resulting T cells exhibit improved T cell function (such as, e.g., increased T cell proliferation; increased cytokine production; and/or increased cytolytic activity) relative to a T cell function of a T cell cultured in the absence of an AKT inhibitor.

[0797] In certain embodiments, a subject in need of a T cell therapy can be conditioned by a method as described in WO2016191756. In some embodiments, such a method can include

administering to the subject in need thereof a dose of cyclophosphamide between 200 mg/m2/day and 2000 mg/m2/day and a dose of fludarabine between 20 mg/m2/day and 900 mg/m^2/day.

[0798] In certain embodiments, ACT includes autologous iPSC-based vaccines, such as irradiated iPSCs in autologous anti-tumor vaccines (see e.g., Kooreman, Nigel G. et al., Autologous iPSC-Based Vaccines Elicit Anti-tumor Responses In Vivo, Cell Stem Cell 22, 1–13, 2018, doi.org/10.1016/j.stem.2018.01.016).

Other Cell Transfer and Replacement Therapies

[0799] Cells of any type can be used, as appropriate, for cell transfer and replacement therapies. The cell type modified or used in a cell replacement therapy will depend on the disease being treated. For example, if a blood disease caused by a deficiency in a blood cell is to be treated, then a blood cell progenitor or blood cell can be modified and transferred to a subject in need thereof so as to treat the blood disorder. Likewise, if a muscle disease caused by a missing or non-function muscle cell, then a muscle cell or muscle cell progenitor can be modified so as to provide the missing function and subsequently transferred to the subject in need thereof so as to treat and/or prevent the disease, condition, or symptom thereof caused by the non-function muscle cell.

[0800] In some embodiments, the cells modified by the compositions, systems, and formulations thereof described herein for a cell transfer or replacement therapy can be, without limitation, stem and progenitor cells (e.g., pluripotent, multipotent, induced pluripotent stem cells, mesenchymal stem cells, hematopoietic stem cells, adipose stem cells, hair follicle stem cells, etc.), primary cells, differentiated cells, and the like. In some embodiments, the cells are progenitors or are differentiated (before or after transfer to the subject in need thereof) into central nervous system cells and/or peripheral nervous system cells (e.g., neurons, astrocytes, microglial cells, Schwan cells, satellite cells, oligodendrocytes, and ependymal cells), muscle cells (e.g., smooth muscle cells, skeletal muscle cells, and cardiac muscle cells), skin cells (e.g., keratinocytes, and melanocytes), eye cells (e.g., photoreceptor cells (e.g. rod cells and cone cells), retinal ganglion, bipolar cells, retinal horizontal cells, and retinal amacrine cells, corneal epithelial cells, corneal keratocytes), liver cells (e.g., liver parenchymal cells (e.g., hepatocytes) and non-parenchymal cells (e.g., Kupffer cells), pancreatic cells (e.g., islet cells (e.g., alpha cells, beta cells), spleen cells (e.g. white pulp cells and red pulp cells), kidney cells (e.g., glomerular cells (e.g. glomerular endothelia cells, glomerular basement membrane cells)

macula densa cells, podocytes, tubule epithelial cells, tubular epithelial cells, intercalated cells, and mesangial cells), stomach cells (e.g. stomach epithelial cells, mucous producing cells, parietal cells, chief cells, and neuroendocrine cells (e.g. G-cells)), intestinal cells (e.g., enterocytes, neuroendocrine cells, Paneth cells, and goblet cells), fat cells (e.g. white adipocytes, brown fat adipocytes, and beige fat adipocytes), blood cells (e.g., red blood cells (e.g., erythrocytes), platelets, white blood cells (e.g., monocytes, lymphocytes, neutrophils, eosinophils, basophils, and macrophages), bone marrow cells (e.g., mesenchymal stem cells, hematopoietic stem cells), bone cells (e.g., osteoclasts, osteoblasts, and osteocytes), cartilage cells (e.g., chondrocytes), tendon and ligament cells (e.g., tenocytes and fibrocytes), gallbladder cells (e.g., gallbladder epitheial cells), ear cells (e.g., hair cells and non-sensory supporting cells) (see e.g., Wan et al., 2013. Semin Cell Dev Biol. 24(5):448-459), nasal cells (e.g., basal epithelial cells, supporting cells, and olfactory receptor cells), blood vessel cells, adrenal gland cells, immune cells (e.g., lymphocytes (e.g., T-cells and B-cells), monocytes, neutrophils, eosinophils, basophils, mast cells, dendritic cells, natural killer cells, macrophages (see also ACT previously described), testicular cells (e.g., Sertoli cells and interstitial cells of Leydig), endometrial cells, adrenal gland cells, lung cells, thyroid cells, and combinations thereof.

[0801] Diseases and conditions that can be treated by cell replacement therapies include, without limitation, spinal cord injuries, type 1 diabetes, Parkinson's disease, amyotrophic lateral sclerosis, Alzheimer's disease, other neurodegenerative diseases, heart disease, stroke, burns, chronic wounds, cancer, osteoarthritis, obesity, multiple sclerosis, kidney diseases, liver diseases, irritable bowel diseases, blood diseases, hair loss, hearing loss, macular degeneration and other retinal disorders and diseases, muscular dystrophies and injuries, skin diseases, fibrotic diseases, tendon and ligament lesions, age related diseases, periodontal diseases, soft tissue diseases and disorders, tissue reconstructions, and others (see e.g., Genc et al. 2019. Adv Exp Med Biol. 2019. 1084:145-174; Loretelli et al. 2020. Expert Opin Biol Ther. 20(8):887-897; Katolikova et al. Mol Biol (Mosk.) 2020. 54(6):939-954; Schuetz et al. 2018. Transplantation. 102(2)215-229; Gao et al. 2018. Cell Tranplant. 27(12)1809-1824); Zeri and Abbaszadeh. 2019. Curr Stem Cell Res Ther. 14(3):244-248; Müller et al. 2018. Cell Physiol Biochem. 48(6): 2607-2655; Schmeer et al. 2012. Cell Tissue Res. 349(1): 363-374; González et al. 2016. Curr Stem Cell Res Ther. 11(7):561-568; Buzhor et al., 2014. Regen Med. 9(5):649-672; Sonntag et al. 2018. Prog Neurobiol. 168:1-20; Barthélém and Wein. 2018. Neuromuscul

Disord. 28(10):803-824; Gagliardi et al. 2019. Prog Retin Eye Res. 71:1-25; Hao et al. 2014. Biomed Res Int. 2014:468748. doi: 10.1155/2014/468748; Owczarczyk-Saczonek et al., 2017. Int J Mol Sci. Oct 20;18(10):2182. doi: 10.3390/ijms18102182, Goldman and Windrem. Philos Trans R Soc Lond B Biol Sci. 2006 Sep 29;361(1473):1463-75; Hu and Ulfendahl. 2006. Stem Cells Dev. 2006 Jun;15(3):449-59. doi: 10.1089/scd.2006.15.449; Rickels and Robertson. Endocr Rev. 2019 Apr 1;40(2):631-668; Protze et al., Cell Stem Cell. 2019 Sep 5;25(3):311-327; Rosser and Svendsen. Mov Disord. 2014 Sep 15;29(11):1446-54; Levin et al., Ophthalmology. 2017 Jul;124(7):926-934; Chakrabarty et al., Stem Cell Res Ther. 2018 Oct 25;9(1):287; Neves et al., Cell Stem Cell. 2017 Feb 2;20(2):161-175; Gennai et al., Br J Anaesth. 2015 Aug;115(2):203-12; Miltner et al., Dev Dyn. 2019 Jan;248(1):118-128; Badner et al., Expert Opin Biol Ther. 2017 May;17(5):529-541; Yuan and Fan. Regen Med. 2015 May;10(4):495-504; Matsui et al., Exp Neurol. 2014 Oct;260:12-8. doi: 10.1016/j.expneurol.2012.09.016. Epub 2012 Oct 1.; Ruben et al., 2013. Lancet. Sep 28;382(9898):1077-9. doi: 10.1016/S0140-6736(13)61744-4; Larson et al., J Dent Res. 2016 Mar;95(3):255-66; Qazi et al., J Cachexia Sarcopenia Muscle. 2019 Jun;10(3):501-516; Brayfield et al., Handchir Mikrochir Plast Chir. 2010 Apr;42(2):124-8; Karagoz et al., Curr Opin Organ Transplant. 2019 Oct;24(5):598-603; and Parkin et al. Stem Cells Int. 2017;2017:3750829. doi: 10.1155/2017/3750829.

[0802] In some embodiments, cells are obtained and/or enriched from a subject in need thereof, modified as is indicated by the disease and/or condition to be treated, and transferred by any suitable method back into the subject in need thereof. In other embodiments, cells from a non-recipient donor, can be optionally modified so as to reduce GVHD, modified as is indicated by the disease and/or condition to be treated, and transferred by any suitable method into a non-donor subject in need thereof. Other approaches to reduce GVHD are described elsewhere herein, such as in connection with ACT, and can be applied in this context. In some embodiments, the cells are sorted and/or enriched by a suitable method of sorting or enrichment. Various general cell sorting and/or enriching techniques (such as those based on biomarker detection etc.) are described elsewhere herein and are generally known in the art and can be applied in this context to sort and/or enrich populations of cells for cell transfer before and/or after *ex vivo* modification.

