



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2021/04/30
(87) Date publication PCT/PCT Publication Date: 2021/11/04
(85) Entrée phase nationale/National Entry: 2022/10/21
(86) N° demande PCT/PCT Application No.: US 2021/030298
(87) N° publication PCT/PCT Publication No.: 2021/222831
(30) Priorités/Priorities: 2020/05/01 (US63/019,221);
2020/08/05 (US63/061,517)

(51) Cl.Int./Int.Cl. *C07K 11/00* (2006.01),
C12N 15/86 (2006.01), *C12N 7/00* (2006.01),
C12N 7/01 (2006.01)
(71) Demandeurs/Applicants:
THE BROAD INSTITUTE, INC., US;
PRESIDENT AND FELLOWS OF HARVARD COLLEGE,
US;
MASSACHUSETTS INSTITUTE OF TECHNOLOGY, US
(72) Inventeurs/Inventors:
SABETI, PARDIS, US;
TABEBORDBAR, MOHAMMADSHARIF, US;
YE, SIMON, US
(74) Agent: ROBIC

(54) Titre : COMPOSITIONS DE SYSTEME NERVEUX CENTRAL GENETIQUEMENT MODIFIEES
(54) Title: ENGINEERED CENTRAL NERVOUS SYSTEM COMPOSITIONS

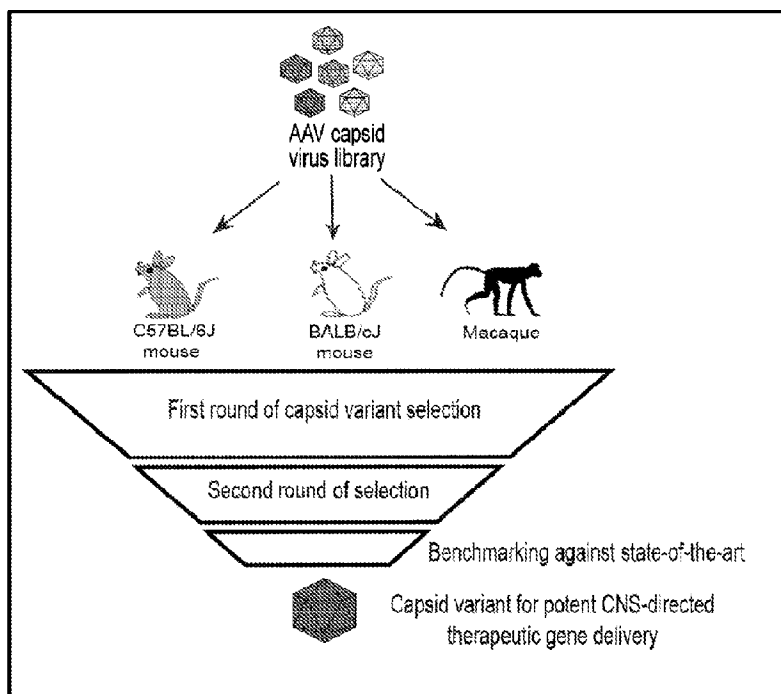


FIG. 6A

(57) **Abrégé/Abstract:**

Described in several exemplary embodiments are compositions including a targeting moiety effective to target a central nervous system cell and formulations thereof. In certain embodiments, the targeting moiety is composed of a n-mer motif, P motif, or both. Also described in certain example embodiments are vector systems configured to generate polypeptides containing the one or more targeting moieties. Also described herein are methods of generating a targeting moiety effective to target a central nervous system cell and using the compositions containing the targeting moieties described herein, such as to deliver a cargo to a subject and/or treat a central nervous system disease, disorder, or system thereof.

Date Submitted: 2022/10/21

CA App. No.: 3176506

Abstract:

Described in several exemplary embodiments are compositions including a targeting moiety effective to target a central nervous system cell and formulations thereof. In certain embodiments, the targeting moiety is composed of a n-mer motif, P motif, or both. Also described in certain example embodiments are vector systems configured to generate polypeptides containing the one or more targeting moieties. Also described herein are methods of generating a targeting moiety effective to target a central nervous system cell and using the compositions containing the targeting moieties described herein, such as to deliver a cargo to a subject and/or treat a central nervous system disease, disorder, or system thereof.

ENGINEERED CENTRAL NERVOUS SYSTEM COMPOSITIONS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 63/019,221, filed May 1, 2020 and U.S. Provisional Application No. 63/061,517, filed August 5, 2020. The entire contents of the above-identified applications are hereby fully incorporated herein by reference.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0002] This application contains a sequence listing filed in electronic form as an ASCII.txt file entitled BROD-5125WP.txt, created on April 30, 2021, and having a size of 144,000 bytes. The content of the sequence listing is incorporated herein in its entirety.

TECHNICAL FIELD

[0003] The subject matter disclosed herein is generally directed to engineered central nervous system targeting compositions including, but not limited to, recombinant adeno-associated virus (AAV) vectors, and systems, compositions, and uses thereof.

BACKGROUND

[0004] Recombinant AAVs (rAAVs) are the most commonly used delivery vehicles for gene therapy and gene editing. Nonetheless, rAAVs that contain natural capsid variants have limited cell tropism. Indeed, rAAVs used today mainly infect the liver after systemic delivery. Further, the transduction efficiency of conventional rAAVs in other cell-types, tissues, and organs by these conventional rAAVs with natural capsid variants is limited. Therefore, AAV-mediated polynucleotide delivery for diseases that affect cells, tissues, and organs other than the liver, such as the central nervous system) typically requires an injection of a large dose of virus (typically about 2×10^{14} vg/kg), which often results in liver toxicity. Furthermore, because large doses are required when using conventional rAAVs, manufacturing sufficient amounts of a therapeutic rAAV needed to dose adult patients is extremely challenging. Additionally, due to differences in gene expression and physiology, mouse and primate models respond differently to viral capsids. Transduction efficiency of different virus particles varies between different species, and as a result, preclinical studies in mice often do not accurately

reflect results in primates, including humans. As such there exists a need for improved rAAVs for use in the treatment of various genetic diseases.

[0005] 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

[0006] Described in certain example embodiments herein are compositions comprising a targeting moiety effective to target a central nervous system (CNS) cell, wherein the targeting moiety comprises one or more P-motifs or wherein the at least one P-motif comprises the amino acid sequence $PX_1QGTX_2RX_n$ (SEQ ID NO: 2), wherein X_1 , X_2 , X_n , are each independently selected from any amino acid and wherein n is 0, 1, 2, 3, 4, 5, 6, or 7, or one or more n -mer inserts selected from the group of any one of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, and 313, and 318-329, or one or more n -mer inserts selected from the group of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, 313, and 318-329 and one or more P-motifs, and optionally a cargo, wherein the cargo is coupled to or is otherwise associated with the targeting moiety.

[0007] In certain example embodiments, the targeting moiety comprises both an n -mer insert and a P-motif and wherein the P-motif is optionally part of or the entirety of the n -mer insert.

[0008] In certain example embodiments, the one or more n -mer inserts, each of the P-motifs, or both are each 3-15 amino acids in length.

[0009] In certain example embodiments, wherein

- a. X_1 is S, T, or A,
- b. X_2 is L, V, F, or I, or
- c. both.

[0010] In certain example embodiments, wherein the n -mer insert and/or P motif is as in **Table 1** (e.g., SEQ ID NOS: 65-199).

[0011] In certain example embodiments, the n -mer insert and/or P motif is as in any one of **Tables 2-3** (e.g., SEQ ID NOS: 200, 202, 204, 206, 208, 210, 212, 214 (**Table 2**) and/or 300, 303, 306, 308, 311, 313 (**Table 3**)).

[0012] In certain example embodiments, the n -mer insert and/or P motif is as in **Table 7** (e.g., SEQ ID NOS: 318-329).

[0013] In certain example embodiments, the *n*-mer insert is immediately preceded by AQ or DG.

[0014] In certain example embodiments, wherein

(a) the *n*-mer insert polypeptide is immediately preceded by AQ and wherein the *n*-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSSG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESGH (SEQ ID NO: 19), or GENSARW (SEQ ID NO: 20); or

(b) the *n*-mer insert polypeptide is immediately preceded by DG and wherein the *n*-mer insert is REQQKLW (SEQ ID NO: 21), ASNPGRW (SEQ ID NO: 22), WTLESGH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25), or RMQRTLY (SEQ ID NO: 26).

[0015] In certain example embodiments, wherein the targeting moiety comprises a polypeptide, a polynucleotide, a lipid, a polymer, a sugar, or a combination thereof.

[0016] In certain example embodiments, the targeting moiety comprises a viral protein.

[0017] In certain example embodiments, the viral protein is a capsid protein.

[0018] In certain example embodiments, the *n*-mer insert(s), is located between two amino acids of the viral protein such that the *n*-mer insert is external to a viral capsid.

[0019] In certain example embodiments, the viral protein is an adeno associated virus (AAV) protein.

[0020] In certain example embodiments, the AAV protein is an AAV capsid protein.

[0021] In certain example embodiments, the one or more *n*-mer inserts and/or P motif are each inserted between any two contiguous amino acids between amino acids independently selected from 262-269, 327-332, 382-386, 452-460, 488-505, 527-539, 545-558, 581-593, 704-714, or any combination thereof in an AAV9 capsid polypeptide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.

[0022] In certain example embodiments, at least one of the one or more *n*-mer inserts is inserted between amino acids 588 and 589 in an AAV9 capsid polypeptide or in an analogous

position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.

[0023] In certain example embodiments, the AAV capsid protein is an engineered AAV capsid protein having reduced or eliminated uptake in a non-CNS cell as compared to a corresponding wild-type AAV capsid polypeptide.

[0024] In certain example embodiments, the non-CNS cell is a liver cell.

[0025] In certain example embodiments, the wild-type capsid polypeptide is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 capsid polypeptide.

[0026] In certain example embodiments, the engineered AAV capsid protein comprises one or more mutations that result in reduced or eliminated uptake in a non-CNS cell.

[0027] In certain example embodiments, the one or more mutations are

- a. in position 267,
- b. in position 269,
- c. in position 504,
- d. in position 505,
- e. in position 590,
- f. or any combination thereof

in the AAV9 capsid protein (SEQ ID NO: 1) or in one or more positions corresponding thereto in a non-AAV9 capsid polypeptide.

[0028] In certain example embodiments, the non-AAV9 capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, or AAV rh.10 capsid polypeptide.

[0029] In certain example embodiments, the mutation in position 267 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X mutation to A, wherein X is any amino acid.

[0030] In certain example embodiments, the mutation in position 269 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is an S or X to T mutation, wherein X is any amino acid.

[0031] In certain example embodiments, the mutation in position 504 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X to A mutation, wherein X is any amino acid.

[0032] In certain example embodiments, the mutation in position 505 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a P or X to A mutation, wherein X is any amino acid.

[0033] In certain example embodiments, the mutation in position 590 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a Q or X to A mutation, wherein X is any amino acid.

[0034] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 267, position 269 or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 267 is a G to A mutation and wherein the mutation at position 269 is an S to T mutation.

[0035] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 590 of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 509 is a Q to A mutation.

[0036] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 504, position 505, or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 504 is a G to A mutation and wherein the mutation at position 505 is a P to A mutation.

[0037] In certain example embodiments, the composition is an engineered viral particle.

[0038] In certain example embodiments, the engineered viral particle is an engineered AAV viral particle.

[0039] In certain example embodiments, the AAV viral particle is an engineered AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 viral particle.

[0040] In certain example embodiments, the optional cargo is capable of treating or preventing a CNS disease or disorder.

[0041] Described in certain example embodiments herein are vector systems that include a vector comprising one or more polynucleotides, wherein at least one of the one or more polynucleotides encodes all or part of a targeting moiety effective to target a central nervous system (CNS) cell, wherein the targeting moiety comprises at least one P-motifs or wherein the at least one P-motif comprises the amino acid sequence $PX_1QGTX_2RX_n$ (SEQ ID NO: 2), wherein X_1 , X_2 , X_n , are each independently selected from any amino acid and wherein n is 0,

1, 2, 3, 4, 5, 6, or 7, or at least one *n*-mer insert selected from the group consisting of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, and 313, and 318-329, or at least one *n*-mer insert selected from the group consisting of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, 313, and 318-329 and at least one P-motif, wherein at least one of the one or more polynucleotides encodes the at least one *n*-mer insert, the at least one P-motif, or both; and optionally, a regulatory element operatively coupled to one or more of the one or more polynucleotides.

[0042] In certain example embodiments, the targeting moiety comprises both an *n*-mer insert and a P-motif and wherein the P-motif is optionally part of or the entirety of the *n*-mer insert.

[0043] In certain example embodiments, the one or more *n*-mer inserts, each of the P-motifs, or both are each 3-15 amino acids in length.

[0044] In certain example embodiments, wherein

- a. X_1 is S, T, or A,
- b. X_2 is L, V, F, or I, or
- c. both.

[0045] In certain example embodiments, the *n*-mer insert and/or P motif is any one in Table 1 (e.g., SEQ ID NOs: 65-199).

[0046] In certain example embodiments, the *n*-mer insert and/or P motif is as in any one of Tables 2-3 (e.g., SEQ ID NOs: 200, 202, 204, 206, 208, 210, 212, 214 (Table 2) and/or 300, 303, 306, 308, 311, 313 (Table 3)).

[0047] In certain example embodiments, the *n*-mer insert and/or P motif is as in Table 7 (e.g., SEQ ID NOS: 318-329).

[0048] In certain example embodiments, the *n*-mer insert is immediately preceded by AQ or DG.

[0049] In certain example embodiments,

(a) the *n*-mer insert polypeptide is immediately preceded by AQ and wherein the *n*-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSGG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO:

17), QIAHGYST (SEQ ID NO: 18), WTLESGH (SEQ ID NO: 19), or GENSARW (SEQ ID NO: 20); or

(b) the *n*-mer insert polypeptide is immediately preceded by DG and wherein the *n*-mer insert is REQQKLW (SEQ ID NO: 21), ASNPGRW (SEQ ID NO: 22), WTLESGH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25), or RMQRTLY (SEQ ID NO: 26).

[0050] In certain example embodiments, the vector system further includes a cargo.

[0051] In certain example embodiments, the cargo is a cargo polynucleotide and is optionally operatively coupled to one or more of the one or more polynucleotides encoding the targeting moiety.

[0052] In certain example embodiments, the vector system is capable of producing virus particles, virus particles that contain the cargo, or both.

[0053] In certain example embodiments, the vector system is capable of producing a polypeptide comprising one or more of the targeting moieties.

[0054] In certain example embodiments, the polypeptide is a viral polypeptide.

[0055] In certain example embodiments, the viral polypeptide is a capsid polypeptide.

[0056] In certain example embodiments, the capsid polypeptide is an adeno associated virus (AAV) capsid polypeptide.

[0057] In certain example embodiments, the virus particles are AAV virus particles.

[0058] In certain example embodiments, the AAV virus particles or AAV capsid polypeptide are engineered AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 viral particles or polypeptides.

[0059] In certain example embodiments, the at least one polynucleotide encoding the at least one *n*-mer inserts is inserted between two codons corresponding to two amino acids of a viral polypeptide such that the *n*-mer insert(s) is external to a viral capsid of the virus particles.

[0060] In certain example embodiments, the at least one polynucleotide is inserted between two codons corresponding to any two contiguous amino acids between amino acids 262-269, 327-332, 382-386, 452-460, 488-505, 527-539, 545-558, 581-593, 704-714, or any combination thereof in an AAV9 capsid polypeptide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.

[0061] In certain example embodiments, the at least one polynucleotide is inserted between the codons corresponding to amino acid 588 and 589 in the AAV9 capsid polynucleotide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.

[0062] In certain example embodiments, the AAV capsid protein is an engineered AAV capsid protein having reduced or eliminated uptake in a non-CNS cell as compared to a corresponding wild-type AAV capsid polypeptide.

[0063] In certain example embodiments, the non-CNS cell is a liver cell.

[0064] In certain example embodiments, the wild-type capsid polypeptide is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 capsid polypeptide.

[0065] In certain example embodiments, the engineered AAV capsid protein comprises one or more mutations that result in reduced or eliminated uptake in a non-CNS cell.

[0066] In certain example embodiments, the one or more mutations are

- a. in position 267,
- b. in position 269,
- c. in position 504,
- d. in position 505,
- e. in position 590,
- f. or any combination thereof

in the AAV9 capsid protein (SEQ ID NO: 1) or in one or more positions corresponding thereto in a non-AAV9 capsid polypeptide.

[0067] In certain example embodiments, the non-AAV9 capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, or AAV rh.10 capsid polypeptide.

[0068] In certain example embodiments, the mutation in position 267 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X mutation to A, wherein X is any amino acid.

[0069] In certain example embodiments, the mutation in position 269 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is an S or X to T mutation, wherein X is any amino acid.

[0070] In certain example embodiments, the mutation in position 504 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X to A mutation, wherein X is any amino acid.

[0071] In certain example embodiments, the mutation in position 505 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a P or X to A mutation, wherein X is any amino acid.

[0072] In certain example embodiments, the mutation in position 590 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a Q or X to A mutation, wherein X is any amino acid.

[0073] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 267, position 269 or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 267 is a G to A mutation and wherein the mutation at position 269 is an S to T mutation.

[0074] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 590 of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 509 is a Q to A mutation.

[0075] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 504, position 505, or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 504 is a G to A mutation and wherein the mutation at position 505 is a P to A mutation.

[0076] In certain example embodiments, the vector comprising the one or more polynucleotides does not comprise splice regulatory elements.

[0077] In certain example embodiments, the vector system further comprises a polynucleotide that encodes a viral rep protein.

[0078] In certain example embodiments, the viral rep protein is an AAV rep protein.

[0079] In certain example embodiments, the polynucleotide that encodes the viral rep protein is on the same vector or a different vector as the one or more polynucleotides.

[0080] In certain example embodiments, the polynucleotide that encodes the viral rep protein is operatively coupled to a regulatory element.

[0081] In certain example embodiments, the vector system is capable of producing a composition or portion thereof as described in any one of the preceding paragraphs and/or elsewhere herein.

[0082] Described in certain example embodiments herein are polypeptides encoded produced, or both by a vector system as described in any one of the preceding paragraphs and/or elsewhere herein.

[0083] In certain example embodiments, the polypeptide is a viral polypeptide.

[0084] In certain example embodiments, the viral polypeptide is an AAV polypeptide.

[0085] In certain example embodiments, the polypeptide is coupled to or otherwise associated with a cargo.

[0086] Described in certain example embodiments herein are particles produced by a vector system as described in any one of the preceding paragraphs and/or elsewhere herein, optionally including a polypeptide as described in any one of the preceding paragraphs and/or elsewhere herein.

[0087] In certain example embodiments, the particle is a viral particle.

[0088] In certain example embodiments, the viral particle is an adeno-associated virus (AAV) particle, lentiviral particle, or a retroviral particle.

[0089] In certain example embodiments, the particle comprises a cargo.

[0090] In certain example embodiments, the viral particle has a central nervous system (CNS) tropism.

[0091] In certain example embodiments, the polypeptide as described in any one of the preceding paragraphs and/or elsewhere herein, or the particle as described in any one of the preceding paragraphs and/or elsewhere herein, wherein the cargo is capable of preventing a CNS disease or disorder.

[0092] Described in certain example embodiments herein are cells comprising:

- a. a composition as described in any one of the preceding paragraphs and/or elsewhere herein;
- b. a vector system as described in any one of the preceding paragraphs and/or elsewhere herein;
- c. a polypeptide as described in any one of the preceding paragraphs and/or elsewhere herein;

- d. a particle as described in any one of the preceding paragraphs and/or elsewhere herein; or
- e. a combination thereof.

[0093] In certain example embodiments, the cell is prokaryotic.

[0094] In certain example embodiments, the cell is eukaryotic.

[0095] Described in certain example embodiments herein are pharmaceutical formulations comprising:

- a. a composition as described in any one of the preceding paragraphs and/or elsewhere herein;
- b. a vector system as described in any one of the preceding paragraphs and/or elsewhere herein;
- c. a polypeptide as described in any one of the preceding paragraphs and/or elsewhere herein;
- d. a particle as described in any one of the preceding paragraphs and/or elsewhere herein;
- e. a cell as described in any one of the preceding paragraphs and/or elsewhere herein; or
- f. a combination thereof; and
a pharmaceutically acceptable carrier.

[0096] Described in certain example embodiments herein are methods of treating a central nervous system disease, disorder, or a symptom thereof comprising:

administering, to the subject in need thereof,

- a. a composition as described in any one of the preceding paragraphs and/or elsewhere herein;
- b. a vector system as described in any one of the preceding paragraphs and/or elsewhere herein;
- c. a polypeptide as described in any one of the preceding paragraphs and/or elsewhere herein;
- d. a particle as described in any one of the preceding paragraphs and/or elsewhere herein;
- e. a cell as described in any one of the preceding paragraphs and/or elsewhere herein;
- f. a pharmaceutical formulation as described in any one of the preceding paragraphs and/or elsewhere herein; or
- g. a combination thereof.

[0097] In certain example embodiments, the central nervous system disease or disorder comprises a secondary muscle disease, disorder, or symptom thereof.

[0098] In certain example embodiments, the central nervous system disease or disorder is Friedreich's Ataxia, Dravet Syndrome, Spinocerebellar Ataxia Type 3, Niemann Pick Type C, Huntington's Disease, Pompe Disease, Myotonic Dystrophy Type 1, Glut1 Deficiency Syndrome (De Vivo Syndrome), Tay-Sachs, Spinal Muscular Atrophy, Alzheimer's disease, Amyotrophic lateral sclerosis (ALS), Danon disease, Rett Syndrome, Angleman Syndrome, or a combination thereof. 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

[0099] 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:

[0100] FIG. 1 shows the adeno-associated virus (AAV) transduction mechanism, which results in production of mRNA from the transgene.

[0101] FIG. 2 shows a graph that can demonstrate that mRNA-based selection of AAV variants can be more stringent than DNA-based selection. The virus library was expressed under the control of a CMV promoter.

[0102] FIGS. 3A-3B show graphs that can demonstrate a correlation between the virus library and vector genome DNA (FIG. 3A) and mRNA (FIG. 3B) in the liver.

[0103] FIGS. 4A-4F show graphs that can demonstrate capsid variants present at the DNA level, and expressed at the mRNA level identified in different tissues. For this experiment, the virus library was expressed under the control of a CMV promoter.

[0104] FIGS. 5A-5C show graphs that can demonstrate capsid mRNA expression in different tissues under the control of cell-type specific promoters (as noted on x-axis). CMV was included as an exemplary constitutive promoter. CK8 is a muscle-specific promoter. MHCK7 is a muscle-specific promoter. hSyn is a neuron specific promoter. Expression levels from the cell type-specific promoters have been normalized based on expression levels from the constitutive CMV promoter in each tissue.

[0105] FIGS. 6A-6B show (FIG. 6A) a schematic demonstrating embodiments of a method of producing and selecting capsid variants for tissue-specific gene delivery across species and (FIG. 6B) a schematic demonstrating benchmarking of the top selected capsids.

[0106] FIG. 7 shows a schematic demonstrating embodiments of generating an AAV capsid variant library, particularly insertion of a random *n*-mer ($n=3-15$ amino acids) into a wild-type AAV, e.g., AAV9.

[0107] FIG. 8 shows a schematic demonstrating embodiments of generating an AAV capsid variant library, particularly variant AAV particle production. Each capsid variant encapsulates its own coding sequence as the vector genome.

[0108] FIG. 9 shows schematic vector maps of representative AAV capsid plasmid library vectors (see e.g., FIG. 8) that can be used in an AAV vector system to generate an AAV capsid variant library.

[0109] FIG. 10 shows a graph that can demonstrate the viral titer (calculated as AAV9 vector genome/15 cm dish) produced by constructs containing different constitutive and cell-type specific mammalian promoters.

[0110] FIGS. 11A-11P show results from benchmarking the top selected capsids from the first and second round of selection.

[0111] The figures herein are for illustrative purposes only and are not necessarily drawn to scale.

DETAILED DESCRIPTION OF THE EXAMPLE EMBODIMENTS

General Definitions

[0112] 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).

[0113] As used herein, the singular forms “a”, “an”, and “the” include both singular and plural referents unless the context clearly dictates otherwise.

[0114] 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.

[0115] The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

[0116] 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.

[0117] 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.

[0118] 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.

[0119] 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.

[0120] 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

[0121] Embodiments disclosed herein provide central nervous system (CNS)-specific targeting moieties that can be coupled to or otherwise associated with a cargo and/or delivery vehicle or system. Embodiments disclosed herein provide polypeptides and particles that can incorporate one or more of the CNS-specific targeting moieties. The polypeptides and/or particles can be coupled to, attached to, encapsulate, or otherwise incorporate a cargo, thereby associating the cargo with the targeting moiety(ies). Embodiments disclosed herein provide CNS-specific targeting moieties that can contain one or more of an *n*-mer insert as further described herein. Some embodiments disclosed herein provide engineered adeno-associated virus (AAV) capsids that can be engineered to confer cell-specific and/or species-specific tropism, such as CNS specific tropism, to an engineered AAV particle.

[0122] In some embodiments, the *n*-mer motif(s) is as in any one of **Tables 1-3** and/or **7** and/or contains or is a P-motif, where the P motif contains or is the amino acid sequence $PX_1QGTX_2RX_n$ (SEQ ID NO: 2), where X_1 , X_2 , X_n , are each selected from any amino acid and where *n* is 0, 1, 2, 3, 4, 5, 6, or 7.

[0123] Embodiments disclosed herein also provide methods of generating recombinant AAVs (rAAVs) having engineered capsids that can involve systematically directing the generation of diverse libraries of variants of modified surface structures, such as variant capsid proteins. Embodiments of the method of generating rAAVs having engineered capsids can also

include stringent selection of capsid variants capable of targeting a specific cell, tissue, and/or organ type. Embodiments of the method of generating rAAVs having engineered capsids can include stringent selection of capsid variants capable of efficient and/or homogenous transduction in at least two or more species.

[0124] Embodiments disclosed herein provide vectors and systems thereof capable of producing an engineered AAV described herein.

[0125] Embodiments disclosed herein provide cells that can be capable of producing the engineered AAV particles described herein. In some embodiments, the cells include one or more vectors or system thereof described herein.

[0126] Embodiments disclosed herein provide engineered AAVs that can include an engineered capsid described herein. In some embodiments, the engineered AAV can include a cargo polynucleotide to be delivered to a cell. In some embodiments, the cargo polynucleotide is a gene modification polynucleotide.

[0127] Embodiments disclosed herein provide formulations that can contain an engineered AAV vector or system thereof, an engineered AAV capsid, engineered AAV particles including an engineered AAV capsid described herein, and/or an engineered cell described herein that contains an engineered AAV capsid, and/or an engineered AAV vector or system thereof. In some embodiments, the formulation can also include a pharmaceutically acceptable carrier. The formulations described herein can be delivered to a subject in need thereof or a cell.

[0128] Embodiments disclosed herein also provide kits that contain one or more of the one or more of the polypeptides, polynucleotides, vectors, engineered AAV capsids, engineered AAV particles, cells, or other components described herein and combinations thereof and pharmaceutical formulations described herein. In embodiments, one or more of the polypeptides, polynucleotides, vectors, engineered AAV capsids, engineered AAV particles cells, and combinations thereof described herein can be presented as a combination kit.

[0129] Embodiments disclosed herein provide methods of using the engineered AAVs having a cell-specific tropism described herein to deliver, for example, a therapeutic polynucleotide to a cell. In this way, the engineered AAVs described herein can be used to treat and/or prevent a disease in a subject in need thereof. Embodiments disclosed herein also provide methods of delivering the engineered AAV capsids, engineered AAV virus particles, engineered AAV vectors or systems thereof and/or formulations thereof to a cell. Also provided

herein are methods of treating a subject in need thereof by delivering an engineered AAV particle, engineered AAV capsid, engineered AAV capsid vector or system thereof, an engineered cell, and/or formulation thereof to the subject.

[0130] Additional features and advantages of the embodiments engineered AAVs and methods of making and using the engineered AAVs are further described herein.

CNS-SPECIFIC TARGETING MOIETIES AND COMPOSITIONS

[0131] Generally, described herein are compositions containing one or more CNS-specific targeting moieties that can effectively target CNS cells. In some embodiments, one or more CNS-specific targeting moieties can be incorporated into a delivery vehicle, agent, or system thereof so as to provide CNS specific targeting capability to the delivery vehicle, agent, or system thereof. Exemplary delivery vehicles include, without limitation, viral particles, (e.g., AAV viral particles), micelles, liposomes, exosomes, and the like. Exemplary delivery vehicles in which the CNS targeting-moieties can be incorporated are described in greater detail elsewhere herein. In some embodiments, the CNS-targeting moieties can be indirectly or directly coupled to a cargo and thus provide CNS specificity to the coupled cargo. In some embodiments, the composition can be specific for a CNS-cell (e.g., as conferred by the CNS-Specific targeting moieties described herein) and have reduced specificity for a non-CNS cell (including but not limited to a liver cell). In some embodiments, the CNS targeting moiety can specifically interact with or otherwise associate with one or more AAV receptors on CNS cells, thus providing CNS specificity (or tropism). Methods of generating and identifying CNS-specific targeting moieties are described in greater detail elsewhere herein.

CNS-Specific Targeting Moieties

[0132] Described herein are targeting moieties capable of specifically targeting, binding, associating with, or otherwise interacting specifically with a CNS cell. In some embodiments, the targeting moiety can be or include an n-mer motif described herein. In one example embodiment, the n-mer motif(s) is as in any one of **Tables 1-3** and/or **7**. In another example embodiment, the n-mer motif is or contains a P-motif. The term “P-motif” as used herein refers to an n-mer motif that contains or is the amino acid sequence $PX_1QGTX_2RX_n$ (SEQ ID NO: 2), where X_1 , X_2 , X_n , are each selected from any amino acid and where n is 0, 1, 2, 3, 4, 5, 6, or 7. n-mer motifs are described in greater detail elsewhere herein. It will be appreciated that in the context of n-mer motifs, the term “motif” and “insert” are used interchangeably herein. Generally, n-mer motifs are short (e.g., about 3 to about 15, 20, or 25) amino acid sequences

where each amino acid of the n -mer motif can be selected from any amino acid. In some embodiments, the n -mer motif (with or without a P motif) is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids in length.

[0133] In some embodiments, the n -mer insert can include an “RGD” insert (also referred to interchangeably herein as an “RGD motif”). An “RGD” insert refers to the presence of the amino acids RGD within the n -mer insert. In some embodiments, the RGD is the first three amino acids of the n -mer insert. Thus, in some embodiments the n -mer can have a sequence of RGD or RGDX_n , where n can be 3-15 amino acids and X, where each amino acid present can each be independently selected from the others and can be selected from the group of any amino acid. In some embodiments, the n -mer insert can be RGD (3-mer), RGDX_1 (4-mer), RGDX_1X_2 (5-mer), $\text{RGDX}_1\text{X}_2\text{X}_3$ (6-mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4$ (7 mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5$ (8 mer), or $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6$ (9-mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6\text{X}_7$ (10-mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6\text{X}_7\text{X}_8$ (11-mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6\text{X}_7\text{X}_8\text{X}_9$ (12-mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6\text{X}_7\text{X}_8\text{X}_9\text{X}_{10}$ (13-mer), $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6\text{X}_7\text{X}_8\text{X}_9\text{X}_{10}\text{X}_{11}$ (14-mer), or $\text{RGDX}_1\text{X}_2\text{X}_3\text{X}_4\text{X}_5\text{X}_6\text{X}_7\text{X}_8\text{X}_9\text{X}_{10}\text{X}_{11}\text{X}_{12}$ (15-mer), where $\text{X}_1, \text{X}_2, \text{X}_3, \text{X}_4, \text{X}_5, \text{X}_6, \text{X}_7, \text{X}_8, \text{X}_9, \text{X}_{10}, \text{X}_{11}, \text{X}_{12}$ can each be independently selected and can be any amino acid. In some embodiments, X_1 can be L, T, A, M, V, Q, or M. In some embodiments, X_2 can be T, M, S, N, L, A, or I. In some embodiments, X_3 can be T, E, N, O, S, Q, Y, A, or D. In some embodiments, X_4 can be P, Y, K, L, H, T, or S.

[0134] In certain example embodiments, the RGD motif has a formula of X_mRGDX_n , wherein m is 0-4 amino acids, wherein n is 0-15 amino acids, wherein X is any amino acid, and wherein each X amino acid present is independently selected from the others from the group consisting of: any amino acid. In certain example embodiments, the RGD motif has the formula RGDX_n , wherein n is 4 or 5, wherein X is any amino acid, and wherein each X amino acid present is independently selected from the others from the group consisting of: any amino acid or any specific combinations described elsewhere herein.

[0135] In some embodiments, n -mers including the RGD insert is included in a CNS specific targeting moiety and can facilitate muscle targeting by the targeting moiety in addition to CNS targeting. As will be appreciated in view of the present disclosure, such targeting moieties with CNS and muscle targeting capabilities can be advantageous for use in compositions and formulations for treating CNS diseases with muscle cell involvement or pathologies. In some exemplary embodiments and as further discussed herein the targeting moiety can be a viral capsid such as an AAV viral capsid.

[0136] In some embodiments, the n-mer motif does not include and RGD insert.

[0137] In some embodiments, the n-mer insert and/or P-motif is immediately preceded by an AQ a DG, such as when the n-mer insert and/or P-motif is inserted into another polypeptide. In some embodiments, the n-mer insert polypeptide is immediately preceded by AQ and wherein the n-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSGG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESGH (SEQ ID NO: 19); or GENSARW (SEQ ID NO: 20). In some embodiments, n-mer insert polypeptide is immediately preceded by DG and wherein the n-mer insert is REQQKLW (SEQ ID NO: 21), ASNPGRW (SEQ ID NO: 22), WTLESGH (SEQ ID NO: 23), REQQKLW (SEQ ID NO: 25), ERLLVQL (SEQ ID NO: 25); or RMQRTLY (SEQ ID NO: 26). In some embodiments, the CNS-specific n-mer motif can be as in **Table 1**.

[0138] In some embodiments, the CNS-specific n-mer motif can be or include a P-motif. In some embodiments, X₁ of the P-motif is S, T, or A. In some embodiments, X₂ of the P-motif L, V, F, or I. In some embodiments, X_n of the P-motif is 0. In some embodiments, the CNS-specific n-mer motif is as in any of **Tables 2-3**.

[0139] In some embodiments, the CNS-specific n-mer insert and/or P-motif is any one of the n-mer inserts and/or P motif in **Table 7** (SEQ ID NOs.: 321-329). In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-324. In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-325. In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-327. In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-324 and 329. In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-324. In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-324 and 326-327. In some embodiments the CNS-specific n-mer insert and/or P-motif is

any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-324 and 326-328. In some embodiments the CNS-specific n-mer insert and/or P-motif is any one or more of the n-mer inserts selected from the group of SEQ ID NOs.: 322-324 and 328.

[0140] In some embodiments, the CNS-specific n-mer insert and/or P-motif is species specific. In other words, in some embodiments, the CNS-specific n-mer insert and/or P-motif can facilitate CNS targeting in one species better than another species. In some embodiments the CNS-specific n-mer insert is specific for primates. In some embodiments, the CNS-specific n-mer insert is specific for human and/or non-human primates.

[0141] In some embodiments, the CNS-specific n-mer insert is capable of targeting one or more cell and/or tissue types over others within the CNS. In some embodiments, the CNS-specific insert is not effective or is less effective at targeting the dorsal root ganglion cells than one or more other cells and/or tissue types of the CNS.

[0142] In some embodiments, the targeting moiety can include more than one n-mer motifs, such as a CNS-specific n-mer motif described herein. In some embodiments, the targeting moiety can include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more n-mer motifs. In some embodiments, all the n-motifs included in the targeting moiety can be the same. In some embodiments where more than one n-mer motif is included, at least two of the n-mer motifs are different from each other. In some embodiments where more than one n-mer motif is included, all the n-mer motifs are different from each other.

[0143] In some embodiments, the targeting moiety, e.g., the CNS-specific targeting moiety, can be coupled to or otherwise associated with a cargo. In some embodiments, one or more muscle-specific targeting moieties described herein is directly attached to the cargo. In some embodiments, one or more muscle-specific targeting moieties described herein is indirectly coupled to the cargo, such as via a linker molecule. In some embodiments, one or more one or more muscle-specific targeting moieties described herein is coupled to associated with a polypeptide or other particle that is coupled to, attached to, encapsulates, and/or contains a cargo.

[0144] Exemplary particles include, without limitation, viral particles (e.g., viral capsids, which is inclusive of bacteriophage capsids), polysomes, liposomes, nanoparticles, microparticles, exosomes, micelles, and the like. The term “nanoparticle” as used herein includes a nanoscale deposit of a homogenous or heterogeneous material. Nanoparticles may be regular or irregular in shape and may be formed from a plurality of co-deposited particles

that form a composite nanoscale particle. Nanoparticles may be generally spherical in shape or have a composite shape formed from a plurality of co-deposited generally spherical particles. Exemplary shapes for the nanoparticles include, but are not limited to, spherical, rod, elliptical, cylindrical, disc, and the like. In some embodiments, the nanoparticles have a substantially spherical shape.

[0145] As used herein, the term “specific” when used in relation to described an interaction between two moieties, refers to non-covalent physical association of a first and a second moiety wherein the association between the first and second moieties is at least 2 times as strong, at least 5 times as strong as, at least 10 times as strong as, at least 50 times as strong as, at least 100 times as strong as, or stronger than the association of either moiety with most or all other moieties present in the environment in which binding occurs. Binding of two or more entities may be considered specific if the equilibrium dissociation constant, K_d , is 10^{-3} M or less, 10^{-4} M or less, 10^{-5} M or less, 10^{-6} M or less, 10^{-7} M or less, 10^{-8} M or less, 10^{-9} M or less, 10^{-10} M or less, 10^{-11} M or less, or 10^{-12} M or less under the conditions employed, e.g., under physiological conditions such as those inside a cell or consistent with cell survival. In some embodiments, specific binding can be accomplished by a plurality of weaker interactions (e.g., a plurality of individual interactions, wherein each individual interaction is characterized by a K_d of greater than 10^{-3} M). In some embodiments, specific binding, which can be referred to as “molecular recognition,” is a saturable binding interaction between two entities that is dependent on complementary orientation of functional groups on each entity. Examples of specific interactions include primer-polynucleotide interaction, aptamer-aptamer target interactions, antibody-antigen interactions, avidin-biotin interactions, ligand-receptor interactions, metal-chelate interactions, hybridization between complementary nucleic acids, etc.