[0803] In some embodiments, the cell therapy is an HSC or iPSC-based cell thereapy. In some embodiments, HSCs or iPSCs can be modified ex vivo for treatment of a disease or

disorder, such as a blood disease or disorder. Other diseases or disorders that can be treated with a modified HSC or iPSC are described in greater detail elsewhere herein. In some embodiments, after ex vivo modification of HSCs or iPCS can be expanded prior to administration to the subject. Expansion of HSCs can be via any suitable method such as that described by, Lee, "Improved ex vivo expansion of adult hematopoietic stem cells by overcoming CUL4-mediated degradation of HOXB4." Blood. 2013 May 16;121(20):4082-9. doi: 10.1182/blood-2012-09-455204. Epub 2013 Mar 21. In some embodiments, the HSCs or iPSCs modified can be autologous. In some embodiments, the HSCs or iPSCs can be allogenic. In addition to the modification of the disease gene(s), allogenic cells can be further modified using the engineered systems of the present invention (such as to deliver a CRISPR-Cas system) to reduce the immunogenicity of the cells when delivered to the recipient. Such techniques are described elsewhere herein and e.g., Cartier, "MINI-SYMPOSIUM: X-Linked Adrenoleukodystrophypa, Hematopoietic Stem Cell Transplantation and Hematopoietic Stem Cell Gene Therapy in X-Linked Adrenoleukodystrophy," Brain Pathology 20 (2010) 857–862, which can be adapted for use with the engineered delivery vesicle generation compositions and systems of the present invention.

[0804] In some embodiments, the cell therapy is a cell therapy for inner ear cells or other cells of the ear. In general, the cell therapy methods described in US patent application 20120328580 can be used to promote complete or partial differentiation of a cell to or towards a mature cell type of the inner ear (e.g., a hair cell) in vitro. Cells resulting from such methods can then be transplanted or implanted into a patient in need of such treatment. The cell culture methods required to practice these methods, including methods for identifying and selecting suitable cell types, methods for promoting complete or partial differentiation of selected cells, methods for identifying complete or partially differentiated cell types, and methods for implanting complete or partially differentiated cells are described below.

[0805] Cells suitable for use in the present invention include, but are not limited to, cells that are capable of differentiating completely or partially into a mature cell of the inner ear, e.g., a hair cell (e.g., an inner and/or outer hair cell), when contacted, e.g., in vitro, with one or more of the compounds described herein. Exemplary cells that are capable of differentiating into a hair cell include, but are not limited to stem cells (e.g., inner ear stem cells, adult stem cells, bone marrow derived stem cells, embryonic stem cells, mesenchymal stem cells, skin stem cells, iPS cells, and fat derived stem cells), progenitor cells (e.g., inner ear progenitor

cells), support cells (e.g., Deiters' cells, pillar cells, inner phalangeal cells, tectal cells and Hensen's cells), and/or germ cells. The use of stem cells for the replacement of inner ear sensory cells is described in Li et al., (U.S. Publication No. 2005/0287127) and Li et al., (U.S. patent Ser. No. 11/953,797). The use of bone marrow derived stem cells for the replacement of inner ear sensory cells is described in Edge et al., PCT/US2007/084654. iPS cells are described, e.g., at Takahashi et al., Cell, Volume 131, Issue 5, Pages 861-872 (2007); Takahashi and Yamanaka, Cell 126, 663-76 (2006); Okita et al., Nature 448, 260-262 (2007); Yu, J. et al., Science 318(5858):1917-1920 (2007); Nakagawa et al., Nat. Biotechnol. 26:101-106 (2008); and Zaehres and Scholer, Cell 131(5):834-835 (2007). Such suitable cells can be identified by analyzing (e.g., qualitatively or quantitatively) the presence of one or more tissue specific genes and/or proteins. Suitable methods of measuring gene and/or protein expression and using such expression profiles to identify cell type are generally known in the art.

Xenotransplantation

[0806] In some embodiments, the engineered compositions, systems, and formulations thereof and/or cargos delivered therefrom of the present disclosure described herein can be used to generate modified cells, tissues, and/or organs, and/or systems for transplantation. In other words, the modified cells, tissues, and/or organs described in greater detail elsewhere herein can be used for xenotransplatation. For example, RNA-guided DNA nucleases may be used to knockout, knockdown or disrupt selected genes in an animal, such as a transgenic pig (such as the human heme oxygenase-1 transgenic pig line), for example by disrupting expression of genes that encode epitopes recognized by the human immune system, i.e., xenoantigen genes. Candidate porcine genes for disruption may for example include α(l,3)-galactosyltransferase and cytidine monophosphate-N-acetylneuraminic acid hydroxylase genes (see PCT Patent Publication WO 2014/066505). In addition, genes encoding endogenous retroviruses may be disrupted, for example the genes encoding all porcine endogenous retroviruses (see Yang et al., 2015, Genome-wide inactivation of porcine endogenous retroviruses (PERVs), Science 27 November 2015: Vol. 350 no. 6264 pp. 1101-1104). In addition, RNA-guided DNA nucleases may be used to target a site for integration of additional genes in xenotransplant donor animals, such as a human CD55 gene to improve protection against hyperacute rejection.

[0807]

Bioreactors

[0808] Engineered, cells, tissues, organoids, plants, and non-human animals can be modified using the engineered compositions, systems, and formulations thereof and/or cargos delivered therefrom of the present disclosure, such that they produce one or more cell, tissue, or organ products that can be secreted, excreted, or otherwise harvested from the cells tissues and/or organ by any suitable methods. Such cells, tissues, organoids, plants, and non-human animals are described in greater detail elsewhere herein. It will be appreciated that the product produced by an engineered cell, tissue, organoid, or non-human animal can be native (e.g., a product normally produced by the cell, tissue, organ the organoid is based on, or non-human animal) or non-native (e.g., a product normally produced by a different cell type, a cell of the same type but a different state, different species, etc.). Where a native product is product, the native product can be produced in the engineered cell, tissue, organoid, or non-human animal at an increased amount, concentration, and/or more readily harvested for subsequent downstream use.

[0809] Products produced can be nucleic acids, proteins (including antibodies and proteins for replacement therapies), exosomes, small molecules, fats, organic molecules, and combinations thereof. The products can have therapeutic and/or industrial applications. For example, the product produced can be a protein therapy or small molecule pharmaceutically active agent or be a biofuel. Others are described herein and will be appreciated by those of skill in the art. Exemplary industrial scale production of therapeutic proteins in various cells is described in e.g., Zhu et al., 2017. Handbook of Industrial Chemistry and Biotechnology. May 3: 1639-1669; Huang and McDonald. 2009. Biocehm Eng J. 3:168-184; Vandermies and Fickers. 2019. Microorganisms 7:40; Wurm. Nat Biotechnol. 2004 Nov;22(11):1393-8. doi: 10.1038/nbt1026;

[0810] In some embodiments, engineered plants and parts thereof described in greater detail elsewhere herein are used as bioreactors to produce one or more proteins and other biological molecules (plant metabolites etc.), small molecule therapeutics, and organic molecules (e.g., volatiles), that can be used as therapeutics, feed and nutritional supplements, and/or for other industrial applications (such as an energy source). See e.g., Balfour. 2020. European Pharmaceutical Review. Burnett and Burnett 2020. Plants People Planet. 2:121-132; Tekoah et al., 2015. Plant Biotech J. 13(8): 1199-1208; Yusibov et al., Annu Rev Plant Biol. 2016 Apr 29;67:669-701; Hidalgo et al., Curr Med Chem. 2018;25(30):3577-3596; Tiwari et

al., Biotechnol Adv. 2009 Jul-Aug;27(4):449-67; Lao et al., Biotechnol Adv. 2009 Nov-Dec;27(6):1015-1022; Vianna et al., Genet Mol Res. 2011;10(3):1733-52; Sharma et al., Biotechnol Adv. 2009 Nov-Dec;27(6):811-832; Yao et al., Int J Mol Sci. 2015 Dec 2;16(12):28549-65; Ahmad et al., Recent Pat Biotechnol. 2010 Nov;4(3):242-59; Donini and Marusic. Biotechnol Lett. 2019 Mar;41(3):335-346; Desai et al., Biotechnol Adv. 2010 Jul-Aug;28(4):427-35; Zhang et al., J Ind Microbiol Biotechnol. 2017 May;44(4-5):773-784; Shakoor et al., Acta Virol. 2019;63(3):245-252; Singh et al., Curr Mol Biol Rep. 2017;3(4):306-316; Lee et al., Biomol Ther (Seoul). 2017 Jul 1;25(4):345-353; Buyel. Front Plant Sci. 2019 Jan 18;9:1893; Rybicki. Hum Vaccin Immunother. 2017 Dec 2;13(12):2912-2917; Dhama et al., Hum Vaccin Immunother. 2020 Dec 1;16(12):2913-2920; Miele. 1 Trends Biotechnol. 1997 Feb;15(2):45-50; Yao et al., Int J Mol Sci. 2015 Dec; 16(12): 28549–28565; Paul and Ma. 2011. Biotech and App. Biochem. 58(1):58-67; Huebbers et al. Biotechnol Adv. 2021 Jan-Feb;46:107681; Hood. F1000Research 2016, 5(F1000 Faculty Rev):185; Sticklen. Nature Reviews Genetics volume 9, pages433–443(2008); Shen et al. Biotechnology for Biofuels 2013 6:71; Kolotilin et al., Biotechnology for Biofuels 2013 6:65; Petersen et al., Biotechnology for Biofuels 2012 5:84; Yee et al., Biotechnology for Biofuels 2012 5:81.

[0811] In some embodiments, the modified plants, algae, yeast or other non-animal organisms can be used to produce a desirable gene product. The desirable gene product can then be harvested after production and used accordingly.