[0146] In some embodiments, in addition to the n-mer motif(s) the targeting moiety can include a polypeptide, a polynucleotide, a lipid, a polymer, a sugar, or a combination thereof.

[0147] In some embodiments, the targeting moiety is incorporated into a viral protein, such as a capsid protein, including but not limited to lentiviral, adenoviral, AAV, bacteriophage, retroviral proteins. In some embodiments, n-mer motif is located between two amino acids of the viral protein such that the n-mer motif is external (i.e., is presented on the surface of) to a viral capsid.

[0148] In some embodiments, the composition containing one or more of the muscle-specific targeting moieties described herein has increased muscle cell potency, muscle cell specificity, reduced immunogenicity, or any combination thereof.

[0149] Cargos can include any molecule that is capable of being coupled to or associated with the muscle-specific targeting moieties described herein. Cargos can include, without limitation, nucleotides, oligonucleotides, polynucleotides, amino acids, peptides, polypeptides, riboproteins, lipids, sugars, pharmaceutically active agents (e.g., drugs, imaging and other diagnostic agents, and the like), chemical compounds, and combinations thereof. In some embodiments, the cargo is DNA, RNA, amino acids, peptides, polypeptides, antibodies, aptamers, ribozymes, guide sequences for ribozymes that inhibit translation or transcription of essential tumor proteins and genes, hormones, immunomodulators, antipyretics, anxiolytics, antipsychotics, analgesics, antispasmodics, anti-inflammatories, anti-histamines, anti-infectives, radiation sensitizers, chemotherapeutics, radioactive compounds, imaging agents, and combinations thereof.

[0150] The CNS-specific n-mer motifs and targeting moieties can be encoded in whole or in part by a polynucleotide. The encoding polynucleotides can be included in one or more vectors (or vector systems) that can be used to generate targeting moieties and compositions thereof that include the CNS-specific n-mer motif(s) and/or P-motif(s). Exemplary encoding polynucleotides, vectors, vector systems, and recombinant engineering techniques are described in greater detail herein and/or are generally known in the art and can be adapted for use with the targeting moieties and compositions thereof described herein.

[0151] In some embodiments, the cargo is capable of treating or preventing a CNS disease or disorder. Exemplary CNS diseases and disorders are described elsewhere herein.

Cargos

[0152] The targeting moiety effective to target a CNS cell can optionally be coupled to a cargo. In this way, the cargo can be delivered selectively to CNS cells when incorporated into a composition described herein. 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

[0153] 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.

[0154] As used herein, “deoxyribonucleic acid (DNA)” and “ribonucleic acid (RNA)” can generally refer to any polyribonucleotide or polydeoxyribonucleotide, 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).

[0155] 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

[0156] 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.

[0157] 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.

[0158] 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)”.

[0159] 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.

[0160] 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. In some embodiments, the disease or condition is a

disease or condition of or affecting the CNS or cell thereof. Exemplary diseases and disorders of and/or affecting the CNS are described in greater detail elsewhere herein.

[0161] 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.

[0162] 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.

[0163] 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

[0164] In certain example embodiments, the one or more polynucleotides may encode one or more interference RNAs. Interference 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).

[0165] 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.

[0166] 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 TGF β pathway, e.g., TGF β , TGF β 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 TGF β pathway. TGF β pathway genes may comprise one or more of ACVR1, ACVR1C, ACVR2A, ACVR2B, ACVRL1, AMH, AMHR2, BMP2, BMP4, BMP5, BMP6, BMP7, BMP8A, BMP8B, BMPRI1A, BMPRI1B, BMPRI2, 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, TGFB1R1, TGFB1R2, THBS1, THBS2, THBS3, THBS4, TNF, ZFYVE16, and/or ZFYVE9.

[0167] 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.

[0168] 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

[0169] 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.

[0170] 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).

[0171] In some embodiments, the cargo polynucleotide is a DNA or RNA (e.g., a mRNA) vaccine.

Aptamers

[0172] 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.

[0173] 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.

[0174] 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.

[0175] 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'-O-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.

[0176] In certain other example embodiments, the polynucleotide may be a ribozyme or other enzymatically active polynucleotide.

Biologically active agents

[0177] 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.

[0178] 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).

[0179] One or two or more different cargoes may be delivered by the delivery particles described herein.

[0180] 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 (GGG)₃ (SEQ ID NO: 27).

[0181] In some embodiments, the cargo comprises a ribonucleoprotein. In specific embodiments, the cargo comprises a genetic modulating agent.

[0182] 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.

Genetic Modifying Systems

[0183] 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.

[0184] 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

[0185] 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](https://doi.org/10.1016/j.molcel.2015.10.008).

Class 1 Systems

[0186] 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 Rossmann 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

[0187] 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.

[0188] 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.

[0189] 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.

[0190] 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.

[0191] 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

[0192] 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.

[0193] 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.

[0194] 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).

[0195] 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.

[0196] 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).

[0197] 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.

[0198] 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.

[0199] 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.

[0200] 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.

[0201] 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.

[0202] 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.

[0203] 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.

[0204] 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

[0205] 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.

[0206] 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.

[0207] 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

[0208] 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.

[0209] 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 4* (from Gleditzsch et al. 2019) below shows several Cas polypeptides and the PAM sequence they recognize.

Table 4 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'

[0210] 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.

[0211] Further, engineering of the PAM Interacting (PI) domain on the Cas protein may allow programming 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.

[0212] 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).

[0213] 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., Gleditsch et al. 2019. RNA Biology. 16(4):504-517.

[0214] 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., Gleditsch et al. 2019. RNA Biology. 16(4):504-517.

[0215] 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

[0216] 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).

[0217] 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: 28) or PKKKRKVEAS (SEQ ID NO: 29); the NLS from nucleoplasmin (e.g., the nucleoplasmin bipartite NLS with the sequence KRPAATKKAGQAKKKK (SEQ ID NO: 30)); the c-myc NLS having the amino acid sequence PAAKRVKLD (SEQ ID NO: 31) or RQRRNELKRSP (SEQ ID NO: 32); the hRNPA1 M9 NLS having the sequence NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 33); the sequence RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 34) of the IBB domain from importin-alpha; the sequences VSRKRPRP (SEQ ID NO: 35) and PPKKARED (SEQ ID NO: 36) of the myoma T protein; the sequence PQPKKKPL (SEQ ID NO: 37) of human p53; the sequence SALIKKKKKMAP (SEQ ID NO: 38) of mouse c-abl IV; the sequences DRLRR (SEQ ID NO: 39) and PKQKKRK (SEQ ID NO: 40) of the influenza virus NS1; the sequence RKLKKKIKKL (SEQ ID NO: 41) of the Hepatitis virus delta antigen; the sequence REKKKFLKRR (SEQ ID NO: 42) of the mouse Mx1 protein; the sequence KRKGDEV DGVDEVAKKKSKK (SEQ ID NO: 42) of the human poly(ADP-ribose) polymerase; and the sequence RKCLQAGMNLEARKTKK (SEQ ID NO: 44) 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.

[0218] 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.

[0219] 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.

[0220] 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.

[0221] 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.

[0222] 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.

[0223] 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 nucleus of a cell, such as a target cell.

Donor Templates

[0224] 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.

[0225] 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.

[0226] 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.

[0227] 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.

[0228] 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.

[0229] 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, 11, 12 or more nucleotides of the target sequence.

[0230] 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, 110+/- 10, 120+/- 10, 130+/- 10, 140+/- 10, 150+/- 10, 160+/- 10, 170+/- 10, 180+/- 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, 110+/-20, 120+/-20, 130+/-20, 140+/-20, 150+/-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 to 300, 50 to 200, or 50 to 100 nucleotides in length.

[0231] 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.

[0232] 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.

[0233] 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.

[0234] 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.

[0235] 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.

[0236] 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).

[0237] In certain embodiments, a template nucleic acid for correcting a mutation may be 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.

[0238] 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

[0239] 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 functional 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.

[0240] 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).

[0241] 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.

[0242] Other suitable functional domains can be found, for example, in International Patent Publication No. WO 2019/018423.

Split CRISPR-Cas systems

[0243] 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 is 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

[0244] 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.

[0245] 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. Genet. 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.

[0246] 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.

[0247] 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.

[0248] An example method for delivery of base-editing systems, including use of a split-intein approach to divide CBE and ABE into reconstitutible 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

[0249] 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 facilitate 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 CRISPR-Cas systems along with few byproducts and greater or similar efficiency as compared to traditional CRISPR-Cas systems.

[0250] 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.

[0251] 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.

[0252] 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,

[0253] 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

[0254] 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 Class I 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 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

[0255] 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.

[0256] 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.

[0257] 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 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.

[0258] 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.

[0259] 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 5**.

Table 5.		
No.	Proteins	Sequences
1	IscB(-HNH) EFH81386	MSTDATLIRTPSHAEADATDTLVATPLMPRRVISPWPGGEGQSLMRIPVVDIRG MALMPCTPAKARHLLKSGNARPKRNKLGIFYVQLSYEQEPDNQSLVAGVDPGSKF EGLSVVGTKDTVNLNLMVEAPDHVKGA VQTRRTMRRARRQRKWRPFRFHNRLNR MQRIIPPSTRSRWEAKARIVAHRLTILPFTDVVVEDVQAVTRKGGKGGTWNNGSFSPVQ VGKEHL YRLLRAMGLTLHLREGWQTKELREQHGLKKTSSKSKQSFESHAVDSWV LAASISGAEHPTCTRLWYMPAILHRRQLHRLQASKGGVRKPYGGTRSLGVKRGTL VEHKKYGRCTVGGVDRKNTISLHEYRTNTRLTQAAK VETCRVLTWLSWRSWLLR GKRTSSKGGKSHSS (SEQ ID NO: 45)
2	IscB(+HNH) TAE54104.1	MQPAKQQNWVFQINGDKQPLDMINPGRCRELQNRGKLASFRRFPYVVIQQQTINP QTIKEYILKIDPGSQWTGFAIQCNDILFRAELNHRGEAIKFDLVKRAWFRGRSRN LRYRKKRLNRAKPEGWLAPSIRHRVLT VETWIKRFMR YCPIAWIEIEQVRFDTQKLA NPEIDGVEYQQGELQGYEVREYLLQKWGRKCAYCGTENVPLEVEHIQSKSKGGSSR IGNLTLACHVCNVKKGNDVDRDFL AKSPDILNOV LENS TKPLKD AA AVNSTRYA IVKMAKSICENVKCSSGARTKMNRVRQGLEKTHSLDAACVGESGASIRVLTDRPLLI TCKGHGSRQSIRVNASGFPVKNKAKTVFTHIAAGDVVRFTIGKDRKKAQAGTYTAR VKTPTPKGFEVLIDGAR ISLSTMSNVVVFVHRSDGYGY EL (SEQ ID NO: 46)
3	IscB(+HNH) WP_038093640.1	MAVFVIDKHKRPLMPCSEKRRARLLERGRAVVRQVVFVIRLKDRTVQHSVAVQPLR VALDPGSRATGMALVREKNTVDTGTGEVYRERIALNLFELVHRGHRIREQLDQRRN FRRRRRGANLRYRAPRFDNRRRPPGWLAPSLQHRVDTTMAWVRRLCRWAPASAI IETVRFDQRLQNPESIGVEYQQGALAGCEVREYLLLEKWGRKCAYCGAENVPLEIE HIVPKSRGGSDRVSNLALACRACNQAKGNRDVRAFLADQPERLARILAQAKAPLK DAAAVNATR WALYRALVD TGLPVEAGTGGRTKWNTRLGLPKTHALDALCVGQ VDQVRHWRVPVLGIR CAGRGSYRRTRLTRHGFPRGYLTRNKSAGFGFOTGDILRAV TKGKKAGTYLGRIAIRASGSFNIQTPMGVVQGIHHRFCTLLQORADGYGYFVQP KPTEAALSSPRLKAGVSSAG N (SEQ ID NO: 47)
4	IscB(+HNH) WP_052490348.1	MTTNVVFVIDTNQKPLQPCSAAVARKLLRGKAAMFRRYPVAILKKEVDSVGGPK IELRIDPGSKYTGFALVDSKDNADFIIWGTELEHRRGAICKELTKRSAIRSRNRKT RYRKKRFERRKPEGWLAPSLQHRVDTTTLTWVKRICKFVPIMSISVEQVKFDLQKLE NSDIQGIEYQQGTLAGYTLREALLEHWGRKCAYCDVENVFLEIEHIYPKSKGGSDKF SNLTLACHKCNINKGNKSIDEFLLSDHKRLEQIKLHQKKTLDAAA V NATRKKLV TLQEKTFNLNVLVSDGASTKMTRLSSSLAKRHWDAGCVNTLIVILKTLQPLQVKCN GHGKQFVTMDAYGFPRKSYEPKKVRKDWKAGDIIRVTKKDGTMLMGRVKKAA KKLVYIPFGGKEASFSS ENAKAIHRSDGYRYSFAAID SELLQKMAT (SEQ ID NO: 48)
5	IscB(+HNH) WP_015325818.1	MPNKYAFVLDSKGLLDPTKSKKAWYLIRK GKASLVEEYPLIILKREVPKDQVNS DKLILGIDDGTTKVGFALVQKCTKNKVLFAVMEQRQDVSKKMEERRGYRRYR RSHKRYRPARFDNRSSSKRKGRIIPSIQKKQAILRVVNKLKKYIRIDKIVLEDVSI DKLREGRELYNWEYQESNRLDENLRKATLYRDDCTCQLCGTTETMLHAHHIMPRR DGGADSIYNLITLCKACHKDKVDNNEYQYKDQFLAIDSKELSDLKSASHVMQGKT WLRDKLSKIAQLEITSGGNTANKRIDYEIEKSHSNDACITGLLPVDNIDDIKEYIIP LRKKSKAKIKELKCFRQDLVKYTKRNGETYTYGITSRLIKNNKYNKVCNFSTLK GKIFRGYGFRNLTLNRPKGLMIV (SEQ ID NO: 49)

6	sp G3ECR1 CAS 9_STRTR	MLFNKCIISINLDFSNKEKCMTPKYSIGLDIGTNSVGWAVITDNYKVPSKKMKVLG NTSKKYIKKNLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRILYLQEIFSTEMATL DDAFFQRLDDSFVLPDDKRDSPYIFGNLVEEKVYHDEFPTIYHLRKYLAADSTKKA DLRLVYLAALAHMIKYRGHFLIEGEFNSKNNDIQNFQDLDTYNIAIFESDLSLENSK QLEEVKDKISKLEKKDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASL HFSKESYDEDLETLGGYIGDDYSDVFLKAKKLYDAILLSGFLTVDNETEAPLSSAM IKRYNEHKEDLALLKEYIRNISLKTYNVFKDDDTKNGYAGYIDNGKTNQEDFYVYLK NLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDQAKFYPFLA KNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFIN RMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFAESMRDYQFLDSKQKKDIV RLYFKDKRKVTDKDIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLNINIINDKEFLD DSSNEAIIIEIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWGLSAKLI NGIRDEKSGNTILDYLIIDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEV VKSPLGSPAIAKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGSNSQQR LKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYL YLQNGKDMYTGDLDI DRLSNYDIDHIIPQAFKLDNSIDNKVLVSSASNRGKSDDFPSLEVVKRKRFTWYQLL KSKLISQRFKFDNLTKAERGGLLPEDKAGFIQRQLVETROITKHVARLLDEKFNKKD ENNRAVRTVKIITLKLSTLVSQFRKDFELYKVRINDFHHAHDAYLNAVIAALLKKY PKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKISLADNGKDMYTGDLDI EETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSK PKPNSNENLVGAKEYLDPKKYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQGIS ILDRINIRKDKLNFLEKGYKDIELIIEPKYSLFELSDGSRMLASILSTNNKRGEIH KGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKFEFELFYIIEFNENYVGA KKNGKLLNSAFQSQWNHISIDELCSSFIGPTGSEKGLFELTSRGSAADEFELGVK IPRYRDTYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG (SEQ ID NO: 50)
7	sp J7RUA5 CAS 9_STAAU	MKNRYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLK RRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSSEEFSAALLHLAK RRGVHNVNEVEEDTGNELSTKEQISRNASKALEEKYVAELQLERLKKDGEVRGSINR FKTSDYVKEAKQLLVQKAYHQLDQSFIDTYIDLETRRTTYEGPGEFSPFGWKDI KEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNLNITRDENEKLEYEYEF QIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKYGHDIDITARK EIIENAELLQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYTGTHNLSLKA NLILDELWHTNDNQIAIFNRLKLVKKVDLSQKKEIPTTLVDDFILSPVVKRSFIQSIK VINAIKKYGLPNDIIIEAREKNSKDAQKMINEMQKRNQRTNERIEEIRTTGKENAK YLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNVYVDHIIPRSVSDFNFNKVLVK QEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKKEYLLEERDINR FSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVK SINGGFTSFLRRWKFK KKERNKGKHHAEADALIANDFIFKEWKLDKAKKVMENQMFEKQAESMPEIE TEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVN NLNGLYDKDNDKLKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDENPLYKY EETGNYLTKYSSKDNPNVVIKKIKYYGNKLNALHDITDDYPNSNRNKVVKLSLKPYP DVYLDNGVYKFVTVKNLVDVICKENYEVNSKCYEEAKKLKISNQAEIFASFYNN DLIKINGELYRVIGVNNDDLNRNIEVNMIDITYREYLENMNDKRPRIIKTIASTQSIK KYSTDILGNLYEVKSKKHPOIKKG (SEQ ID NO: 51)
8	Streptococcus_p yogenes_SF370	KYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFHRLLEESFLVEEDKKHERHP IFGNIVDEVAAYHEKYPTIYHLRKKLVDSIDKADRLIYLAALAHMIKFRGHFLIEGDL NPDNSDVDKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGE KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR TFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYF TVYNELTKVKYVTEGMKPAFLSGEQKKAIVDILLFKTNRKVTYVQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLSDGFA NRNFMQLIHDDSLTFKEDIQKAQVSGQDSLHEHIANLAGSPAIAKKGILQTVKVVDE LVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEKGELGQSILKEHPVE NTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYVDVHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDK AGFIKQRLVETROITKHVAQILDSRMNTKYDENDKLIREVKVITLKSLSLSDFRKDF QFYKVRINNYYHHAHDAYLNAVVGTAIIKKYPKLESEFYDYKVVYDVRKMIAS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPRIETNGETGEIVWDKGRDFATV RKVL SMPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTV

		AYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLII KLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDRDKPIREQAENII HLF TL TN LGAPAAFKYFDTTIDRKRYTSTKEVL DATL IHQSITGLYETRIDLSQLGGD (SEQ ID NO: 52)
--	--	---

Table 5 (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 J7RUΛ5 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

[0260] In some embodiments, the IscB proteins comprise an X domain, e.g., at its N-terminal.

[0261] In certain embodiments, the X domain include the X domains in **Table 5**. 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 5**.

[0262] 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

[0263] In some embodiments, the IscB proteins comprise a Y domain, e.g., at its C-terminal.

[0264] In certain embodiments, the X domain include Y domains in **Table 5**. 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 5**.

RuvC domain

[0265] 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.

[0266] 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.

[0267] In certain embodiments, examples of the RuvC domain include those in **Table 5**. 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 65%, 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 5**.

Bridge helix

[0268] 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.

[0269] 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.

[0270] In certain embodiments, examples of the BH domain include those in **Table 5**. 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 65%, 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 5**.

HNH domain

[0271] 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.

[0272] 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.

[0273] In certain embodiments, examples of the HNH domain include those in **Table 5**. 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 65%, 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 5**.

hRNA

[0274] 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.

[0275] 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

[0276] 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.

[0277] 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.

[0278] 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).

[0279] 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.

[0280] 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.

[0281] 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.

[0282] 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.

[0283] An exemplary amino acid sequence of a N-terminal capping region is:

MDPIRSRTSPSPARELLSGPQPDGVQPTADRGVSPPAGGPLDGL
LPARRTMSRTRLPSPPAPSPAFSADSFSDLLRQFDPSLNFNTSL
FDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVT
AARPPRAKPAARRRAAQPSDASPAQVDLRTLGYSSQQQKEK
IKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAV
KYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELR
GPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN (SEQ
ID NO: 53)

[0284] An exemplary amino acid sequence of a C-terminal capping region is:

RPALESIVAQLSRPDPALAAALTNDHLVALACLGGRPALDAVK
KGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHS
HPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARS GTLPP
ASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLER
DLDA P S P M H E G D Q T R A S (SEQ ID NO: 54)

[0285] 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.

[0286] 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.

[0287] 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.

[0288] 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.

[0289] 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.

[0290] 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.

[0291] 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.

[0292] 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.

[0293] 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.

[0294] 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

[0295] Zinc Finger proteins 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

[0296] 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

[0297] 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%.

[0298] 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.

[0299] 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).

[0300] 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.

[0301] 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.

[0302] 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

[0303] 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

[0304] 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.

[0305] 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.

[0306] 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., α 1-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.), α 2-macroglobulin, α 1-antitrypsin, antithrombin III, protein S, protein C, plasminogen, α 2-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, adrenocorticotrophic 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

[0307] 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.

[0308] 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.

[0309] 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, Th $\alpha\beta$, CD4+, CD8+, effector Th, memory Th, regulatory Th, CD4+/CD8+ thymocytes, CD4-/CD8- thymocytes, $\gamma\delta$ 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.

[0310] 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, IL-5, 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.

[0311] 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.

[0312] 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, TNF α and IFN γ 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).

[0313] 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 (IFN γ). 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.

[0314] 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

[0315] 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 PPAR α , PPAR γ is a dual PPAR α/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

[0316] In certain example embodiments, the one or more polypeptides may comprise 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, Ptf1a1, 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

[0317] In certain embodiments, the one or more 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), adrenocorticotrophic 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.

[0318] 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

[0319] 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

[0320] 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, cathelicidin 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

[0321] 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: 55)), P5 (VVVVV (SEQ ID NO: 56)), KR7 (KPWWPRR (SEQ ID NO: 57)), NK9 (NIVNVSLVK (SEQ ID NO: 58)), iAb5p (Leu-Pro-Phe-Phe-Asp (SEQ ID NO: 59)), KLVF (SEQ ID NO: 60) 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: 61)), RI-OR2-TAT, cyclo(17, 21)-(Lys17, Asp21)A₁(1-28), SEN304, SEN1576, D3, R8-A β (25-35), human γ D-crystallin (HGD), poly-lysine, heparin, poly-Asp, poly-Gl, poly-L-lysine, poly-L-glutamic acid, LVEALYL (SEQ ID NO: 62), RGFFYT (SEQ ID NO: 63), 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 A β 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

[0322] 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.

[0323] 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.

[0324] 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.

[0325] It is intended that the term "antibody" encompass any Ig class or any Ig subclass (e.g., the IgG1, IgG2, IgG3, and IgG4 subclasses of IgG) obtained from any source (e.g., humans and non-human primates, and in rodents, lagomorphs, caprines, bovines, equines, ovines, etc.).

[0326] 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.

[0327] 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, $\gamma 1$ - $\gamma 4$, 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).

[0328] 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.

[0329] 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.

[0330] 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).

[0331] 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).

[0332] "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 $1 \times 10^7 \text{ M}^{-1}$ (or a dissociation coefficient of $1 \mu\text{M}$ or less or a dissociation coefficient of 1 nm 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, 100 nM or less, 75 nM or less, 50 nM or less, 25 nM or less, for example 10 nM or less, 5 nM or less, 1 nM or less, or in embodiments 500 pM or less, 100 pM or less, 50 pM or less or 25 pM 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.

[0333] 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, K_d , and the association constant, K_a , are quantitative measures of affinity.

[0334] 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.

[0335] 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.

[0336] "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.

[0337] 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 C_{H1} 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).

[0338] 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).

[0339] 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.

[0340] 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).

[0341] 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, phosphorylation, 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

[0342] 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

[0343] In some embodiments, the cargo molecule is a small molecule. Techniques and methods of coupling peptides to small molecule agents are generally known in the art and can be applied here to couple a targeting moiety effective to target a CNS cell to a small molecule cargo. Small molecules include, without limitation, hormones, immunomodulators, antipyretics, anxiolytics, antipsychotics, analgesics, antispasmodics, anti-inflammatories, anti-histamines, anti-infectives, radiation sensitizers, chemotherapeutics.

[0344] 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 testosterone 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- α , 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).

[0345] 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 salicylate, and sodium salicylate), paracetamol/acetaminophen, metamizole, nabumetone, phenazone, and quinine.

[0346] Suitable anxiolytics include, but are not limited to, benzodiazepines (e.g., alprazolam, bromazepam, chlordiazepoxide, clonazepam, clorazepate, diazepam, flurazepam, lorazepam, oxazepam, temazepam, triazolam, and tofisopam), serotonergic antidepressants (e.g. selective serotonin reuptake inhibitors, tricyclic antidepressants, and monoamine oxidase inhibitors), mebicar, afobazole, selank, bromantane, emoxypine, azapirones, barbiturates, hydroxyzine, pregabalin, validol, and beta blockers.

[0347] 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, flupenthixol, tiotixene, zuclopenthixol, clotiapine, loxapine, prothipendyl, caripramine, 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 ziconapine.

[0348] 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).

[0349] Suitable antispasmodics include, but are not limited to, mebeverine, papverine, cyclobenzaprine, carisoprodol, orphenadrine, tizanidine, metaxalone, methocarbamol, 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).

[0350] 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, levocetirizine, loratadine, meclozine, mirtazapine, olopatadine, orphenadrine, phenindamine, pheniramine, phenyltoloxamine, promethazine, pyrilamine, quetiapine, rupatadine, tripeleminamine, and triprolidine), H2-receptor antagonists (e.g., cimetidine, famotidine, lafutidine, nizatidine, ranitidine, and roxatidine), tritoqualine, catechin, cromoglicate, nedocromil, and p2-adrenergic agonists.

[0351] 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, albendazole, thiabendazole, oxfamiquine), 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/proguanil, quinine, hydroxychloroquine, mefloquine, chloroquine, doxycycline, pyrimethamine, and halofantrine), antituberculosis agents (e.g., aminosaliclates (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, avacavir, zidovudine, stavudine, emtricitabine, xalcitabine, telbivudine, simeprevir, boceprevir, telaprevir, lopinavir/ritonavir, fosamprenvir, dranuavir, ritonavir, tipranavir, atazanavir, nelfinavir, amprenavir, indinavir, sawuinaur, ribavirin, valcyclovir, acyclovir, famciclovir, ganciclovir, and valganciclovir), carbapenems (e.g., doripenem, meropenem, ertapenem, and cilastatin/imipenem), cephalosporins (e.g., cefadroxil, cephadrine, cefazolin, cephalixin, cefepime, ceftaroline, 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), glycylicyclines (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, erythromycin, 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, oxacillin, dicloxacillin, and nafcillin), quinolones (e.g., lomefloxacin, norfloxacin, ofloxacin, gatifloxacin, 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/salicylic 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).

[0352] 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, Cytosan, cyclophosphamide, decitabine, dexamethasone, docetaxel, hydroxyurea, decarbazine, leuprolide, epirubicin, oxaliplatin, asparaginase, estramustine, cetuximab, vismodegib, asparaginase *Erwinia chrysanthemi*, amifostine, etoposide, flutamide, toremifene, fulvestrant, letrozole, degarelix, pralatrexate, methotrexate, floxuridine, obinutuzumab, gemcitabine, afatinib, imatinib mesylate, 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, octetide, dasatinib, regorafenib, histrelin, sunitinib, siltuximab, omacetaxine, thioguanine (tioguanine), dabrafenib, erlotinib, bexarotene, temozolomide, thiotepa, thalidomide, BCG, temsirolimus, bendamustine hydrochloride, triptorelin, arsenic 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 Viral Capsids and Encoding Polynucleotides

[0353] Described herein are exemplary embodiments of engineered viral proteins, (e.g., capsid proteins), such as adeno-associated virus (AAV) viral proteins (e.g., capsid proteins), that can be engineered to confer cell-specific tropism to an engineered viral particle (AAV particle) that contains the engineered viral protein(s). The engineered viral protein(s) (e.g., capsid(s)) can be included in an engineered virus particle, and can confer cell-specific tropism, such as CNS-specific tropism, reduced immunogenicity, or both to the engineered viral (e.g., an AAV) particle. As is described elsewhere herein, the particles can include a cargo. In this

way, the particles can be a cell-specific delivery vehicle for a cargo. The engineered viral capsids described herein can include one or more engineered viral capsid proteins described herein. Engineered viral capsid proteins can be lentiviral, retroviral, adenoviral, or AAV. Engineered capsids can contain one or more of the viral capsid proteins. Engineered virus particles can include one or more of the engineered viral capsid proteins and thus contain an engineered viral capsid. The engineered viral capsid proteins, capsids, and/or viral particles that contain one or more CNS-specific targeting moieties containing or composed of one or more n-mer motifs described elsewhere herein. In some embodiments, the engineered viral capsid proteins, viral capsids, and/or viral particles can have a CNS-specific tropism conferred to it by the one or more n-mer motifs contained therein.

[0354] The CNS-specific n-mer motifs and targeting moieties can be encoded in whole or in part by a polynucleotide. The engineered viral capsid and/or viral capsid proteins can be encoded by one or more engineered viral capsid polynucleotides. In some embodiments, the engineered viral capsid polynucleotide is an engineered AAV capsid polynucleotide, engineered lentiviral capsid polynucleotide, engineered retroviral capsid polynucleotide, or engineered adenovirus capsid polynucleotide. In some embodiments, an engineered viral capsid polynucleotide (e.g., an engineered AAV capsid polynucleotide, engineered lentiviral capsid polynucleotide, engineered retroviral capsid polynucleotide, or engineered adenovirus capsid polynucleotide) can include a 3' polyadenylation signal. The polyadenylation signal can be an SV40 polyadenylation signal.

[0355] The engineered AAV capsids can be variants of wild-type AAV capsids. In some embodiments, the wild-type AAV capsids can be composed of VP1, VP2, VP3 capsid proteins or a combination thereof. In other words, the engineered AAV capsids can include one or more variants of a wild-type VP1, wild-type VP2, and/or wild-type VP3 capsid proteins. In some embodiments, the serotype of the reference wild-type AAV capsid can be AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, AAV-8, AAV-9 or any combination thereof. In some embodiments, the serotype of the wild-type AAV capsid can be AAV-9. The engineered AAV capsids can have a different tropism than that of the reference wild-type AAV capsid.

[0356] The engineered AAV capsid can contain 1-60 engineered capsid proteins. In some embodiments, the engineered AAV capsids can contain 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, or 60 engineered

capsid proteins. In some embodiments, the engineered AAV capsid can contain 0-59 wild-type AAV capsid proteins. In some embodiments, the engineered AAV capsid can contain 0, 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, or 59 wild-type AAV capsid proteins.

[0357] In some embodiments, the engineered AAV capsid protein can have an *n*-mer amino acid insert (also referred interchangeably herein as an “*n*-mer motif” or “*n*-mer insert”), where *n* can be at least 3 amino acids. In some embodiments, *n* can be 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids. In some embodiments, the engineered AAV capsid can have a 6-mer or 7-mer amino acid insert. In some embodiments, the *n*-mer amino acid insert can be inserted between two amino acids in the wild-type viral protein (VP) (or capsid protein). In some embodiments, the *n*-mer insert can be inserted between two amino acids in a variable amino acid region in an AAV capsid protein. The core of each wild-type AAV viral protein contains an eight-stranded beta-barrel motif (betaB to betaI) and an alpha-helix (alphaA) that are conserved in autonomous parovirus capsids (see e.g., DiMattia et al. 2012. J. Virol. 86(12):6947-6958). Structural variable regions (VRs) occur in the surface loops that connect the beta-strands, which cluster to produce local variations in the capsid surface. AAVs have 12 variable regions (also referred to as hypervariable regions) (see e.g., Weitzman and Linden. 2011. “Adeno-Associated Virus Biology.” In Snyder, R.O., Moullier, P. (eds.) Totowa, NJ: Humana Press). In some embodiments, one or more *n*-mer inserts can be inserted between two amino acids in one or more of the 12 variable regions in the wild-type AAV capsid proteins. In some embodiments, the one or more *n*-mer inserts can be each be inserted between two amino acids in VR-I, VR-II, VR-III, VR-IV, VR-V, VR-VI, VR-VII, VR-III, VR-IX, VR-X, VR-XI, VR-XII, or a combination thereof. In some embodiments, the *n*-mer can be inserted between two amino acids in the VR-III of a capsid protein. In some embodiments, the engineered capsid can have an *n*-mer inserted between any two contiguous amino acids between amino acids 262 and 269, between any two contiguous amino acids between amino acids 327 and 332, between any two contiguous amino acids between amino acids 382 and 386, between any two contiguous amino acids between amino acids 452 and 460, between any two contiguous amino acids between amino acids 488 and 505, between any two contiguous amino acids between amino acids 545 and 558, between any two contiguous amino acids between amino acids 581 and 593, between any two contiguous amino acids between amino

acids 704 and 714 of an AAV9 viral protein. In some embodiments, the engineered capsid can have an *n*-mer inserted between amino acids 588 and 589 of an AAV9 viral protein. In some embodiments, the engineered capsid can have a 7-mer insert inserted between amino acids 588 and 589 of an AAV9 viral protein. SEQ ID NO: 1 is a reference AAV9 capsid sequence for at least referencing the insertion sites discussed above. It will be appreciated that *n*-mers can be inserted in analogous positions in AAV viral proteins of other serotypes. In some embodiments as previously discussed, the *n*-mer(s) can be inserted between any two contiguous amino acids within the AAV viral protein and in some embodiments the insertion is made in a variable region.

[0358] SEQ ID NO: 1 AAV9 capsid (wild-type) reference Sequence:

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLG
 PGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDT
 SFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQ
 PAKKRLNFGQTGDTEVPDPQPIGEPPAAPSGVGSMTMASGGGAPVADNNEGADGV
 GSSSGNWHCDSQWLGDREVTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGY
 STPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTI
 ANNLSTTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVG
 RSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSY AHSQSLDRLMNPLIDQYLYLS
 KTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWP
 GASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVM
 TNEEEIKTTNPVATESYGVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQ
 GPIWAKIPHTDGNFHPSPLMGGFGMKHPPQILIKNTVPADPPTAFNKDKLNSFITQ
 YSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTR
 YLTRNL

[0359] In some embodiments, an AAV capsid can contain one or more targeting moieties having one or more *n*-mer motifs that contain a P-motifs. P-motifs are described in greater detail elsewhere herein. In some embodiments, an AAV capsid can contain one or more targeting moieties having one or more *n*-mer motifs that are each immediately preceded by AQ and wherein the *n*-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLG DAS (SEQ ID NO: 5), WVLPSSG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13),

VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESGH (SEQ ID NO: 19); or GENSARW (SEQ ID NO: 20). In some embodiments, an AAV capsid can contain one or more targeting moieties having one or more *n*-mer motifs that are each immediately preceded by DG and wherein the *n*-mer insert is REQQKLW (SEQ ID NO: 21), ASNPGRW (SEQ ID NO: 22), WTLESGH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25); or RMQRTLY (SEQ ID NO: 26). In some embodiments, the *n*-mer in an AAV capsid, such as an CNS-specific AAV capsid can be any one or more as set forth in **Tables 1-3**. In some embodiments, insertion of the *n*-mer in an AAV capsid can result in cell, tissue, organ, specific engineered AAV capsids. In some embodiments, the engineered viral protein, engineered viral capsid protein, engineered viral capsid, and/or engineered viral particle can have a specificity for bone tissue and/or cells, lung tissue and/or cells, liver tissues and/or cells, bladder tissue and/or cells, kidney tissue and/or cells, cardiac tissue and/or cells, skeletal muscle tissue and/or cells, smooth muscle and/or cells, neuronal tissue and/or cells, intestinal tissue and/or cells, pancreases tissue and/or cells, adrenal gland tissue and/or cells, brain tissue and/or cells, tendon tissues or cells, skin tissues and/or cells, spleen tissue and/or cells, eye tissue and/or cells, blood cells, synovial fluid cells, immune cells (including specificity for particular types of immune cells), and combinations thereof.

[0360] In some embodiments, the engineered viral protein, engineered viral capsid protein, engineered viral capsid, and/or engineered viral particle can have a specificity for CNS cells and/or tissue.

[0361] In some embodiments, the CNS AAV capsid contains an RGD insert. In some embodiments, the CNS AAV capsid does not contain an RGD insert. RGD motifs are described in greater detail elsewhere herein.

[0362] In some embodiments, the *n*-mer inserts includes a “P motif” (also referred to interchangeably herein as a “P insert”). As used herein, the term “P motif” can refer to the motif PX_1QGTX_2R (SEQ ID NO: 64 or 2), where X_1 and X_2 can each be selected from any amino acid. In some embodiments, X_1 is S, T, or A. In some embodiments, X_2 is L, V, F, or I. In some embodiments, X_1 is S, T, or A and X_2 is L, V, F, or I. Exemplary, non-limiting P motifs are shown at least in e.g., **Table 3**. P-inserts are further described elsewhere herein.

[0363] In some embodiments, one or more *n*-mer inserts can be as set forth in any one or more of **Tables 1, 2, or 3** can be included in a CNS specific engineered capsid.

[0364] As is demonstrated in **Table 1** and the Working Examples, the n-mer insert (such as a 7-mer insert) can be inserted into an AAV vector between two contiguous amino acids where the amino acids in the AAV vector immediately preceding the n-mer insert can be DG or AQ. It will be appreciated that the DG and AQ are not considered part of the n-mer insert as the term is used herein. In other words, the first amino acid of the n-mer insert is the third amino acid in each of the sequences shown in **Table 1**. In each case, these two amino acids are either AQ or DG. Each n-mer insert was tested in both configurations (e.g., with AQ and DG as amino acids 587 and 588 of the AAV). **Table 1** represents exemplary variants having CNS transduction efficiency. As further discussed in the Working Examples herein, these engineered AAV variants were able to transduce cells from multiple strains of mice. This is in contrast to other AAVs, which at least in some cases, can only transduce certain strains of mice. In some embodiments, amino acids 587 and 588 of the AAV or analogous amino acids thereto are DG. In some embodiments, amino acids 587 and 588 of the AAV or analogous amino acids thereto are AQ.

[0365] In some embodiments, amino acids 587 and 588 of the AAV or analogous amino acids thereto are AQ and are followed by a 7-mer amino acid insert. In some embodiments, amino acids 587 and 588 of the AAV or analogous amino acids thereto are AQ and are followed by a 7-mer amino acid insert, where the 7-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSSG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESQH (SEQ ID NO: 19); or GENSARW (SEQ ID NO: 20).

[0366] In some embodiments, amino acids 587 and 588 of the AAV or analogous amino acids thereto are DG and are followed by a 7-mer amino acid insert. In some embodiments, amino acids 587 and 588 of the AAV or analogous amino acids thereto are DG and are followed by a 7-mer amino acid insert, where the 7-mer insert is REQKLY (SEQ ID NO: 64), ASNPGRW (SEQ ID NO: 22), WTLESQH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25); or RMQRTLY (SEQ ID NO: 26).

[0367] In some embodiments, the AAV capsids can be CNS-specific. In some embodiments, CNS-specificity of the engineered AAV capsid is conferred by a CNS specific

n-mer insert incorporated in the engineered AAV capsid. While not intending to be bound by theory, it is believed that the n-mer insert confers a 3D structure to or within a domain or region of the engineered AAV capsid such that the interaction of an engineered AAV containing said engineered AAV capsid has increased or improved interactions (e.g., increased affinity) with a cell surface receptor and/or other molecule on the surface of an endothelial and/or a CNS cell. In some embodiments the cell surface receptor is AAV receptor (AAVR). In some embodiments, the cell surface receptor is a CNS cell specific AAV receptor. In some embodiments, a CNS specific engineered AAV containing the CNS-specific capsid can have an increased transduction rate, efficiency, amount, or a combination thereof in a CNS cell as compared to other cell types and/or other AAVs that do not contain a muscle-specific engineered AAV capsid as described herein.

Table 1. Exemplary CNS <i>n</i>-mer inserts		
Initial "AQ" or "DG" correspond to the two amino acids in the targeting moiety that immediately precede the "<i>n</i>-mer insert" in a targeting moiety or composition (e.g., AA 587 and 588 of an AAV9 that has an <i>n</i>-mer insert placed between AA 588 and 589).	CNS Transduction efficiency Score	SEQ ID NO:
AQRSVGSVY	46000	65
AQKTVGTVY	45980	66
AQRYLGDAS	40592	67
DGREQQKLW	39151	68
AQWVLPSGG	37597	69
AQVTVGSIY	32968	70
AQVRGSSIL	32330	71
AQRHHGDAA	32171	72
AQVIQAMKL	32127	73
AQLTYGMAQ	31956	74
AQLRIGLSQ	31710	75
AQGDYSMIV	31497	76
AQVNYSVAL	31271	77
AQRYSGDAS	31198	78
AQRYSGDSV	29860	79
AQRHIADAS	29554	80
AQRYLGDAT	29527	81
AQQRVGFAQ	29454	82
AQIAHGYST	28216	83
AQWTLESGH	27471	84
AQGENSARW	27287	85
DGASNPGRW	24583	86
AQLAVGQKW	24445	87
AQVKLGYSQ	23912	88
AQEAGSARW	23888	89
AQLNYSVSL	21972	90
AQW AISDGY	21970	91
AQRGPGLSQ	21738	92
DGWTLESGH	20534	93
AQRYVGESS	19635	94
DGREQKKLW	17695	95
AQFTLTTPK	15607	96
DGERLLVQL	15513	97
AQEDLLRLR	14920	98

AQPIIEHAV	12837	99
DGRMQRTLY	12453	100
DGWAISDGY	10828	101
AQRYISDSA	10788	102
AQWSTSSGF	10614	103
AQWSLGSGH	10498	104
AQWSQSSGY	10258	105
DGVRGSSIL	9714	106
AQIMLGYST	9404	107
DGKLADSVP	9356	108
AQASNPGRW	9173	109
AQHVENWHI	8680	110
AQVAGSSIL	8645	111
DGRQQQKLW	8393	112
DGTVNNDRF	8028	113
DGMSANERT	8000	114
AQATVAGQF	7885	115
DGRDQQKLW	7761	116
AQGKSPGVW	7685	117
DGGASNGGT	7674	118
AQSLVTSST	6782	119
AQLLYGYSS	6779	120
DGVTELTKF	6655	121
AQALVQNGV	6638	122
AQVLESNPR	6572	123
AQPASHEVL	6460	124
AQAGVQNAL	6452	125
DGKEISVSV	6420	126
AQGLNERVA	6410	127
DGQVAQQGA	6393	128
DGGVAGTNT	6386	129
DGASAQ GAL	6382	130
AQAGVSSQT	6357	131
AQKNRRHSV	6312	132
AQKVDSAQL	6311	133
AQYTLSQGW	6310	134
DGQSVDRSK	6293	135
AQASASSPR	6266	136
DGRYVGESS	6252	137
DGLGHNAGV	6239	138
AQPNERINV	6142	139
AQVMSGTSH	6122	140
DGVLVSPGP	6000	141

DGVGISSGV	6000	142
DGSGETLRI	5977	143
DGSTEGAAL	5954	144
AQTSLSQDR	5943	145
AQSANPVVT	5937	146
DGVLASNGP	5898	147
AQAHLDNAP	5893	148
DGVVQVTGR	5875	149
DGFAVRLSS	5855	150
DGLVRDTKT	5811	151
DSGESLSR	5804	152
AQTNEQAQR	5796	153
DGTLANSQR	5746	154
AQLLADKSV	5680	155
DGSQEQRAR	5679	156
AQVNGNTTY	5655	157
AQALAEAGA	5624	158
DGSREGGNV	5580	159
AQMGDSVTI	5574	160
DGLGGSSMG	5565	161
AQGVRTDNI	5562	162
DGSGSTDKL	5556	163
AQASQNSTV	5493	164
AQGGTSSGH	5462	165
AQAADSSVR	5404	166
AQAANSSVR	5387	167
AQWADSKDQ	5374	168
AQPTQGTVR	5353	169
AQGSTDFTK	5344	170
AQVDHGGVV	5342	171
AQGEQQKGW	5322	172
DGIANLAAS	5311	173
DGAGGVRDR	5299	174
DGGSGSGGL	5252	175
DGTLANSER	5237	176
AQKGASVTL	5236	177
AQSNVALTG	5235	178
DGVNYSVAL	5206	179
AQGLNEHGA	5193	180
DGKNPGVYT	5173	181
DGQREAARI	5173	182
AQGLVDSSR	5168	183
DGNGSEGDR	5157	184

DGNVGVVQL	5144	185
AQVTDGVRS	5109	186
AQVIASNEH	5109	187
AQMSVGQSW	5098	188
DGHSLQTSA	5096	189
AQQDGYGTR	5093	190
AQLSNGQGP	5071	191
AQPVTDSKM	5068	192
AQNGTAADR	5057	193
AQIIVDNGS	5024	194
AQEADNHGR	5023	195
AQAADSSGR	4995	196
AQVVDSNNL	4986	197
DGSGANLSY	4985	198
DGKAHDGEV	4978	199

Table 2. Exemplary CNS n-mer inserts and/or P-motifs

Rank	N-mer insert	SEQ ID NO:	Encoding sequence	SEQ ID NO:
1	PSQGTLR	200	CCTTCTCAGGGGACGCTTCGG	201
2	TDALTTK	202	ACTGATGCGCTTACGACTAAG	203
3	PTQGTVR	204	CCCACACAAGGCACAGTCCGT	205
4	PTQGTLR	206	CCTACTCAGGGGACGCTTCGG	207
5	PTQGTVR	208	CCTACTCAGGGGACGGTTTCGG	209
6	STIPTMK	210	AGTACTATTCCTACTATGAAG	211
7	TDAGDGK	212	ACAGACGCGGGGGACGGCAAA	213
8	YQRTESE	214	TATCAGAGGACGGAGTCTCTG	215
9	RVDPSGL	216	AGAGTCGACCCCAGTGGACTA	217
10	SLVTSST	218	TCGCTTGTTACTTCTAGTACG	219
11	LLAGADR	220	TTGCTTGCTGGTGCTGATCGT	221
12	STDRESR	222	TCCACGGACCGTGAAAGCCGA	223
13	NGYTEGR	224	AATGGGTATACGGAGGGGCGT	225
14	PTQGTFR	226	CCGACACAAGGAACATTCAGG	227
15	MTGISIV	228	ATGACAGGCATCTCTATCGTA	229
16	DGRAELR	230	GATGGGCGGGCGGAGTTGCGT	231
17	AADSSAR	232	GCCGCTGACTCATCGGCCCGT	233
18	PTQGTIR	234	CCTACTCAGGGGACGATTCGG	235
19	LSRGEEK	236	CTTTCGAGGGGTGAGGAGAAG	237
20	AIVSIAQ	238	GCGATTGTGTCGATTGCTCAG	239
21	LTSGLAA	240	TTGACGTCTGGTTTGGCGGCG	241
22	PTQGTFR	242	CCTACTCAGGGGACGTTTCGG	243

23	TLAISGR	244	ACTTTGGCGATTTCTGGGCGG	245
24	VHSQDVS	246	GTCCACAGTCAAGACGTTTCC	247
25	FQVEQVK	248	TTTCAGGTTGAGCAGGTTAAG	249
26	NRELALG	250	AACCGCGAACTCGCACTCGGG	251
27	SIGDLGK	252	AGTATCGGTGACCTAGGTAAA	253
28	TVGHDK	254	ACCGTAGGACACGACAACAAA	255
29	HSKGFY	256	CACAGTAAAGGTTTCGACTAC	257
30	HTQGTLR	258	CATACTCAGGGGACGCTTCGG	259
31	PAQGTLR	260	CCGGCGCAAGGAACACTACGA	261
32	AGGGDPR	262	GCTGGTGGAGGTGACCCCCGA	263
33	LGKADPV	264	TTGGGAAAAGCTGACCCAGTA	265
34	ALNEHVA	266	GCTCTGAATGAGCATGTGGCG	267
35	GSGGVSV	268	GGTTCGGGTGGTGTAGTGTG	269
36	PSQGTLR	270	CCGTCCCAAGGAACACTCAGG	271
37	TGGRDQY	272	ACTGGTGGTCGGGATCAGTAT	273
38	YLVTTEN	274	TATTTGGTTACTACTGAGAAT	275
39	LSRDVAV	276	TTGTTCGAGGGATGTGGCGGTT	277
40	RIVDSVP	278	AGGATTGTGGATAGTGTTCCG	279
41	KGYDTPM	280	AAAGGCTACGACACACCCATG	281
42	TSREEQW	282	ACTTCTCGTGAGGAGCAGTGG	283
43	RASADV	284	AGGGCGAGTGCGGATGTTGTG	285
44	NLGAALS	286	AACCTTGGGGCTGCCCTATCG	287
45	SVTDIKH	288	TCGGTGACGGACATAAAACAC	289
46	FQDTIGV	290	TTTCAGGATACGATTGGGGTG	291
47	PNERLAV	292	CCTAACGAACGATTGGCAGTC	293
48	HTIAASM	294	CACACCATAGCCGCAAGTATG	295
49	NSDLMGR	296	AACAGTGACCTAATGGGCCGA	297
50	AGVSASL	298	GCGGGTGTTTCTGCGTCGTTG	299

Table 3. Exemplary P-motifs			
n-mer motif	SEQ ID NO:	Encoding Sequence(s)	SEQ ID NO:
PSQGTLR	300	CCTTCTCAGGGGACGCTTCGG;	301
		CCGTCCCAAGGAACACTCAGG	302
PTQGTVR	303	CCCACACAAGGCACAGTCCGT;	304
		CCTACTCAGGGGACGGTTCGG	305
PTQGTLR	306	CCTACTCAGGGGACGCTTCGG	307

PTQGTFR	308	CCGACACAAGGAACATTCAGG; CCTACTCAGGGGACGTTTCGG	309 310
PTQGTIR	311	CCTACTCAGGGGACGATTCGG	312
PAQGTLR	313	CCGGCGCAAGGAACACTACGA	314

[0368] Also described herein are polynucleotides that encode the engineered targeting moieties, viral proteins (e.g., capsid proteins), and other polypeptides described herein, including but not limited to, the engineered AAV capsids described herein. In some embodiments, the engineered AAV capsid encoding polynucleotide can be included in a polynucleotide that is configured to be an AAV genome donor in an AAV vector system that can be used to generate engineered AAV particles described elsewhere herein.

[0369] In some embodiments, the AAV capsids or other viral capsids or compositions can be CNS-specific. In some embodiments, CNS-specificity of the engineered AAV or other viral capsid or other composition is conferred by one or more CNS specific n-mer motifs incorporated in the engineered AAV or other viral capsid or other composition described herein. While not intending to be bound by theory, it is believed that the n-mer motif confers a 3D structure to or within a domain or region of the engineered AAV capsid or other viral capsid or other composition such that the interaction of the viral particle or other composition containing the engineered AAV capsid or other viral capsid or other composition described herein has increased or improved interactions (e.g., increased affinity) with a cell surface receptor and/or other molecule on the surface of a CNS cell. In some embodiments, the cell surface receptor is AAV receptor (AAVR). In some embodiments, the cell surface receptor is a CNS cell specific AAV receptor. In some embodiments, the cell surface receptor or other molecule is a cell surface receptor or other molecule selectively expressed on the surface of a CNS cell.

[0370] In some embodiments the engineered viral (e.g., AAV) capsid encoding polynucleotide can be operably coupled to a poly adenylation tail. In some embodiments, the poly adenylation tail can be an SV40 poly adenylation tail. In some embodiments, the viral (e.g., AAV) capsid encoding polynucleotide can be operably coupled to a promoter. In some embodiments, the promoter can be a tissue specific promoter. In some embodiments, the tissue specific promoter is specific for muscle (e.g., cardiac, skeletal, and/or smooth muscle), neurons and supporting cells (e.g., astrocytes, glial cells, Schwann cells, etc.), fat, spleen, liver, kidney,

immune cells, spinal fluid cells, synovial fluid cells, skin cells, cartilage, tendons, connective tissue, bone, pancreas, adrenal gland, blood cell, bone marrow cells, placenta, endothelial cells, and combinations thereof. In some embodiments the promoter can be a constitutive promoter. Suitable tissue specific promoters and constitutive promoters are discussed elsewhere herein and are generally known in the art and can be commercially available.

[0371] Suitable neuronal tissue/cell specific promoters include, but are not limited to, GFAP promoter (astrocytes), SYN1 promoter (neurons), and NSE/RU5⁺ (mature neurons).

[0372] Other suitable CNS specific promoters can include, but are not limited to, neuroactive peptide cholecystokinin (CCK) (see e.g., Chhatawl et al. Gene Therapy volume 14, pages575–583(2007)), a brain specific DNA MiniPromoter (such as any of those identified for brain or pan-neronal expression as in de Leeuw et al. Mol. Therapy. 1(5): 2014. doi:10.1038/mtm.2013.5), myelin basic promoter (MBP) (see e.g., von Jonquieres, G., Mersmann, N., Klugmann, C. B., Harasta, A. E., Lutz, B., Teahan, O., et al. (2013). Glial promoter selectivity following AAV-delivery to the immature brain. PLoS One 8 (6), e65646. doi: 10.1371/journal.pone.0065646), glial fibrillary acid protein (GFAP) for expression in astrocytes (see e.g., Smith-Arica, J. R., Morelli, A. E., Larregina, A. T., Smith, J., Lowenstein, P. R., Castro, M. G. (2000). Cell-type-specific and regulatable transgenesis in the adult brain: adenovirus-encoded combined transcriptional targeting and inducible transgene expression. Mol. Ther. 2 (6), 579–587. doi: 10.1006/mthe.2000.0215 and Lee, Y., Messing, A., Su, M., Brenner, M. (2008). GFAP promoter elements required for region-specific and astrocyte-specific expression. Glia 56 (5), 481–493. doi: 10.1002/glia.20622), human myelin associated glycoprotein promoter (full-length or truncated) (see e.g., von Jonquieres, G., Frohlich, D., Klugmann, C. B., Wen, X., Harasta, A. E., Ramkumar, R., et al. (2016). Recombinant human myelin-associated glycoprotein promoter drives selective AAV-mediated transgene expression in oligodendrocytes. Front. Mol. Neurosci. 9, 13. doi: 10.3389/fnmol.2016.00013), F4/80 promoter (see e.g., Rosario, A. M., Cruz, P. E., Ceballos-Diaz, C., Strickland, M. R., Sieminski, Z., Pardo, M., et al. (2016). Microglia-specific targeting by novel capsid-modified AAV6 vectors. Mol. Ther. Methods Clin. Dev. 3, 16026. doi: 10.1038/mtm.2016.26), phosphate-activated glutaminase (PAG) or the vesicular glutamate transporter (vGLUT) promoter (for about 90% glutamatergic neuron-specific expression) (see e.g., Rasmussen, M., Kong, L., Zhang, G. R., Liu, M., Wang, X., Szabo, G., et al. (2007). Glutamatergic or GABAergic neuron-specific, long-term expression in neocortical neurons from helper virus-

free HSV-1 vectors containing the phosphate-activated glutaminase, vesicular glutamate transporter-1, or glutamic acid decarboxylase promoter. Brain Res. 1144, 19–32. doi: 10.1016/j.brainres.2007.01.125), glutamic acid decarboxylase (GAD) promoter (for about 90% GABAergic neuron-specific expression) (see e.g., Rasmussen, M., Kong, L., Zhang, G. R., Liu, M., Wang, X., Szabo, G., et al. (2007). Glutamatergic or GABAergic neuron-specific, long-term expression in neocortical neurons from helper virus-free HSV-1 vectors containing the phosphate-activated glutaminase, vesicular glutamate transporter-1, or glutamic acid decarboxylase promoter. Brain Res. 1144, 19–32. doi: 10.1016/j.brainres.2007.01.125), MeCP2 promoter (see e.g., Gray et al. Hum Gene Ther. 2011 Sep;22(9):1143-53. doi: 10.1089/hum.2010.245), and retinoblastoma gene promoter (see e.g., Jiang et al., J. Biol. Chem. 2001. 276, 593-600).

[0373] Suitable constitutive promoters include, but are not limited to CMV, RSV, SV40, EF1alpha, CAG, and beta-actin.

AAVs with reduced Non-CNS cell Specificity

[0374] In some embodiments, the n-mer insert(s) and/or P-motif(s) are inserted into an AAV protein (e.g., an AAV capsid protein) that has reduced specificity (or no detectable, measurable, or clinically relevant interaction) for one or more non-CNS cell types. Exemplary non-CNS cell types include, but are not limited to, liver, kidney, lung, heart, spleen, muscle (skeletal and cardiac), bone, immune, stomach, intestine, eye, skin cells and the like. In some embodiments, the non-CNS cells are liver cells.

[0375] In certain example embodiments, the AAV capsid protein is an engineered AAV capsid protein having reduced or eliminated uptake in a non-CNS cell as compared to a corresponding wild-type AAV capsid polypeptide.

[0376] In certain example embodiments, the non-CNS cell is a liver cell.

[0377] In certain example embodiments, the wild-type capsid polypeptide is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 capsid polypeptide.

[0378] In certain example embodiments, the engineered AAV capsid protein comprises one or more mutations that result in reduced or eliminated uptake in a non-CNS cell.

[0379] In certain example embodiments, the one or more mutations are

- a. in position 267,
- b. in position 269,

- c. in position 504,
- d. in position 505,
- e. in position 590,
- f. or any combination thereof

in the AAV9 capsid protein (SEQ ID NO: 1) or in one or more positions corresponding thereto in a non-AAV9 capsid polypeptide.

[0380] In certain example embodiments, the non-AAV9 capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, or AAV rh.10 capsid polypeptide.

[0381] In certain example embodiments, the mutation in position 267 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X mutation to A, wherein X is any amino acid.

[0382] In certain example embodiments, the mutation in position 269 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is an S or X to T mutation, wherein X is any amino acid.

[0383] In certain example embodiments, the mutation in position 504 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X to A mutation, wherein X is any amino acid.

[0384] In certain example embodiments, the mutation in position 505 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a P or X to A mutation, wherein X is any amino acid.

[0385] In certain example embodiments, the mutation in position 590 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a Q or X to A mutation, wherein X is any amino acid.

[0386] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 267, position 269 or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 267 is a G to A mutation and wherein the mutation at position 269 is an S to T mutation.

[0387] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 590 of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 509 is a Q to A mutation.

[0388] In certain example embodiments, the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 504, position 505, or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 504 is a G to A mutation and wherein the mutation at position 505 is a P to A mutation.

[0389] In some embodiments, the AAV capsid protein in which the n-mer motif(s) and/or P motif(s) can be inserted can be 80-100 (e.g., 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 to SEQ ID NO: 4 or SEQ ID NO: 5 of International Patent Application Publication WO 2019/217911, which is incorporated by reference as if expressed in its entirety herein. These sequences are also incorporated herein as SEQ ID NOS: 330 and 331 respectively. It will be appreciated that when considering variants of these AAV9 capsid proteins with reduced liver specificity, that residues 267 and/or 269 must contain the relevant mutations or equivalents.

[0390] In some embodiments, the AAV capsid protein in which the n-mer motif(s) and/or P motif(s) can be inserted can be 80-100 (e.g., 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 to any of those described in Adachi et al., (Nat. Comm. 2014. 5:3075, DOI: 10.1038/ncomms4075) that have reduced specificity for a non-CNS cell, particularly a liver cell. Adachi et al., (Nat. Comm. 2014. 5:3075, DOI: 10.1038/ncomms4075) is incorporated by reference herein as if expressed in its entirety.

[0391] In some embodiments, the modified AAV can have about a 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, 100 percent or fold reduction in specificity for a non-CNS cells as compared to a wild-type AAV or control. In some embodiments, the modified AAV can have no measurable or detectable uptake and/or expression in one or more non-CNS cells.

Methods of Generating Engineered AAV Capsids

[0392] Also provided herein are methods of generating engineered AAV capsids. The engineered AAV capsid variants can be variants of wild-type AAV capsids. **FIGS. 6A-8** can illustrate various embodiments of methods capable of generating engineered AAV capsids described herein. Generally, an AAV capsid library can be generated by expressing engineered

capsid vectors each containing an engineered AAV capsid polynucleotide previously described in an appropriate AAV producer cell line. See e.g., **FIG. 8**. It will be appreciated that although **FIG. 8** shows a helper-dependent method of AAV particle production, it will be appreciated that this can be done via a helper-free method as well. This can generate an AAV capsid library that can contain one more desired cell-specific engineered AAV capsid variant. As shown in **FIG. 6** the AAV capsid library can be administered to various non-human animals for a first round of mRNA-based selection. As shown in **FIG. 1**, the transduction process by AAVs and related vectors can result in the production of an mRNA molecule that is reflective of the genome of the virus that transduced the cell. As is at least demonstrated in the Examples herein, mRNA based-selection can be more specific and effective to determine a virus particle capable of functionally transducing a cell because it is based on the functional product produced as opposed to just detecting the presence of a virus particle in the cell by measuring the presence of viral DNA.

[0393] After first-round administration, one or more engineered AAV virus particles having a desired capsid variant can then be used to form a filtered AAV capsid library. Desirable AAV virus particles can be identified by measuring the mRNA expression of the capsid variants and determining which variants are highly expressed in the desired cell type(s) as compared to non-desired cells type(s). Those that are highly expressed in the desired cell, tissue, and/or organ type are the desired AAV capsid variant particles. In some embodiments, the AAV capsid variant encoding polynucleotide is under control of a tissue-specific promoter that has selective activity in the desired cell, tissue, or organ.

[0394] The engineered AAV capsid variant particles identified from the first round can then be administered to various non-human animals. In some embodiments, the animals used in the second round of selection and identification are not the same as those animals used for first round selection and identification. Similar to round 1, after administration the top expressing variants in the desired cell, tissue, and/or organ type(s) can be identified by measuring viral mRNA expression in the cells. The top variants identified after round two can then be optionally barcoded and optionally pooled. In some embodiments, top variants from the second round can then be administered to a non-human primate to identify the top cell-specific variant(s), particularly if the end use for the top variant is in humans. Administration at each round can be systemic.

[0395] In some embodiments, the method of generating an AAV capsid variant can include the steps of: (a) expressing a vector system described herein that contains an engineered AAV capsid polynucleotide in a cell to produce engineered AAV virus particle capsid variants; (b) harvesting the engineered AAV virus particle capsid variants produced in step (a); (c) administering engineered AAV virus particle capsid variants to one or more first subjects, wherein the engineered AAV virus particle capsid variants are produced by expressing an engineered AAV capsid variant vector or system thereof in a cell and harvesting the engineered AAV virus particle capsid variants produced by the cell; and (d) identifying one or more engineered AAV capsid variants produced at a significantly high level by one or more specific cells or specific cell types in the one or more first subjects. In this context, “significantly high” can refer to a titer that can range from between about 2×10^{11} to about 6×10^{12} vector genomes per 15 cm dish.

[0396] The method can further include the steps of: (e) administering some or all engineered AAV virus particle capsid variants identified in step (d) to one or more second subjects; and (f) identifying one or more engineered AAV virus particle capsid variants produced at a significantly high level in one or more specific cells or specific cell types in the one or more second subjects. The cell in step (a) can be a prokaryotic cell or a eukaryotic cell. In some embodiments, the administration in step (c), step (e), or both is systemic. In some embodiments, one or more first subjects, one or more second subjects, or both, are non-human mammals. In some embodiments, one or more first subjects, one or more second subjects, or both, are each independently selected from the group consisting of: a wild-type non-human mammal, a humanized non-human mammal, a disease-specific non-human mammal model, and a non-human primate.

Engineered Vectors and Vector Systems

[0397] Also provided herein are vectors and vector systems that can contain one or more of the engineered polynucleotides, (e.g., an AAV capsid polynucleotide) described herein. As used in this context, engineered viral (e.g., AAV) capsid polynucleotides refers to any one or more of the polynucleotides described herein capable of encoding an engineered viral (e.g., AAV) capsid as described elsewhere herein and/or polynucleotide(s) capable of encoding one or more engineered viral (e.g., AAV) capsid proteins described elsewhere herein. Further, where the vector includes an engineered viral (e.g., AAV) capsid polynucleotide described herein, the vector can also be referred to and considered an engineered vector or system thereof

although not specifically noted as such. In embodiments, the vector can contain one or more polynucleotides encoding one or more elements of an engineered viral (e.g., AAV) capsid described herein. The vectors can be useful in producing bacterial, fungal, yeast, plant cells, animal cells, and transgenic animals that can express one or more components of the engineered viral (e.g., AAV) capsid described herein. Within the scope of this disclosure are vectors containing one or more of the polynucleotide sequences described herein. One or more of the polynucleotides that are part of the engineered viral (e.g., AAV) capsid and system thereof described herein can be included in a vector or vector system.

[0398] In some embodiments, the vector can include an engineered viral (e.g., AAV) capsid polynucleotide having a 3' polyadenylation signal. In some embodiments, the 3' polyadenylation is an SV40 polyadenylation signal. In some embodiments the vector does not have splice regulatory elements. In some embodiments, the vector includes one or more minimal splice regulatory elements. In some embodiments, the vector can further include a modified splice regulatory element, wherein the modification inactivates the splice regulatory element. In some embodiments, the modified splice regulatory element is a polynucleotide sequence sufficient to induce splicing, between a rep protein polynucleotide and the engineered viral (e.g., AAV) capsid protein variant polynucleotide. In some embodiments, the polynucleotide sequence can be sufficient to induce splicing is a splice acceptor or a splice donor. In some embodiments, the viral (e.g., AAV) capsid polynucleotide is an engineered viral (e.g., AAV) capsid polynucleotide as described elsewhere herein.

[0399] The vectors and/or vector systems can be used, for example, to express one or more of the engineered viral (e.g., AAV) capsid polynucleotides in a cell, such as a producer cell, to produce engineered viral (e.g., AAV) particles containing an engineered viral (e.g., AAV) capsid described elsewhere herein. Other uses for the vectors and vector systems described herein are also within the scope of this disclosure. In general, and throughout this specification, the term is a tool that allows or facilitates the transfer of an entity from one environment to another. In some contexts which will be appreciated by those of ordinary skill in the art, "vector" can be a term of art to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. A vector can be a replicon, such as a plasmid, phage, or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment. Generally, a vector is capable of replication when associated with the proper control elements.

[0400] Vectors include, but are not limited to, nucleic acid molecules that are single-stranded, double-stranded, or partially double-stranded; nucleic acid molecules that comprise one or more free ends, no free ends (e.g., circular); nucleic acid molecules that comprise DNA, RNA, or both; and other varieties of polynucleotides known in the art. One type of vector is a “plasmid,” which refers to a circular double stranded DNA loop into which additional DNA segments can be inserted, such as by standard molecular cloning techniques. Another type of vector is a viral vector, wherein virally-derived DNA or RNA sequences are present in the vector for packaging into a virus (e.g., retroviruses, replication defective retroviruses, adenoviruses, replication defective adenoviruses, and adeno-associated viruses (AAVs)). Viral vectors also include polynucleotides carried by a virus for transfection into a host cell. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as “expression vectors.” Common expression vectors of utility in recombinant DNA techniques are often in the form of plasmids.

[0401] Recombinant expression vectors can be composed of a nucleic acid (e.g., a polynucleotide) of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory elements, which can be selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, “operably linked” and “operatively-linked” are used interchangeably herein and further defined elsewhere herein. In the context of a vector, the term “operably linked” is intended to mean that the nucleotide sequence of interest is linked to the regulatory element(s) in a manner that allows for expression of the nucleotide sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). Advantageous vectors include adeno-associated viruses, and types of such vectors can also be selected for targeting particular types of cells, such as those engineered viral (e.g., AAV) vectors containing an engineered viral (e.g., AAV) capsid polynucleotide with a desired

cell-specific tropism. These and other embodiments of the vectors and vector systems are described elsewhere herein.

[0402] In some embodiments, the vector can be a bicistronic vector. In some embodiments, a bicistronic vector can be used for one or more elements of the engineered viral (e.g., AAV) capsid system described herein. In some embodiments, expression of elements of the engineered viral (e.g., AAV) capsid system described herein can be driven by a suitable constitutive or tissue specific promoter. Where the element of the engineered viral (e.g., AAV) capsid system is an RNA, its expression can be driven by a Pol III promoter, such as a U6 promoter. In some embodiments, the two are combined.

Cell-based Vector Amplification and Expression

[0403] Vectors can be designed for expression of one or more elements of the engineered targeting moieties, polypeptides, viral (e.g., AAV) capsid system described herein (e.g., nucleic acid transcripts, proteins, enzymes, and combinations thereof), etc. in a suitable host cell. In some embodiments, the suitable host cell is a prokaryotic cell. Suitable host cells include, but are not limited to, bacterial cells, yeast cells, insect cells, and mammalian cells. The vectors can be viral-based or non-viral based. In some embodiments, the suitable host cell is a eukaryotic cell. In some embodiments, the suitable host cell is a suitable bacterial cell. Suitable bacterial cells include, but are not limited to, bacterial cells from the bacteria of the species *Escherichia coli*. Many suitable strains of *E. coli* are known in the art for expression of vectors. These include, but are not limited to Pir1, Stbl2, Stbl3, Stbl4, TOP10, XL1 Blue, and XL10 Gold. In some embodiments, the host cell is a suitable insect cell. Suitable insect cells include those from *Spodoptera frugiperda*. Suitable strains of *S. frugiperda* cells include, but are not limited to, Sf9 and Sf21. In some embodiments, the host cell is a suitable yeast cell. In some embodiments, the yeast cell can be from *Saccharomyces cerevisiae*. In some embodiments, the host cell is a suitable mammalian cell. Many types of mammalian cells have been developed to express vectors. Suitable mammalian cells include, but are not limited to, HEK293, Chinese Hamster Ovary Cells (CHOs), mouse myeloma cells, HeLa, U2OS, A549, HT1080, CAD, P19, NIH 3T3, L929, N2a, MCF-7, Y79, SO-Rb50, HepG 2, DIKX-X11, J558L, Baby hamster kidney cells (BHK), and chicken embryo fibroblasts (CEFs). Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990).

[0404] In some embodiments, the vector can be a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerevisiae* include pYepSec1 (Baldari, et al., 1987. EMBO J. 6: 229-234), pMFa (Kuijan and Herskowitz, 1982. Cell 30: 933-943), pJRY88 (Schultz et al., 1987. Gene 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.). 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 can 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, 2 μ plasmids, yeast integrative plasmids, yeast replicative plasmids, shuttle vectors, and episomal plasmids.

[0405] In some embodiments, the vector is a baculovirus vector or expression vector and can be suitable for expression of polynucleotides and/or proteins in insect cells. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., SF9 cells) include the pAc series (Smith, et al., 1983. Mol. Cell. Biol. 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. Virology 170: 31-39). rAAV (recombinant Adeno-associated viral) vectors are preferably produced in insect cells, e.g., *Spodoptera frugiperda* Sf9 insect cells, grown in serum-free suspension culture. Serum-free insect cells can be purchased from commercial vendors, e.g., Sigma Aldrich (EX-CELL 405).

[0406] In some embodiments, the vector is a mammalian expression vector. In some embodiments, the mammalian expression vector is capable of expressing one or more polynucleotides and/or polypeptides in a mammalian cell. Examples of mammalian expression vectors include, but are not limited to, pCDM8 (Seed, 1987. Nature 329: 840) and pMT2PC (Kaufman, et al., 1987. EMBO J. 6: 187-195). The mammalian expression vector can include

one or more suitable regulatory elements capable of controlling expression of the one or more polynucleotides and/or proteins in the mammalian cell. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, simian virus 40, and others disclosed herein and known in the art. More detail on suitable regulatory elements is described elsewhere herein.

[0407] For other suitable expression vectors and vector systems for both prokaryotic and eukaryotic cells see, e.g., Chapters 16 and 17 of Sambrook, et al., **MOLECULAR CLONING: A LABORATORY MANUAL**. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

[0408] In some embodiments, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, et al., 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Baneiji, et al., 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, et al., 1985. *Science* 230: 912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and the α -fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546). With regards to these prokaryotic and eukaryotic vectors, mention is made of U.S. Patent 6,750,059, the contents of which are incorporated by reference herein in their entirety. Other embodiments can utilize viral vectors, with regards to which mention is made of U.S. Patent application 13/092,085, the contents of which are incorporated by reference herein in their entirety. Tissue-specific regulatory elements are known in the art and in this regard, mention is made of U.S. Patent 7,776,321, the contents of which are incorporated by reference herein in their entirety. In some embodiments, a regulatory element can be operably linked to one or more elements of an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system so as to drive expression of the one or more

elements of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein.

[0409] Vectors may be introduced and propagated in a prokaryote or prokaryotic cell. In some embodiments, a prokaryote is used to amplify copies of a vector to be introduced into a eukaryotic cell or as an intermediate vector in the production of a vector to be introduced into a eukaryotic cell (e.g., amplifying a plasmid as part of a viral vector packaging system). In some embodiments, a prokaryote is used to amplify copies of a vector and express one or more nucleic acids, such as to provide a source of one or more proteins for delivery to a host cell or host organism.

[0410] In some embodiments, the vector can be a fusion vector or fusion expression vector. In some embodiments, fusion vectors add a number of amino acids to a protein encoded therein, such as to the amino terminus, carboxy terminus, or both of a recombinant protein. Such fusion vectors can serve one or more purposes, such as: (i) to increase expression of recombinant protein; (ii) to increase the solubility of the recombinant protein; and (iii) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. In some embodiments, expression of polynucleotides (such as non-coding polynucleotides) and proteins in prokaryotes can be carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polynucleotides and/or proteins. In some embodiments, the fusion expression vector can include a proteolytic cleavage site, which can be introduced at the junction of the fusion vector backbone or other fusion moiety and the recombinant polynucleotide or protein to enable separation of the recombinant polynucleotide or protein from the fusion vector backbone or other fusion moiety subsequent to purification of the fusion polynucleotide or protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Example fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY* 185, Academic Press, San Diego, Calif. (1990) 60-89).

[0411] In some embodiments, one or more vectors driving expression of one or more elements of an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein are introduced into a host cell such that expression of the elements of the engineered delivery system described herein direct formation of an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein (including but not limited to an engineered gene transfer agent particle, which is described in greater detail elsewhere herein). For example, different elements of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein can each be operably linked to separate regulatory elements on separate vectors. RNA(s) of different elements of the engineered delivery system described herein can be delivered to an animal or mammal or cell thereof to produce an animal or mammal or cell thereof that constitutively or inducibly or conditionally expresses different elements of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein that incorporates one or more elements of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein or contains one or more cells that incorporates and/or expresses one or more elements of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system described herein.

[0412] In some embodiments, two or more of the elements expressed from the same or different regulatory element(s), can be combined in a single vector, with one or more additional vectors providing any components of the system not included in the first vector. Engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system polynucleotides that are combined in a single vector may be arranged in any suitable orientation, such as one element located 5' with respect to ("upstream" of) or 3' with respect to ("downstream" of) a second element. The coding sequence of one element may be located on the same or opposite strand of the coding sequence of a second element, and oriented in the same or opposite direction. In some embodiments, a single promoter drives expression of a transcript encoding one or more engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid proteins, embedded within one or more intron sequences (e.g., each in a different intron, two or more in at least one intron, or all in a single intron). In some embodiments, the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides can be operably linked to and expressed from the same promoter.

Vector Features

[0413] 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 therefrom, 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

[0414] In embodiments, the polynucleotides and/or vectors thereof described herein (such as the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides of the present invention) can include one or more regulatory elements that can be operatively linked to the polynucleotide. The term “regulatory element” is intended to include promoters, enhancers, internal ribosomal entry sites (IRES), and other expression control elements (e.g., transcription termination signals, such as polyadenylation signals and poly-U sequences). 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 EF1 α 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).

[0415] 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 PCT publication 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, 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.

[0416] 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-1 α , β -actin, RSV, and 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.

[0417] 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. In some embodiments, the regulated promoter is a tissue specific promoter as previously discussed elsewhere herein. 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, CNS tissue and cell specific promoters.

[0418] Suitable neuronal tissue/cell specific promoters include, but are not limited to, GFAP promoter (astrocytes), SYN1 promoter (neurons), and NSE/RU5' (mature neurons).