[0812] In some embodiments, the engineered plants and parts thereof are used as bioreactors for the production of small molecules, such as small molecule therapeutics. See e.g., Buyel. Front Plant Sci. 2019 Jan 18;9:1893.

Biological System Characterization and Screening

[0813] In some embodiments, the modified non-human animals and plants can be modified such that they allow for identification, characterization, and/or an understanding, such as a function, biological role, and or mechanism, of an endogenous or exogenous system. In particular embodiments, the modified non-human animals and/or plants are modified in such a way that they allow for visualization of genetic element dynamics. For example, CRISPR imaging can visualize either repetitive or non-repetitive genomic sequences, report telomere length change and telomere movements and monitor the dynamics of gene loci throughout the cell cycle (Chen et al., Cell, 2013). These methods may also be applied to plants and non-human animals.

[0814] The engineered compositions, systems, and methods may be used for screening genes (e.g., endogenous, mutations) of interest in a plant or non-human animal or in cells of a plant, human, or non-human animal. In some examples, genes of interest include those encoding enzymes involved in the production of a component of added nutritional value or generally genes affecting agronomic traits of interest, across species, phyla, and plant or animal kingdom. By selectively targeting e.g., genes encoding enzymes of metabolic pathways, the genes responsible for certain nutritional, immunological, growth, and/or other performance or production aspects of a plant or animal can be identified. Similarly, by selectively targeting genes which may affect a desirable agronomic or production trait, the relevant genes can be identified. Accordingly, the present invention encompasses screening methods for genes encoding enzymes involved in the production of compounds with a particular nutritional, immunological, growth or other performance value and/or agronomic traits.

[0815] In particular embodiments, the systems, and preferably the engineered compositions, systems, vesicles etc. of the present, can be used to purify a specific portion of the chromatin and identify the associated proteins, thus elucidating their regulatory roles in transcription (Waldrip et al., Epigenetics, 2014). These methods may also be applied to plants and non-human animals.

[0816] The polynucleotide modifying agent(s) and systems can be used to generate loss of function plants, algae, yeast, and other non-animal organisms, which can allow for functional analysis of genomic material. Ma et al. (Mol Plant. 2015 Aug 3;8(8):1274-84. doi: 10.1016/j.molp.2015.04.007) reports robust CRISPR-Cas9 vector system, utilizing a plant codon optimized Cas9 gene, for convenient and high-efficiency multiplex genome editing in monocot and dicot plants. Ma et al. designed PCR-based procedures to rapidly generate multiple sgRNA expression cassettes, which can be assembled into the binary CRISPR-Cas9 vectors in one round of cloning by Golden Gate ligation or Gibson Assembly. With this system, Ma et al. edited 46 target sites in rice with an average 85.4% rate of mutation, mostly in biallelic and homozygous status. Ma et al. provide examples of loss-of-function gene mutations in T0 rice and T1Arabidopsis plants by simultaneous targeting of multiple (up to eight) members of a gene family, multiple genes in a biosynthetic pathway, or multiple sites in a single gene. The methods of Ma et al. can be applied to the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention.

[0817] The engineered compositions, systems, cells, methods, modified plants and non-human animals may be used to identify, screen, introduce or remove mutations or sequences lead to genetic variability that give rise to susceptibility to certain pathogens, e.g., host specific pathogens. Such approach may generate plants that are non-host resistance, e.g., the host and pathogen are incompatible or there can be partial resistance against all races of a pathogen, typically controlled by many genes and/or also complete resistance to some races of a pathogen but not to other races.

[0818]

Model Systems

[0819] In some embodiments, the engineered compositions, systems, formulations, delivery vesicles, and/or cargos of the present invention can be used to generate a cellular, tissue, organ, organoid, tumor, non-human animal or non-animal organism model system. Modified non-human non-animal organisms or cells thereof can be modified to express one or more heterologous genes, such as genes from a human, non-human animal, or plant. Such model systems can be used for example, to determine response to disease, environmental toxins, pharmaceutical agents, or other stimuli. Other uses for such model systems will be appreciated by those of ordinary skill in the art.

[0820] The modified non-human animals described herein can be a platform to model a disease or disorder of an animal, including but not limited to mammals. In some of these embodiments, the mammal can be a human. In certain embodiments, such models and platforms are rodent based, in non-limiting examples rat or mouse. Such models and platforms can take advantage of distinctions among and comparisons between inbred rodent strains. In certain embodiments, such models and platforms primate, horse, cattle, sheep, goat, swine, dog, cat or bird-based, for example to directly model diseases and disorders of such animals or to create modified and/or improved lines of such animals. Advantageously, in certain embodiments, an animal-based platform or model is created to mimic a human disease or disorder. For example, the similarities of swine to humans make swine an ideal platform for modeling human diseases. Compared to rodent models, development of swine models has been costly and time intensive. On the other hand, swine and other animals are much more similar to humans genetically, anatomically, physiologically and pathophysiologically. The present invention provides a high efficiency platform for targeted gene and genome editing, gene and genome modification and gene and genome regulation to be used in such animal platforms and

models. Though ethical standards block development of human models and in many cases, models based on non-human primates, the present invention is used with in vitro systems, including but not limited to cell culture systems, three dimensional models and systems, and organoids to mimic, model, and investigate genetics, anatomy, physiology and pathophysiology of structures, organs, and systems of humans. The platforms and models provide manipulation of single or multiple targets.

[0821] In certain embodiments, the present invention is applicable to disease models like that of Schomberg et al. (FASEB Journal, April 2016; 30(1):Suppl 571.1). To model the inherited disease neurofibromatosis type 1 (NF-1) Schomberg used CRISPR-Cas9 to introduce mutations in the swine neurofibromin 1 gene by cytosolic microinjection of CRISPR/Cas9 components into swine embryos. CRISPR guide RNAs (gRNA) were created for regions targeting sites both upstream and downstream of an exon within the gene for targeted cleavage by Cas9 and repair was mediated by a specific single-stranded oligodeoxynucleotide (ssODN) template to introduce a 2500 bp deletion. The systems were also used to engineer swine with specific NF-1 mutations or clusters of mutations, and further can be used to engineer mutations that are specific to or representative of a given human individual. Such techniques and modifications can be adapted for and used with the modifying agent(s) and systems thereof described herein to generate a modified non-human animal or cell thereof. In some embodiments, the polynucleotide modifying agent(s) or systems thereof can be similarly used to develop animal models, including but not limited to swine models, of human multigenic diseases. In some embodiments, multiple genetic loci in one gene or in multiple genes are simultaneously targeted using multiplexed guides and optionally one or multiple templates.

Other Uses of Engineered Non-Human Animals and Plants

[0822] As is also described elsewhere herein, the engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used in the generation of "improved plants, algae, fungi, yeast, etc." in that they have one or more desirable traits compared to the wildtype plant. In particular embodiments, the plants, algae, fungi, yeast, etc., cells or parts obtained are transgenic plants, comprising an exogenous DNA sequence incorporated into the genome of all or part of the cells. In particular embodiments, non-transgenic genetically modified plants, algae, fungi, yeast, etc., parts or cells are obtained, in that no exogenous DNA sequence is incorporated into the genome of any of the cells of the plant. In such embodiments, the improved plants, algae, fungi, yeast, etc. are non-transgenic.

Where only the modification of an endogenous gene is ensured and no foreign genes are introduced or maintained in the plant, algae, fungi, yeast, etc. genome, the resulting genetically modified crops contain no foreign genes and can thus basically be considered non-transgenic. The different applications of the systems for plant, algae, fungi, yeast, etc. genome editing include, but are not limited to introduction of one or more foreign genes to confer an agricultural trait of interest; editing of endogenous genes to confer an agricultural trait of interest; modulating of endogenous genes by the systems to confer an agricultural trait of interest. Exemplary genes conferring agronomic traits include, but are not limited to, genes that confer resistance to pests or diseases; genes involved in plant diseases, such as those listed in WO 2013046247; genes that confer resistance to herbicides, fungicides, or the like; genes involved in (abiotic) stress tolerance. Other aspects of the use of the systems include, but are not limited to create (male) sterile plants; increasing the fertility stage in plants/algae etc.; generate genetic variation in a crop of interest; affect fruit-ripening; increasing storage life of plants/algae etc.; reducing allergen in plants/algae etc.; ensure a value-added trait (e.g. nutritional improvement); Screening methods for endogenous genes of interest; biofuel, fatty acid, organic acid, etc. production.

[0823] Also described herein are modified non-animal organisms (e.g. plants, algae, and yeast cells) obtainable and obtained by the methods provided herein that can be improved in at least one aspect as compared to an unmodified plant. The improved non-animal organisms obtained by the methods described herein may be useful in one or more fields (e.g. food or feed production) through expression of genes or alleles which, for instance ensure tolerance to infectious agents, pests, herbicides, drought, low or high temperatures, excessive water, toxins, etc.

[0824] The improved plants obtained by the methods described herein, especially crops and algae may be useful in food or feed production through expression of, for instance, higher protein, carbohydrate, nutrient or vitamin levels than would normally be seen in the wildtype. In this regard, improved plants, especially pulses and tubers are preferred.

[0825] Improved algae or other plants such as rape may be particularly useful in the production of vegetable oils or biofuels such as alcohols (especially methanol and ethanol), for instance. These may be engineered to express or overexpress high levels of oil or alcohols for use in the oil or biofuel industries.