[0419] Other suitable CNS specific promoters can include, but are not limited to, neuroactive peptide cholecystokinin (CCK) (see e.g., Chhatawl et al. *Gene Therapy* volume 14, pages 575–583 (2007)), a brain specific DNA MiniPromoter (such as any of those identified for brain or pan-neronal expression as in de Leeuw et al. *Mol. Therapy*. 1(5): 2014. doi:10.1038/mtm.2013.5), myelin basic promoter (MBP) (see e.g., von Jonquieres, G., Mersmann, N., Klugmann, C. B., Harasta, A. E., Lutz, B., Teahan, O., et al. (2013). Glial promoter selectivity following AAV-delivery to the immature brain. *PLoS One* 8 (6), e65646. doi: 10.1371/journal.pone.0065646), glial fibrillary acid protein (GFAP) for expression in astrocytes (see e.g., Smith-Arica, J. R., Morelli, A. E., Larregina, A. T., Smith, J., Lowenstein, P. R., Castro, M. G. (2000). Cell-type-specific and regulatable transgenesis in the adult brain: adenovirus-encoded combined transcriptional targeting and inducible transgene expression. *Mol. Ther.* 2 (6), 579–587. doi: 10.1006/mthe.2000.0215 and Lee, Y., Messing, A., Su, M., Brenner, M. (2008). GFAP promoter elements required for region-specific and astrocyte-specific expression. *Glia* 56 (5), 481–493. doi: 10.1002/glia.20622), human myelin associated glycoprotein promoter (full-length or truncated) (see e.g., von Jonquieres, G., Frohlich, D., Klugmann, C. B., Wen, X., Harasta, A. E., Ramkumar, R., et al. (2016). Recombinant human myelin-associated glycoprotein promoter drives selective AAV-mediated transgene expression in oligodendrocytes. *Front. Mol. Neurosci.* 9, 13. doi: 10.3389/fnmol.2016.00013), F4/80 promoter (see e.g., Rosario, A. M., Cruz, P. E., Ceballos-Diaz, C., Strickland, M. R., Siemienski, Z., Pardo, M., et al. (2016). Microglia-specific targeting by novel capsid-modified AAV6 vectors. *Mol. Ther. Methods Clin. Dev.* 3, 16026. doi: 10.1038/mtm.2016.26), phosphate-activated glutaminase (PAG) or the vesicular glutamate transporter (vGLUT) promoter (for about 90% glutamatergic neuron-specific expression) (see e.g., Rasmussen, M., Kong, L., Zhang, G. R., Liu, M., Wang, X., Szabo, G., et al. (2007). Glutamatergic or GABAergic neuron-specific, long-term expression in neocortical neurons from helper virus-free HSV-1 vectors containing the phosphate-activated glutaminase, vesicular glutamate transporter-1, or glutamic acid decarboxylase promoter. *Brain Res.* 1144, 19–32. doi: 10.1016/j.brainres.2007.01.125), glutamic acid decarboxylase (GAD) promoter (for about 90% GABAergic neuron-specific expression) (see e.g., Rasmussen, M., Kong, L., Zhang, G. R., Liu, M., Wang, X., Szabo, G., et al. (2007). Glutamatergic or GABAergic neuron-specific, long-term expression in neocortical neurons from helper virus-free HSV-1 vectors containing the phosphate-activated glutaminase, vesicular glutamate transporter-1, or glutamic acid

decarboxylase promoter. *Brain Res.* 1144, 19–32. doi: 10.1016/j.brainres.2007.01.125), MeCP2 promoter (see e.g., Gray et al. *Hum Gene Ther.* 2011 Sep;22(9):1143-53. doi: 10.1089/hum.2010.245), and retinoblastoma gene promoter (see e.g., Jiang et al., *J. Biol. Chem.* 2001. 276, 593-600).

[0420] Other tissue and/or cell specific promoters are discussed elsewhere herein and can be generally known in the art and are within the scope of this disclosure.

[0421] 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 promoter 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.

[0422] Where expression in a plant cell is desired, the components of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid 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. In some embodiments, inclusion of an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system vector in a plant can be for AAV vector production purposes.

[0423] 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 engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system components 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 engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid 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.

[0424] 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 targeting moiety, polypeptide, viral (e.g., AAV) capsid 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 PCT publication WO 2014/018423 and US Publications, 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.

[0425] In some embodiments, transient or inducible expression can be achieved by including, for example, chemical-regulated promoters, 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 In2-2 promoter, activated by benzene sulfonamide herbicide safeners (De Veylder et al., (1997) *Plant Cell Physiol* 38:568-77), the maize GST promoter (GST-II-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.

[0426] In some embodiments, the vector or system thereof can include one or more elements capable of translocating and/or expressing an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid 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.

Selectable Markers and Tags

[0427] One or more of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid 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 polypeptide encoding a polypeptide selectable marker can be incorporated in the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system polynucleotide such that the selectable marker polypeptide, when translated, is inserted between two amino acids between the N- and C- terminus of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polypeptide or at the N- and/or C-terminus of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polypeptide. In some embodiments, the selectable marker or tag is a polynucleotide barcode or unique molecular identifier (UMI).

[0428] 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 targeting moiety, polypeptide, viral (e.g., AAV) capsid 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.

[0429] 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.

[0430] Selectable markers and tags can be operably linked to one or more components of the engineered AAV capsid system described 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: 315) or (GGGGS)₃ (SEQ ID NO: 316). Other suitable linkers are described elsewhere herein.

[0431] 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 targeting moiety, polypeptide, viral (e.g., AAV) capsid 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 targeting

moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) to specific cells, tissues, organs, etc.

Cell-free Vector and Polynucleotide Expression

[0432] In some embodiments, the polynucleotide encoding one or more features of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system can be expressed from a vector or suitable polynucleotide in a cell-free in vitro system. In other words, the polynucleotide can be transcribed and optionally translated in vitro. In vitro transcription/translation systems and appropriate vectors are generally known in the art and commercially available. Generally, in vitro transcription and in vitro translation systems replicate the processes of RNA and protein synthesis, respectively, outside of the cellular environment. Vectors and suitable polynucleotides for in vitro transcription can include T7, SP6, T3, promoter regulatory sequences that can be recognized and acted upon by an appropriate polymerase to transcribe the polynucleotide or vector.

[0433] In vitro translation can be stand-alone (e.g., translation of a purified polyribonucleotide) or linked/coupled to transcription. In some embodiments, the cell-free (or in vitro) translation system can include extracts from rabbit reticulocytes, wheat germ, and/or E. coli. The extracts can include various macromolecular components that are needed for translation of exogenous RNA (e.g., 70S or 80S ribosomes, tRNAs, aminoacyl-tRNA, synthetases, initiation, elongation factors, termination factors, etc.). Other components can be included or added during the translation reaction, including but not limited to, amino acids, energy sources (ATP, GTP), energy regenerating systems (creatine phosphate and creatine phosphokinase (eukaryotic systems)) (phosphoenol pyruvate and pyruvate kinase for bacterial systems), and other co-factors (Mg^{2+} , K^{+} , etc.). As previously mentioned, in vitro translation can be based on RNA or DNA starting material. Some translation systems can utilize an RNA template as starting material (e.g., reticulocyte lysates and wheat germ extracts). Some translation systems can utilize a DNA template as a starting material (e.g., E coli-based systems). In these systems transcription and translation are coupled and DNA is first transcribed into RNA, which is subsequently translated. Suitable standard and coupled cell-free translation systems are generally known in the art and are commercially available.

Codon Optimization of Vector Polynucleotides

[0434] As described elsewhere herein, the polynucleotide encoding one or more embodiments of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system

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 targeting moiety, polypeptide, viral (e.g., AAV) capsid system 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.or.jp/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.

[0435] 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, CNS epithelial cells (including but not limited to the cells lining the brain ventricles), nerve cells (nerves, brain cells, spinal column cells, nerve support cells (e.g., astrocytes, glial cells, Schwann cells etc.), connective tissue cells of the CNS (fat and other soft tissue padding cells of the CNS such as the meninges), stem cells and other progenitor cells, CNS immune 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, CNS tissue and/or cells 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 organ. Such organs include, but are not limited to, the brain. Such codon optimized sequences are within the ambit of the ordinary skilled artisan in view of the description herein.

[0436] 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.

Non-Viral Vectors and Carriers

[0437] In some embodiments, the vector is a non-viral vector or carrier. In some embodiments, non-viral vectors can have the advantage(s) of reduced toxicity and/or immunogenicity and/or increased bio-safety as compared to viral vectors. The terms of art “Non-viral vectors and carriers” and as used herein in this context refers to molecules and/or compositions that are not based on one or more component of a virus or virus genome (excluding any nucleotide to be delivered and/or expressed by the non-viral vector) that can be capable of attaching to, incorporating, coupling, and/or otherwise interacting with an

engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide of the present invention and can be capable of ferrying the polynucleotide to a cell and/or expressing the polynucleotide. It will be appreciated that this does not exclude the inclusion of a virus-based polynucleotide that is to be delivered. For example, if a gRNA to be delivered is directed against a virus component and it is inserted or otherwise coupled to an otherwise non-viral vector or carrier, this would not make said vector a “viral vector”. Non-viral vectors and carriers include naked polynucleotides, chemical-based carriers, polynucleotide (non-viral) based vectors, and particle-based carriers. It will be appreciated that the term “vector” as used in the context of non-viral vectors and carriers refers to polynucleotide vectors and “carriers” used in this context refers to a non-nucleic acid or polynucleotide molecule or composition that be attached to or otherwise interact with a polynucleotide to be delivered, such as an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide of the present invention.

Naked Polynucleotides

[0438] In some embodiments one or more engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides described elsewhere herein can be included in a naked polynucleotide. The term of art “naked polynucleotide” as used herein refers to polynucleotides that are not associated with another molecule (e.g., proteins, lipids, and/or other molecules) that can often help protect it from environmental factors and/or degradation. As used herein, associated with includes, but is not limited to, linked to, adhered to, adsorbed to, enclosed in, enclosed in or within, mixed with, and the like. Naked polynucleotides that include one or more of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides described herein can be delivered directly to a host cell and optionally expressed therein. The naked polynucleotides can have any suitable two- and three-dimensional configurations. By way of non-limiting examples, naked polynucleotides can be single-stranded molecules, double stranded molecules, circular molecules (e.g., plasmids and artificial chromosomes), molecules that contain portions that are single stranded and portions that are double stranded (e.g., ribozymes), and the like. In some embodiments, the naked polynucleotide contains only the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention. In some embodiments, the naked polynucleotide can contain other nucleic acids and/or polynucleotides in addition to the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention. The naked polynucleotides can

include one or more elements of a transposon system. Transposons and system thereof are described in greater detail elsewhere herein.

Non-Viral Polynucleotide Vectors

[0439] In some embodiments, one or more of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides can be included in a non-viral polynucleotide vector. Suitable non-viral polynucleotide vectors include, but are not limited to, transposon vectors and vector systems, plasmids, bacterial artificial chromosomes, yeast artificial chromosomes, AR(antibiotic resistance)-free plasmids and miniplasmids, circular covalently closed vectors (e.g., minicircles, minivectors, miniknots,), linear covalently closed vectors (“dumbbell shaped”), MIDGE (minimalistic immunologically defined gene expression) vectors, MiLV (micro-linear vector) vectors, Ministrings, mini-intronic plasmids, PSK systems (post-segregationally killing systems), ORT (operator repressor titration) plasmids, and the like. See e.g., Hardee et al. 2017. *Genes*. 8(2):65.

[0440] In some embodiments, the non-viral polynucleotide vector can have a conditional origin of replication. In some embodiments, the non-viral polynucleotide vector can be an ORT plasmid. In some embodiments, the non-viral polynucleotide vector can have a minimalistic immunologically defined gene expression. In some embodiments, the non-viral polynucleotide vector can have one or more post-segregationally killing system genes. In some embodiments, the non-viral polynucleotide vector is AR-free. In some embodiments, the non-viral polynucleotide vector is a minivector. In some embodiments, the non-viral polynucleotide vector includes a nuclear localization signal. In some embodiments, the non-viral polynucleotide vector can include one or more CpG motifs. In some embodiments, the non-viral polynucleotide vectors can include one or more scaffold/matrix attachment regions (S/MARs). See e.g., Mirkovitch et al. 1984. *Cell*. 39:223-232, Wong et al. 2015. *Adv. Genet.* 89:113-152, whose techniques and vectors can be adapted for use in the present invention. S/MARs are AT-rich sequences that play a role in the spatial organization of chromosomes through DNA loop base attachment to the nuclear matrix. S/MARs are often found close to regulatory elements such as promoters, enhancers, and origins of DNA replication. Inclusion of one or S/MARs can facilitate a once-per-cell-cycle replication to maintain the non-viral polynucleotide vector as an episome in daughter cells. In embodiments, the S/MAR sequence is located downstream of an actively transcribed polynucleotide (e.g., one or more engineered AAV capsid polynucleotides of the present invention) included in the non-viral polynucleotide

vector. In some embodiments, the S/MAR can be a S/MAR from the beta-interferon gene cluster. See e.g., Verghese et al. 2014. *Nucleic Acid Res.* 42:e53; Xu et al. 2016. *Sci. China Life Sci.* 59:1024-1033; Jin et al. 2016. 8:702-711; Koirala et al. 2014. *Adv. Exp. Med. Biol.* 801:703-709; and Nehlsen et al. 2006. *Gene Ther. Mol. Biol.* 10:233-244, whose techniques and vectors can be adapted for use in the present invention.

[0441] In some embodiments, the non-viral vector is a transposon vector or system thereof. As used herein, “transposon” (also referred to as transposable element) refers to a polynucleotide sequence that is capable of moving from location in a genome to another. There are several classes of transposons. Transposons include retrotransposons and DNA transposons. Retrotransposons require the transcription of the polynucleotide that is moved (or transposed) in order to transpose the polynucleotide to a new genome or polynucleotide. DNA transposons are those that do not require reverse transcription of the polynucleotide that is moved (or transposed) in order to transpose the polynucleotide to a new genome or polynucleotide. In some embodiments, the non-viral polynucleotide vector can be a retrotransposon vector. In some embodiments, the retrotransposon vector includes long terminal repeats. In some embodiments, the retrotransposon vector does not include long terminal repeats. In some embodiments, the non-viral polynucleotide vector can be a DNA transposon vector. DNA transposon vectors can include a polynucleotide sequence encoding a transposase. In some embodiments, the transposon vector is configured as a non-autonomous transposon vector, meaning that the transposition does not occur spontaneously on its own. In some of these embodiments, the transposon vector lacks one or more polynucleotide sequences encoding proteins required for transposition. In some embodiments, the non-autonomous transposon vectors lack one or more Ac elements.

[0442] In some embodiments a non-viral polynucleotide transposon vector system can include a first polynucleotide vector that contains the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention flanked on the 5' and 3' ends by transposon terminal inverted repeats (TIRs) and a second polynucleotide vector that includes a polynucleotide capable of encoding a transposase coupled to a promoter to drive expression of the transposase. When both are expressed in the same cell the transposase can be expressed from the second vector and can transpose the material between the TIRs on the first vector (e.g., the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention) and integrate it into one or more positions in the

host cell's genome. In some embodiments the transposon vector or system thereof can be configured as a gene trap. In some embodiments, the TIRs can be configured to flank a strong splice acceptor site followed by a reporter and/or other gene (e.g., one or more of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention) and a strong poly A tail. When transposition occurs while using this vector or system thereof, the transposon can insert into an intron of a gene and the inserted reporter or other gene can provoke a mis-splicing process and as a result it in activates the trapped gene.

[0443] Any suitable transposon system can be used. Suitable transposon and systems thereof can include, but are not limited to, Sleeping Beauty transposon system (Tc1/mariner superfamily) (see e.g., Ivics et al. 1997. Cell. 91(4): 501-510), piggyBac (piggyBac superfamily) (see e.g., Li et al. 2013 110(25): E2279-E2287 and Yusa et al. 2011. PNAS. 108(4): 1531-1536), Tol2 (superfamily hAT), Frog Prince (Tc1/mariner superfamily) (see e.g., Miskey et al. 2003 Nucleic Acid Res. 31(23):6873-6881) and variants thereof.

Chemical Carriers

[0444] In some embodiments the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) can be coupled to a chemical carrier. Chemical carriers that can be suitable for delivery of polynucleotides can be broadly classified into the following classes: (i) inorganic particles, (ii) lipid-based, (iii) polymer-based, and (iv) peptide based. They can be categorized as (1) those that can form condensed complexes with a polynucleotide (such as the engineered targeting moiety, polypeptide, viral (e.g. AAV) capsid polynucleotide(s) of the present invention), (2) those capable of targeting specific cells, (3) those capable of increasing delivery of the polynucleotide (such as the engineered targeting moiety, polypeptide, viral (e.g. AAV) capsid polynucleotide(s) of the present invention) to the nucleus or cytosol of a host cell, (4) those capable of disintegrating from DNA/RNA in the cytosol of a host cell, and (5) those capable of sustained or controlled release. It will be appreciated that any one given chemical carrier can include features from multiple categories. The term "particle" as used herein, refers to any suitable sized particles for delivery of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system components described herein. Suitable sizes include macro-, micro-, and nano-sized particles.

[0445] In some embodiments, the non-viral carrier can be an inorganic particle. In some embodiments, the inorganic particle, can be a nanoparticle. The inorganic particles can be configured and optimized by varying size, shape, and/or porosity. In some embodiments, the

inorganic particles are optimized to escape from the reticuloendothelial system. In some embodiments, the inorganic particles can be optimized to protect an entrapped molecule from degradation. The Suitable inorganic particles that can be used as non-viral carriers in this context can include, but are not limited to, calcium phosphate, silica, metals (e.g., gold, platinum, silver, palladium, rhodium, osmium, iridium, ruthenium, mercury, copper, rhenium, titanium, niobium, tantalum, and combinations thereof), magnetic compounds, particles, and materials, (e.g., supermagnetic iron oxide and magnetite), quantum dots, fullerenes (e.g., carbon nanoparticles, nanotubes, nanostrings, and the like), and combinations thereof. Other suitable inorganic non-viral carriers are discussed elsewhere herein.

[0446] In some embodiments, the non-viral carrier can be lipid-based. Suitable lipid-based carriers are also described in greater detail herein. In some embodiments, the lipid-based carrier includes a cationic lipid or an amphiphilic lipid that is capable of binding or otherwise interacting with a negative charge on the polynucleotide to be delivered (e.g., such as an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide of the present invention). In some embodiments, chemical non-viral carrier systems can include a polynucleotide such as the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention) and a lipid (such as a cationic lipid). These are also referred to in the art as lipoplexes. Other embodiments of lipoplexes are described elsewhere herein. In some embodiments, the non-viral lipid-based carrier can be a lipid nano emulsion. Lipid nano emulsions can be formed by the dispersion of an immiscible liquid in another stabilized emulsifying agent and can have particles of about 200 nm that are composed of the lipid, water, and surfactant that can contain the polynucleotide to be delivered (e.g., the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention). In some embodiments, the lipid-based non-viral carrier can be a solid lipid particle or nanoparticle.

[0447] In some embodiments, the non-viral carrier can be peptide-based. In some embodiments, the peptide-based non-viral carrier can include one or more cationic amino acids. In some embodiments, 35 to 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 99 or 100 % of the amino acids are cationic. In some embodiments, peptide carriers can be used in conjunction with other types of carriers (e.g., polymer-based carriers and lipid-based carriers to functionalize these carriers). In some embodiments, the functionalization is targeting a host cell. Suitable polymers that can be included in the polymer-based non-viral carrier can include,

but are not limited to, polyethylenimine (PEI), chitosan, poly (DL-lactide) (PLA), poly (DL-Lactide-co-glycoside) (PLGA), dendrimers (see e.g., US Pat. Pub. 2017/0079916 whose techniques and compositions can be adapted for use with the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides of the present invention), polymethacrylate, and combinations thereof.

[0448] In some embodiments, the non-viral carrier can be configured to release an engineered delivery system polynucleotide that is associated with or attached to the non-viral carrier in response to an external stimulus, such as pH, temperature, osmolarity, concentration of a specific molecule or composition (e.g., calcium, NaCl, and the like), pressure and the like. In some embodiments, the non-viral carrier can be a particle that is configured includes one or more of the engineered AAV capsid polynucleotides describe herein and an environmental triggering agent response element, and optionally a triggering agent. In some embodiments, the particle can include a polymer that can be selected from the group of polymethacrylates and polyacrylates. In some embodiments, the non-viral particle can include one or more embodiments of the compositions microparticles described in US Pat. Pubs. 20150232883 and 20050123596, whose techniques and compositions can be adapted for use in the present invention.

[0449] In some embodiments, the non-viral carrier can be a polymer-based carrier. In some embodiments, the polymer is cationic or is predominantly cationic such that it can interact in a charge-dependent manner with the negatively charged polynucleotide to be delivered (such as the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide(s) of the present invention). Polymer-based systems are described in greater detail elsewhere herein.

Viral Vectors

[0450] In some embodiments, the vector is a viral vector. 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 an engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotide of the present invention, 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 components of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system 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 adenoviral-based vectors, adeno associated vectors, helper-dependent adenoviral (HdAd) vectors, hybrid adenoviral vectors, and the like. 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.

Adenoviral vectors, Helper-dependent Adenoviral vectors, and Hybrid Adenoviral Vectors

[0451] In some embodiments, the vector can be an adenoviral vector. In some embodiments, the adenoviral vector can include elements such that the virus particle produced using the vector or system thereof can be serotype 2, 5, or 9. In some embodiments, the polynucleotide to be delivered via the adenoviral particle can be up to about 8 kb. Thus, in some embodiments, an adenoviral vector can include a DNA polynucleotide to be delivered that can range in size from about 0.001 kb to about 8 kb. Adenoviral vectors have been used successfully in several contexts (see e.g., Teramoto et al. 2000. *Lancet*. 355:1911-1912; Lai et al. 2002. *DNA Cell. Biol.* 21:895-913; Flotte et al., 1996. *Hum. Gene. Ther.* 7:1145-1159; and Kay et al. 2000. *Nat. Genet.* 24:257-261. The engineered AAV capsids can be included in an adenoviral vector to produce adenoviral particles containing said engineered AAV capsids.

[0452] In some embodiments the vector can be a helper-dependent adenoviral vector or system thereof. These are also referred to in the field as “gutless” or “guttled” vectors and are a modified generation of adenoviral vectors (see e.g., Thrasher et al. 2006. *Nature*. 443:E5-7). In embodiments of the helper-dependent adenoviral vector system one vector (the helper) can contain all the viral genes required for replication but contains a conditional gene defect in the packaging domain. The second vector of the system can contain only the ends of the viral genome, one or more engineered AAV capsid polynucleotides, and the native packaging recognition signal, which can allow selective packaged release from the cells (see e.g., Cideciyan et al. 2009. *N Engl J Med.* 361:725-727). Helper-dependent Adenoviral vector systems have been successful for gene delivery in several contexts (see e.g., Simonelli et al. 2010. *J Am Soc Gene Ther.* 18:643-650; Cideciyan et al. 2009. *N Engl J Med.* 361:725-727; Crane et al. 2012. *Gene Ther.* 19(4):443-452; Alba et al. 2005. *Gene Ther.* 12:18-S27; Croyle et al. 2005. *Gene Ther.* 12:579-587; Amalfitano et al. 1998. *J. Virol.* 72:926-933; and Morral et al. 1999. *PNAS.* 96:12816-12821). The techniques and vectors described in these

publications can be adapted for inclusion and delivery of the engineered AAV capsid polynucleotides described herein. In some embodiments, the polynucleotide to be delivered via the viral particle produced from a helper-dependent adenoviral vector or system thereof can be up to about 38 kb. Thus, in some embodiments, an adenoviral vector can include a DNA polynucleotide to be delivered that can range in size from about 0.001 kb to about 37 kb (see e.g., Rosewell et al. 2011. J. Genet. Syndr. Gene Ther. Suppl. 5:001).

[0453] In some embodiments, the vector is a hybrid-adenoviral vector or system thereof. Hybrid adenoviral vectors are composed of the high transduction efficiency of a gene-deleted adenoviral vector and the long-term genome-integrating potential of adeno-associated, retroviruses, lentivirus, and transposon based-gene transfer. In some embodiments, such hybrid vector systems can result in stable transduction and limited integration site. See e.g., Balague et al. 2000. Blood. 95:820-828; Morral et al. 1998. Hum. Gene Ther. 9:2709-2716; Kubo and Mitani. 2003. J. Virol. 77(5): 2964-2971; Zhang et al. 2013. PloS One. 8(10) e76771; and Cooney et al. 2015. Mol. Ther. 23(4):667-674), whose techniques and vectors described therein can be modified and adapted for use in the engineered AAV capsid system of the present invention. In some embodiments, a hybrid-adenoviral vector can include one or more features of a retrovirus and/or an adeno-associated virus. In some embodiments the hybrid-adenoviral vector can include one or more features of a spuma retrovirus or foamy virus (FV). See e.g., Ehrhardt et al. 2007. Mol. Ther. 15:146-156 and Liu et al. 2007. Mol. Ther. 15:1834-1841, whose techniques and vectors described therein can be modified and adapted for use in the engineered AAV capsid system of the present invention. Advantages of using one or more features from the FVs in the hybrid-adenoviral vector or system thereof can include the ability of the viral particles produced therefrom to infect a broad range of cells, a large packaging capacity as compared to other retroviruses, and the ability to persist in quiescent (non-dividing) cells. See also e.g., Ehrhardt et al. 2007. Mol. Ther. 15:146-156 and Shuji et al. 2011. Mol. Ther. 19:76-82, whose techniques and vectors described therein can be modified and adapted for use in the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid system of the present invention.

Adeno Associated Vectors

[0454] In an embodiment, the engineered vector or system thereof can be an adeno-associated vector (AAV). See, e.g., West et al., Virology 160:38-47 (1987); U.S. Pat. No. 4,797,368; WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); and Muzyczka, J.

Clin. Invest. 94:1351 (1994). Although similar to adenoviral vectors in some of their features, AAVs have some deficiency in their replication and/or pathogenicity and thus can be safer than adenoviral vectors. In some embodiments the AAV can integrate into a specific site on chromosome 19 of a human cell with no observable side effects. In some embodiments, the capacity of the AAV vector, system thereof, and/or AAV particles can be up to about 4.7 kb. The AAV vector or system thereof can include one or more engineered capsid polynucleotides described herein.

[0455] The AAV vector or system thereof can include one or more regulatory molecules. In some embodiments the regulatory molecules can be promoters, enhancers, repressors and the like, which are described in greater detail elsewhere herein. In some embodiments, the AAV vector or system thereof can include one or more polynucleotides that can encode one or more regulatory proteins. In some embodiments, the one or more regulatory proteins can be selected from Rep78, Rep68, Rep52, Rep40, variants thereof, and combinations thereof. In some embodiments, the promoter can be a tissue specific promoter as previously discussed. In some embodiments, the tissue specific promoter can drive expression of an engineered capsid AAV capsid polynucleotide described herein.

[0456] The AAV vector or system thereof can include one or more polynucleotides that can encode one or more capsid proteins, such as the engineered AAV capsid proteins described elsewhere herein. The engineered capsid proteins can be capable of assembling into a protein shell (an engineered capsid) of the AAV virus particle. The engineered capsid can have a cell-, tissue- and/or organ-specific tropism.

[0457] In some embodiments, the AAV vector or system thereof can include one or more adenovirus helper factors or polynucleotides that can encode one or more adenovirus helper factors. Such adenovirus helper factors can include, but are not limited, E1A, E1B, E2A, E4ORF6, and VA RNAs. In some embodiments, a producing host cell line expresses one or more of the adenovirus helper factors.

[0458] The AAV vector or system thereof can be configured to produce AAV particles having a specific serotype. In some embodiments, the serotype can be AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, AAV-8, AAV-9 or any combinations thereof. In some embodiments, the AAV can be AAV1, AAV-2, AAV-5, AAV-9 or any combination thereof. One can select the AAV of the AAV with regard to the cells to be targeted; e.g., one can select AAV serotypes 1, 2, 5, 9 or a hybrid capsid AAV-1, AAV-2, AAV-5, AAV-9 or any

combination thereof for targeting brain and/or neuronal cells; and one can select AAV-4 for targeting cardiac tissue; and one can select AAV-8 for delivery to the liver. Thus, in some embodiments, an AAV vector or system thereof capable of producing AAV particles capable of targeting the brain and/or neuronal cells can be configured to generate AAV particles having serotypes 1, 2, 5 or a hybrid capsid AAV-1, AAV-2, AAV-5 or any combination thereof. In some embodiments, an AAV vector or system thereof capable of producing AAV particles capable of targeting cardiac tissue can be configured to generate an AAV particle having an AAV-4 serotype. In some embodiments, an AAV vector or system thereof capable of producing AAV particles capable of targeting the liver can be configured to generate an AAV having an AAV-8 serotype. See also Srivastava. 2017. *Curr. Opin. Virol.* 21:75-80.

[0459] It will be appreciated that while the different serotypes can provide some level of cell, tissue, and/or organ specificity, each serotype still is multi-tropic and thus can result in tissue-toxicity if using that serotype to target a tissue that the serotype is less efficient in transducing. Thus, in addition to achieving some tissue targeting capacity via selecting an AAV of a particular serotype, it will be appreciated that the tropism of the AAV serotype can be modified by an engineered AAV capsid described herein. As described elsewhere herein, variants of wild-type AAV of any serotype can be generated via a method described herein and determined to have a particular cell-specific tropism, which can be the same or different as that of the reference wild-type AAV serotype. In some embodiments, the cell, tissue, and/or specificity of the wild-type serotype can be enhanced (e.g., made more selective or specific for a particular cell type that the serotype is already biased towards). For example, wild-type AAV-9 is biased towards muscle and brain in humans (see e.g., Srivastava. 2017. *Curr. Opin. Virol.* 21:75-80.) By including an engineered AAV capsid and/or capsid protein variant of wild-type AAV-9 as described herein, the bias for e.g., muscle (or other non-CNS tissue or cell) can be reduced or eliminated and/or the CNS tissue or cell specificity increased such that the muscle (or other non-CNS tissue or cell) specificity appears reduced in comparison, thus enhancing the specificity for the CNS tissue or cell as compared to the wild-type AAV-9. As previously mentioned, inclusion of an engineered capsid and/or capsid protein variant of a wild-type AAV serotype can have a different tropism than the wild-type reference AAV serotype. For example, an engineered AAV capsid and/or capsid protein variant of AAV-9 can have specificity for a tissue other than muscle or brain in humans.

[0460] In some embodiments, the AAV vector is a hybrid AAV vector or system thereof. Hybrid AAVs are AAVs that include genomes with elements from one serotype that are packaged into a capsid derived from at least one different serotype. For example, if it is the rAAV2/5 that is to be produced, and if the production method is based on the helper-free, transient transfection method discussed above, the 1st plasmid and the 3rd plasmid (the adeno helper plasmid) will be the same as discussed for rAAV2 production. However, the 2nd plasmid, the pRepCap will be different. In this plasmid, called pRep2/Cap5, the Rep gene is still derived from AAV2, while the Cap gene is derived from AAV5. The production scheme is the same as the above-mentioned approach for AAV2 production. The resulting rAAV is called rAAV2/5, in which the genome is based on recombinant AAV2, while the capsid is based on AAV5. It is assumed the cell or tissue-tropism displayed by this AAV2/5 hybrid virus should be the same as that of AAV5. It will be appreciated that wild-type hybrid AAV particles suffer the same specificity issues as with the non-hybrid wild-type serotypes previously discussed.

[0461] Advantages achieved by the wild-type based hybrid AAV systems can be combined with the increased and customizable cell-specificity that can be achieved with the engineered AAV capsids can be combined by generating a hybrid AAV that can include an engineered AAV capsid described elsewhere herein. It will be appreciated that hybrid AAVs can contain an engineered AAV capsid containing a genome with elements from a different serotype than the reference wild-type serotype that the engineered AAV capsid is a variant of. For example, a hybrid AAV can be produced that includes an engineered AAV capsid that is a variant of an AAV-9 serotype that is used to package a genome that contains components (e.g., rep elements) from an AAV-2 serotype. As with wild-type based hybrid AAVs previously discussed, the tropism of the resulting AAV particle will be that of the engineered AAV capsid.

[0462] A tabulation of certain wild-type AAV serotypes as to these cells can be found in Grimm, D. et al, J. Virol. 82: 5887-5911 (2008) reproduced below as **Table 6**. Further tropism details can be found in Srivastava. 2017. Curr. Opin. Virol. 21:75-80 as previously discussed.

Table 6								
Cell Line	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2.5	0.0	0.1	10	0.7	0.0
HEK293	25	100	2.5	0.1	0.1	5	0.7	0.1
HeLa	3	100	2.0	0.1	6.7	1	0.2	0.1
HepG2	3	100	16.7	0.3	1.7	5	0.3	ND
Hep1A	20	100	0.2	1.0	0.1	1	0.2	0.0
911	17	100	11	0.2	0.1	17	0.1	ND
CHO	100	100	14	1.4	333	50	10	1.0
COS	33	100	33	3.3	5.0	14	2.0	0.5
MeWo	10	100	20	0.3	6.7	10	1.0	0.2
NIH3T3	10	100	2.9	2.9	0.3	10	0.3	ND
A549	14	100	20	ND	0.5	10	0.5	0.1
HT1180	20	100	10	0.1	0.3	33	0.5	0.1
Monocytes	1111	100	ND	ND	125	1429	ND	ND
Immature DC	2500	100	ND	ND	222	2857	ND	ND
Mature DC	2222	100	ND	ND	333	3333	ND	ND

[0463] In some embodiments, the AAV vector or system thereof is AAV rh.74 or AAV rh.10.

[0464] In some embodiments, the AAV vector or system thereof is configured as a “gutless” vector, similar to that described in connection with a retroviral vector. In some embodiments, the “gutless” AAV vector or system thereof can have the cis-acting viral DNA elements involved in genome amplification and packaging in linkage with the heterologous sequences of interest (e.g., the engineered AAV capsid polynucleotide(s)).

Vector Construction

[0465] 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. Application publication No. US 2004-0171156 A1. Other suitable methods and techniques are described elsewhere herein.

[0466] 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. AAV vectors are discussed elsewhere herein.

[0467] In some embodiments, the vector can have 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.

[0468] Delivery vehicles, vectors, particles, nanoparticles, formulations and components thereof for expression of one or more elements of an engineered AAV capsid system described herein are as used in the foregoing documents, such as WO 2014/093622 (PCT/US2013/074667) and are discussed in greater detail herein.

Virus Particle Production from Viral Vectors

AAV Particle Production

[0469] There are two main strategies for producing AAV particles from AAV vectors and systems thereof, such as those described herein, which depend on how the adenovirus helper factors are provided (helper v. helper free). In some embodiments, a method of producing AAV particles from AAV vectors and systems thereof can include adenovirus infection into cell lines that stably harbor AAV replication and capsid encoding polynucleotides along with AAV vector containing the polynucleotide to be packaged and delivered by the resulting AAV particle (e.g., the engineered AAV capsid polynucleotide(s)). In some embodiments, a method of producing AAV particles from AAV vectors and systems thereof can be a “helper free” method, which includes co-transfection of an appropriate producing cell line with three vectors (e.g., plasmid vectors): (1) an AAV vector that contains a polynucleotide of interest (e.g., the engineered AAV capsid polynucleotide(s)) between 2 ITRs; (2) a vector that carries the AAV Rep-Cap encoding polynucleotides; and (helper polynucleotides. One of skill in the art will appreciate various methods and variations thereof that are both helper and -helper free and as well as the different advantages of each system.

[0470] The engineered AAV vectors and systems thereof described herein can be produced by any of these methods.

Vector and Virus Particle Delivery

[0471] A vector (including non-viral carriers) described herein can be introduced into host cells to thereby produce transcripts, proteins, or peptides, including fusion proteins or peptides encoded by nucleic acids as described herein (e.g., engineered AAV capsid system transcripts, proteins, enzymes, mutant forms thereof, fusion proteins thereof, etc.), and virus particles (such as from viral vectors and systems thereof).

[0472] One or more engineered AAV capsid polynucleotides can be delivered using adeno associated virus (AAV), adenovirus or other plasmid or viral vector types as previously described, in particular, using formulations and doses from, for example, US Patents Nos. 8,454,972 (formulations, doses for adenovirus), 8,404,658 (formulations, doses for AAV) and 5,846,946 (formulations, doses for DNA plasmids) and from clinical trials and publications regarding the clinical trials involving lentivirus, AAV and adenovirus. For examples, for AAV, the route of administration, formulation and dose can be as in US Patent No. 8,454,972 and as in clinical trials involving AAV. For Adenovirus, the route of administration, formulation and dose can be as in US Patent No. 8,404,658 and as in clinical trials involving adenovirus.

[0473] For plasmid delivery, the route of administration, formulation and dose can be as in US Patent No 5,846,946 and as in clinical studies involving plasmids. In some embodiments, doses can be based on or extrapolated to an average 70 kg individual (e.g., a male adult human), and can be adjusted for patients, subjects, mammals of different weight and species. Frequency of administration is within the ambit of the medical or veterinary practitioner (e.g., physician, veterinarian), depending on usual factors including the age, sex, general health, other conditions of the patient or subject and the particular condition or symptoms being addressed. The viral vectors can be injected into or otherwise delivered to the tissue or cell of interest.

[0474] In terms of in vivo delivery, AAV is advantageous over other viral vectors for a couple of reasons such as low toxicity (this may be due to the purification method not requiring ultra-centrifugation of cell particles that can activate the immune response) and a low probability of causing insertional mutagenesis because it doesn't integrate into the host genome.

[0475] The vector(s) and virus particles described herein can be delivered into a host cell in vitro, in vivo, and or ex vivo. Delivery can occur by any suitable method including, but not limited to, physical methods, chemical methods, and biological methods. Physical delivery methods are those methods that employ physical force to counteract the membrane barrier of

the cells to facilitate intracellular delivery of the vector. Suitable physical methods include, but are not limited to, needles (e.g., injections), ballistic polynucleotides (e.g., particle bombardment, micro projectile gene transfer, and gene gun), electroporation, sonoporation, photoporation, magnetofection, hydroporation, and mechanical massage. Chemical methods are those methods that employ a chemical to elicit a change in the cells membrane permeability or other characteristic(s) to facilitate entry of the vector into the cell. For example, the environmental pH can be altered which can elicit a change in the permeability of the cell membrane. Biological methods are those that rely and capitalize on the host cell's biological processes or biological characteristics to facilitate transport of the vector (with or without a carrier) into a cell. For example, the vector and/or its carrier can stimulate an endocytosis or similar process in the cell to facilitate uptake of the vector into the cell.

[0476] Delivery of engineered AAV capsid system components (e.g., polynucleotides encoding engineered AAV capsid and/or capsid proteins) to cells via particles. The term “particle” as used herein, refers to any suitable sized particles for delivery of the engineered AAV capsid system components described herein. Suitable sizes include macro-, micro-, and nano-sized particles. In some embodiments, any of the of the engineered AAV capsid system components (e.g., polypeptides, polynucleotides, vectors, and combinations thereof described herein) can be attached to, coupled to, integrated with, otherwise associated with one or more particles or component thereof as described herein. The particles described herein can then be administered to a cell or organism by an appropriate route and/or technique. In some embodiments, particle delivery can be selected and be advantageous for delivery of the polynucleotide or vector components. It will be appreciated that in embodiments, particle delivery can also be advantageous for other engineered capsid system molecules and formulations described elsewhere herein.

Engineered Virus Particles Including an Engineered Viral Capsid

[0477] Also described herein are engineered virus particles (also referred to here and elsewhere herein as “engineered viral particles”) that can contain an engineered viral (e.g., AAV) capsid as described in detail elsewhere herein. Viral particles with an engineered AAV capsid are referred to herein as engineered AAV particles. It will be appreciated that the engineered viral (e.g., AAV) particles can be adenovirus-based particles, helper adenovirus-based particles, AAV-based particles, or hybrid adenovirus-based particles that contain at least one engineered AAV capsid proteins as previously described. An engineered AAV capsid is

one that contains one or more engineered AAV capsid proteins as are described elsewhere herein. In some embodiments, the engineered AAV particles can include 1-60 engineered AAV capsid proteins described herein. In some embodiments, the engineered AAV particles can contain 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, or 60 engineered capsid proteins. In some embodiments, the engineered AAV particles can contain 0-59 wild-type AAV capsid proteins. In some embodiments, the engineered AAV particles can contain 0, 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, or 59 wild-type AAV capsid proteins. The engineered AAV particles can thus include one or more *n*-mer inserts as is previously described.

[0478] The engineered AAV particle can include one or more cargo polynucleotides. Cargo polynucleotides are discussed in greater detail elsewhere herein. Methods of making the engineered AAV particles from viral and non-viral vectors are described elsewhere herein. Formulations containing the engineered virus particles are described elsewhere herein.

[0479] The engineered viral (e.g., AAV) capsid polynucleotides, other viral (e.g., AAV) polynucleotide(s), and/or vector polynucleotides can contain one or more cargo polynucleotides. The cargo polynucleotides can encode one or more polypeptides. Exemplary cargos are described in greater detail elsewhere herein. It will be appreciated that when a cargo polypeptide is described that its encoding polynucleotide can be a cargo polynucleotide described in this context. In some embodiments, the one or more cargo polynucleotides can be operably linked to the engineered viral (e.g., AAV) capsid polynucleotide(s) and can be part of the engineered viral (e.g., AAV) genome of the viral (e.g., AAV) system of the present invention. The cargo polynucleotides can be packaged into an engineered viral (e.g., AAV) particle, which can be delivered to, e.g., a cell. In some embodiments, the cargo polynucleotide can be capable of modifying a polynucleotide (e.g., gene or transcript) of a cell to which it is delivered. As used herein, “gene” can refer to a hereditary unit corresponding to a sequence of DNA that occupies a specific location on a chromosome and that contains the genetic instruction for a characteristic(s) or trait(s) in an organism. The term gene can refer to translated and/or untranslated regions of a genome. “Gene” can refer to the specific sequence of DNA that is transcribed into an RNA transcript that can be translated into a polypeptide or be a

catalytic RNA molecule, including but not limited to, tRNA, siRNA, piRNA, miRNA, long-non-coding RNA and shRNA. Polynucleotide, gene, transcript, etc. modification includes all genetic engineering techniques including, but not limited to, gene editing as well as conventional recombinational gene modification techniques (e.g., whole or partial gene insertion, deletion, and mutagenesis (e.g., insertional and deletional mutagenesis) techniques.

Engineered Cells and Organisms Expressing said engineered Viral capsids

[0480] Described herein are engineered cells that can include one or more of the engineered targeting moieties, polypeptides, viral (e.g., AAV) capsid polynucleotides, polypeptides, vectors, and/or vector systems described in greater detail elsewhere herein. In some embodiments, one or more of the engineered viral (e.g., AAV) capsid polynucleotides can be expressed in the engineered cells. In some embodiments, the engineered cells can be capable of producing engineered viral (e.g., AAV) capsid proteins and/or engineered viral (e.g., AAV) capsid particles that are described elsewhere herein. Also described herein are modified or engineered organisms that can include one or more engineered cells described herein. The engineered cells can be engineered to express a cargo molecule (e.g., a cargo polynucleotide) dependently or independently of an engineered viral (e.g., AAV) capsid polynucleotide as described elsewhere herein.

[0481] A wide variety of animals, plants, algae, fungi, yeast, etc. and animal, plant, algae, fungus, yeast cell or tissue systems may be engineered to express one or more nucleic acid constructs of the engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid system described herein using various transformation methods mentioned elsewhere herein. This can produce organisms that can produce engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid particles, such as for production purposes, engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid design and/or generation, and/or model organisms. In some embodiments, the polynucleotide(s) encoding one or more components of the engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid system described herein can be stably or transiently incorporated into one or more cells of a plant, animal, algae, fungus, and/or yeast or tissue system. In some embodiments, one or more of engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid system polynucleotides are genomically incorporated into one or more cells of a plant, animal, algae, fungus, and/or yeast or tissue system. Further embodiments of the modified organisms and systems are described elsewhere herein. In some embodiments, one or more components of the engineered targeting

moiety, polypeptide, vector, viral (e.g., AAV) capsid system described herein are expressed in one or more cells of the plant, animal, algae, fungus, yeast, or tissue systems.

Engineered Cells

[0482] Described herein are various embodiments of engineered cells that can include one or more of the engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid system polynucleotides, polypeptides, vectors, and/or vector systems described elsewhere herein. In some embodiments, the cells can express one or more of the engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid polynucleotides and can produce one or more engineered targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid particles, which are described in greater detail herein. Such cells are also referred to herein as “producer cells”. 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 cells (i.e. they do not make engineered viral (e.g., AAV) particles) unless they include one or more of the engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid polynucleotides, engineered targeting moiety, polypeptide, viral (e.g., AAV) capsid vectors or other vectors described herein that render the cells capable of producing an engineered viral (e.g., AAV) capsid particle or other particles described herein. Modified cells can be recipient cells of an engineered viral (e.g., AAV) capsid particles and can, in some embodiments, be modified by the engineered viral (e.g., AAV) capsid particle(s) and/or a cargo polynucleotide delivered to the recipient cell. Modified cells are discussed in greater detail elsewhere herein. 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 targeting moiety, polypeptide, viral (e.g., AAV) capsid molecule.

[0483] In an embodiment, the invention provides a non-human eukaryotic organism; for example, a multicellular eukaryotic organism, including a eukaryotic host cell containing one or more components of an engineered delivery system described herein according to any of the described embodiments. In other embodiments, the invention provides a eukaryotic organism; preferably a multicellular eukaryotic organism, comprising a eukaryotic host cell containing one or more components of an engineered delivery system described herein according to any of the described embodiments. In some embodiments, the organism is a host of a virus (e.g., an AAV).

[0484] 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.

[0485] 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).

[0486] 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, 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) (Manassas, Va.)).

[0487] In some embodiments, the engineered cell is a muscle cell (e.g., cardiac muscle, skeletal muscle, and/or smooth muscle), bone cell, blood cell, immune cell (including but not limited to B cells, macrophages, T-cells, CAR-T cells, and the like), kidney cells, bladder cells, lung cells, heart cells, liver cells, brain cells, neurons, skin cells, stomach cells, neuronal support cells, intestinal cells, epithelial cells, endothelial cells, stem or other progenitor cells, adrenal gland cells, cartilage cells, and combinations thereof.

[0488] In some embodiments, the engineered cell can 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.

[0489] 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*).

[0490] 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.

[0491] 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.

[0492] 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.

[0493] In some embodiments, the engineered cell is a cell obtained from a 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 targeting moiety, polypeptide, vector, viral (e.g., AAV) capsid particle is produced it can package one or more cargo polynucleotides that can be related to the desired physiological and/or biological characteristic and/or capable of modifying the desired physiological and/or biological characteristic. Thus, the cargo polynucleotides of the produced engineered viral (e.g., AAV) or other particle can be capable of transferring the desired characteristic to a recipient cell. In some embodiments, the cargo polynucleotides are capable of modifying a polynucleotide of the engineered cell such that the engineered cell has a desired physiological and/or biological characteristic.

[0494] In some embodiments, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences.

[0495] The engineered cells can be used to produce engineered targeting moieties, polypeptides, viral (e.g., AAV) capsid polynucleotides, vectors, and/or particles. In some embodiments, the engineered targeting moieties, polypeptides, viral (e.g., AAV) capsid polynucleotides, vectors, and/or particles are produced, harvested, and/or delivered to a subject in need thereof. In some embodiments, the engineered cells are delivered to a subject. Other uses for the engineered cells are described elsewhere herein. In some embodiments, the engineered cells can be included in formulations and/or kits described elsewhere herein.

[0496] The engineered cells can be stored short-term or long-term for use at a later time. Suitable storage methods are generally known in the art. Further, methods of restoring the stored cells for use (such as thawing, reconstitution, and otherwise stimulating metabolism in the engineered cell after storage) at a later time are also generally known in the art.

Formulations

[0497] Component(s) of the engineered targeting moieties, polypeptides, viral (e.g., AAV) capsid system, engineered cells, engineered viral (e.g., AAV) particles, and/or combinations thereof can be included in a formulation that can be delivered to a subject or a cell. In some embodiments, the formulation is a pharmaceutical formulation. One or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein can be

provided to a subject in need thereof or a cell alone or as an active ingredient, such as in a pharmaceutical formulation. As such, also described herein are pharmaceutical formulations containing an amount of one or more of the polypeptides, polynucleotides, vectors, cells, or combinations thereof described herein. In some embodiments, the pharmaceutical formulation can contain an effective amount of the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein. The pharmaceutical formulations described herein can be administered to a subject in need thereof or a cell.

[0498] In some embodiments, the amount of the one or more of the polypeptides, polynucleotides, vectors, cells, virus particles, nanoparticles, other delivery particles, and combinations thereof 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. The amount of the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein in the pharmaceutical formulation can range from about 1 pg to about 10 g, from about 10 nL to about 10 mL. In embodiments where the pharmaceutical formulation contains one or more cells, the amount can range from about 1 cell to 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} or more cells. In embodiments where the pharmaceutical formulation contains one or more cells, the amount can range from about 1 cell to 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} or more cells per nL, μ L, mL, or L.

[0499] In embodiments, where engineered AAV capsid particles are included in the formulation, the formulation can contain 1 to 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , 1×10^{18} , 1×10^{19} , or 1×10^{20} transducing units (TU)/mL of the engineered AAV capsid particles. In some embodiments, the formulation can be 0.1 to 100 mL in volume and can contain 1 to 1×10^1 , 1×10^2 , 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^7 , 1×10^8 , 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} , 1×10^{14} , 1×10^{15} , 1×10^{16} , 1×10^{17} , 1×10^{18} , 1×10^{19} , or 1×10^{20} transducing units (TU)/mL of the engineered AAV capsid particles.

Pharmaceutically Acceptable Carriers and Auxiliary Ingredients and Agents

[0500] In embodiments, the pharmaceutical formulation containing an amount of one or more of the polypeptides, polynucleotides, vectors, cells, virus particles, nanoparticles, other delivery particles, and combinations thereof described herein can further 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.

[0501] The pharmaceutical formulations can be sterilized, and if desired, mixed with auxiliary 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 composition.

[0502] In addition to an amount of one or more of the polypeptides, polynucleotides, vectors, cells, engineered viral (e.g. AAV) capsids, viral (e.g. AAV) or other particles, nanoparticles, other delivery particles, and combinations thereof described herein, the pharmaceutical formulation can also include an effective amount of an auxiliary active agent, including 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.

[0503] 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 testosterone 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- α , 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.

[0504] 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 salicylate, and sodium salicylate), paracetamol/acetaminophen, metamizole, nabumetone, phenazone, and quinine.

[0505] Suitable anxiolytics include, but are not limited to, benzodiazepines (e.g., alprazolam, bromazepam, chlordiazepoxide, clonazepam, clorazepate, diazepam, flurazepam, lorazepam, oxazepam, temazepam, triazolam, and tofisopam), serotonergic antidepressants (e.g., selective serotonin reuptake inhibitors, tricyclic antidepressants, and monoamine oxidase inhibitors), mebicar, afobazole, selank, bromantane, emoxypine, azapirones, barbiturates, hydroxyzine, pregabalin, validol, and beta blockers.

[0506] 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 ziconapine.

[0507] 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).

[0508] Suitable antispasmodics include, but are not limited to, mebeverine, papaverine, cyclobenzaprine, carisoprodol, orphenadrine, tizanidine, metaxalone, methocarbamol, 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).

[0509] 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, dexbrompheniramine, dexchlorpheniramine, dimenhydrinate, dimetindene, diphenhydramine, doxylamine, ebastine, embramine, fexofenadine, hydroxyzine, levocetirizine, loratadine, meclozine, mirtazapine, olopatadine, orphenadrine, phenindamine, pheniramine, phenyltoloxamine, promethazine, pyrilamine, quetiapine, rupatadine, tripeleminamine, and triprolidine), H₂-receptor antagonists (e.g., cimetidine, famotidine, lafutidine, nizatidine, ranitidine, and roxatidine), tritoqualine, catechin, cromoglicate, nedocromil, and p₂-adrenergic agonists.

[0510] 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, albendazole, thiabendazole, oxamniquine), antifungals (e.g., azole antifungals (e.g. itraconazole, fluconazole, parconazole, 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, abacavir/lamivudine/zidovudine, lamivudine/zidovudine, emtricitabine/tenofovir, emtricitabine/lopinavir/ritonavir/tenofovir, interferon alfa-2v/ribavirin, peginterferon alfa-2b, maraviroc, raltegravir, dolutegravir, enfuvirtide, foscarnet, fomivirsen, oseltamivir, zanamivir, nevirapine, efavirenz, etravirine, rilpivirine, delavirdine, nevirapine, entecavir, lamivudine, adefovir, sofosbuvir, didanosine, tenofovir, abacavir, zidovudine, stavudine, emtricitabine,

zalcitabine, telbivudine, simeprevir, boceprevir, telaprevir, lopinavir/ritonavir, boceprevir, darunavir, ritonavir, tipranavir, atazanavir, nelfinavir, amprenavir, indinavir, sawuina vir, ribavirin, valacyclovir, acyclovir, famciclovir, ganciclovir, and valganciclovir), carbapenems (e.g., doripenem, meropenem, ertapenem, and cilastatin/imipenem), cephalosporins (e.g., cefadroxil, cephadrine, cefazolin, cephalexin, cefepime, cefazoline, loracarbef, cefotetan, cefuroxime, cefprozil, loracarbef, cefoxitin, cefaclor, ceftibuten, ceftriaxone, cefotaxime, cefpodoxime, cefdinir, cefixime, cefditoren, ceftizoxime, and ceftazidime), glycopeptide antibiotics (e.g. vancomycin, dalbavancin, oritavancin, and telavancin), glycyclcyclines (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, erythromycin, azithromycin, clarithromycin, dirithromycin, and troleandomycin), linezolid, sulfamethoxazole/trimethoprim, rifaximin, chloramphenicol, Fosfomycin, metronidazole, aztreonam, bacitracin, penicillin (amoxicillin, ampicillin, bacampicillin, carbenicillin, piperacillin, ticarcillin, amoxicillin/clavulanate, ampicillin/sulbactam, piperacillin/tazobactam, clavulanate/ticarcillin, penicillin, procaine penicillin, oxacillin, dicloxacillin, and nafcillin), quinolones (e.g., lomefloxacin, norfloxacin, ofloxacin, qatifloxacin, moxifloxacin, ciprofloxacin, levofloxacin, gemifloxacin, moxifloxacin, cinoxacin, nalidixic acid, enoxacin, grepafloxacin, gatifloxacin, trovalfloxacin, and sparfloxacin), sulfonamides (e.g., sulfamethoxazole/trimethoprim, sulfasalazine, and sulfasoxazole), tetracyclines (e.g., doxycycline, demeclocycline, minocycline, doxycycline/salicylic 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).

[0511] 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, Cytosan, cyclophosphamide, decitabine, dexamethasone, docetaxel, hydroxyurea, decarbazine, leuprolide, epirubicin, oxaliplatin, asparaginase, estramustine, cetuximab, vismodegib, asparaginase *Erwinia chrysanthemi*, amifostine, etoposide, flutamide, toremifene, fulvestrant, letrozole, degarelix, pralatrexate, methotrexate, floxuridine,

obinutuzumab, gemcitabine, afatinib, imatinib mesylate, 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, octetide, dasatinib, regorafenib, histrelin, sunitinib, siltuximab, omacetaxine, thioguanine (tioguanine), dabrafenib, erlotinib, bexarotene, temozolomide, thiotepa, thalidomide, BCG, temsirolimus, bendamustine hydrochloride, triptorelin, arsenic 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.

[0512] In embodiments where there is an auxiliary active agent contained in the pharmaceutical formulation in addition to the one or more of the polypeptides, polynucleotides, compositions, vectors, cells, virus particles, nanoparticles, other delivery particles, and combinations thereof described herein, amount, such as an effective amount, of the auxiliary active agent will vary depending on the auxiliary active agent. In some embodiments, the amount of the auxiliary active agent ranges from 0.001 micrograms to about 1 milligram. In other embodiments, the amount of the auxiliary active agent ranges from about 0.01 IU to about 1000 IU. In further embodiments, the amount of the auxiliary active agent ranges from 0.001 mL to about 1 mL. In yet other embodiments, the amount of the auxiliary active agent ranges from about 1 % w/w to about 50% w/w of the total pharmaceutical formulation. In additional embodiments, the amount of the auxiliary active agent ranges from about 1 % v/v to about 50% v/v of the total pharmaceutical formulation. In still other embodiments, the amount of the auxiliary active agent ranges from about 1 % w/v to about 50% w/v of the total pharmaceutical formulation.

Dosage Forms

[0513] In some embodiments, the pharmaceutical formulations described herein may be in a dosage form. 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, epidural, intracranial, intraocular, inhaled, intranasal, topical (including buccal, sublingual, or transdermal), vaginal, intraurethral, parenteral, intracranial, subcutaneous, intramuscular, intravenous, intraperitoneal, intradermal, intraosseous, intracardiac, intraarticular, intracavernous, intrathecal, intravitreal, intracerebral, gingival, subgingival, intracerebroventricular, and intradermal. Such formulations may be prepared by any method known in the art.

[0514] Dosage forms adapted for oral administration can be 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 foam, spray, or liquid solution. In some embodiments, the oral dosage form can contain about 1 ng to 1000 g of a pharmaceutical formulation containing a therapeutically effective amount or an appropriate fraction thereof of the targeted effector fusion protein and/or complex thereof or composition containing the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein. The oral dosage form can be administered to a subject in need thereof.

[0515] Where appropriate, the dosage forms described herein can be microencapsulated.

[0516] The dosage form can also be prepared to prolong or sustain the release of any ingredient. In some embodiments, the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein can be the ingredient whose release is delayed. In other embodiments, the release of an optionally included auxiliary ingredient 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 & Wilkins, Baltimore, MD, 2000, and "Pharmaceutical dosage forms and drug delivery systems", 6th Edition, Ansel et al., (Media, PA: Williams and Wilkins, 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.

[0517] 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.

[0518] 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.

[0519] 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, the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein can be formulated with a paraffinic or water-miscible ointment base. In some embodiments, the 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.

[0520] Dosage forms adapted for nasal or inhalation administration include aerosols, solutions, suspension drops, gels, or dry powders. In some embodiments, the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein is contained 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 D50 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 ingredient (e.g., the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein and/or auxiliary active agent), which may be generated by various types of metered dose pressurized aerosols, nebulizers, or insufflators.

[0521] In some embodiments, the dosage forms can be aerosol formulations suitable for administration by inhalation. In some of these embodiments, the aerosol formulation can contain a solution or fine suspension of the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein 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.

[0522] 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 one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein. In further embodiments, the aerosol formulation can also contain 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, or 3 doses are delivered each time.

[0523] For some dosage forms suitable and/or adapted for inhaled administration, the pharmaceutical formulation is a dry powder inhalable formulation. In addition to the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein, an auxiliary active ingredient, and/or pharmaceutically acceptable salt thereof, such a dosage form can contain a powder base such as lactose, glucose, trehalose, mannitol, and/or starch. In some of these embodiments, the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein 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.

[0524] In some embodiments, the aerosol dosage forms can be arranged so that each metered dose of aerosol contains a predetermined amount of an active ingredient, such as the one or more of the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein.

[0525] 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.

[0526] Dosage forms adapted for parenteral administration and/or adapted for any type of injection (e.g., intravenous, intraperitoneal, subcutaneous, intramuscular, intradermal, intraosseous, epidural, intracardiac, intraarticular, intracavernous, gingival, subgingival, intrathecal, intravireal, intracerebral, and intracerebroventricular) can include aqueous and/or non-aqueous sterile injection solutions, which can contain anti-oxidants, 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 resuspended 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.

[0527] Dosage forms adapted for ocular administration can include aqueous and/or nonaqueous sterile solutions that can optionally be adapted for injection, and which can optionally contain anti-oxidants, buffers, bacteriostats, solutes that render the composition isotonic with the eye or fluid contained therein or around the eye of the subject, and aqueous and nonaqueous sterile suspensions, which can include suspending agents and thickening agents.

[0528] For some embodiments, the dosage form contains a predetermined amount of the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein per unit dose. In some embodiments, the predetermined amount of the Such

unit doses may therefore be administered once or more than once a day. Such pharmaceutical formulations may be prepared by any of the methods well known in the art.

Kits

[0529] Also described herein are kits that contain one or more of the one or more of the polypeptides, polynucleotides, vectors, cells, or other components described herein and combinations thereof and pharmaceutical formulations described herein. In embodiments, one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof described herein can be presented as a combination kit. As used herein, the terms "combination kit" or "kit of parts" refers to the compounds, or formulations and additional components that are used to package, screen, test, 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. The combination kit can contain one or more of the components (e.g., one or more of the one or more of the polypeptides, polynucleotides, vectors, cells, and combinations thereof) or formulation thereof can be provided in a single formulation (e.g., a liquid, lyophilized powder, etc.), or in separate formulations. The separate components or formulations can be contained in a single package or in separate packages within the kit. The kit can also include instructions in a tangible medium of expression that can contain information and/or directions regarding the content of the components and/or formulations contained therein, safety information regarding the content of the components(s) and/or formulation(s) contained therein, information regarding the amounts, dosages, indications for use, screening methods, component design recommendations and/or information, recommended treatment regimen(s) for the components(s) and/or formulations contained therein. As used herein, "tangible medium of expression" refers to a medium that is physically tangible or accessible and is not a mere abstract thought or an unrecorded spoken word. "Tangible medium of expression" includes, but is not limited to, words on a cellulosic or plastic material, or data stored in a suitable computer readable memory form. The data can be stored on a unit device, such as a flash memory drive or CD-ROM or on a server that can be accessed by a user via, e.g., a web interface.

[0530] In one embodiment, the invention provides a kit comprising one or more of the components described herein. In some embodiments, the kit comprises a vector system and instructions for using the kit. In some embodiments, the vector system includes a regulatory

element operably linked to one or more engineered targeting moiety, polypeptide, viral (e.g., AAV) delivery system polynucleotides, as described elsewhere herein and, optionally, a cargo molecule, which can optionally be operably linked to a regulatory element. The one or more engineered targeting moiety, polypeptide, viral (e.g., AAV) delivery system polynucleotides, can be included on the same or different vectors as a cargo molecule capable of being delivered by the engineered targeting moiety, polypeptide, viral (e.g., AAV) delivery system described herein in embodiments containing a cargo molecule within the kit.

[0531] In some embodiments, the kit comprises a vector system and instructions for using the kit. In some embodiments, the vector system comprises (a) a first regulatory element operably linked to a direct repeat sequence and one or more insertion sites for inserting one or more guide sequences up- or downstream (whichever applicable) of the direct repeat sequence, wherein when expressed, the guide sequence directs sequence-specific binding of a Cas9 CRISPR complex to a target sequence in a eukaryotic cell, wherein the Cas9 CRISPR complex comprises a Cas9 enzyme complexed with the guide sequence that is hybridized to the target sequence; and/or (b) a second regulatory element operably linked to an enzyme-coding sequence encoding said Cas9 enzyme comprising a nuclear localization sequence. Where applicable, a tracr sequence may also be provided. In some embodiments, the kit comprises components (a) and (b) located on the same or different vectors of the system. In some embodiments, component (a) further comprises two or more guide sequences operably linked to the first regulatory element, wherein when expressed, each of the two or more guide sequences direct sequence specific binding of a CRISPR complex to a different target sequence in a eukaryotic cell. In some embodiments, the Cas9 enzyme comprises one or more nuclear localization sequences of sufficient strength to drive accumulation of said CRISPR enzyme in a detectable amount in the nucleus of a eukaryotic cell. In some embodiments, the CRISPR enzyme is a type V or VI CRISPR system enzyme. In some embodiments, the CRISPR enzyme is a Cas9 enzyme. In some embodiments, the Cas9 enzyme is derived from *Francisella tularensis* 1, *Francisella tularensis* subsp. *novicida*, *Prevotella albensis*, *Lachnospiraceae* bacterium MC2017_1, *Butyrivibrio proteoclasticus*, *Peregrinibacteria* bacterium GW2011_GWA2_33_10, *Parcubacteria* bacterium GW2011_GWC2_44_17, *Smithella* sp. SCADC, *Acidaminococcus* sp. BV3L6, *Lachnospiraceae* bacterium MA2020, *Candidatus Methanoplasma termitum*, *Eubacterium eligens*, *Moraxella bovoculi* 237, *Leptospira inadai*, *Lachnospiraceae* bacterium ND2006, *Porphyromonas crevioricanis* 3, *Prevotella disiens*, or

Porphyromonas macacae Cas9 (e.g., modified to have or be associated with at least one DD), and may include further alteration or mutation of the Cas9, and can be a chimeric Cas9. In some embodiments, the DD-CRISPR enzyme is codon-optimized for expression in a eukaryotic cell. In some embodiments, the DD-CRISPR enzyme directs cleavage of one or two strands at the location of the target sequence. In some embodiments, the DD-CRISPR enzyme lacks or substantially DNA strand cleavage activity (e.g., no more than 5% nuclease activity as compared with a wild-type enzyme or enzyme not having the mutation or alteration that decreases nuclease activity). In some embodiments, the first regulatory element is a polymerase III promoter. In some embodiments, the second regulatory element is a polymerase II promoter. In some embodiments, the guide sequence is at least 16, 17, 18, 19, 20, 25 nucleotides, or between 16-30, or between 16-25, or between 16-20 nucleotides in length.

METHODS OF USE

General Discussion

[0532] The compositions containing the CNS-specific targeting moieties described herein (e.g., the engineered targeting moiety system polynucleotides, polypeptides, vector(s), engineered cells, engineered viral (e.g., AAV) capsids, and viral and other particles) can be used generally to package and/or deliver one or more cargo polynucleotides to a recipient cell. In some embodiments, delivery, is done in a cell-specific manner based upon the specificity of the targeting moiety(ies). In some embodiments, the cell-specificity is conferred via the n-mer motif(s) included in the targeting moiety as previously discussed. In some embodiments, delivery is done in cell-specific manner based upon the tropism of the engineered viral (e.g., AAV) capsid. In some embodiments, engineered targeting moiety(ies), polypeptides, viral (e.g., AAV) capsids, particles, viral (e.g., AAV) particles, compositions thereof, and/or cells discussed herein can be administered to a subject or a cell, tissue, and/or organ and facilitate the transfer and/or integration of the cargo polynucleotide to the recipient cell. In other embodiments, engineered cells capable of producing engineered targeting moiety(ies), polypeptides, viral (e.g., AAV) capsids, particles, viral (e.g., AAV) particles and/or compositions thereof can be generated from engineered targeting moiety system molecules (e.g., polynucleotides, vectors, and vector systems, etc.). In some embodiments, the engineered targeting moiety(ies), polypeptides, viral (e.g., AAV) capsids, particles, viral (e.g., AAV) particles and/or compositions thereof can be delivered to a subject or a cell, tissue, and/or organ. When delivered to a subject, they engineered delivery system molecule(s) can transform

a subject's cell in vivo or ex vivo to produce an engineered cell that can be capable of making an engineered targeting moiety(ies), polypeptides, viral (e.g., AAV) capsids, particles, viral (e.g., AAV) particles and/or compositions thereof, which can be released from the engineered cell and deliver cargo molecule(s) to a recipient cell in vivo or produce personalized engineered polypeptides, viral (e.g., AAV) particles, and/or other particles for reintroduction into the subject from which the recipient cell was obtained. In some embodiments, an engineered cell can be delivered to a subject, where it can release produced engineered targeting moieties, polypeptides, viral (e.g., AAV) particles, and/or other particles such that they can then deliver a cargo (e.g., cargo polynucleotide(s)) to a recipient cell. These general processes can be used in a variety of ways to treat and/or prevent disease or a symptom thereof in a subject, generate model cells, generate modified organisms, provide cell selection and screening assays, in bioproduction, and in other various applications.

[0533] In some embodiments, the engineered targeting moieties, polypeptides, viral (e.g., AAV) particles, and/or other particles, polynucleotides, vectors, and systems thereof can be used to generate engineered AAV capsid variant libraries that can be mined for variants with a desired cell-specificity, such as CNS specificity. The description provided herein as supported by the various Examples can demonstrate that one having a desired cell-specificity in mind could utilize the present invention as described herein to obtain a capsid with the desired cell-specificity, such as CNS specificity.

[0534] The subject invention may be used as part of a research program wherein there is transmission of results or data. A computer system (or digital device) may be used to receive, transmit, display and/or store results, analyze the data and/or results, and/or produce a report of the results and/or data and/or analysis. A computer system may be understood as a logical apparatus that can read instructions from media (e.g., software) and/or network port (e.g., from the internet), which can optionally be connected to a server having fixed media. A computer system may comprise one or more of a CPU, disk drives, input devices such as key board and/or mouse, and a display (e.g., a monitor). Data communication, such as transmission of instructions or reports, can be achieved through a communication medium to a server at a local or a remote location. The communication medium can include any means of transmitting and/or receiving data. For example, the communication medium can be a network connection, a wireless connection, or an internet connection. Such a connection can provide for communication over the World Wide Web. It is envisioned that data relating to the present

invention can be transmitted over such networks or connections (or any other suitable means for transmitting information, including but not limited to mailing a physical report, such as a print-out) for reception and/or for review by a receiver. The receiver can be but is not limited to an individual, or electronic system (e.g., one or more computers, and/or one or more servers). In some embodiments, the computer system comprises one or more processors. Processors may be associated with one or more controllers, calculation units, and/or other units of a computer system, or implanted in firmware as desired. If implemented in software, the routines may be stored in any computer readable memory such as in RAM, ROM, flash memory, a magnetic disk, a laser disk, or other suitable storage medium. Likewise, this software may be delivered to a computing device via any known delivery method including, for example, over a communication channel such as a telephone line, the internet, a wireless connection, etc., or via a transportable medium, such as a computer readable disk, flash drive, etc. The various steps may be implemented as various blocks, operations, tools, modules, and techniques which, in turn, may be implemented in hardware, firmware, software, or any combination of hardware, firmware, and/or software. When implemented in hardware, some or all of the blocks, operations, techniques, etc. may be implemented in, for example, a custom integrated circuit (IC), an application specific integrated circuit (ASIC), a field programmable logic array (FPGA), a programmable logic array (PLA), etc. A client-server, relational database architecture can be used in embodiments of the invention. A client-server architecture is a network architecture in which each computer or process on the network is either a client or a server. Server computers are typically powerful computers dedicated to managing disk drives (file servers), printers (print servers), or network traffic (network servers). Client computers include PCs (personal computers) or workstations on which users run applications, as well as example output devices as disclosed herein. Client computers rely on server computers for resources, such as files, devices, and even processing power. In some embodiments of the invention, the server computer handles all of the database functionality. The client computer can have software that handles all the front-end data management and can also receive data input from users. A machine-readable medium comprising computer-executable code may take many forms, including but not limited to, a tangible storage medium, a carrier wave medium or physical transmission medium. Non-volatile storage media include, for example, optical or magnetic disks, such as any of the storage devices in any computer(s) or the like, such as may be used to implement the databases, etc. shown in the drawings. Volatile storage

media include dynamic memory, such as main memory of such a computer platform. Tangible transmission media include coaxial cables; copper wire and fiber optics, including the wires that comprise a bus within a computer system. Carrier-wave transmission media may take the form of electric or electromagnetic signals, or acoustic or light waves such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media therefore include for example: a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD or DVD-ROM, any other optical medium, punch cards paper tape, any other physical storage medium with patterns of holes, a RAM, a ROM, a PROM and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave transporting data or instructions, cables or links transporting such a carrier wave, or any other medium from which a computer may read programming code and/or data. Many of these forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to a processor for execution. Accordingly, the invention comprehends performing any method herein-discussed and storing and/or transmitting data and/or results therefrom and/or analysis thereof, as well as products from performing any method herein-discussed, including intermediates.

Therapeutics

[0535] In some embodiments, one or more molecules of the engineered delivery system, engineered targeting moieties, polypeptides, viral (e.g., AAV) particles, and/or other particles, polynucleotides, vectors, systems thereof, engineered cells, and/or formulations thereof described herein can be delivered to a subject in need thereof as a therapy for one or more diseases. In some embodiments, the disease to be treated is a genetic or epigenetic based disease. In some embodiments, the disease to be treated is not a genetic or epigenetic based disease. In some embodiments, one or more molecules of the engineered delivery system, engineered targeting moieties, polypeptides, viral (e.g., AAV) particles, and/or other particles, polynucleotides, vectors, and systems thereof, engineered cells, and/or formulations thereof described herein can be delivered to a subject in need thereof as a treatment or prevention (or as a part of a treatment or prevention) of a disease. It will be appreciated that the specific disease to be treated and/or prevented by delivery of an engineered cell and/or engineered can be dependent on the cargo molecule packaged into an engineered AAV capsid particle.

[0536] Generally, the compositions described herein can be used in a therapy for treating a CNS disease, disorder, or a symptom thereof. It will be appreciated that a CNS disease or

disorder refers to any disease or disorder whose pathology involves or affects one or more cell types of the central nervous system. In some embodiments, the CNS disease or disorder is one whose primary pathology involves one or more cell types of the CNS. In some embodiments, one or more other cell types outside of the CNS are involved in the pathology of the CNS diseases, such as a muscle cell or peripheral nervous system cell. In some embodiments, the CNS disease or disorder can be caused by one or more genetic abnormalities. In some embodiments, the CNS disease or disorder is not caused by a genetic abnormality. Non-genetic cause of diseases include infection, cancer, physical trauma and others that will be appreciated by those of skill in the art. It also will be appreciated that gene modification approaches to treating disease can be applied to treat and/or prevent both genetic diseases and non-genetic diseases. For example, in the case of non-genetic diseases, a gene therapy approach can be used to modify the cause of the non-genetic disease (e.g., a cancer or infectious organism) such that the cause is no longer disease causing (e.g., by eliminating or rendering non-functional the cancer cells or infectious organism).

[0537] Exemplary CNS diseases and disorders include, without limitation, Friedreich's Ataxia, Dravet Syndrome, Spinocerebellar Ataxia Type 3, Niemann Pick Type C, Huntington's Disease, Pompe Disease, Myotonic Dystrophy Type 1, Glut1 Deficiency Syndrome (De Vivo Syndrome), Tay-Sachs, Spinal Muscular Atrophy, Alzheimer's disease, Amyotrophic lateral sclerosis (ALS), Danon disease, Rett Syndrome, Angelman Syndrome, or a combination thereof. Others are described elsewhere herein and/or will be appreciated by those of ordinary skill in the art in view of the description provided herein).

[0538] Genetic diseases that can be treated are discussed in greater detail elsewhere herein (see e.g., discussion on Gene-modification based-therapies below). Other diseases can include, but are not limited to, any of the following: cancer (such as glioblastoma or other brain or CNS cancers), Acubetivacter infections, actinomycosis, African sleeping sickness, AIDS/HIV, ameobiasis, Anaplasmosis, Angiostrongyliasis, Anisakiasis, Anthrax, Acranobacterium haemolyticum infection, Argentine hemorrhagic fever, Ascariasis, Aspergillosis, Astrovirus infection, Babesiosis, Bacterial meningitis, Bacterial pneumonia, Bacterial vaginosis, Bacteroides infection, balantidiasis, Bartonellosis, Baylisascaris infection, BK virus infection, Black Piedra, Blastocytosis, Blastomycosis, Bolivian hemorrhagic fever, Botulism, Brazillian hemorrhagic fever, brucellosis, Bubonic plague, Burkholderia infection, buruli ulcer, calicivirus invention, campylobacteriosis, Candidiasis, Capillariasis, Carrion's disease, Cat-

scratch disease, cellulitis, Chagas Disease, Chancroid, Chickenpox, Chikungunya, Chlamydia, Chlamydia pneumoniae, Cholera, Chromoblastomycosis, Chytridiomycosis, Clonochiasis, Clostridium difficile colitis, Coccidioidomycosis, Colorado tick fever, rhinovirus/coronavirus infection (common cold), Creutzfeldt-Jakob disease, Crimean-congo hemorrhagic fever, Cryptococcosis, Cryptosporidiosis, Cutaneous larva migrans (CLM), cyclosporiasis, cysticercosis, cytomegalovirus infection, Dengue fever, Desmodermis infection, Dientamoebiasis, Diphtheria, Diphylobothriasis, Dracunculiasis, Ebola, Echinococcosis, Ehrlichiosis, Enterobiasis, Enterococcus infection, Enterovirus infection, Epidemic typhus, Erythema Infectiosum, Exanthema subitum, Fascioliasis, Fasciolopsiasis, fatal familial insomnia, filariasis, Clostridium perfringens infection, Fusobacterium infection, Gas gangrene (clostridial myonecrosis), Geotrichosis, Gerstmann-Strausler-Scheinker syndrome, Giardiasis, Glanders, Gnathostomiasis, Gonorrhea, Granuloma inguinale, Group A streptococcal infection, Group B streptococcal infection, Haemophilus influenzae infection, Hand, foot, and mouth disease, hanta virus pulmonary syndrome, heartland virus disease, helicobacter pylori infection, hemorrhagic fever with renal syndrome, Hendra virus infection, Hepatitis (all groups A, B, C, D, E), hepatitis simplex, histoplasmosis, hookworm infection, human bocavirus infection, human ewingii ehrlichiosis, Human granulocytic anaplasmosis, human metapneumovirus infection, human monocytic ehrlichiosis, human papilloma virus, Hymenolepiasis, Epstein-Barr infection, mononucleosis, influenza, isosporiasis, Kawasaki disease, Kingella kingae infection, Kuru, Lassa fever, Legionellosis (Legionnaires's disease and Potomac Fever), Leishmaniasis, Leprosy, Leptospirosis, Listeriosis, Lyme disease, lymphatic filariasis, lymphocytic choriomeningitis, Malaria, Marburg hemorrhagic fever, measles, Middle East respiratory syndrome, Melioidosis, meningitis, Meningococcal disease, Metagonimiasis, Microsporidiosis, Molluscum contagiosum, Monkeypox, Mumps, Murine typhus, Mycoplasma pneumoniae, Mycoplasma genitalium infection, Mycetoma, Myiasis, Conjunctivitis, Nipah virus infection, Norovirus, Variant Creutzfeldt-Jakob disease, Nocardiosis, Onchocerciasis, Opisthorchiasis, Paracoccidioidomycosis, Paragonimiasis, Pasteurellosis, Pediculosis capitis, Pediculosis corporis, Pediculosis pubis, pelvic inflammatory disease, pertussis, plague, pneumococcal infection, pneumocystis pneumonia, pneumonia, poliomyelitis, prevotella infection, primary amoebic meningoencephalitis, progressive multifocal leukoencephalopathy, Psittacosis, Q fever, rabies, relapsing fever, respiratory syncytial virus infection, rhinovirus infection, rickettsial infection, Rickettsialpox, Rift Valley Fever, Rocky Mountain Spotted Fever, Rotavirus

infection, Rubella, Salmonellosis, SARS, Scabies, Scarlet fever, Schistosomiasis, Sepsis, Shigellosis, Shingles, Smallpox, Sporotrichosis, Staphylococcal infection (including MRSA), strongyloidiasis, subacute sclerosing panencephalitis, Syphilis, Taeniasis, tetanus, Trichophyton species infection, Toxocariasis, Toxoplasmosis, Trachoma, Trichinosis, Trichuriasis, Tuberculosis, Tularemia, Typhoid Fever, Typhus Fever, Ureaplasma urealyticum infection, Valley fever, Venezuelan equine encephalitis, Venezuelan hemorrhagic fever, Vibrio species infection, Viral pneumonia, West Nile Fever, White Piedra, Yersinia pseudotuberculosis, Yersiniosis, Yellow fever, Zoonosis, Zika fever, Zygomycosis and combinations thereof.