[0826] Also described herein are improved parts of a plant. Plant parts include, but are not limited to, leaves, stems, roots, tubers, seeds, endosperm, ovule, and pollen. Plant parts as envisaged herein may be viable, nonviable, regeneratable, and/or non- regeneratable. The improved part of the plant can, for example, result in earlier fruit, higher content of one or more molecules involved in fruit taste, color, maturity, ripening, etc. or have other desired characteristics. In one embodiment, the method described in Soyk et al. (Nat Genet. 2017 Jan;49(1):162-168), which used CRISPR-Cas9 mediated mutation targeting flowering repressor SP5G in tomatoes to produce early yield tomatoes can be modified and adapted for use with the polynucleotide modifying agent(s) and systems thereof described herein.

[0827] Improved animals can provide increased performance for sport and commercial production.

Engineered Plants

[0828] As is also described elsewhere herein, the engineered compositions, systems, and methods described herein can be used to perform gene or genome interrogation or editing or manipulation in plants and fungi. For example, the applications include investigation and/or selection and/or interrogations and/or comparison and/or manipulations and/or transformation of plant genes or genomes; e.g., to create, identify, develop, optimize, or confer trait(s) or characteristic(s) to plant(s) or to transform a plant or fugus genome. There can accordingly be improved production of plants, new plants with new combinations of traits or characteristics or new plants with enhanced traits. The compositions, systems, and methods can be used with regard to plants in Site-Directed Integration (SDI) or Gene Editing (GE) or any Near Reverse Breeding (NRB) or Reverse Breeding (RB) techniques.

[0829] The engineered compositions, systems, and methods herein may be used to confer desired traits (e.g., enhanced nutritional quality, increased resistance to diseases and resistance to biotic and abiotic stress, and increased production of commercially valuable plant products or heterologous compounds) on essentially any plants and fungi, and their cells and tissues. The compositions, systems, and methods may be used to modify endogenous genes or to modify their expression without the permanent introduction into the genome of any foreign gene.

[0830] In some embodiments, compositions, systems, and methods may be used in genome editing in plants or where RNAi or similar genome editing techniques have been used previously; see, e.g., Nekrasov, "Plant genome editing made easy: targeted mutagenesis in

model and crop plants using the CRISPR-Cas system," Plant Methods 2013, 9:39 (doi:10.1186/1746-4811-9-39); Brooks, "Efficient gene editing in tomato in the first generation using the CRISPR-Cas9 system," Plant Physiology September 2014 pp 114.247577; Shan, "Targeted genome modification of crop plants using a CRISPR-Cas system," Nature Biotechnology 31, 686-688 (2013); Feng, "Efficient genome editing in plants using a CRISPR/Cas system," Cell Research (2013) 23:1229–1232. doi:10.1038/cr.2013.114; published online 20 August 2013; Xie, "RNA-guided genome editing in plants using a CRISPR-Cas system," Mol Plant. 2013 Nov;6(6):1975-83. doi: 10.1093/mp/sst119. Epub 2013 Aug 17; Xu, "Gene targeting using the Agrobacterium tumefaciens-mediated CRISPR-Cas system in rice," Rice 2014, 7:5 (2014), Zhou et al., "Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate: CoA ligase specificity and Redundancy," New Phytologist (2015) (Forum) 1-4 (available online only at www.newphytologist.com); Caliando et al, "Targeted DNA degradation using a CRISPR device stably carried in the host genome, NATURE COMMUNICATIONS 6:6989, DOI: 10.1038/ncomms7989, www.nature.com/naturecommunications DOI: 10.1038/ncomms7989; US Patent No. 6,603,061 - Agrobacterium-Mediated Plant Transformation Method; US Patent No. 7,868,149 - Plant Genome Sequences and Uses Thereof and US 2009/0100536 - Transgenic Plants with Enhanced Agronomic Traits, Morrell et al "Crop genomics: advances and applications," Nat Rev Genet. 2011 Dec 29;13(2):85-96, all the contents and disclosure of each of which are herein incorporated by reference in their entirety. Aspects of utilizing the compositions, systems, and methods may be analogous to the use of the CRISPR-Cas (e.g. CRISPR-Cas9) system in plants, and mention is made of the University of Arizona website "CRISPR-PLANT" (www.genome.arizona.edu/crispr/) (supported by Penn State and AGI).

[0831] In some embodiments, the modified plants are configured such that a targeted gene disruption results positive-selection screening in vitro and in vivo (see e.g., Malina et al., Genes and Development, 2013).

[0832] In some embodiments, the modified plants are modified such that they have epigenomic changes. For example, the teachings of Rusk et al. Nature Methods. 2014 can be applied to engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom of the present invention. In some embodiments, a cargo is or can be fused with a histone-modifying enzyme and/or compound. In some embodiments, an engineered delivery vesicle can contain a histone modification.

[0833] In particular embodiments, present invention could be used to alter genome complexity. In further particular embodiment, the systems, and preferably the systems described herein, can be used to disrupt or alter chromosome number and generate haploid plants, which only contain chromosomes from one parent. Such plants can be induced to undergo chromosome duplication and converted into diploid plants containing only homozygous alleles (Karimi-Ashtiyani et al., PNAS, 2015; Anton et al., Nucleus, 2014). These methods may also be adapted for use with the present invention.

[0834] Similarly, the engineered compositions, systems, formulations, cargos, and delivery vesicles of the present invention can be configured and used to induce mutations to allow for genome wide screening for mutations, alleles, and variants that have a desired characteristic (e.g., heat tolerance, cold tolerance, fast growth, pest resistance, etc.) and also used to generate plants with the identified and desired allele(s).

[0835] The plants can be modified so as to have one or more modified, such as improved, agronomic traits so as to improve field or production performance. Examples of the traits include improved agronomic traits such as herbicide resistance, disease resistance, abiotic stress tolerance, high yield, and superior quality, pesticide-resistance, disease resistance, insect and nematode resistance, resistance against parasitic weeds, drought tolerance, nutritional value, stress tolerance, self-pollination voidance, forage digestibility biomass, and grain yield.

[0836] The plants can be modified so as to have improved pest or disease resistance. Where the unmodified plant has endogenous genes that confer pest or disease resistance, the expression, activity, and/or function of these can be enhanced in the modified plant (e.g., by introducing extra copies, modifications that enhance expression and/or activity). In some embodiments, the modified plant is modified so as to have heterologous genes that confer pest or disease resistance. Non-limiting exemplary genes that can confer resistance include e.g., Cf-9, Pto, RSP2, SlDMR6-1), genes conferring resistance to a pest (e.g., those described in WO96/30517), Bacillus thuringiensis proteins, lectins, Vitamin-binding proteins (e.g., avidin), enzyme inhibitors (e.g., protease or proteinase inhibitors or amylase inhibitors), insect-specific hormones or pheromones (e.g., ecdysteroid or a juvenile hormone, variant thereof, a mimetic based thereon, or an antagonist or agonist thereof) or genes involved in the production and regulation of such hormone and pheromones, insect-specific peptides or neuropeptide, Insect-specific venom (e.g., produced by a snake, a wasp, etc., or analog thereof), enzymes responsible for a hyperaccumulation of a monoterpene, a sesquiterpene, a steroid, hydroxamic acid, a

phenylpropanoid derivative or another nonprotein molecule with insecticidal activity, enzymes involved in the modification of biologically active molecule (e.g., a glycolytic enzyme, a proteolytic enzyme, a lipolytic enzyme, a nuclease, a cyclase, a transaminase, an esterase, a hydrolase, a phosphatase, a kinase, a phosphorylase, a polymerase, an elastase, a chitinase and a glucanase, whether natural or synthetic), molecules that stimulates signal transduction, Vvral-invasive proteins or a complex toxin derived therefrom, developmental-arrestive proteins produced in nature by a pathogen or a parasite, a developmental-arrestive protein produced in nature by a plant, or any combination thereof.

[0837] As is also described elsewhere herein, the modified plant can be modified so as to have one or more genes involved in disease so as to improve disease resistance or treat disease in a plant. Exemplary disease related genes include, without limitation, those described in [0045]-[0080] of US20140213619A1, which is incorporated by reference herein in its entirety.

[0838] The modified plants can be modified so as to have resistance to herbicides so as to e.g., increase efficiency of plant propagation. Examples of genes that confer resistance to herbicides include genes conferring resistance to herbicides that inhibit the growing point or meristem, such as an imidazolinone or a sulfonylurea, genes conferring glyphosate tolerance (e.g., resistance conferred by, e.g., mutant 5-enolpyruvylshikimate-3- phosphate synthase genes, aroA genes and glyphosate acetyl transferase (GAT) genes, respectively), or resistance to other phosphono compounds such as by glufosinate (phosphinothricin acetyl transferase (PAT) genes from Streptomyces species, including Streptomyces hygroscopicus and Streptomyces viridichromogenes), and to pyridinoxy or phenoxy proprionic acids and cyclohexones by ACCase inhibitor-encoding genes), genes conferring resistance to herbicides that inhibit photosynthesis (such as a triazine (psbA and gs+ genes) or a benzonitrile (nitrilase gene), and glutathione S-transferase), genes encoding enzymes detoxifying the herbicide or a mutant glutamine synthase enzyme that is resistant to inhibition, genes encoding a detoxifying enzyme is an enzyme encoding a phosphinothricin acetyltransferase (such as the bar or pat protein from Streptomyces species), genes encoding hydroxyphenylpyruvatedioxygenases (HPPD) inhibitors, e.g., naturally occurring HPPD resistant enzymes, and genes encoding a mutated or chimeric HPPD enzyme.

[0839] The modified plants can be modified so as to have improved abiotic stress tolerance Examples of genes include those capable of reducing the expression and/or the activity of poly(ADP-ribose) polymerase (PARP) gene, transgenes capable of reducing the expression

and/or the activity of the PARG encoding genes, genes coding for a plant-functional enzyme of the nicotineamide adenine dinucleotide salvage synthesis pathway including nicotinamidase, nicotinate phosphoribosyltransferase, nicotinic acid mononucleotide adenyl transferase, nicotinamide adenine dinucleotide synthetase or nicotine amide phosphorybosyltransferase, enzymes involved in carbohydrate biosynthesis, enzymes involved in the production of polyfructose (e.g., the inulin and levan-type), the production of alpha-1,6 branched alpha-1,4-glucans, the production of alternan, the production of hyaluronan. In some embodiments, genes that improve drought resistance may be introduced to plants. Examples of genes Ubiquitin Protein Ligase protein (UPL) protein (UPL3), DR02, DR03, ABC transporter, and DREB1A.

Nutritionally Improved Plants

[0840] The modified plants can be modified so as to have improved or modified nutritional content or profile. Such plants or products therefrom (such as fruits, nuts, etc.) are also referred to in the art as "functional foods". A functional food is a food that provide a health or other benefit beyond the traditional nutrients or content that is provided by a native or wild-type plant or product therefrom. In certain examples, such plants may provide nutraceuticals foods, e.g., substances that may be considered a food or part of a food and provides health benefits, including the prevention and treatment of disease. The nutraceutical foods may be useful in the prevention and/or treatment of diseases in animals and humans, e.g., cancers, diabetes, cardiovascular disease, and hypertension.

[0841] An improved plant may naturally produce one or more desired compounds and the modification may enhance the level or activity or quality of the compounds. In some cases, the improved plant may not naturally produce the compound(s), while the modification enables the plant to produce such compound(s). In some cases, the compositions, systems, and methods used to modify the endogenous synthesis of these compounds indirectly, e.g. by modifying one or more transcription factors that controls the metabolism of this compound.

[0842] Examples of nutritionally improved plants include plants comprising modified protein quality, content and/or amino acid composition, essential amino acid contents, oils and fatty acids, carbohydrates, vitamins and carotenoids, functional secondary metabolites, and minerals. In some examples, the improved plants may comprise or produce compounds with health benefits. Examples of nutritionally improved plants include those described in Newell-McGloughlin, Plant Physiology, July 2008, Vol. 147, pp. 939–953.

[0843] Examples of compounds that can be produced include carotenoids (e.g., α-Carotene or β-Carotene), lutein, lycopene, Zeaxanthin, Dietary fiber (e.g., insoluble fibers, β-Glucan, soluble fibers, fatty acids (e.g., ω-3 fatty acids, Conjugated linoleic acid, and GLA), Flavonoids (e.g., Hydroxycinnamates, flavonols, catechins and tannins), Glucosinolates, indoles, isothiocyanates (e.g., Sulforaphane), Phenolics (e.g., stilbenes, caffeic acid and ferulic acid, epicatechin), Plant stanols/sterols, Fructans, inulins, fructo-oligosaccharides, Saponins, Soybean proteins, Phytoestrogens (e.g., isoflavones, lignans), Sulfides and thiols such as diallyl sulphide, Allyl methyl trisulfide, dithiolthiones, Tannins, such as proanthocyanidins, or any combination thereof.

[0844] The compositions, systems, and methods may also be used to modify protein/starch functionality, shelf life, taste/aesthetics, fiber quality, and allergen, antinutrient, and toxin reduction traits.

[0845] Examples of genes and nucleic acids that can be modified to introduce the traits include stearyl-ACP desaturase, DNA associated with the single allele which may be responsible for maize mutants characterized by low levels of phytic acid, Tf RAP2.2 and its interacting partner SINAT2, Tf Dof1, and DOF Tf AtDof1.1 (OBP2).

Polyploid Plants

[0846] In some embodiments, modified polyploid plants can be generated using the engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom. The Polyploid plants carry duplicate copies of their genomes (e.g. as many as six, such as in wheat). In some cases, the compositions, systems, and methods may be can be multiplexed to affect all copies of a gene, or to target dozens of genes at once. For instance, the compositions, systems, and methods may be used to simultaneously ensure a loss of function mutation in different genes responsible for suppressing defenses against a disease. The modification may be simultaneous suppression the expression of the TaMLO-Al, TaMLO-Bl and TaMLO-Dl nucleic acid sequence in a wheat plant cell and regenerating a wheat plant therefrom, in order to ensure that the wheat plant is resistant to powdery mildew (e.g., as described in WO2015109752).

Controlled Fruit Ripening

[0847] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to modify plants such that they have altered fruit ripening as compared to non-engineered plants. Such plants can be used to decrease losses due to overripe

fruits. Ripening is a normal phase in the maturation process of fruits and vegetables. Only a few days after it starts it may render a fruit or vegetable inedible, which can bring significant losses to both farmers and consumers.

[0848] In some embodiments, the compositions, systems, and methods are used to reduce ethylene production. In some examples, the compositions, systems, and methods may be used to suppress the expression and/or activity of ACC synthase, insert a ACC deaminase gene or a functional fragment thereof, insert a SAM hydrolase gene or functional fragment thereof, suppress ACC oxidase gene expression

[0849] Alternatively or additionally, the compositions, systems, and methods may be used to modify ethylene receptors (e.g., suppressing ETR1) and/or Polygalacturonase (PG). Suppression of a gene may be achieved by introducing a mutation, an antisense sequence, and/or a truncated copy of the gene to the genome.

Increased Plant Storage Life

[0850] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to modify plants such that they contain modified genes involved in the production of compounds which affect storage life of the plant or plant part. The modification may be in a gene that prevents the accumulation of reducing sugars in potato tubers. Upon high-temperature processing, these reducing sugars react with free amino acids, resulting in brown, bitter-tasting products and elevated levels of acrylamide, which is a potential carcinogen. In particular embodiments, the methods provided herein are used to reduce or inhibit expression of the vacuolar invertase gene (VInv), which encodes a protein that breaks down sucrose to glucose and fructose.

Plants with Reduced Allergens

[0851] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to modify plants such that they reduced or eliminated amounts of one or more allergens. To this end, the compositions, systems, formulations, vesicles, cells, methods and the like of the present invention can be used to identify and modify (e.g., suppress) one or more genes responsible for the production of plant allergens. Examples of such genes include Lol p5, as well as those in peanuts, soybeans, lentils, peas, lupin, green beans, mung beans, such as those described in Nicolaou et al., Current Opinion in Allergy and Clinical Immunology 2011;11(3):222), which is incorporated by reference herein in its entirety.

Generation of male sterile plants

[0852] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to generate male sterile plants. Hybrid plants typically have advantageous agronomic traits compared to inbred plants. However, for self-pollinating plants, the generation of hybrids can be challenging. In different plant types (e.g., maize and rice), genes have been identified which are important for plant fertility, more particularly male fertility. Plants that are as such genetically altered can be used in hybrid breeding programs.

[0853] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to modify genes involved male fertility, e.g., inactivating (such as by introducing mutations to) genes required for male fertility. Examples of the genes involved in male fertility include cytochrome P450-like gene (MS26) or the meganuclease gene (MS45), and those described in Wan X et al., Mol Plant. 2019 Mar 4;12(3):321-342; and Kim YJ, et al., Trends Plant Sci. 2018 Jan;23(1):53-65.

Increasing the fertility stage in plants

[0854] In some embodiments, the engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to prolong the fertility stage of a plant such as of a rice. For instance, a rice fertility stage gene such as Ehd3 can be targeted in order to generate a mutation in the gene and plantlets can be selected for a prolonged regeneration plant fertility stage.

Production of early yield of products

[0855] In some embodiments, the engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to produce early yield of the product. For example, flowering process may be modulated, e.g., by mutating flowering repressor gene such as SP5G. Examples of such approaches include those described in Soyk S, et al., Nat Genet. 2017 Jan;49(1):162-168.

Oil and biofuel production

[0856] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to generate plants for oil and biofuel production. Biofuels include fuels made from plant and plant-derived resources. Biofuels may be extracted from organic matter whose energy has been obtained through a process of carbon fixation or are made through the use or conversion of biomass. This biomass can be used directly for biofuels or can be converted to convenient energy containing substances by thermal conversion,

chemical conversion, and biochemical conversion. This biomass conversion can result in fuel in solid, liquid, or gas form. Biofuels include bioethanol and biodiesel. Bioethanol can be produced by the sugar fermentation process of cellulose (starch), which may be derived from maize and sugar cane. Biodiesel can be produced from oil crops such as rapeseed, palm, and soybean. Biofuels can be used for transportation.

Generation of plants for production of vegetable oils and biofuels

[0857] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to generate algae (e.g., diatom) and other plants (e.g., grapes) that express or overexpress high levels of oil or biofuels.

[0858] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to modify genes involved in the modification of the quantity of lipids and/or the quality of the lipids. Examples of such genes include those involved in the pathways of fatty acid synthesis, e.g., acetyl-CoA carboxylase, fatty acid synthase, 3-ketoacyl_acyl- carrier protein synthase III, glycerol-3-phospate deshydrogenase (G3PDH), Enoyl-acyl carrier protein reductase (Enoyl-ACP-reductase), glycerol-3-phosphate acyltransferase, lysophosphatidic acyl transferase or diacylglycerol acyltransferase, phospholipid:diacylglycerol acyltransferase, phoshatidate phosphatase, fatty acid thioesterase such as palmitoyi protein thioesterase, or malic enzyme activities.