[0539] Other diseases and disorders that can be treated using embodiments of the present invention include, but are not limited to, endocrine diseases (e.g., Type I and Type II diabetes, gestational diabetes, hypoglycemia, Glucagonoma, Goitre, Hyperthyroidism, hypothyroidism, thyroiditis, thyroid cancer, thyroid hormone resistance, parathyroid gland disorders, Osteoporosis, osteitis deformans, rickets, osteomalacia, hypopituitarism, pituitary tumors, etc.), skin conditions of infectious and non-infectious origin, eye diseases of infectious or non-infectious origin, gastrointestinal disorders of infectious or non-infectious origin, cardiovascular diseases of infectious or non-infectious origin, brain and neuron diseases of infectious or non-infectious origin, nervous system diseases of infectious or non-infectious origin, muscle diseases of infectious or non-infectious origin, bone diseases of infectious or non-infectious origin, reproductive system diseases of infectious or non-infectious origin, renal system diseases of infectious or non-infectious origin, blood diseases of infectious or non-infectious origin, lymphatic system diseases of infectious or non-infectious origin, immune system diseases of infectious or non-infectious origin, mental-illness of infectious or non-infectious origin and the like.

[0540] In some embodiments, the disease to be treated is a CNS or CNS related disease or disorder, such as a genetic CNS disease or disorder. Such CNS or CNS related disease (including genetic CNS disease or disorders are described in greater detail elsewhere herein).

[0541] Other diseases and disorders will be appreciated by those of skill in the art.

Adoptive Cell Therapies

[0542] Generally speaking, adoptive cell transfer involves the transfer of cells (autologous, allogeneic, and/or xenogeneic) to a subject. The cells may or may not be modified and/or otherwise manipulated prior to delivery to the subject. Manipulation can include genetic

modification by one or more gene modifying agents. Exemplary gene modifying agents and systems are described in greater detail elsewhere herein and will be appreciated by those of ordinary skill in the art. Such gene or other modification compositions or systems can be delivered to a cell to be modified for adoptive therapy by one or more of the compositions described herein containing a CNS specific targeting moiety.

[0543] In some embodiments, an engineered cell as described herein can be included in an adoptive cell transfer therapy. In some embodiments, an engineered cell as described herein can be delivered to a subject in need thereof. In some embodiments, the cell can be isolated from a subject, manipulated in vitro such that it is capable of generating an engineered AAV capsid particle described herein to produce an engineered cell and delivered back to the subject in an autologous manner or to a different subject in an allogeneic or xenogeneic manner. The cell isolated, manipulated, and/or delivered can be a eukaryotic cell. The cell isolated, manipulated, and/or delivered can be a stem cell. The cell isolated, manipulated, and/or delivered can be a differentiated cell. The cell isolated, manipulated, and/or delivered can be an immune cell, a blood cell, an endocrine cell, a brain cell, a nervous system cell, a vascular cell, a muscle cell, a soft tissue cell, a neuron, a glial cell, an astrocyte, a Schwann cell, a microglial cell, or other neuron support cell, or a combination thereof. Other specific cell types will instantly be appreciated by one of ordinary skill in the art.

[0544] In some embodiments, the isolated cell can be manipulated such that it becomes an engineered cell as described elsewhere herein (e.g., contain and/or express one or more engineered delivery system molecules or vectors described elsewhere herein). Methods of making such engineered cells are described in greater detail elsewhere herein.

[0545] The administration of the cells or population of cells according to the present invention may be carried out in any convenient manner, including 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, by intravenous or intralymphatic injection, or intraperitoneally. In one embodiment, the cell compositions of the present invention are preferably administered by intravenous injection.

[0546] The administration of the cells or population of cells can be or involve the administration of 10^4 - 10^9 cells per kg body weight including all integer values of cell numbers within those ranges. In some embodiments, 10^5 to 10^6 cells/kg are delivered. Dosing in adoptive

cell therapies may for example involve administration of from 10^6 to 10^9 cells/kg, with or without a course of lymphodepletion, for example with cyclophosphamide. 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.

[0547] In another embodiment, the effective amount of cells or composition comprising those cells are administrated parenterally. The administration can be an intravenous administration. The administration can be directly done by injection within a tissue. In some embodiments, the tissue can be a tumor.

[0548] To guard against possible adverse reactions, engineered cells can be equipped with 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 the engineered cell similar to that discussed in 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)).

[0549] Methods of modifying isolated cells to obtain the engineered cells with the desired properties are described elsewhere herein. In some embodiments, the methods can include genome modification, including, but not limited to, genome editing using a CRISPR-Cas system to modify the cell. This can be in addition to introduction of an engineered AAV capsid system molecule describe elsewhere herein.

[0550] 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 cells, such as engineered cells described herein. 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 a particular embodiment, the present invention further comprises a step of modifying the engineered cells 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 engineered cells for adoptive cell therapy by inactivating the target of the immunosuppressive agent in engineered 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.

[0551] 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 is the programmed death-1 (PD-1 or CD279) gene (PDCD1). In other embodiments, the immune checkpoint targeted 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 further additional embodiments, the immune checkpoint targeted is a member of the TNFR superfamily such as CD40, OX40, CD137, GITR, CD27 or TIM-3.

[0552] Additional immune checkpoints include 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). SHP-1 is a widely expressed inhibitory protein tyrosine phosphatase (PTP). In T-cells, it is a negative regulator of antigen-dependent activation and proliferation. It is a cytosolic protein, and therefore not amenable to antibody-mediated therapies, but its role in activation and proliferation makes it an attractive target for genetic manipulation in adoptive transfer strategies, such as chimeric antigen receptor (CAR) T cells. Immune checkpoints may also include 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).

[0553] WO2014172606 relates to the use of MT1 and/or MT1 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 by gene editing in adoptively transferred T cells.

[0554] In certain embodiments, targets of gene editing may be at least one targeted locus involved in the expression of an immune checkpoint protein. 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 or TIM-3. In some embodiments, the gene locus involved in the expression of PD-1 or CTLA-4 genes is targeted. In some embodiments, combinations of genes are targeted, such as but not limited to PD-1 and TIGIT.

[0555] In some embodiments, at least two genes are edited. Pairs of genes may include, but are not limited to PD1 and TCR α , PD1 and TCR β , CTLA-4 and TCR α , CTLA-4 and TCR β ,

LAG3 and TCR α , LAG3 and TCR β , Tim3 and TCR α , Tim3 and TCR β , BTLA and TCR α , BTLA and TCR β , BY55 and TCR α , BY55 and TCR β , TIGIT and TCR α , TIGIT and TCR β , B7H5 and TCR α , B7H5 and TCR β , LAIR1 and TCR α , LAIR1 and TCR β , SIGLEC10 and TCR α , SIGLEC10 and TCR β , 2B4 and TCR α , 2B4 and TCR β .

[0556] Whether prior to or after genetic or other modification of the engineered cells (such as engineered T cells (e.g., the isolated cell is a T cell), the engineered cells can be activated and expanded generally using methods as 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. The engineered cells can be expanded *in vitro* or *in vivo*.

[0557] In some embodiments, the method comprises editing the engineered cells *ex vivo* by a suitable gene modification method described elsewhere herein (e.g., gene editing via a CRISPR-Cas or IscBsystem) to eliminate potential alloreactive TCRs or other receptors to allow allogeneic adoptive transfer. In some embodiments, T cells are edited *ex vivo* by a CRISPR-Cas system or other suitable genome modification technique to knock-out or knock-down an endogenous gene encoding a TCR (e.g., an $\alpha\beta$ TCR) or other relevant receptor to avoid graft-versus-host-disease (GVHD). In some embodiments, where the engineered cells are T cells, the engineered cells are edited *ex vivo* by CRISPR or other appropriate gene modification method to mutate the TRAC locus. In some embodiments, T cells are edited *ex vivo* via a CRISPR-Cas system using one or more guide sequences targeting the first exon of TRAC. See Liu et al., Cell Research 27:154-157 (2017). In some embodiments, the first exon of TRAC is modified using another appropriate gene modification method. In some embodiments, the method comprises use of CRISPR or other appropriate method to knock-in an exogenous gene encoding a CAR or a TCR into the TRAC locus, while simultaneously knocking-out the endogenous TCR (e.g., with a donor sequence encoding a self-cleaving P2A peptide following the CAR cDNA). See Eyquem et al., Nature 543:113-117 (2017). In some embodiments, the exogenous gene comprises a promoter-less CAR-encoding or TCR-encoding sequence which is inserted operably downstream of an endogenous TCR promoter.

[0558] In some embodiments, the method comprises editing the engineered cell, e.g., engineered T cells, *ex vivo* via a CRISPR-Cas system to knock-out or knock-down an endogenous gene encoding an HLA-I protein to minimize immunogenicity of the edited cells, e.g. engineered T cells. In some embodiments, engineered T cells can be edited *ex vivo* via a

CRISPR-Cas system to mutate the beta-2 microglobulin (B2M) locus. In some embodiments, engineered cell, e.g., engineered T cells, are edited ex vivo via a CRISPR-Cas system using one or more guide sequences targeting the first exon of B2M. The first exon of B2M can also be modified using another appropriate modification method. See Liu et al., Cell Research 27:154-157 (2017). The first exon of B2M can also be modified using another appropriate modification method, which will be appreciated by those of ordinary skill in the art. In some embodiments, the method comprises use a CRISPR-Cas system to knock-in an exogenous gene encoding a CAR or a TCR into the B2M locus, while simultaneously knocking-out the endogenous B2M (e.g., with a donor sequence encoding a self-cleaving P2A peptide following the CAR cDNA). See Eyquem et al., Nature 543:113-117 (2017). This can also be accomplished using another appropriate modification method, which will be appreciated by those of ordinary skill in the art. In some embodiments, the exogenous gene comprises a promoter-less CAR-encoding or TCR-encoding sequence which is inserted operably downstream of an endogenous B2M promoter.

[0559] In some embodiments, the method comprises editing the engineered cell, e.g., engineered T cells, ex vivo via a CRISPR-Cas system to knock-out or knock-down an endogenous gene encoding an antigen targeted by an exogenous CAR or TCR. This can also be accomplished using another appropriate modification method, which will be appreciated by those of ordinary skill in the art. In some embodiments, the engineered cells, such as engineered T cells, are edited ex vivo via a CRISPR-Cas system to knock-out or knock-down the expression of a tumor antigen selected from 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 or cyclin (DI) (see WO2016/011210). This can also be accomplished using another appropriate modification method, which will be appreciated by those of ordinary skill in the art. In some embodiments, the engineered cells, such as engineered T cells are edited ex vivo via a CRISPR-Cas system to knock-out or knock-down the expression of an antigen selected from B cell maturation antigen (BCMA), transmembrane activator and CAML Interactor (TACI), or B-cell activating factor receptor (BAFF-R), CD38, CD138, CS-1, CD33, CD26, CD30, CD53, CD92, CD100, CD148, CD150, CD200, CD261, CD262, or CD362 (see WO2017/011804). This can

also be accomplished using another appropriate modification method, which will be appreciated by those of ordinary skill in the art.

Gene Drives

[0560] The present invention also contemplates use of the engineered delivery system molecules, vectors, engineered cells, and/or engineered AAV capsid particles described herein to generate a gene drive via delivery of one or more cargo polynucleotides or production of engineered AAV capsid particles with one or more cargo polynucleotides capable of producing a gene drive. In some embodiments, the gene drive can be a Cas-mediated RNA-guided gene drive e.g., Cas- to provide 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 (such as gene drive engineered delivery system 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).

Transplantation and Xenotransplantation

[0561] The engineered AAV capsid system molecules, vectors, engineered cells, and/or engineered delivery particles described herein, can be used to deliver cargo polynucleotides and/or otherwise be involved in modifying tissues for transplantation between two different persons (transplantation) or between species (xenotransplantation). Such techniques for generation of transgenic animals are described elsewhere herein. Interspecies transplantation techniques are generally known in the art. For example, RNA-guided DNA nucleases can be delivered using via engineered AAV capsid polynucleotides, vectors, engineered cells, and/or engineered AAV capsid particles described herein and can be used to knockout, knockdown or disrupt selected genes in an organ for transplant (e.g. ex vivo (e.g. after harvest but before transplantation) or in vivo (in donor or recipient)), 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 $\alpha(1,3)$ -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.

[0562] Where it is interspecies transplantation (such as human to human) the engineered AAV capsid system molecules, vectors, engineered cells, and/or engineered delivery particles

described herein, can be used to deliver cargo polynucleotides and/or otherwise be involved to modify the tissue to be transplanted. In some embodiments, the modification can include modifying one or more HLA antigens or other tissue type determinants, such that the immunogenic profile is more similar or identical to the recipient's immunogenic profile than to the donor's so as to reduce the occurrence of rejection by the recipient. Relevant tissue type determinants are known in the art (such as those used to determine organ matching) and techniques to determine the immunogenic profile (which is made up of the expression signature of the tissue type determinants) are generally known in the art.

[0563] In some embodiments, the donor (such as before harvest) or recipient (after transplantation) can receive one or more of the engineered AAV capsid system molecules, vectors, engineered cells, and/or engineered delivery particles described herein that are capable of modifying the immunogenic profile of the transplanted cells, tissue, and/or organ. In some embodiments, the transplanted cells, tissue, and/or organ can be harvested from the donor and the engineered AAV capsid system molecules, vectors, engineered cells, and/or engineered delivery particles described herein capable of modifying the harvested cells, tissue, and/or organ to be, for example, less immunogenic or be modified to have some specific characteristic when transplanted in the recipient can be delivered to the harvested cells, tissue, and/or organ *ex vivo*. After delivery the cells, tissue, and/or organs can be transplanted into the donor.

Gene Modification and Treatment of Diseases with Genetic or Epigenetic Embodiments that affect the CNS, brain, and/or neurons

[0564] The engineered delivery system molecules, vectors, engineered cells, and/or engineered delivery particles described herein (e.g., those with one or more targeting moieties, such as a CNS-specific targeting moiety described herein) can be used to modify genes or other polynucleotides and/or treat diseases of the CNS, brain, and/or neurons with genetic and/or epigenetic embodiments. As described elsewhere herein the cargo molecule can be a polynucleotide that can be delivered to a cell and, in some embodiments, be integrated into the genome of the cell. In some embodiments, the cargo molecule(s) can be one or more CRISPR-Cas system components. In some embodiments, the CRISPR-Cas components, when delivered by an engineered AAV capsid particles described herein can be optionally expressed in the recipient cell and act to modify the genome of the recipient cell in a sequence specific manner. In some embodiments, the cargo molecules that can be packaged and delivered by the engineered AAV capsid particles described herein can facilitate/mediate genome modification

via a method that is not dependent on CRISPR-Cas. Such non-CRISPR-Cas genome modification systems will instantly be appreciated by those of ordinary skill in the art and are also, at least in part, described elsewhere herein. In some embodiments, modification is at a specific target sequence. In other embodiments, modification is at locations that appear to be random throughout the genome.

[0565] Examples of CNS, brain, and/or neuronal disease-associated genes and polynucleotides that can be modified using the engineered delivery AAV delivery system molecules, vectors, capsids, engineered cells, and/or engineered delivery particles described herein are described below.

[0566] In some embodiments, a therapeutic or preventive, such as the engineered AAV capsids and systems thereof as described elsewhere herein, can be delivered to a subject in need thereof or a cell thereof to treat a brain, neuron, neurological, and/or central nervous system disease or disorder (CNS). In some embodiments the brain, neuron, neurological, and/or CNS disease or disorder can be caused, directly or indirectly, by one or mutations in one or more of the following genes as compared to normal or non-pathological variant of the same: in the case of Amyotrophic lateral sclerosis (ALS): SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b, VEGF-c); in the case of Alzheimer's disease: E1, CHIP, UCH, UBB, Tau, LRP, PICALM, Clusterin, PS1, SORL1, CR1, Vldlr, Uba1, Uba3, CHIP28, Aqp1, Uchl1, Uchl3, APP, AAA, CVAP, AD1, APOE, AD2, PSEN2, AD4, STM2, APBB2, FE65L1, NOS3, PLA2, URK, ACE, DCPI, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3); in the case of Autism: Mecp2, BZRAP1, MDGA2, Sema5A, Neurexin 1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4, KIAA1260, AUTSX2; in the case of Fragile X Syndrome: FMR2, FXR1, FXR2, mGLUR5; in the case of Huntington's disease and disease like disorders: HD, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); in the case of Parkinson's disease: NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA, NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1, PARK6, UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN, PARK2, PDJ, DBH, NDUFV2, PINK1, x-synuclein); in the case of Rett syndrome: MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-Synuclein, DJ-1; in the case of Schizophrenia: 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)); in

the case of Secretase Related Disorders (APH-1 (alpha and beta), Presenilin (Psen1), nicastrin, (Ncstn), PEN-2, Nos1, Parp1, Nat1, Nat2); in the case of Trinucleotide Repeat Disorders (HTT (Huntington's Dx), SBMA/SMAX1/AR (Kennedy's Dx), FXN/X25 (Friedrich's Ataxia), ATX3 (Machado- Joseph's Dx), ATXN1 and ATXN2 (spinocerebellar ataxias), DMPK (myotonic dystrophy), Atrophin-1 and Atn1 (DRPLA Dx), CBP (Creb-BP - global instability), VLDLR (Alzheimer's), Atxn7, Atxn10); in the case of diseases or disorders associated with or involving aberrant or abnormal axonal guidance signaling in the brain, neurons, and/or CNS: PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12; IGF1; RAC1; RAPIA; 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; in the case of diseases or disorders associated with or involving aberrant or abnormal actin cytoskeleton signaling in the brain, neurons, and/or CNS: 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; in the case of diseases or disorders associated with or involving 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal apoptosis regulation and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal leukocyte extravasation signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal integrin signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal acute phase response signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal PTEN signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L1; 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; in the case of diseases or disorders

associated with or involving aberrant, pathologic, and/or abnormal p53 signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal aryl hydrocarbon receptor signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal xenobiotic metabolism signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal SAPK/JNK signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal PPAR/RXR signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal NF-kappaB signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal neuregulin signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal wnt and beta catenin signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal insulin receptor signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal IL-6 signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal IGF-1 signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal NRF2-mediated oxidative stress response pathway regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; PRDX1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal PPAR (e.g. PPAR alpha, PPAR beta, PPAR delta, and/or PPAR gamma) regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Fc Epsilon RI regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal G-protein coupled receptor regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving

aberrant, pathologic, and/or abnormal inositol phosphate metabolism regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal PDGF regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with involving aberrant, pathologic, and/or abnormal VEGF regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal natural killer cell regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal cell cycle G1/S checkpoint regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal T-cell receptor regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant,

pathologic, and/or abnormal death receptor regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: CRADD; HSPB1; BID; BIRC4; TBK1; IKKBK; FADD; FAS; NFkB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8; DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBK; RELB; CASP9; CHUK; APAF1; NFkB1; CASP2; BIRC2; CASP3; BIRC3; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or FGF regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or GM-CSF regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or amyotrophic lateral sclerosis regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or JAK/Stat regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or nicotinate and nicotinamide metabolism regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or chemokine signaling regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or IL-2 signaling regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving synaptic long term depression in the brain, neurons, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or estrogen receptor regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2; SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1; HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP; MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or protein ubiquitination pathway activity, regulation, and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4; CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7; USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8; USP1; VHL; HSP90AA1; BIRC3; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or IL-10 regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF; IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1; JUN; IL1R1; IL6; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or Vitamin D receptor (VDR) and/or RXR regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1; NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD; RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1; LRP5; CEBPB; FOXO1; PRKCA; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or TGF-beta regulation or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1; FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2;

SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2; MAP2K1; TGFBR1; SMAD4; JUN; SMAD5; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or Toll-like Receptor activity, regulation, and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK; NFKB1; TLR2; JUN; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or p38 MAPK activity, regulation, and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS; CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2; MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1; SRF; STAT1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or neurotrophin/TRK activity, regulation, and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A; RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; CDC42; JUN; ATF4; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or FXR and/or RXR activity, regulation, and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8; APOB; MAPK10; PPARG; MTPP; MAPK9; PPARGC1A; TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or synaptic long term potentiation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1; PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS; PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1; ATF4; PRKCA; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or calcium regulation and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1; CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11; HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4; HDAC6; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or EGF or EGFR regulation and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1; STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1; in the case of diseases or disorders associated with or involving

aberrant, pathologic, and/or LPS/IL-1 mediated inhibition of RXR function, regulation and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1; MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2; TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or LXR/RXR function, regulation and/or signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA; NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1; SREBF1; IL1R1; CCL2; IL6; MMP9; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or amyloid processing in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2; CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1; PSEN1; CSNK1A1; GSK3B; AKT3; APP; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal IL-4 activity, signaling, and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1; PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1; FRAP1; AKT3; RPS6KB1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal cell cycle: G2/M DNA damage checkpoint regulation activity, signaling, and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC; CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A; PRKDC; ATM; SFN; CDKN2A; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal purine metabolism signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4; PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C; NT5E; POLD1; NME1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal cAMP-mediated signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: RAPIA; MAPK1; GNAS; CREB1; CAMK2A; MAPK3; SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal mitochondrial function in the brain, neurons, and/or CNS and/or diseases or disorders thereof: SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9; PARK7; PSEN1; PARK2; APP; CASP3; AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2; in the case of diseases or disorders associated with or involving aberrant,

pathologic, and/or abnormal notch signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2; PSEN1; NOTCH3; NOTCH1; DLL4; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal endoplasmic reticulum stress pathway activity, signaling, and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4; EIF2AK3; CASP3; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal pyrimidine metabolism, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B; NT5E; POLD1; NME1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Parkinson's signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7; PARK2; CASP3; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Glycolysis/Gluconeogenesis activity, signaling, and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal interferon activity, signaling, and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal sonic the hedgehog activity, signaling, and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal glycerophospholipid metabolism, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal phospholipid degradation, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal tryptophan metabolism, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1; in the case of

diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal lysine degradation, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal nucleotide excision repair pathway activity, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ERCC5; ERCC4; XPA; XPC; ERCC1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal nucleotide starch and sucrose metabolism, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: UCHL1; HK2; GCK; GPI; HK1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal aminosugars metabolism, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: NQO1; HK2; GCK; HK1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal arachidonic acid metabolism, signaling thereof, and/or regulation thereof in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRDX6; GRN; YWHAZ; CYP1B1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal circadian rhythm signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: CSNK1E; CREB1; ATF4; NR1D1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or coagulation system activity signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: BDKRB1; F2R; SERPINE1; F3; a PAR (e.g. PAR1, PAR2, etc.) PLC, aPC; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal dopamine receptor signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PPP2R1A; PPP2CA; PPP1CC; PPP2R5C; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Glutathione Metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: IDH2; GSTP1; ANPEP; IDH1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Glycerolipid Metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; GPAM; SPHK1; SPHK2; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Linoleic Acid Metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or

diseases or disorders thereof: PRDX6; GRN; YWHAZ; CYP1B1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Methionine Metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: DNMT1; DNMT3B; AHCY; DNMT3A; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Pyruvate Metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: GLO1; ALDH1A1; PKM2; LDHA; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Arginine and Proline Metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; NOS3; NOS2A; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Eicosanoid signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRDX6; GRN; YWHAZ; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal fructose and mannose metabolism signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: HK2; GCK; HK1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal antigen presentation pathway activity, signaling and/or regulation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: CALR; B2M; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal steroid biosynthesis in the brain, neurons, and/or CNS and/or diseases or disorders thereof: NQO1; DHCR7; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal butanoate metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; NLGN1; in the case of diseases or disorders associated with or involving an aberrant, pathologic, and/or abnormal citrate cycle in the brain, neurons, and/or CNS and/or diseases or disorders thereof: IDH2; IDH1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal fatty acid metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; CYP1B1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Glycerophospholipid metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRDX6; CHKA; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal histidine metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRMT5; ALDH1A1; in the case of diseases or disorders associated with or

involving aberrant, pathologic, and/or abnormal inositol metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ERO1L; APEX1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Phenylalanine metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRDX6; PRDX1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Seleno amino acid metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRMT5; AHCY; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Sphingolipid metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: SPHK1; SPHK2; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Aminophosphonate metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRMT5; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal androgen and/or estrogen metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRMT5; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Ascorbate and Aldarate metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Cysteine Metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: LDHA; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal fatty acid biosynthesis in the brain, neurons, and/or CNS and/or diseases or disorders thereof: FASN; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal glutamate receptor signaling in the brain, neurons, and/or CNS and/or diseases or disorders thereof: GNB2L1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Pentose Phosphate pathway in the brain, neurons, and/or CNS and/or diseases or disorders thereof: GPI; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal retinol metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Pentose and Glucuronate interconversions in the brain, neurons, and/or CNS and/or diseases or disorders thereof: UCHL1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Riboflavin Metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: TYR; in the case of

diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Tyrosine Metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRMT5, TYR; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Ubiquinone biosynthesis in the brain, neurons, and/or CNS and/or diseases or disorders thereof: PRMT5; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal Valine, leucine and isoleucine degradation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal glycine, serine, and threonine metabolism in the brain, neurons, and/or CNS and/or diseases or disorders thereof: CHKA; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal lysine degradation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: ALDH1A1; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal pain or pain signaling or pain signal generation in the brain, neurons, and/or CNS and/or diseases or disorders thereof: TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2; Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca; Prkacb; Prkar1a; Prkar2a; in the case of diseases or disorders associated with or involving aberrant, pathologic, and/or abnormal brain, neuron, and/or CNS development and/or diseases or disorders thereof: 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 Bm3a); Numb; Reln; in the case of diseases or disorders associated with or involving prion disorders of or in the brain, neuron, and/or CNS and/or diseases or disorders thereof: Prp; in the case of substance or activity additions involving activities of the brain, neuron, and/or CNS: Prkce (alcohol); Drd2; Drd4; ABAT (alcohol); GRIA2; Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1 (alcohol); in the case of diseases or disorders associated with or involving PI3K/AKT signaling and/or regulation thereof in the brain, neuron, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving ERK/MAPK signaling and/or regulation thereof in the brain, neuron, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving glucocorticoid receptor signaling and/or regulation thereof in the brain, neuron, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving ephrin receptor signaling and/or regulation thereof in the brain, neuron, and/or CNS and/or diseases or disorders thereof: 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; in the case of diseases or disorders associated with or involving B cell receptor signaling and/or regulation thereof in the brain, neuron, and/or CNS and/or diseases or disorders thereof: 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.

[0567] Thus, also described herein are methods of inducing one or more mutations in a eukaryotic or prokaryotic cell (in vitro, i.e. in an isolated eukaryotic cell) as herein discussed comprising delivering to cell a vector as described herein. The mutation(s) can include the introduction, deletion, or substitution of one or more nucleotides at a target sequence of cell(s). In some embodiments, the mutations can include the introduction, deletion, or substitution of 1-75 nucleotides at each target sequence of said cell(s). The mutations 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 each target sequence. The mutations 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 each target sequence of said cell(s). The mutations 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 each target sequence of said cell(s). The mutations 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 each target sequence of said cell(s). The mutations can include the introduction, deletion, or substitution of 40, 45, 50, 75, 100, 200, 300, 400 or 500 nucleotides at each target sequence of said cell(s). The mutations 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 nucleotides at each target sequence of said cell(s).

[0568] In some embodiments, the modifications can include the introduction, deletion, or substitution of nucleotides at each target sequence of said cell(s) via nucleic acid components (e.g. guide(s) RNA(s) or sgRNA(s)), such as those mediated by a CRISPR-Cas system.

[0569] In some embodiments, the modifications can include the introduction, deletion, or substitution of nucleotides at a target or random sequence of said cell(s) via a non CRISPR-Cas system or technique. Such techniques are discussed elsewhere herein, such as where engineered cells and methods of generating the engineered cells and organisms are discussed.

[0570] For minimization of toxicity and off-target effect when using a CRISPR-Cas system, it may be important to control the concentration of Cas mRNA and guide RNA delivered. Optimal concentrations of Cas mRNA and guide RNA can be determined by testing different concentrations in a cellular or non-human eukaryote animal model and using deep sequencing to analyze the extent of modification at potential off-target genomic loci. Alternatively, to minimize the level of toxicity and off-target effect, Cas nickase mRNA (for example *S. pyogenes* Cas9-like with the D10A mutation) can be delivered with a pair of guide RNAs targeting a site of interest. Guide sequences and strategies to minimize toxicity and off-target effects can be as in WO 2014/093622 (PCT/US2013/074667); or, via mutation as herein.

[0571] Typically, in the context of an endogenous CRISPR system, formation of a CRISPR complex (comprising a guide sequence hybridized to a target sequence and complexed with one or more Cas proteins) results in cleavage of one or both strands in or near (e.g., within 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, or more base pairs from) the target sequence. Without wishing to be bound by theory, a tracr sequence, which may comprise or consist of all or a portion of a wild-type tracr sequence (e.g. about or more than about 20, 26, 32, 45, 48, 54, 63, 67, 85, or more nucleotides of a wild-type tracr sequence), may also form part of a CRISPR complex, such as by hybridization along at least a portion of the tracr sequence to all or a portion of a tracr mate sequence that is operably linked to a guide sequence.

[0572] In one embodiment, the invention provides a method of modifying a target polynucleotide in a eukaryotic cell. In some embodiments, the method includes delivering an engineered targeting moiety, polypeptide, polynucleotide, vector, vector system, particle, viral (e.g., AAV) particle, cell, or any combination thereof described herein having a CRISPR-Cas molecule as a cargo molecule to a subject and/or cell. The CRISPR-Cas system molecule(s) delivered can complex to bind to the target polynucleotide, e.g., to effect cleavage of said target polynucleotide, thereby modifying the target polynucleotide, wherein the CRISPR complex comprises a CRISPR enzyme complexed with a guide sequence hybridized to a target sequence within said target polynucleotide, wherein said guide sequence can be linked to a tracr mate sequence which in turn hybridizes to a tracr sequence. In some embodiments, said cleavage

comprises cleaving one or two strands at the location of the target sequence by said CRISPR enzyme. In some embodiments, said cleavage results in decreased transcription of a target gene. In some embodiments, the method further comprises repairing said cleaved target polynucleotide by homologous recombination with an exogenous template polynucleotide, wherein said repair results in a mutation comprising an insertion, deletion, or substitution of one or more nucleotides of said target polynucleotide. In some embodiments, said mutation results in one or more amino acid changes in a protein expressed from a gene comprising the target sequence. In some embodiments, the method further comprises delivering one or more vectors to said eukaryotic cell, wherein one or more vectors comprise the CRISPR enzyme and one or more vectors drive expression of one or more of: the guide sequence linked to the tracr mate sequence, and the tracr sequence. In some embodiments, said CRISPR enzyme drive expression of one or more of: the guide sequence linked to the tracr mate sequence, and the tracr sequence. In some embodiments such CRISPR enzyme are delivered to the eukaryotic cell in a subject. In some embodiments, said modifying takes place in said eukaryotic cell in a cell culture. In some embodiments, the method further comprises isolating said eukaryotic cell from a subject prior to said modifying. In some embodiments, the method further comprises returning said eukaryotic cell and/or cells derived therefrom to said subject. In some embodiments, the isolated cells can be returned to the subject after delivery of one or more engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein to the isolated cell. In some embodiments, the isolated cells can be returned to the subject after delivering one or more molecules of the engineered delivery system described herein to the isolated cell, thus making the isolated cells engineered cells as previously described.

Screening and Cell Selection

[0573] The targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein described herein can be used in a screening assay and/or cell selection assay. The engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein can be delivered to a subject and/or cell. In some embodiments, the cell is a eukaryotic cell. The cell can be in vitro, ex vivo, in situ, or in vivo. The targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described

herein can introduce an exogenous molecule or compound, such as a cargo, to subject or cell to which they are delivered. The presence of an exogenous molecule or compound can be detected which can allow for identification of a cell and/or attribute thereof. In some embodiments, the delivered molecules or particles can impart a gene or other nucleotide modification (e.g., mutations, gene or polynucleotide insertion and/or deletion, etc.). In some embodiments the nucleotide modification can be detected in a cell by sequencing. In some embodiments, the nucleotide modification can result in a physiological and/or biological modification to the cell that results in a detectable phenotypic change in the cell, which can allow for detection, identification, and/or selection of the cell. In some embodiments, the phenotypic change can be cell death, such as embodiments where binding of a CRISPR complex to a target polynucleotide results in cell death. Embodiments of the invention allow for selection of specific cells without requiring a selection marker or a two-step process that may include a counter-selection system. The cell(s) may be prokaryotic or eukaryotic cells.

[0574] In one embodiment the invention provides for a method of selecting one or more cell(s) by introducing one or more mutations in a gene in the one or more cell (s), the method comprising: introducing one or more vectors, which can include one or more engineered delivery system molecules or vectors described elsewhere herein, into the cell (s), wherein the one or more vectors can include a CRISPR enzyme and/or drive expression of one or more of: a guide sequence linked to a tracr mate sequence, a tracr sequence, and an editing template; or other polynucleotide to be inserted into the cell and/or genome thereof; wherein, for example that which is being expressed is within and expressed in vivo by the CRISPR enzyme and/or the editing template, when included, comprises the one or more mutations that abolish CRISPR enzyme cleavage; allowing homologous recombination of the editing template with the target polynucleotide in the cell(s) to be selected; allowing a CRISPR complex to bind to a target polynucleotide to effect cleavage of the target polynucleotide within said gene, wherein the CRISPR complex comprises the CRISPR enzyme complexed with (1) the guide sequence that is hybridized to the target sequence within the target polynucleotide, and (2) the tracr mate sequence that is hybridized to the tracr sequence, wherein binding of the CRISPR complex to the target polynucleotide induces cell death, thereby allowing one or more cell(s) in which one or more mutations have been introduced to be selected. In a preferred embodiment, the CRISPR enzyme is a Cas protein. In another embodiment of the invention the cell to be selected may be a eukaryotic cell.

[0575] The screening methods involving the engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein, including but not limited to those that deliver one more CRISPR-Cas system molecules to cell, can be used in detection methods such as fluorescence in situ hybridization (FISH). In some embodiments, one or more components of an engineered CRISPR-Cas system that includes a catalytically inactive Cas protein, can be delivered by an engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g. AAV) particles, cells, or any combination thereof described herein to a cell and used in a FISH method. The CRISPR-Cas system can include an inactivated Cas protein (dCas) (e.g. a dCas9), which lacks the ability to produce DNA double-strand breaks may be fused with a marker, such as fluorescent protein, such as the enhanced green fluorescent protein (eEGFP) and co-expressed with small guide RNAs to target pericentric, centric and teleomeric repeats in vivo. The dCas system can be used to visualize both repetitive sequences and individual genes in the human genome. Such new applications of labelled dCas, dCas CRISPR-Cas systems, engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g. AAV) particles, cells, or any combination thereof described herein can be used in imaging cells and studying the functional nuclear architecture, especially in cases with a small nucleus volume or complex 3-D structures. (Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155(7):1479-91. doi: 10.1016/j.cell.2013.12.001., the teachings of which can be applied and/or adapted to the CRISPR systems described herein. A similar approach involving a polynucleotide fused to a marker (e.g. a fluorescent marker) can be delivered to a cell via an engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g. AAV) particles, cells, or any combination thereof described herein and integrated into the genome of the cell and/or otherwise interact with a region of the genome of a cell for FISH analysis.

[0576] Similar approaches for studying other cell organelles and other cell structures can be accomplished by delivering to the cell (e.g., via an engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g. AAV) particles, cells, or any combination thereof described herein) one or more molecules fused to a marker (such as a fluorescent marker), wherein the molecules fused to the marker are capable of

targeting one or more cell structures. By analyzing the presence of the markers, one can identify and/or image specific cell structures.

[0577] In some embodiments, the engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein can be used in a screening assay inside or outside of a cell. In some embodiments, the screening assay can include delivering a CRISPR-Cas cargo molecule(s) via an engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g. AAV) particles, cells, or any combination thereof described herein.

[0578] Use of the present system in screening is also provided by the present invention, e.g., gain of function screens. Cells which are artificially forced to overexpress a gene are able to down regulate the gene over time (re-establishing equilibrium) e.g., by negative feedback loops. By the time the screen starts the unregulated gene might be reduced again. Other screening assays are discussed elsewhere herein.

[0579] In an embodiment, the invention provides a cell from or of an in vitro method of delivery, wherein the method comprises contacting the delivery system with a cell, optionally a eukaryotic cell, whereby there is delivery into the cell of constituents of the delivery system, and optionally obtaining data or results from the contacting, and transmitting the data or results.