[0859] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to generate plants that that have increased lipid accumulation. This can be achieved by targeting genes that decrease lipid catabolization. Examples of genes include those involved in the activation of triacylglycerol and free fatty acids, β-oxidation of fatty acids, such as genes of acyl-CoA synthetase, 3-ketoacyl-CoA thiolase, acyl-CoA oxidase activity and phosphoglucomutase.

[0860] In some examples, algae may be modified for production of oil and biofuels, including fatty acids (e.g., fatty esters such as acid methyl esters (FAME) and fatty acid ethyl esters (FAEE)). Examples of methods of modifying microalgae include those described in Stovicek et al. Metab. Eng. Comm., 2015; 2:1; US 8945839; and WO 2015086795.

[0861] In some examples, one or more genes may be introduced (e.g., overexpressed) to the plants (e.g., algae) to produce oils and biofuels (e.g., fatty acids) from a carbon source (e.g., alcohol). Examples of the genes include genes encoding acyl-CoA synthases, ester synthases, thioesterases (e.g., tesA, 'tesA, tesB, fatB, fatB2, fatB3, fatAl, or fatA), acyl-CoA synthases

(e.g., fadD, JadK, BH3103, pfl-4354, EAV15023, fadDl, fadD2, RPC_4074,fadDD35, fadDD22, faa39), ester synthases (e.g., synthase/acyl-CoA:diacylglycerl acyltransferase from *Simmondsia chinensis*, *Acinetobacter sp. ADP*, *Alcanivorax borkumensis*, *Pseudomonas aeruginosa*, *Fundibacter jadensis*, *Arabidopsis thaliana*, or *Alkaligenes eutrophus*, or variants thereof).

[0862] In some embodiments, one or more genes in the plants (e.g., algae) may be inactivated (e.g., expression of the genes is decreased). For examples, one or more mutations may be introduced to the genes. Examples of such genes include genes encoding acyl-CoA dehydrogenases (e.g., fade), outer membrane protein receptors, and transcriptional regulator (e.g., repressor) of fatty acid biosynthesis (e.g., fabR), pyruvate formate lyases (e.g., pflB), lactate dehydrogenases (e.g., IdhA).

Organic acid production

[0863] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to produce plants that produce or have increased production of organic acids such as lactic acid. The plants may produce organic acids using sugars, pentose or hexose sugars. To this end, one or more genes may be introduced (e.g., and overexpressed) in the plants. An example of such genes include LDH gene.

[0864] In some examples, one or more genes may be inactivated (e.g., expression of the genes is decreased). For examples, one or more mutations may be introduced to the genes. The genes may include those encoding proteins involved an endogenous metabolic pathway which produces a metabolite other than the organic acid of interest and/or wherein the endogenous metabolic pathway consumes the organic acid.

[0865] Examples of genes that can be modified or introduced include those encoding pyruvate decarboxylases (pdc), fumarate reductases, alcohol dehydrogenases (adh), acetaldehyde dehydrogenases, phosphoenolpyruvate carboxylases (ppc), D-lactate dehydrogenases (d-ldh), L-lactate dehydrogenases (l-ldh), lactate 2-monooxygenases, lactate dehydrogenase, cytochrome-dependent lactate dehydrogenases (e.g., cytochrome B2-dependent L-lactate dehydrogenases).

Enhancing plant properties for biofuel production

[0866] The engineered compositions, systems, formulations, vesicles, cells, and/or cargos delivered therefrom can be used to alter the properties of the cell wall of plants to facilitate access by key hydrolyzing agents for a more efficient release of sugars for fermentation. By

reducing the proportion of lignin in a plant the proportion of cellulose can be increased. In particular embodiments, lignin biosynthesis may be downregulated in the plant so as to increase fermentable carbohydrates.

[0867] In some examples, one or more lignin biosynthesis genes may be down regulated. Examples of such genes include 4-coumarate 3-hydroxylases (C3H), phenylalanine ammonia-lyases (PAL), cinnamate 4-hydroxylases (C4H), hydroxycinnamoyl transferases (HCT), caffeic acid O-methyltransferases (COMT), caffeoyl CoA 3-O-methyltransferases (CCoAOMT), ferulate 5- hydroxylases (F5H), cinnamyl alcohol dehydrogenases (CAD), cinnamoyl CoA-reductases (CCR), 4- coumarate-CoA ligases (4CL), monolignol-lignin-specific glycosyltransferases, and aldehyde dehydrogenases (ALDH), and those described in WO 2008064289.

[0868] In some examples, plant mass that produces lower level of acetic acid during fermentation may be reduced. To this end, genes involved in polysaccharide acetylation (e.g., Cas1L and those described in WO 2010096488) may be inactivated.

Other microorganisms for oils and biofuel production

[0869] In some embodiments, microorganisms other than plants may be used for production of oils and biofuels using the compositions, systems, and methods herein. Examples of the microorganisms include those of the genus of *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Synechococcus*, *Synechoystis*, *Pseudomonas*, *Aspergillus*, *Trichoderma, Neurospora, Fusarium, Humicola, Rhizomucor, Kluyveromyces, Pichia, Mucor, Myceliophtora, Penicillium, Phanerochaete, Pleurotus, Trametes, Chrysosporium, Saccharomyces, Stenotrophamonas, Schizosaccharomyces, Yarrowia, or Streptomyces.*

[0870] In some embodiments, the compositions, systems, and methods may be used for generating modified fungi for biofuel and material productions. For instance, the modified fungi for production of biofuel or biopolymers from fermentable sugars and optionally to be able to degrade plant-derived lignocellulose derived from agricultural waste as a source of fermentable sugars. Foreign genes required for biofuel production and synthesis may be introduced into fungi. In some examples, the genes may encode enzymes involved in the conversion of pyruvate to ethanol or another product of interest, degrade cellulose (e.g., cellulase), endogenous metabolic pathways which compete with the biofuel production pathway.

[0871] In some examples, the compositions, systems, and methods may be used for generating and/or selecting yeast strains with improved xylose or cellobiose utilization, isoprenoid biosynthesis, and/or lactic acid production. One or more genes involved in the metabolism and synthesis of these compounds may be modified and/or introduced to yeast cells. Examples of the methods and genes include lactate dehydrogenase, PDC1 and PDC5, and those described in Ha, S.J., et al. (2011) Proc. Natl. Acad. Sci. USA 108(2):504-9 and Galazka, J.M., et al. (2010) Science 330(6000):84-6; Jakočiūnas T et al., Metab Eng. 2015 Mar;28:213-222; Stovicek V, et al., FEMS Yeast Res. 2017 Aug 1;17(5).

Engineered Non-animal organisms

[0872] The compositions, systems, and methods may be used to study and modify non-human animals, e.g., introducing desirable traits and disease resilience, treating diseases, facilitating breeding, etc. In some embodiments, the compositions, systems, and methods may be used to improve breeding and introducing desired traits, e.g., increasing the frequency of trait-associated alleles, introgression of alleles from other breeds/species without linkage drag, and creation of de novo favorable alleles. Genes and other genetic elements that can be targeted may be screened and identified. Examples of application and approaches include those described in Tait-Burkard C, et al., Livestock 2.0 - genome editing for fitter, healthier, and more productive farmed animals. Genome Biol. 2018 Nov 26;19(1):204; Lillico S, Agricultural applications of genome editing in farmed animals. Transgenic Res. 2019 Aug;28(Suppl 2):57-60; Houston RD, et al., Harnessing genomics to fast-track genetic improvement in aquaculture. Nat Rev Genet. 2020 Apr 16. doi: 10.1038/s41576-020-0227-y, which are incorporated herein by reference in their entireties. Applications described in other sections such as therapeutic, diagnostic, etc. can also be used on the animals herein.

[0873] The compositions, systems, and methods may be used on animals such as fish, amphibians, reptiles, mammals, and birds. The animals may be farm and agriculture animals, or pets. Examples of farm and agriculture animals include horses, goats, sheep, swine, cattle, llamas, alpacas, and birds, e.g., chickens, turkeys, ducks, and geese. The animals may be a non-human primate, e.g., baboons, capuchin monkeys, chimpanzees, lemurs, macaques, marmosets, tamarins, spider monkeys, squirrel monkeys, and vervet monkeys. Examples of pets include dogs, cats, horses, wolfs, rabbits, ferrets, gerbils, hamsters, chinchillas, fancy rats, guinea pigs, canaries, parakeets, and parrots.

[0874] In some embodiments, one or more genes may be introduced (e.g., overexpressed) in the animals to obtain or enhance one or more desired traits. Growth hormones, insulin-like growth factors (IGF-1) may be introduced to increase the growth of the animals, e.g., pigs or salmon (such as described in Pursel VG et al., J Reprod Fertil Suppl. 1990;40:235-45; Waltz E, Nature. 2017;548:148). Fat-1 gene (e.g., from C elegans) may be introduced for production of larger ratio of n-3 to n-6 fatty acids may be induced, e.g. in pigs (such as described in Li M, et al., Genetics. 2018;8:1747–54). Phytase (e.g., from E coli) xylanase (e.g., from Aspergillus niger), beta-glucanase (e.g., from bacillus lichenformis) may be introduced to reduce the environmental impact through phosphorous and nitrogen release reduction, e.g. in pigs (such as described in Golovan SP, et al., Nat Biotechnol. 2001;19:741–5; Zhang X et al., elife. 2018). shRNA decoy may be introduced to induce avian influenza resilience e.g. in chicken (such as described in Lyall et al., Science. 2011;331:223–6). Lysozyme or lysostaphin may be introduced to induce mastitis resilience e.g., in goat and cow (such as described in Maga EA et al., oodborne Pathog Dis. 2006;3:384–92; Wall RJ, et al., Nat Biotechnol. 2005;23:445–51). Histone deacetylase such as HDAC6 may be introduced to induce PRRSV resilience, e.g., in pig (such as described in Lu T., et al., PLoS One. 2017;12:e0169317). CD163 may be modified (e.g., inactivated or removed) to introduce PRRSV resilience in pigs (such as described in Prather RS et al.., Sci Rep. 2017 Oct 17;7(1):13371). Similar approaches may be used to inhibit or remove viruses and bacteria (e.g., Swine Influenza Virus (SIV) strains which include influenza C and the subtypes of influenza A known as H1N1, H1N2, H2N1, H3N1, H3N2, and H2N3, as well as pneumonia, meningitis and oedema) that may be transmitted from animals to humans.

[0875] In some embodiments, one or more genes may be modified or edited for disease resistance and production traits. Myostatin (e.g., GDF8) may be modified to increase muscle growth, e.g., in cow, sheep, goat, catfish, and pig (such as described in Crispo M et al., PLoS One. 2015;10:e0136690; Wang X, et al., Anim Genet. 2018;49:43–51; Khalil K, et al., Sci Rep. 2017;7:7301; Kang J-D, et al., RSC Adv. 2017;7:12541–9). Pc POLLED may be modified to induce horlessness, e.g., in cow (such as described in Carlson DF et al., Nat Biotechnol. 2016;34:479–81). KISS1R may be modified to induce boretaint (hormone release during sexual maturity leading to undesired meat taste), e.g., in pigs. Dead end protein (dnd) may be modified to induce sterility, e.g., in salmon (such as described in Wargelius A, et al., Sci Rep. 2016;6:21284). Nano2 and DDX may be modified to induce sterility (e.g., in surrogate hosts),

e.g., in pigs and chicken (such as described Park K-E, et al., Sci Rep. 2017;7:40176; Taylor L et al., Development. 2017;144:928–34). CD163 may be modified to induce PRRSV resistance, e.g., in pigs (such as described in Whitworth KM, et al., Nat Biotechnol. 2015;34:20–2). RELA may be modified to induce ASFV resilience, e.g., in pigs (such as described in Lillico SG, et al., Sci Rep. 2016;6:21645). CD18 may be modified to induce Mannheimia (Pasteurella) haemolytica resilience, e.g., in cows (such as described in Shanthalingam S, et al., roc Natl Acad Sci U S A. 2016;113:13186–90). NRAMP1 may be modified to induce tuberculosis resilience, e.g., in cows (such as described in Gao Y et al., Genome Biol. 2017;18:13). Endogenous retrovirus genes may be modified or removed for xenotransplantation such as described in Yang L, et al. Science. 2015;350:1101–4; Niu D et al., Science. 2017;357:1303–7). Negative regulators of muscle mass (e.g., Myostatin) may be modified (e.g., inactivated) to increase muscle mass, e.g., in dogs (as described in Zou Q et al., J Mol Cell Biol. 2015 Dec;7(6):580-3).

[0876] Animals such as pigs with severe combined immunodeficiency (SCID) may generated (e.g., by modifying RAG2) to provide useful models for regenerative medicine, xenotransplantation (discussed also elsewhere herein), and tumor development. Examples of methods and approaches include those described Lee K, et al., Proc Natl Acad Sci U S A. 2014 May 20;111(20):7260-5; and Schomberg et al. FASEB Journal, April 2016; 30(1):Suppl 571.1.

[0877] SNPs in the animals may be modified. Examples of methods and approaches include those described Tan W. et al., Proc Natl Acad Sci U S A. 2013 Oct 8;110(41):16526-31; Mali P, et al., Science. 2013 Feb 15;339(6121):823-6.

[0878] Stem cells (e.g., induced pluripotent stem cells) may be modified and differentiated into desired progeny cells, e.g., as described in Heo YT et al., Stem Cells Dev. 2015 Feb 1;24(3):393-402.

[0879] Profile analysis (such as Igenity) may be performed on animals to screen and identify genetic variations related to economic traits. The genetic variations may be modified to introduce or improve the traits, such as carcass composition, carcass quality, maternal and reproductive traits and average daily gain.

Diagnostics

[0880] The engineered compositions, systems, delivery vesicles, cells, and the like of the present invention can be used in one or more diagnostics or screening assays. Exemplary assays are described here and elsewhere herein. For example, single and co-culture systems expressing

one or more of the engineered systems and/or containing engineered delivery vesicles can be used to screen therapeutic candidates. In some embodiments, the cargo delivered by the engineered system and/or vesicles produced therefrom can be a test therapeutic candidate. In some of these embodiments, cells of the assay can produce engineered delivery vesicles containing the test candidate cargo which can be delivered to one or more other cells of the assay. The cells can then be evaluated for genotypic, phenotypic, and/or functional changes by a suitable method, which will be dependent upon the characteristic being evaluated, and test candidates with the desirable response can be selected for further evaluation and/or treatment. In some embodiments, cells are obtained from a subject for which a therapy is being developed for. In some embodiments, the cells are obtained from a tumor of a subject for which a therapy is being developed for.

Co-Culture Systems

[0881] Described in several exemplary embodiments herein are co-culture systems comprising two or more cell types, where at least one, all, or a sub-combination of cell types comprise an engineered delivery system as described in greater detail elsewhere herein, wherein the engineered delivery system is capable of generating one or more delivery vesicles. In general, a co-culture as the term is used herein, is a cell culture system in which two or more different populations of cells are grown with some degree of contact between the two or more different populations. Cell populations within the co-culture can differ in cell type, state, origin, lineage, passage, species of origin, and the like.

[0882] In some embodiments, the engineered delivery system in a given cell population within the co-culture includes a cargo and thus can produce a delivery vesicle comprising a cargo. The delivery vesicle can be released by the cell which produced it into the co-culture where it can then deliver its cargo to another cell, such as a cell of another cell population within the co-culture. This can drive, for example, the development of synthetic interactions between cells of the co-culture, formation of synthetic ecologies, or other complex interactions within the co-culture.

[0883] The co-cultures can be used for studying and/or engineering complex multicellular populations and synthetic systems. In some embodiments, the co-cultures described herein can be configured and used for culturing one or more cell populations, such as traditionally difficult to culture cell populations. In some embodiments, the co-cultures described herein can be configured and used for establishing synthetic interactions between populations. In some

embodiments, the co-cultures described herein can be configured and used for studying natural interactions such as infections and creating model systems and biomimetic environments of natural systems, such as artificial tissues or organs. Such systems can be used in screening assays to study complex reactions to agents of interest, such as therapeutic agents, pathogens, and/or toxins. Additional applications for the co-cultures containing at least one cell population containing an engineered delivery system and capable of generating delivery vesicles therefrom are described in e.g., Goers et al., 2014. J R. Soc. Interface 11:20140065; http://dx.doi.org/10.1098/rsif.20140065.

[0884] Further embodiments are illustrated in the following Examples which are given for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLES

Example 1 – Improved Yield of Lentiviruses from 293T cells

[0885] Lentiviral transfer vectors encoding a gene of interest and vectors encoding effector genes were generated. Effector genes were encoded in plasmid vectors according to SEQ ID NO:1 (encoding PEG10) or SEQ ID NO:2 (encoding RTL1). Transfer vectors and plasmids encoding PEG10, RTL1, or a combination thereof were transfected into 293T cells along with lentiviral envelope and packaging constructs. Three days later, the supernatant was collected and subjected to qPCR to quantify viral RNA copies. Plaque assays were carried out to determine viral titer (**FIG. 1A**). qPCR results showed that addition of plasmids encoding PEG10 led to an 8.3-fold increase in lentiviral RNA copy numbers compared to the control. Addition of RTL1 led to a 19.5-fold increase in lentiviral RNA copy numbers compared to the control (**FIG. 1B**).

***

[0886] Various modifications and variations of the described methods, pharmaceutical compositions, and kits of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments, it will be understood that it is capable of further modifications and that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the art are intended to be within the scope of the invention. This application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures

from the present disclosure come within known customary practice within the art to which the invention pertains and may be applied to the essential features herein before set forth.

CLAIMS

What is claimed is:

1. An engineered retroviral delivery vesicle generation system comprising:
   a. one or more polynucleotides encoding one or more effectors;
   b. one or more polynucleotides encoding a cargo; and
   c. one or more polynucleotides encoding one or more packaging elements, one or more vesicle elements, or both.

2. The system of claim 1, wherein the system is capable of generating 1 to 50 or more fold more vesicles as compared to a system lacking the one or more polynucleotides encoding one or more effectors.

3. The system of any one of claims 1-2, wherein the one or more effectors are retrotransposon-derived genes.

4. The system of claim 1, wherein the one or more effectors are PEG10, RTL1, or both.

5. The system of claim 4, wherein the one or more polynucleotides encoding PEG10 is 80 percent to 100 percent identical to SEQ ID NO: 1.

6. The system of claim 4, wherein the one or more polynucleotides encoding RTL1 is 80 percent to 100 percent identical to SEQ ID NO: 2.