[0580] In an embodiment, the invention provides a cell from or of an in vitro method of delivery, wherein the method comprises contacting the delivery system with a cell, optionally a eukaryotic cell, whereby there is delivery into the cell of constituents of the delivery system, and optionally obtaining data or results from the contacting, and transmitting the data or results; and wherein the cell product is altered compared to the cell not contacted with the delivery system, for example altered from that which would have been wild type of the cell but for the contacting. In an embodiment, the cell product is non-human or animal. In some embodiments, the cell product is human.

[0581] In some embodiments, a host cell is transiently or non-transiently transfected with one or more vectors described herein. In some embodiments, a cell is transfected as it naturally occurs in a subject optionally to be reintroduced therein. In some embodiments, a cell that is transfected is taken from a subject. In some embodiments, the cell obtained from or is derived from cells taken from a subject, such as a cell line. Delivery mechanisms and techniques of the

targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein.

[0582] In some embodiments it is envisaged to introduce one or more of the engineered targeting moieties, polypeptides, polynucleotides, vectors, vector systems, particles, viral (e.g., AAV) particles, cells, or any combination thereof described herein directly to the host cell. For instance, the engineered AAV capsid system molecule(s) can be delivered together with one or more cargo molecules to be packaged into an engineered AAV particle.

[0583] In some embodiments, the invention provides a method of expressing an engineered delivery molecule and cargo molecule to be packaged in an engineered viral (e.g. AAV) particle in a cell that can include the step of introducing the vector according any of the vector delivery systems disclosed herein.

[0584] The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

[0585] 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 – mRNA based detection methods are more stringent for selection of AAV variants.

[0586] **FIG. 1** demonstrates the adeno-associated virus (AAV) transduction mechanism, which results in production of mRNA. As is demonstrated in **FIG. 1**, functional transduction of a cell by an AAV particle can result in the production of an mRNA strand. Non-functional transduction would not produce such a product despite the viral genome being detectable using a DNA-based assay. Thus, mRNA-based detection assays to detect transduction by e.g., an AAV can be more stringent and provide feedback as to the functionality of a virus particle that is able to functionally transduce a cell. **FIG. 2** shows a graph that can demonstrate that mRNA-based selection of AAV variants can be more stringent than DNA-based selection. The virus library was expressed under the control of a CMV promoter.

Example 2 – mRNA based detection methods can be used to detect AAV capsid variants from a capsid variant library.

[0587] **FIGS. 3A-3B** show graphs that can demonstrate a correlation between the virus library and vector genome DNA (**FIG. 3A**) and mRNA (**FIG. 3B**) in the liver. **FIGS. 4A-4F**

show graphs that can demonstrate capsid variants expressed at the mRNA level identified in different tissues.

Example 3 – Capsid mRNA expression can be driven by tissue specific promoters.

[0588] FIGS. 5A-5C show graphs that can demonstrate capsid mRNA expression in different tissues under the control of cell-type specific promoters (as noted on x-axis). CMV was included as an exemplary constitutive promoter. CK8 is a muscle-specific promoter. MHCK7 is a muscle-specific promoter. hSyn is a neuron specific promoter.

Example 4 – Capsid variant library generation, variant screening, and variant identification.

[0589] Generally, an AAV capsid library can be generated by expressing engineered capsid vectors each containing an engineered AAV capsid polynucleotide previously described in an appropriate AAV producer cell line. See e.g., FIG. 8. This can generate an AAV capsid library that can contain one more desired cell-specific engineered AAV capsid variant. FIG. 7 shows a schematic demonstrating embodiments of generating an AAV capsid variant library, particularly insertion of a random *n*-mer (*n*=3-15 amino acids) into a wild-type AAV, e.g., AAV9. In this example, random 7-mers were inserted between aa588-589 of variable region VIII of AAV9 viral protein and used to form the viral genome containing vectors with one variant per vector. As shown in FIG. 8, the capsid variant vector library was used to generate AAV particles where each capsid variant encapsulated its coding sequence as the vector genome. FIG. 9 shows vector maps of representative AAV capsid plasmid library vectors (see e.g., FIG. 8) that can be used in an AAV vector system to generate an AAV capsid variant library. The library can be generated with the capsid variant polynucleotide under the control of a tissue specific promoter or constitutive promoter. The library was also made with capsid variant polynucleotide that included a polyadenylation signal.

[0590] As shown in FIGS. 6A the AAV capsid library can be administered to various non-human animals for a first round of mRNA-based selection. As shown in FIG. 1, the transduction process by AAVs and related vectors can result in the production of an mRNA molecule that is reflective of the genome of the virus that transduced the cell. As is at least demonstrated in the Examples herein, mRNA based selection can be more specific and effective to determine a virus particle capable of functionally transducing a cell because it is based on the functional product produced as opposed to just detecting the presence of a virus particle in the cell by measuring the presence of viral DNA.

[0591] As is further shown in **FIG. 6A**, after first-round administration, one or more engineered AAV virus particles having a desired capsid variant can then be used to form a filtered AAV capsid library. Desirable AAV virus particles can be identified by measuring the mRNA expression of the capsid variants and determining which variants are highly expressed in the desired cell type(s) as compared to non-desired cells type(s). Those that are highly expressed in the desired cell, tissue, and/or organ type are the desired AAV capsid variant particles. In some embodiments, the AAV capsid variant encoding polynucleotide is under control of a tissue-specific promoter that has selective activity in the desired cell, tissue, or organ.

[0592] The engineered AAV capsid variant particles identified from the first round can then be administered to various non-human animals. In some embodiments, the animals used in the second round of selection and identification are not the same as those animals used for first round selection and identification. Similar to round 1, after administration the top expressing variants in the desired cell, tissue, and/or organ type(s) can be identified by measuring viral mRNA expression in the cells. The top variants identified after round two can then be optionally barcoded and optionally pooled. In some embodiments, top variants from the second round can then be administered to a non-human primate to identify the top cell-specific variant(s), particularly if the end use for the top variant is in humans. Administration at each round can be systemic. As further shown in **FIG. 6B** after the second round of selection, a third round of selection, which can optionally include benchmarking against known, control, and/or standard (e.g., benchmark) variants can be performed.

[0593] **FIG. 10** shows a graph that can demonstrate the viral titer (calculated as AAV9 vector genome/15 cm dish) produced by libraries generated using different promoters. As demonstrated in **FIG. 10**, virus titer was not affected significantly by the use of different promoters.

Example 5 – CNS n-mer inserts.

[0594] CNS n-mer inserts were generated as described elsewhere herein and then screened for transduction efficiency in various strains of mice (C57BL/6J and BALB/cJ). **Table 1** shows the top motifs based on CNS transduction. As previously discussed, each n-mer insert's transduction efficacy in CNS cells was tested with both AQ and DG as the aa587 and aa588 (the two amino acids in the AAV immediately preceding the n-mer insert). Some exemplary n-mer inserts that stood out when preceded by AQ are KTVGTVY (SEQ ID NO: 3), RSVGSVY

(SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSSG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESGH (SEQ ID NO: 19), and GENSARW (SEQ ID NO: 20).

[0595] Some exemplary n-mer inserts that stood out when preceded by DG are ASNPGRW (SEQ ID NO: 22), WTLESGH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25), RMQRTLY (SEQ ID NO: 26), and REQQKLW (SEQ ID NO: 21). Engineered AAVs including a CNS n-mer of **Table 1** demonstrated the ability to specifically transduce CNS cells in both strains of mice, which is in contrast to the commonly used in the art CNS AAV. Without being bound by theory, this observation can demonstrate that the engineered AAVs containing an CNS-specific n-mer motif described herein can operate through a different receptor on the surface of CNS cells than the conventional AAV used in the art to achieve CNS specificity. Given that n-mer motifs preceded by AQ with top scores did not necessarily perform the same when preceded by DG can suggest that the 3D structure of the capsid conferred by the n-mer and its interaction with endogenous AAV amino acids can influence the ability of the engineered AAV capsid to transduce a cell and thus, without being bound by theory, can play a role in contributing to the cell-type specificity of the engineered capsids.

Example 6 – CNS n-mer inserts in Non-Human Primates.

[0596] CNS n-mer inserts were generated as described elsewhere herein and then screened for transduction efficiency in non-human primates. **Tables 2-3** show the top n-mer motifs. A general motif was observed across the very top hits (**Table 3**). The motif observed was P-motif having the formula amino acid sequence PX_1QGTX_2R , (SEQ ID NO: 317) wherein X_1 and X_2 are each selected from any amino acid. Exemplary n-mer motif variants containing a P-motif are shown in **Table 3**.

Example 7 – Benchmarking.

[0597] As shown in **FIGs. 6A-6B** shows a general schematic for selecting CNS specific capsid, which includes a benchmarking round which evaluates the performance of selected capsids against currently used capsids for, e.g., delivery to the CNS. **Table 7** shows the selected capsids used in the benchmarking round of selection. Four variants developed using selection

in mice and 8 using selection in NHPs were used for benchmarking. For benchmarking here, capsid variant specific barcodes were included with each variant. Viral particles for each capsid variant was produced individually and viral particles were then pooled. Such barcoding and pooling methodology is described in greater detail elsewhere herein and applied in this context. Pooled viral particles were then injected systemically (via I.V. administration) to the periphery of different mouse strains (C57BL/6J (“C57”) and BALB/c (“BALBc”)) and non-human primates (Macaques) so that the ability for the capsid variants to cross the blood brain barrier in different species could be evaluated. Included in the benchmarking were both engineered capsid variants from mouse and non-human primate selection (rounds 1 and/or2) and currently used capsid variants (AAV-CAP-B10, AAV-CAP-B22, and AAV-PhP.22). mRNA and DNA corresponding to the capsid variants in various tissues were then examined to determine the CNS, strain, and species specificity of the capsid variants.

Table 7.		
Capsid variant	Insert sequence	SEQ ID NO:
Mouse variant 1	RSVGSVY	318
Mouse variant 2	KTVGTVY	319
Mouse variant 3	WVLPSGG	320
Mouse variant 4	(DG)REQQKLW	321
NHP variant 1	PTQGTVR	322
NHP variant 2	PSQGTLR	323
NHP variant 3	PTQGTLR	324
NHP variant 4	RVDPSGL	325
NHP variant 5	VVSDYTV	326
NHP variant 6	TDALTTK	327
NHP variant 7	STIPTMK	328
NHP variant 8	PTQGTFR	329

FIGS. 11A-11P show results from benchmarking the top selected capsids out of the second round of selection. In agreement with the literature, the AAV-CAP-B10, AAV-CAP22, and AAV-PhP.22 capsids demonstrated a species and strain preference, and importantly did not appear to perform well in non-human primates. Indeed, the NHP capsid variants developed

using the methods described and benchmarked herein were successfully delivered to and expressed in one or more CNS tissues. Further, several NHP capsid variants tested here showed increased delivery to the CNS as compared to the capsid variants currently known and alleged to target the CNS and cross the blood brain barrier (AAV-CAP-B10, AAV-CAP22, and AAV-PhP.22). Further, most of the NHP variants were not observed to have strong liver delivery or expression (see e.g., **FIGS. 11O and 11P**). Expression in the dorsal root ganglion can lead to significant toxicity. Several NHP variants showed reduced or negligible delivery and/or expression in the dorsal root ganglion (DRG) (see e.g., **FIG. 11N**).

[0598] 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. A composition comprising:
 - a targeting moiety effective to target a central nervous system (CNS) cell, wherein the targeting moiety comprises:
 - one or more P-motifs, wherein the at least one P-motif comprises the amino acid sequence $PX_1QGTX_2RX_n$ (SEQ ID NO: 2), wherein X_1 , X_2 , X_n , are each independently selected from any amino acid and wherein n is 0, 1, 2, 3, 4, 5, 6, or 7, or
 - one or more n-mer inserts selected from the group consisting of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, 313, and 318-329, or
 - one or more n-mer inserts selected from the group consisting of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, 313, and 318-329 and one or more P-motifs, and
 - optionally a cargo, wherein the cargo is coupled to or is otherwise associated with the targeting moiety.
2. The composition of claim 1, wherein the targeting moiety comprises both an n-mer insert and a P-motif and wherein the P-motif is optionally part of or the entirety of the n-mer insert.
3. The composition of claim 1, wherein the one or more n-mer inserts, each of the P-motifs, or both are each 3-15 amino acids in length.
4. The composition of claim 1, wherein
 - a. X_1 is S, T, or A,
 - b. X_2 is L, V, F, or I, or
 - c. both.

5. The composition of claim 1, wherein the *n*-mer insert and/or P motif is selected from the group consisting of SEQ ID NOs: 65-199.
6. The composition of claim 1, wherein the *n*-mer insert and/or P motif is selected from the group consisting of: SEQ ID NO: 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, and 313.
7. The composition of claim 1, wherein the *n*-mer insert and/or P motif is selected from the group consisting of SEQ ID NOs: 318-329.
8. The composition of claim 1, wherein the *n*-mer insert is immediately preceded by AQ or DG.
9. The composition of claim 8, wherein
 - (a) the *n*-mer insert polypeptide is immediately preceded by AQ and wherein the *n*-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSGG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESGH (SEQ ID NO: 19), or GENSARW (SEQ ID NO: 20); or
 - (b) the *n*-mer insert polypeptide is immediately preceded by DG and wherein the *n*-mer insert is REQQKLW (SEQ ID NO: 21), ASNPGRW (SEQ ID NO: 22), WTLESGH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25), or RMQRTLY (SEQ ID NO: 26).
10. The composition of claim 1, wherein the targeting moiety comprises a polypeptide, a polynucleotide, a lipid, a polymer, a sugar, or a combination thereof.
11. The composition of claim 10, wherein the targeting moiety comprises a viral protein.

12. The composition of claim 11, wherein the viral protein is a capsid protein.
13. The composition of claim 10, wherein the n-mer insert(s), is located between two amino acids of the viral protein such that the n-mer insert is external to a viral capsid.
14. The composition of claim 11, wherein the viral protein is an adeno associated virus (AAV) protein.
15. The composition of claim 14, wherein the AAV protein is an AAV capsid protein.
16. The composition of claim 15, wherein the one or more n-mer inserts and/or P motif are each inserted between any two contiguous amino acids between amino acids independently selected from 262-269, 327-332, 382-386, 452-460, 488-505, 527-539, 545-558, 581-593, 704-714, or any combination thereof in an AAV9 capsid polypeptide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.
17. The composition of claim 16, wherein at least one of the one or more n-mer inserts is inserted between amino acids 588 and 589 in an AAV9 capsid polypeptide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.
18. The composition of claim 15, wherein the AAV capsid protein is an engineered AAV capsid protein having reduced or eliminated uptake in a non-CNS cell as compared to a corresponding wild-type AAV capsid polypeptide.
19. The composition of claim 18, wherein the non-CNS cell is a liver cell.
20. The composition of claim 18, wherein the wild-type capsid polypeptide is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 capsid polypeptide.

21. The composition of claim 18, wherein the engineered AAV capsid protein comprises one or more mutations that result in reduced or eliminated uptake in a non-CNS cell.
22. The composition of claim 21, wherein the one or more mutations are
 - a. in position 267,
 - b. in position 269,
 - c. in position 504,
 - d. in position 505,
 - e. in position 590,
 - f. or any combination thereofin the AAV9 capsid protein (SEQ ID NO: 1) or in one or more positions corresponding thereto in a non-AAV9 capsid polypeptide.
23. The composition of claim 22, wherein the non-AAV9 capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, or AAV rh.10 capsid polypeptide.
24. The composition of claim 22, wherein the mutation in position 267 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X mutation to A, wherein X is any amino acid.
25. The composition of claim 22, wherein the mutation in position 269 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is an S or X to T mutation, wherein X is any amino acid.
26. The composition of claim 22, wherein the mutation in position 504 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X to A mutation, wherein X is any amino acid.
27. The composition of claim 22, wherein the mutation in position 505 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a P or X to A mutation, wherein X is any amino acid.

28. The composition of claim 22, wherein the mutation in position 590 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a Q or X to A mutation, wherein X is any amino acid.
29. The composition of claim 21, wherein the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 267, position 269 or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 267 is a G to A mutation and wherein the mutation at position 269 is an S to T mutation.
30. The composition of claim 21, wherein the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 590 of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 590 is a Q to A mutation.
31. The composition of claim 21, wherein the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 504, position 505, or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 504 is a G to A mutation and wherein the mutation at position 505 is a P to A mutation.
32. The composition of any one of claims 1-31, wherein the composition is an engineered viral particle.
33. The composition of claim 32, wherein the engineered viral particle is an engineered AAV viral particle.
34. The composition of claim 33, wherein the AAV viral particle is an engineered AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 viral particle.

35. The composition of any of claims 1-34, wherein the optional cargo is capable of treating or preventing a CNS disease or disorder.
36. A vector system comprising:
a vector comprising:
one or more polynucleotides, wherein at least one of the one or more polynucleotides encodes all or part of a targeting moiety effective to target a central nervous system (CNS) cell, wherein the targeting moiety comprises
at least one P-motif, wherein the at least one P-motif comprises the amino acid sequence $PX_1QGTX_2RX_n$ (SEQ ID NO: 2), wherein X_1 , X_2 , X_n , are each independently selected from any amino acid and wherein n is 0, 1, 2, 3, 4, 5, 6, or 7, or
at least one n-mer insert selected from the group consisting of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, and 313, and 318-329, or
at least one n-mer insert selected from the group consisting of SEQ ID NO: 65-199, 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, 313, and 318-329 and at least one P-motif,
wherein at least one of the one or more polynucleotides encodes the at least one n-mer insert, the at least one P-motif, or both; and
optionally, a regulatory element operatively coupled to one or more of the one or more polynucleotides.

37. The vector system of claim 36, wherein the targeting moiety comprises both an *n*-mer insert and a P-motif and wherein the P-motif is optionally part of or the entirety of the *n*-mer insert.
38. The vector system of claim 36, wherein the one or more *n*-mer inserts, each of the P-motifs, or both are each 3-15 amino acids in length.
39. The vector system of claim 36, wherein
- a. X₁ is S, T, or A,
 - b. X₂ is L, V, F, or I, or
 - c. both.
40. The vector system of claim 36, wherein the *n*-mer insert and/or P motif is selected from the group consisting of SEQ ID NOs: 65-199.
41. The vector system of claim 36, wherein the *n*-mer insert and/or P motif is selected from the group consisting of: SEQ ID NO: 200, 202, 204, 206, 208, 210, 212, 214, 300, 303, 306, 308, 311, and 313.
42. The vector system of claim 36, wherein the *n*-mer insert and/or P motif is selected from the group consisting of SEQ ID NOs: 318-329..
43. The vector system of claim 36, wherein the *n*-mer insert is immediately preceded by AQ or DG.

44. The vector system of claim 43, wherein
- (a) the n-mer insert polypeptide is immediately preceded by AQ and wherein the n-mer insert is KTVGTVY (SEQ ID NO: 3), RSVGSVY (SEQ ID NO: 4), RYLGDAS (SEQ ID NO: 5), WVLPSSG (SEQ ID NO: 6), VTVGSIY (SEQ ID NO: 7), VRGSSIL (SEQ ID NO: 8), RHHGDAA (SEQ ID NO: 9), VIQAMKL (SEQ ID NO: 10), LTYGMAQ (SEQ ID NO: 11), LRIGLSQ (SEQ ID NO: 12), GDYSMIV (SEQ ID NO: 13), VNYSVAL (SEQ ID NO: 14), RHIADAS (SEQ ID NO: 15), RYLGDAT (SEQ ID NO: 16), QRVGFAQ (SEQ ID NO: 17), QIAHGYST (SEQ ID NO: 18), WTLESQH (SEQ ID NO: 19), or GENSARW (SEQ ID NO: 20); or
- (b) the n-mer insert polypeptide is immediately preceded by DG and wherein the n-mer insert is REQQKLW (SEQ ID NO: 21), ASNPGRW (SEQ ID NO: 22), WTLESQH (SEQ ID NO: 23), REQKKLW (SEQ ID NO: 24), ERLLVQL (SEQ ID NO: 25), or RMQRTLY (SEQ ID NO: 26).
45. The vector system of any one of claims 36-44, further comprising a cargo.
46. The vector system of claim 45, wherein the cargo is a cargo polynucleotide and is optionally operatively coupled to one or more of the one or more polynucleotides encoding the targeting moiety.
47. The vector system of any one of claims 36-46, wherein the vector system is capable of producing virus particles, virus particles that contain the cargo, or both.
48. The vector system of any one of claims 36-47, wherein the vector system is capable of producing a polypeptide comprising one or more of the targeting moieties.
49. The vector system of claim 48, wherein the polypeptide is a viral polypeptide.
50. The vector system of claim 49, wherein the viral polypeptide is a capsid polypeptide.
51. The vector system of claim 50, wherein the capsid polypeptide is an adeno associated virus (AAV) capsid polypeptide.

52. The vector system of any one of claims 49-51, wherein the virus particles are AAV virus particles.
53. The vector system of any one of claims 50-51, wherein the AAV virus particles or AAV capsid polypeptide are engineered AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 viral particles or polypeptides.
54. The vector system of any one of claims 49-51, wherein the at least one polynucleotide encoding the at least one *n*-mer inserts is inserted between two codons corresponding to two amino acids of a viral polypeptide such that the *n*-mer insert(s) is external to a viral capsid of the virus particles.
55. The vector system of claim 53, wherein the at least one polynucleotide is inserted between two codons corresponding to any two contiguous amino acids between amino acids 262-269, 327-332, 382-386, 452-460, 488-505, 527-539, 545-558, 581-593, 704-714, or any combination thereof in an AAV9 capsid polypeptide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.
56. The vector system of claim 54, wherein the at least one polynucleotide is inserted between the codons corresponding to amino acid 588 and 589 in the AAV9 capsid polynucleotide or in an analogous position in an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, AAV rh.10 capsid polypeptide.
57. The vector system of claim 51, wherein the AAV capsid protein is an engineered AAV capsid protein having reduced or eliminated uptake in a non-CNS cell as compared to a corresponding wild-type AAV capsid polypeptide.
58. The vector system of claim 57, wherein the non-CNS cell is a liver cell.

59. The vector system of claim 57, wherein the wild-type capsid polypeptide is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV rh.74, or AAV rh.10 capsid polypeptide.
60. The vector system of claim 57, wherein the engineered AAV capsid protein comprises one or more mutations that result in reduced or eliminated uptake in a non-CNS cell.
61. The vector system of claim 60, wherein the one or more mutations are
- a. in position 267,
 - b. in position 269,
 - c. in position 504,
 - d. in position 505,
 - e. in position 590,
 - f. or any combination thereof
- in the AAV9 capsid protein (SEQ ID NO: 1) or in one or more positions corresponding thereto in a non-AAV9 capsid polypeptide.
62. The vector system of claim 61, wherein the non-AAV9 capsid protein is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV rh.74, or AAV rh.10 capsid polypeptide.
63. The vector system of claim 61, wherein the mutation in position 267 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X mutation to A, wherein X is any amino acid.
64. The vector system of claim 61, wherein the mutation in position 269 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is an S or X to T mutation, wherein X is any amino acid.
65. The vector system of claim 61, wherein the mutation in position 504 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a G or X to A mutation, wherein X is any amino acid.

66. The vector system of claim 61, wherein the mutation in position 505 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a P or X to A mutation, wherein X is any amino acid.
67. The vector system of claim 61, wherein the mutation in position 590 in the AAV9 capsid protein (SEQ ID NO: 1) or position corresponding thereto in a non-AAV9 capsid polypeptide is a Q or X to A mutation, wherein X is any amino acid.
68. The vector system of claim 60, wherein the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 267, position 269 or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 267 is a G to A mutation and wherein the mutation at position 269 is an S to T mutation.
69. The vector system of claim 60, wherein the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 590 of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 509 is a Q to A mutation.
70. The vector system of claim 60, wherein the engineered AAV capsid protein is an engineered AAV9 capsid polypeptide comprising a mutation at position 504, position 505, or both of a wild-type AAV9 capsid protein (SEQ ID NO: 1), wherein the mutation at position 504 is a G to A mutation and wherein the mutation at position 505 is a P to A mutation.
71. The vector system of any one of claims 36-70, wherein the vector comprising the one or more polynucleotides does not comprise splice regulatory elements.
72. The vector system of any one of claims 36-71, further comprising a polynucleotide that encodes a viral rep protein.

- 73. The vector system of claim 72, wherein the viral rep protein is an AAV rep protein.
- 74. The vector system of any one of claims 72-73, wherein the polynucleotide that encodes the viral rep protein is on the same vector or a different vector as the one or more polynucleotides.
- 75. The vector system of any one of claims 72-74, wherein the polynucleotide that encodes the viral rep protein is operatively coupled to a regulatory element.
- 76. The vector system of any one of claims 36-75, wherein the vector system is capable of producing a composition or portion thereof as in any of claims 1-35.
- 77. A polypeptide encoded, produced, or both by a vector system as in any of claims 36-76.
- 78. The polypeptide of claim 77, wherein the polypeptide is a viral polypeptide.
- 79. The polypeptide of claim 78, wherein the viral polypeptide is an AAV polypeptide.
- 80. The polypeptide of any one of claims 77-79, wherein the polypeptide is coupled to or otherwise associated with a cargo.
- 81. A particle produced by a vector system as in any one of claims 36-76, optionally including a polypeptide as in any one of claims 77-80.
- 82. The particle of claim 81, wherein the particle is a viral particle.
- 83. The particle of claim 82, wherein the viral particle is an adeno-associated virus (AAV) particle, lentiviral particle, or a retroviral particle.
- 84. The particle of any one of claims 81-83, wherein the particle comprises a cargo.

85. The particle of any of claims 81-84, wherein the viral particle has a central nervous system (CNS) tropism.
86. The vector system of any one of claims 45-76, the polypeptide as in any one of claims 77-80, or the particle of any one of claims 81-85, wherein the cargo is capable of preventing a CNS disease or disorder.
87. A cell comprising:
- a. a composition as in any of claims 1-35;
 - b. a vector system as in any one of claims 36-76 or 86;
 - c. a polypeptide as in any one of claims 77-80 or 86;
 - d. a particle of any one of claims 81-86; or
 - e. a combination thereof.
88. The cell of claim 87, wherein the cell is prokaryotic.
89. The cell of claim 87, wherein the cell is eukaryotic.
90. A pharmaceutical formulation comprising:
- a. a composition as in any of claims 1-35
 - b. a vector system as in any one of claims 36-76 or 86;
 - c. a polypeptide as in any one of claims 77-80 or 86;
 - d. a particle of any one of claims 81-86;
 - e. a cell as in any one of claims 87-89; or
 - f. a combination thereof; and
- a pharmaceutically acceptable carrier.

91. A method of treating a central nervous system disease, disorder, or a symptom thereof comprising:
- administering, to the subject in need thereof,
- a. a composition as in any of claims 1-35;
 - b. a vector system as in any one of claims 36-76 or 86;
 - c. a polypeptide as in any one of claims 77-80 or 86;
 - d. a particle of any one of claims 81-86;
 - e. a cell as in any one of claims 87-89;
 - f. a pharmaceutical formulation as in claim 90; or
 - g. a combination thereof.
92. The method of claim 91, wherein the central nervous system disease or disorder comprises a secondary muscle disease, disorder, or symptom thereof.
93. The method of claim 91, wherein the central nervous system disease or disorder is Friedreich's Ataxia, Dravet Syndrome, Spinocerebellar Ataxia Type 3, Niemann Pick Type C, Huntington's Disease, Pompe Disease, Myotonic Dystrophy Type 1, Glut1 Deficiency Syndrome (De Vivo Syndrome), Tay-Sachs, Spinal Muscular Atrophy, Alzheimer's disease, Amyotrophic lateral sclerosis (ALS), Danon disease, Rett Syndrome, Angleman Syndrome, or a combination thereof.

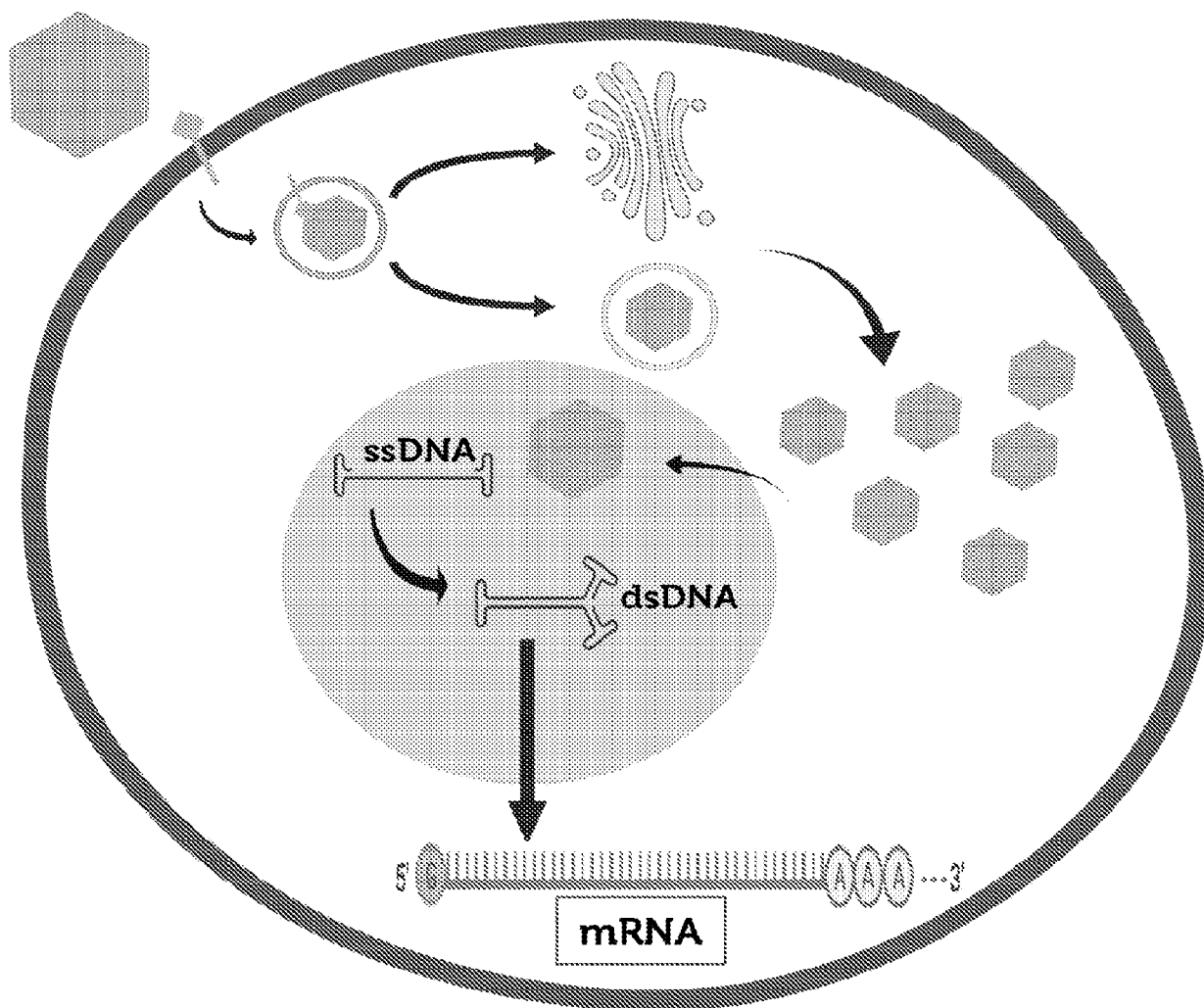
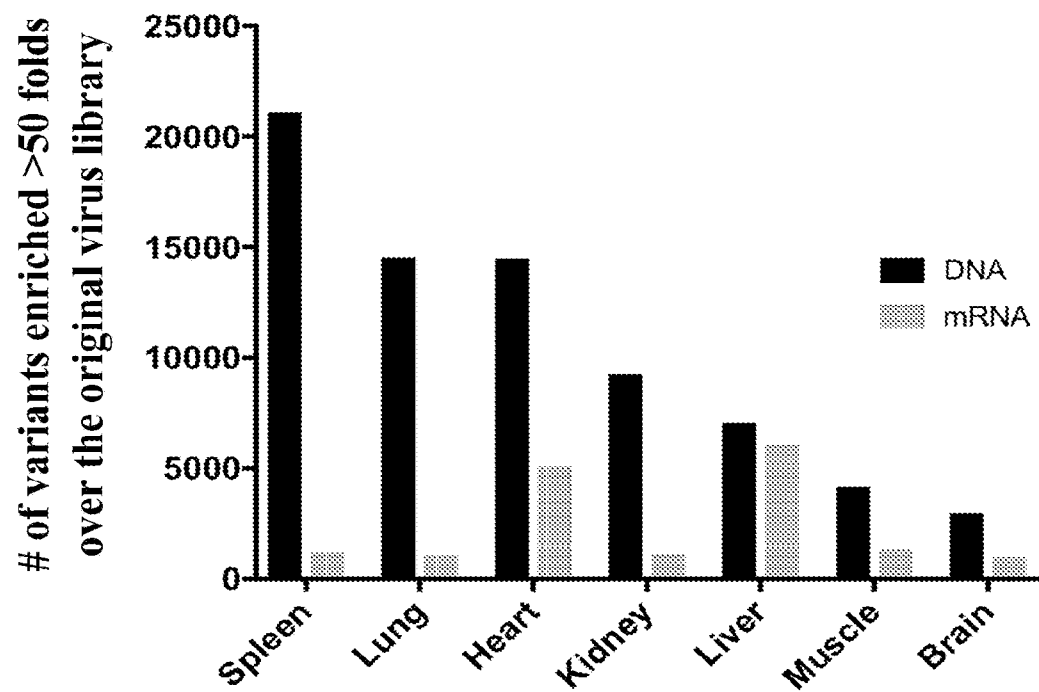
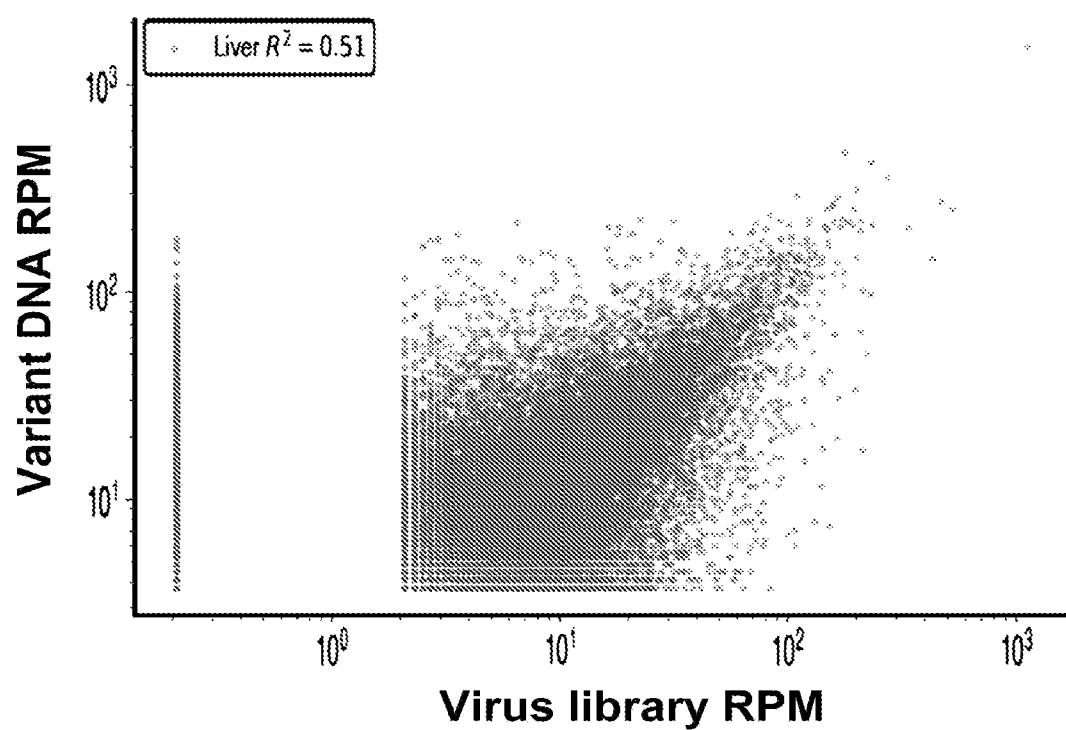
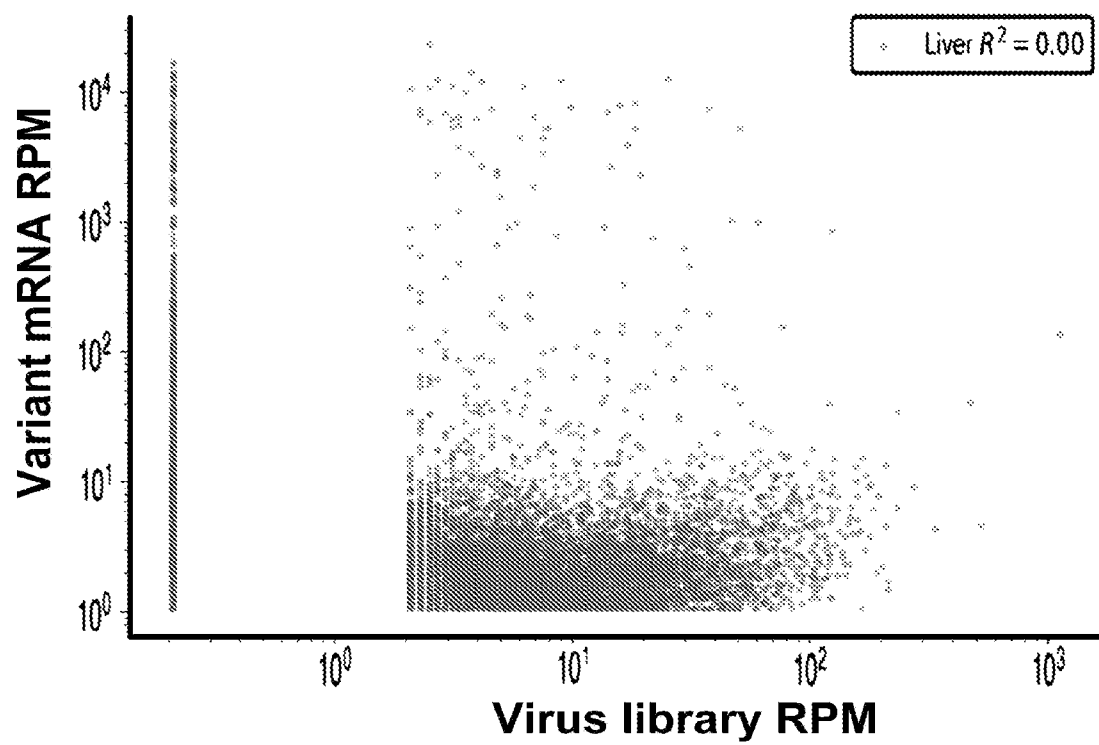