7. The system of claim 6, wherein the retroviral system is a lentiviral system.

8. The system of claim 1, wherein (a), (b), and (c) are included in one or more vectors comprising one or more regulatory elements, and wherein each of the one or more polynucleotides of (a), (b), and (c) are optionally operably coupled a regulatory element.

9. The system of claim 8, wherein (a), (b), (c), or any combination thereof are included on the same vector, are included in different vectors, or any permissible combination thereof.

10. An engineered retroviral delivery vesicle generation system comprising:
a. one or more polynucleotides encoding one or more retroviral polypeptides capable of forming a delivery vesicle and encapsulating one or more cargos therein, wherein at least one of the one or more polypeptides is an effector; and
b. one or more cargos, wherein the one or more cargos are optionally polynucleotide cargos.

11. The system of claim 10, wherein the system is capable of generating 1 to 50 or more fold more delivery vesicles containing one or more cargos as compared to a system lacking the one or more polynucleotides encoding one or more effectors.

12. The system of claim 10, wherein the effector is a retrotransposon-derived effector.

13. The system of claim 10, wherein the effector is PEG10, RTL1, or both.

14. The system of claim 13, wherein PEG10 is encoded by a polynucleotide that is 80 to 100 percent identical to SEQ ID NO: 1.

15. The system of claim 13, wherein RTL1 is encoded by a polynucleotide that is 80 to 100 percent identical to SEQ ID NO: 2.

16. The system of claim 10, wherein the one or more retroviral polypeptides are one or more lentiviral polypeptides.

17. The system of claim 10, wherein the one or more retroviral polypeptides comprises one or more packaging elements, vesicle elements, or both.

18. The system of claim 10, wherein (a), (b), optionally (c), or a combination thereof are included in one or more vectors comprising one or more regulatory elements, and wherein each of the one or more polynucleotides of (a) and (b)are optionally operably coupled a regulatory element.

19. The system of claim 10, wherein (a), (b), optionally (c), or a combination thereof are included on the same vector or are included on different vectors.

20. The system of claim 10, wherein the one or more cargos comprise one or more packaging elements.

21. A method of generating delivery vesicles loaded with one or more cargos comprising: expressing one or more components of the delivery vesicle systems of any one of claims 1-20 in one or more suitable bioreactors under conditions such that vesicles are formed wherein one or more of the vesicles contains one or more cargos.

22. The method of claim 21, wherein the one or more suitable bioreactors are cells.

23. A delivery vesicle comprising one or more cargos generated according to the method of claim 22.

24. The delivery vesicle of claim 23, wherein the delivery vesicle is a retroviral particle.

25. The delivery vesicle of claim 24, wherein the delivery vesicle is a lentiviral particle.

26. The delivery vesicle one of claims 22-25, wherein the one or more cargos is/are polynucleotide cargos.

27. A co-culture system comprising two or more cell types wherein at least one, all, or a sub-combination of cell types comprise an engineered delivery vesicle generation system of any one of claims 1-20.

28. A method of cellular delivery of one or more cargos comprising:

a. delivering an engineered retroviral delivery vesicle generation system to any one of claims 1-20 to a donor cell type, wherein expression of the engineered retroviral delivery vesicle generation system in the donor cell type

results in the generation of one or more delivery vesicles comprising one or more cargos;

b. delivery to or uptake of one or more delivery vesicles generated in (a) to or by a recipient cell.

29. The method of claim 28, wherein the method occurs *in vivo, in vitro*, or *ex vivo*.

30. The method of any one of claims 28-29, wherein the recipient cell is diseased or pathogenic.

31. The method of claim 30, wherein the recipient cell is a eukaryote or a prokaryote.

32. A cell comprising and capable of expressing an engineered delivery system as in any one of claims 1-20.

33. A pharmaceutical formulation comprising:

a. an engineered delivery system as in any one of claims 1-20;

b. a delivery vesicle as in any one of claims 23-26;

c. a cell as in claim 32; or

d. any combination thereof; and

a pharmaceutically acceptable carrier.

34. A kit comprising:

a. an engineered delivery system as in any one of claims 1-20;

b. a delivery vesicle as in any one of claims 23-26;

c. a cell as in claim 32 or population thereof;

d. a pharmaceutical formulation as in claim 33; or

e. a combination thereof.

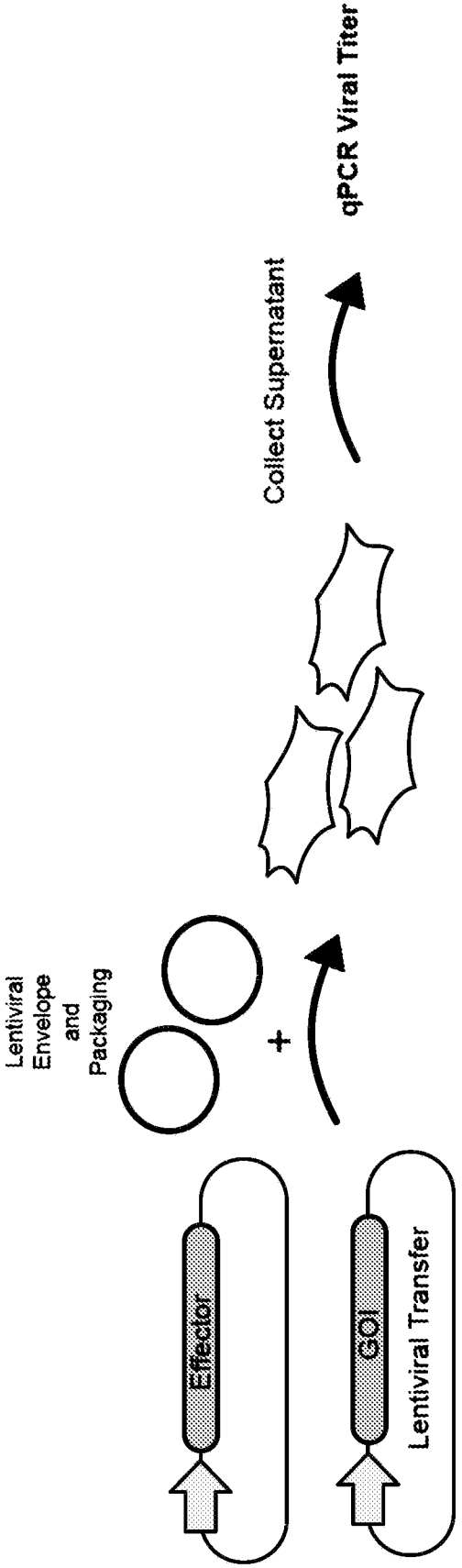
Effector
GOI
Lentiviral Transfer
Lentiviral Envelope and Packaging
+
Collect Supernatant
qPCR Viral Titer

FIG. 1A

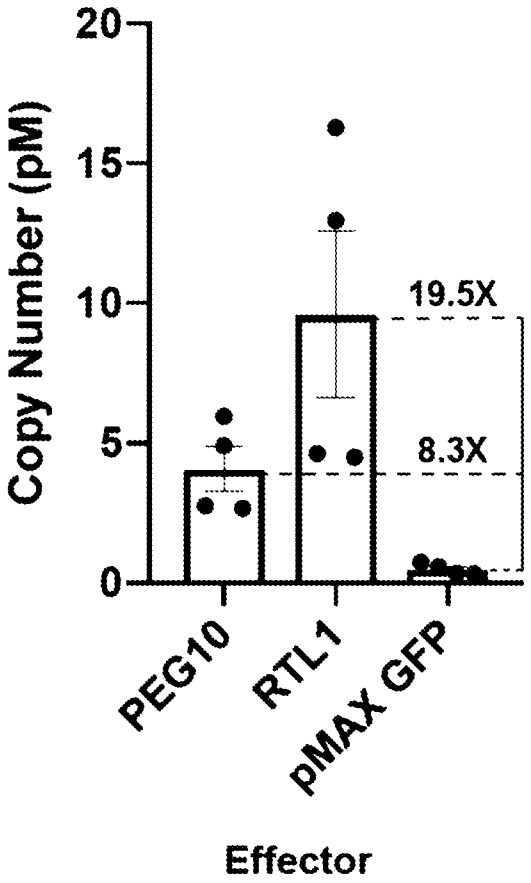
20
15
10
5
0
Copy Number (pM)
19.5X
8.3X
PEG10
RTL1
pMAX GFP
Effector

FIG. 1B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2021/023318

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C07K 14/005; C07K 14/47; C12N 7/00; C12N 9/50; C12N 15/11 (2021.01)
CPC - C07K 2319/40; C12N 15/11; C12N 15/113; C12N 15/1137; C12N 2310/20; C12N 2320/12 (2021.05)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
see Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
see Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
see Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	WO 2020/012335 A1 (ALIA THERAPEUTICS S.R.L.) 16 January 2020 (16.01.2020) entire document	1-3, 8-12, 16-20 ---------- 4, 13
Y	ABED et al. "The Gag protein PEG10 binds to RNA and regulates trophoblast stem cell lineage specification," PLoS One, 05 April 2019 (05.04,2019), Vol. 14, Iss. 4, Pgs. 1-18. entire document	4, 13
A	US 9,328,146 B2 (CHARNEAU et al) 03 May 2016 (03.05.2016) entire document	1-20
A	WO 2019/199689 A1 (THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA) 17 October 2019 (17.10.2019) entire document	1-20
P, A	US 2020/0347100 A1 (THE BROAD INSTITUTE, INC. et al) 05 November 2020 (05.11.2020) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 June 2021	JUL 20 2021
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Harry Kim Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2021/023318

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs:1-2 were searched.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2021/023318

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 21-34
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.