FIG. 1

2/18

**FIG. 2**

3/18

**FIG. 3A****FIG. 3B**

4/18

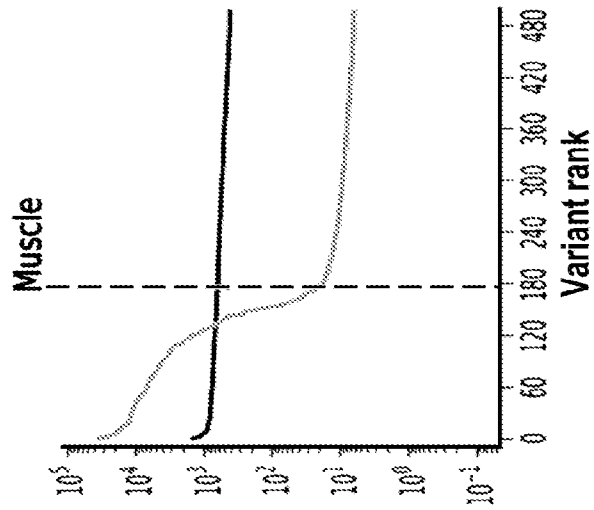


FIG. 4C

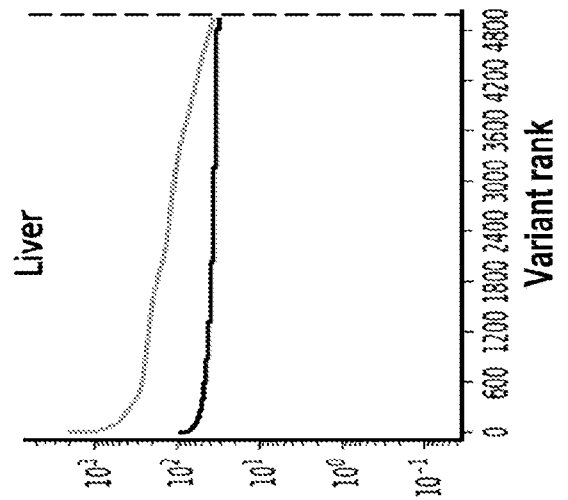


FIG. 4F

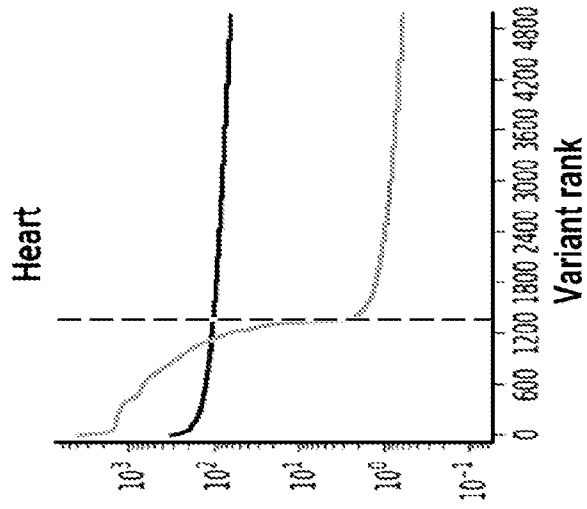


FIG. 4B

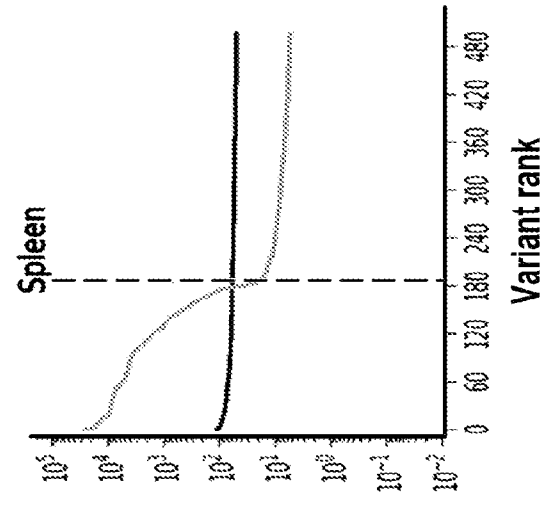


FIG. 4E

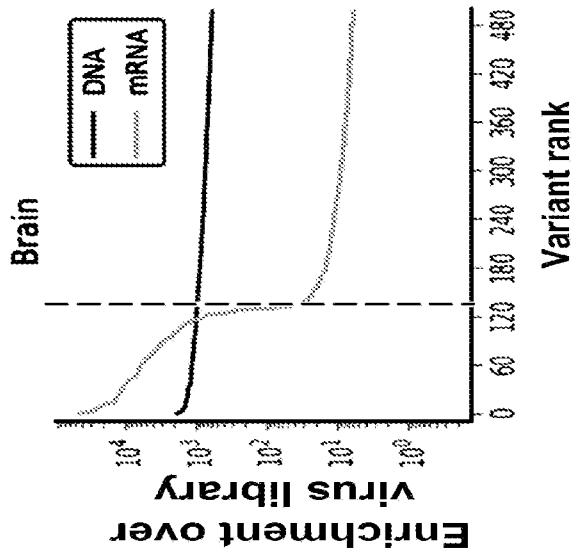


FIG. 4A

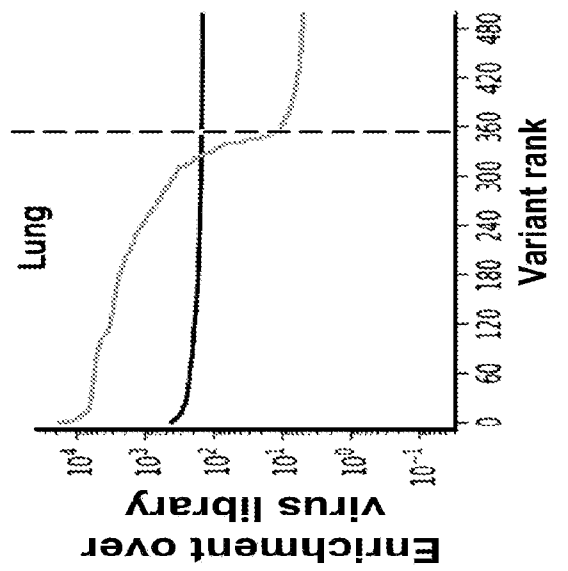
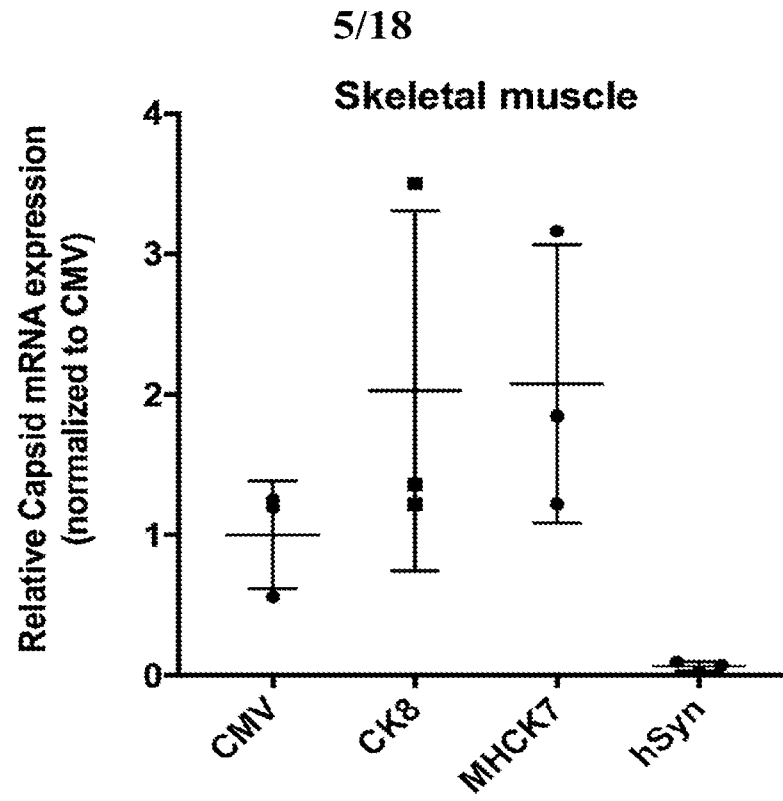
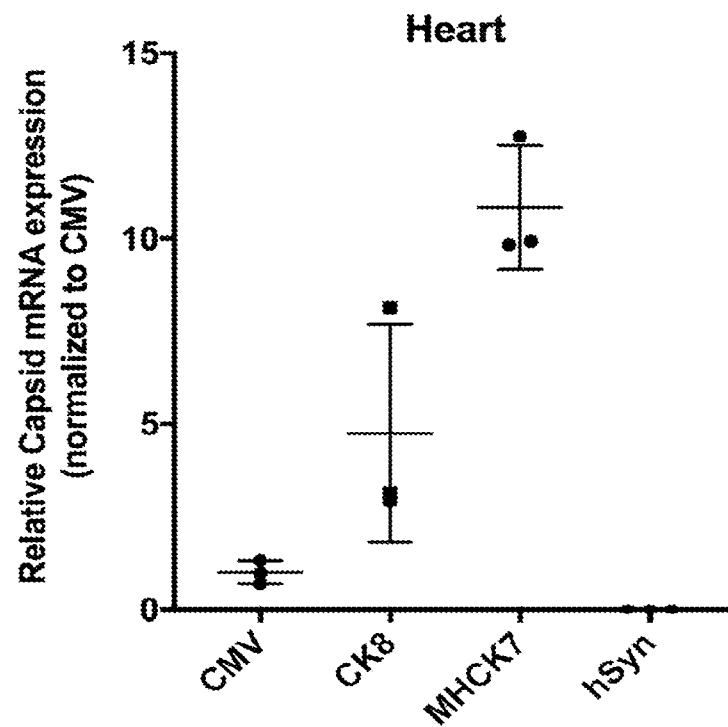
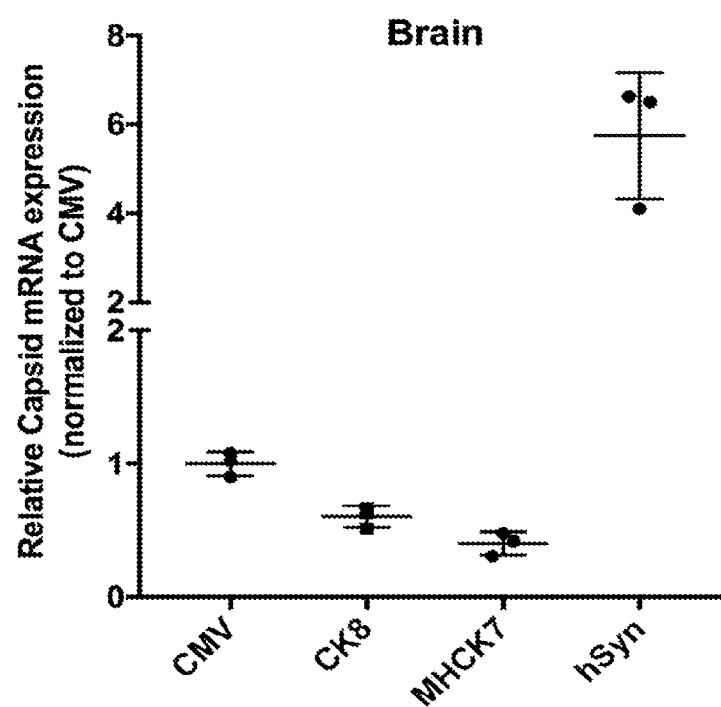


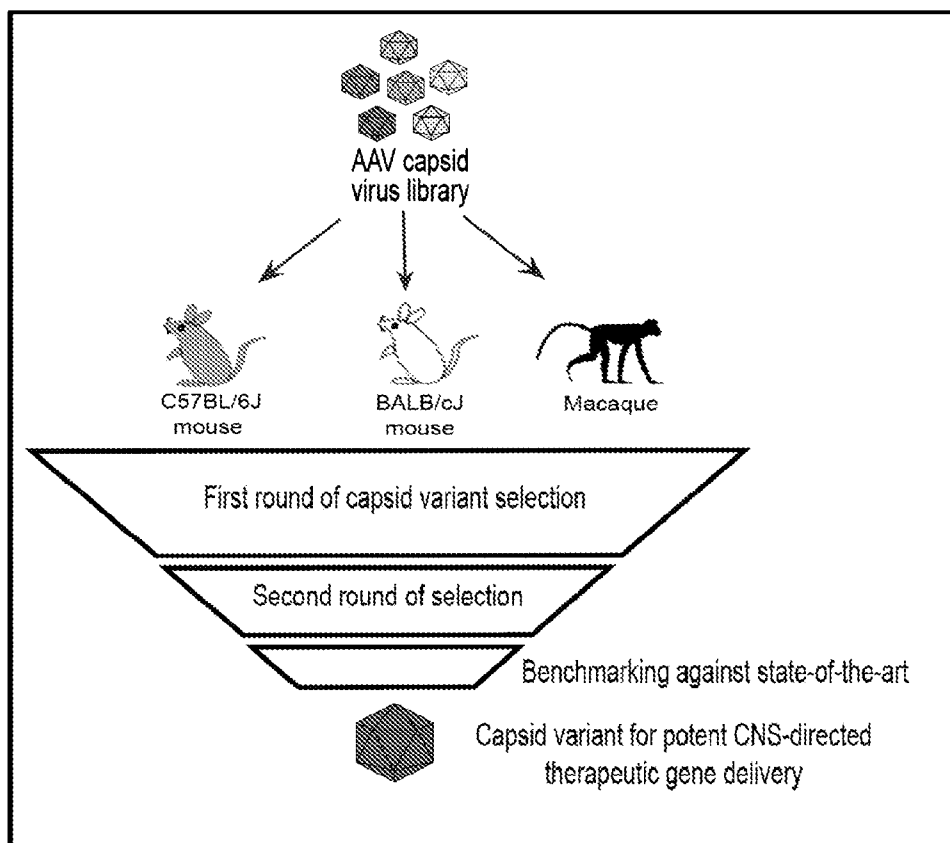
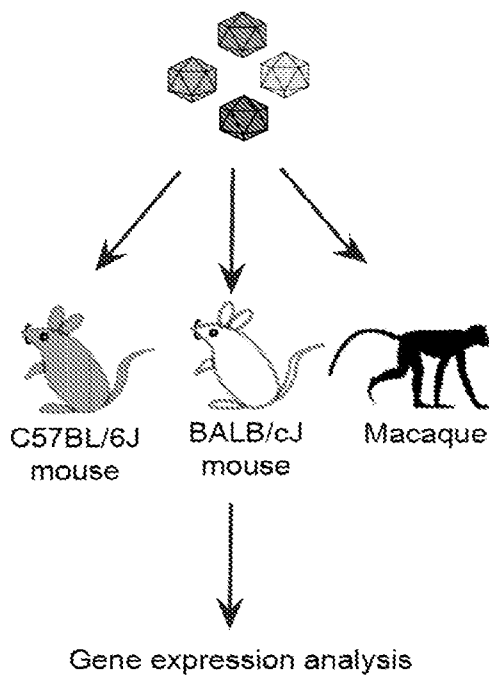
FIG. 4D

**FIG. 5A****FIG. 5B**

6/18

**FIG. 5C**

7/18

**FIG. 6A****FIG. 6B**

8/18

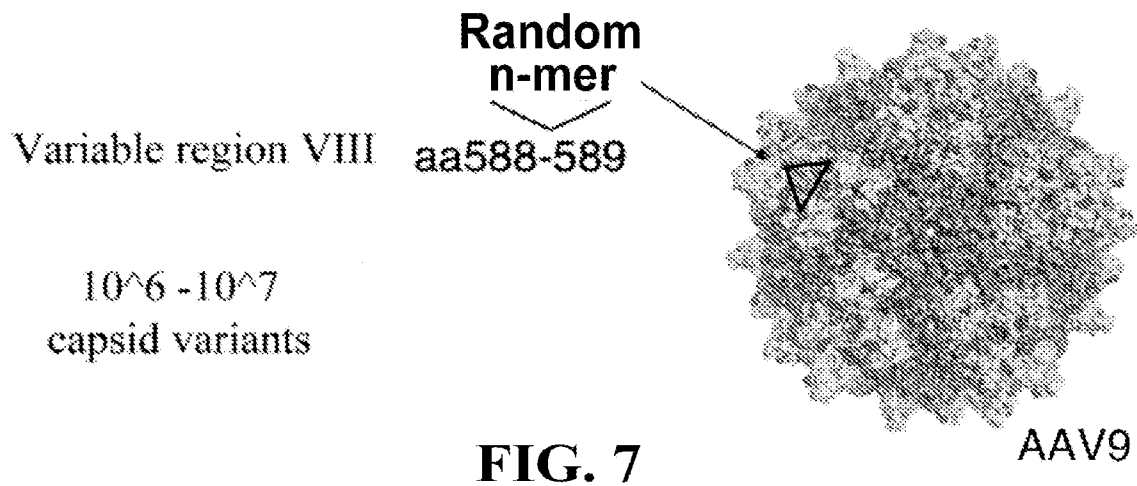


FIG. 7

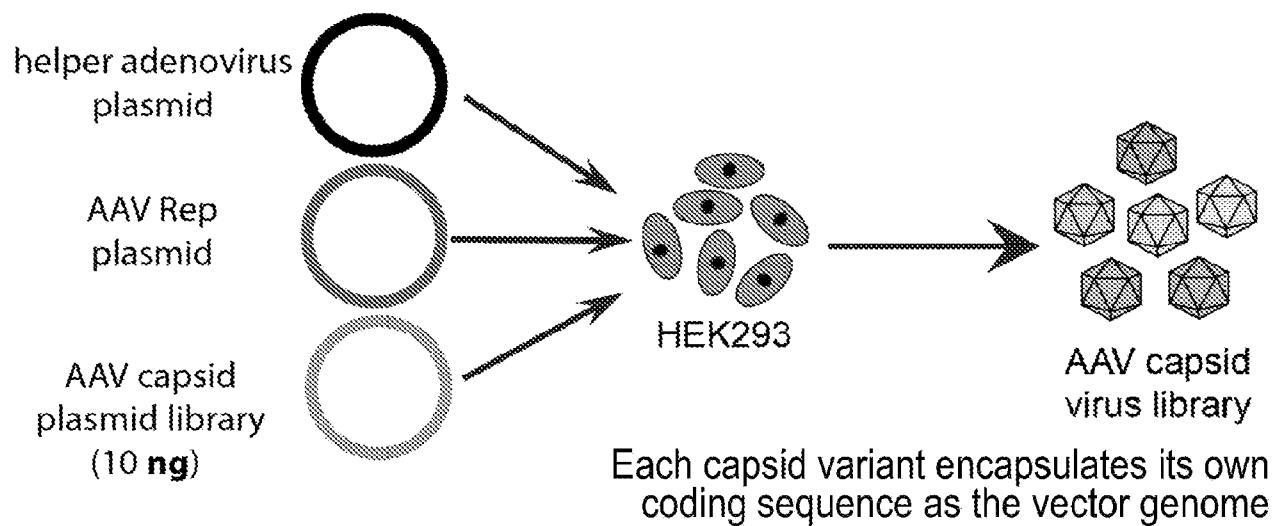


FIG. 8

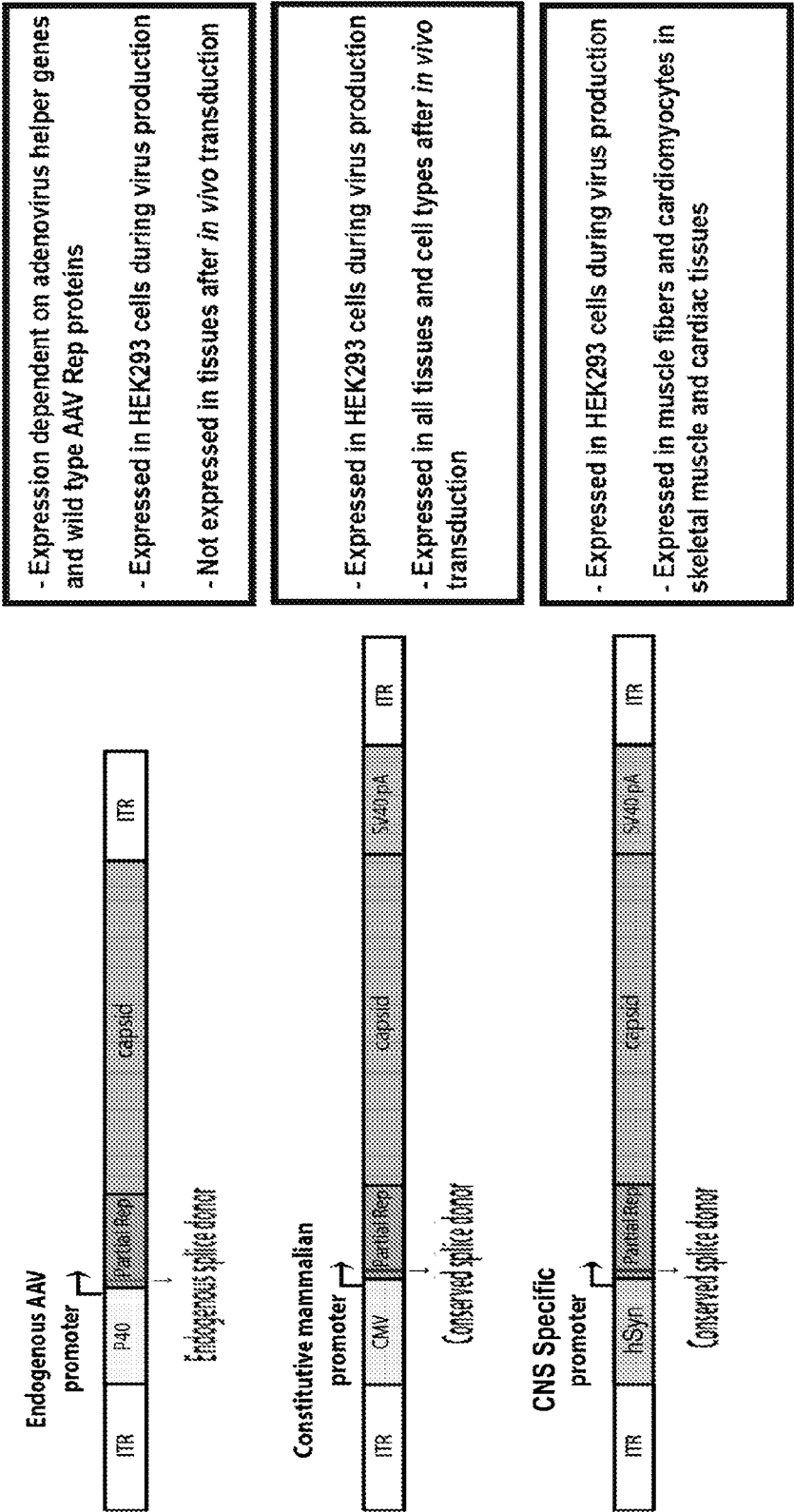
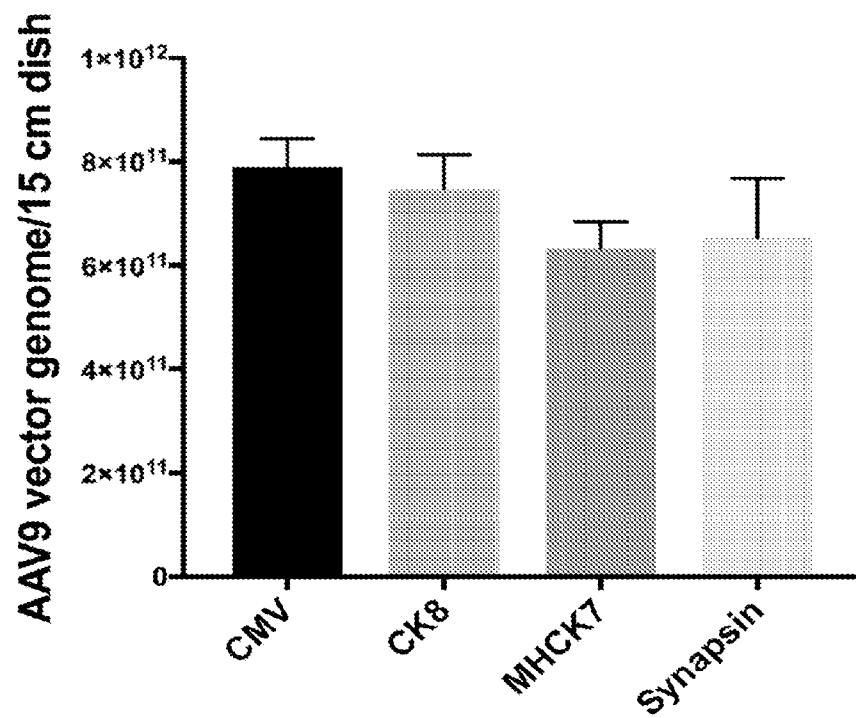


FIG. 9

10/18

**FIG. 10**

11/18

Mouse Brain mRNA

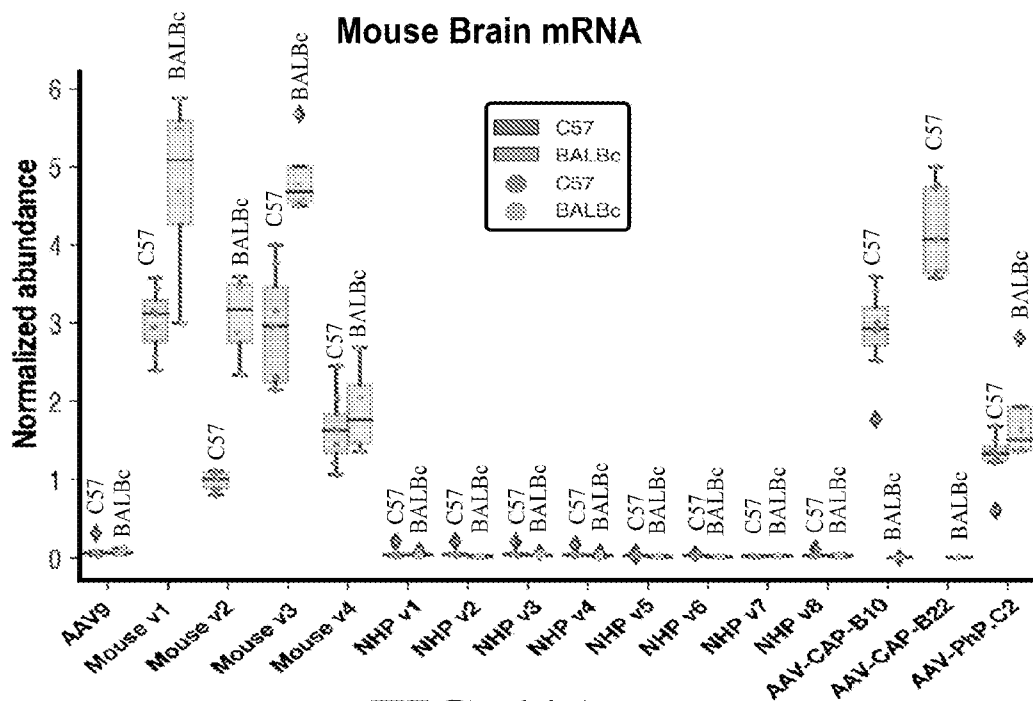


FIG. 11A

Mouse Brain DNA

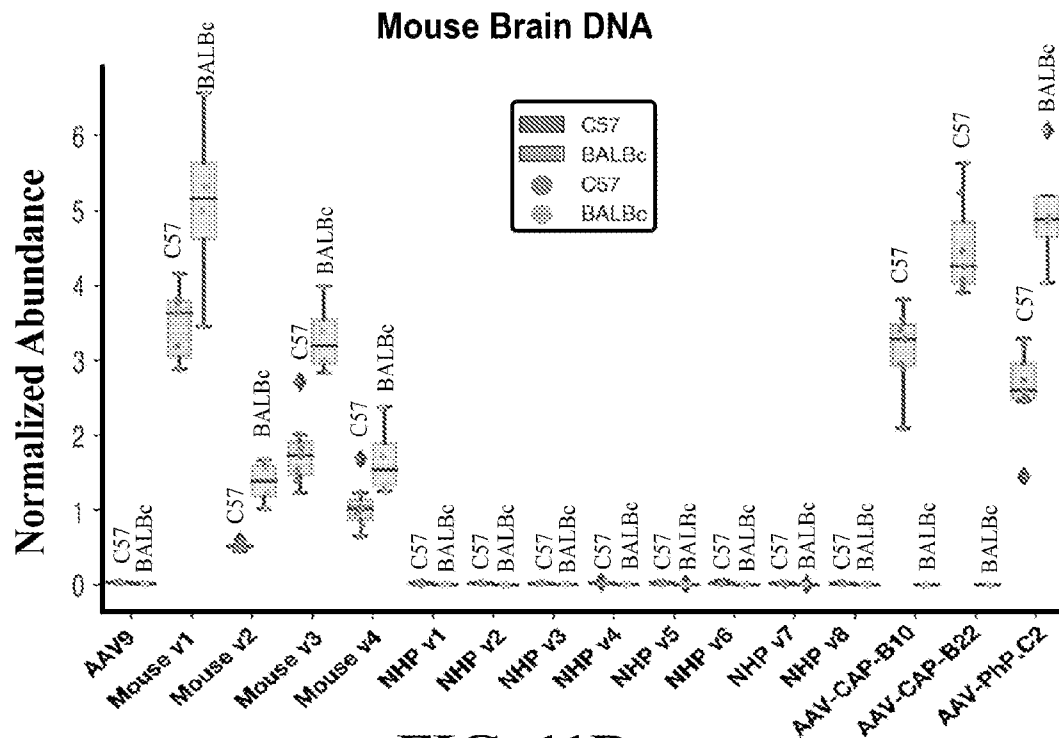


FIG. 11B

12/18

Mouse spinal cord mRNA

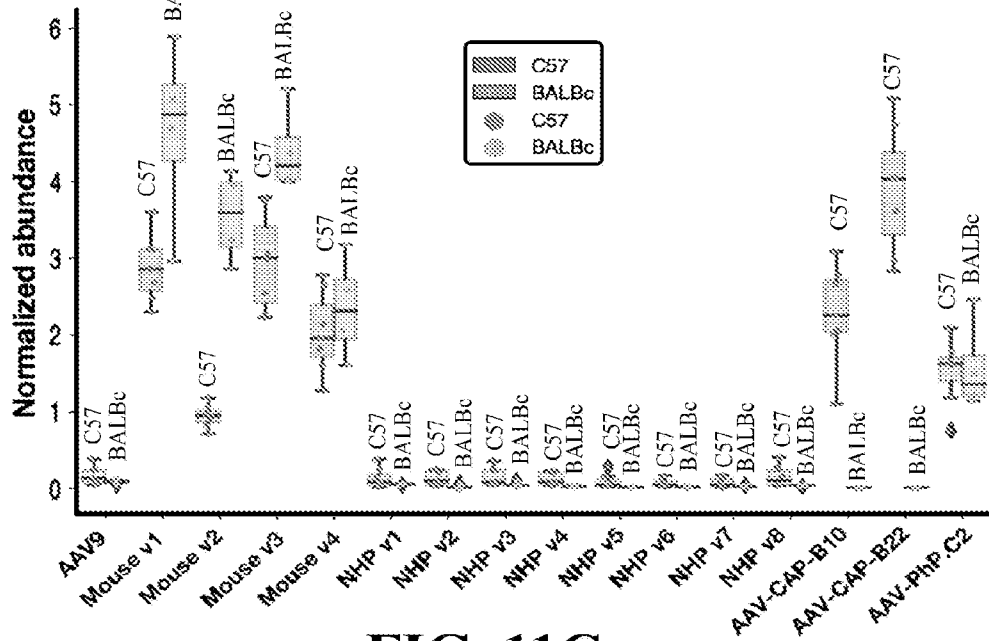


FIG. 11C

Mouse spinal cord DNA

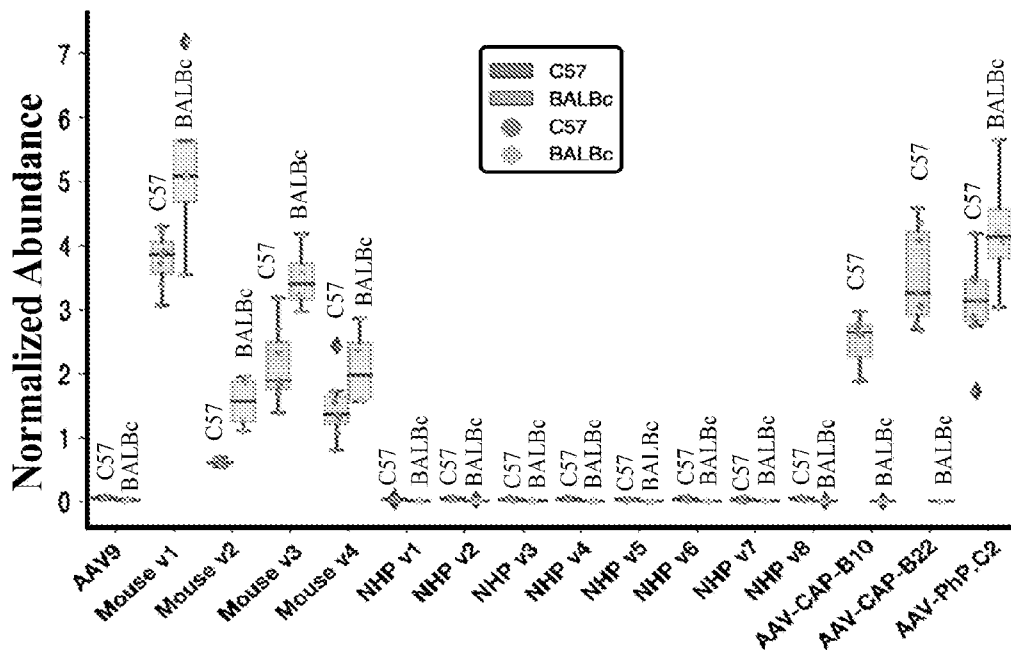
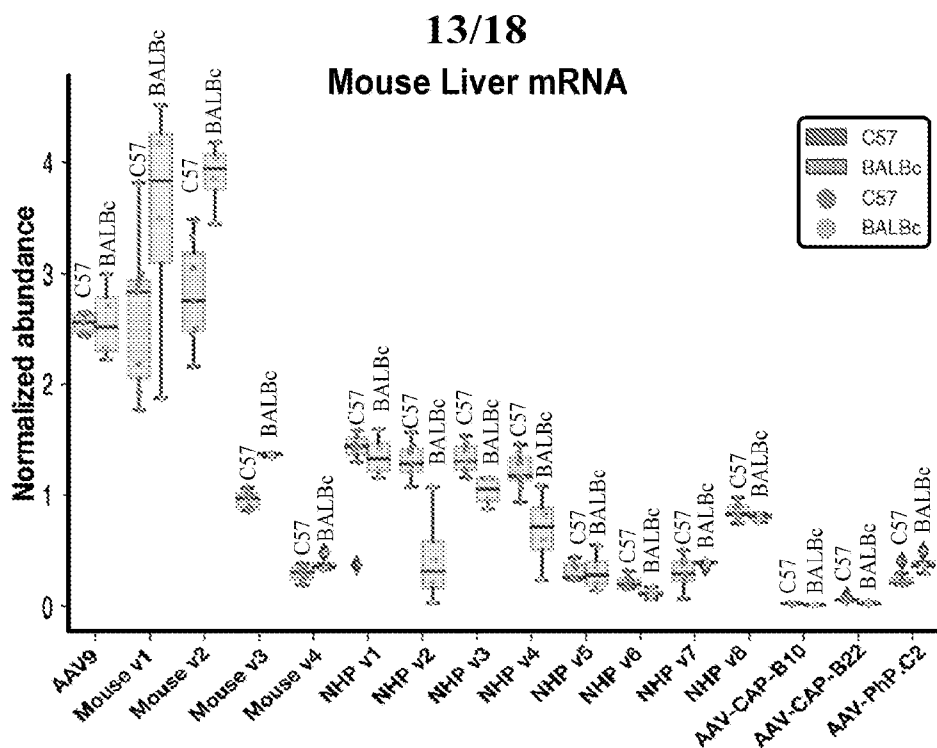
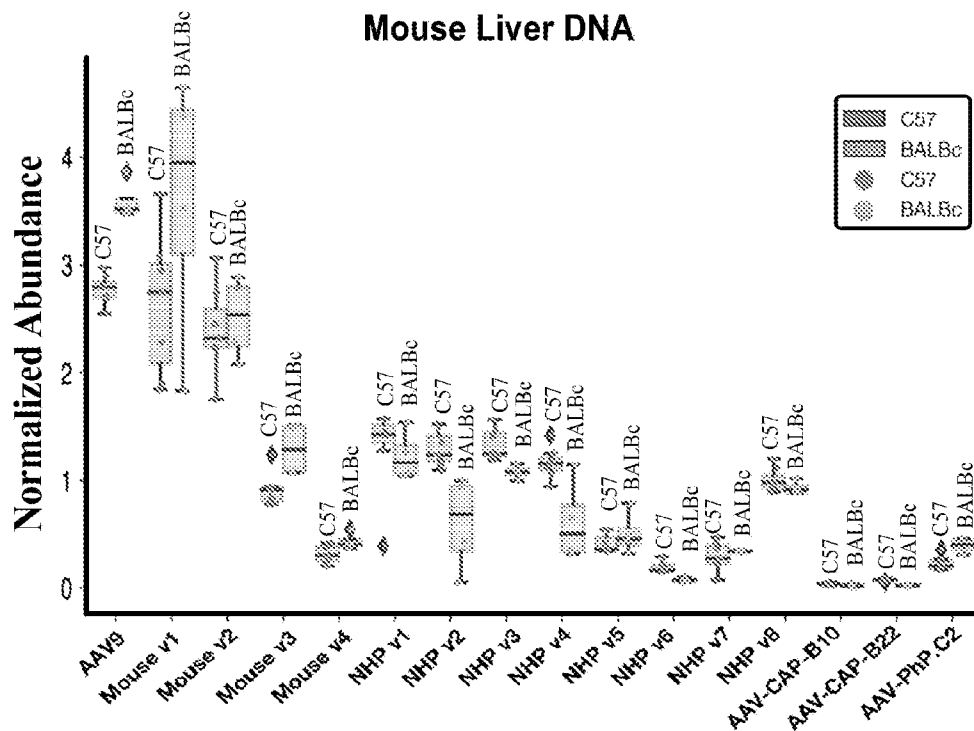


FIG. 11D

**FIG. 11E****FIG. 11F**

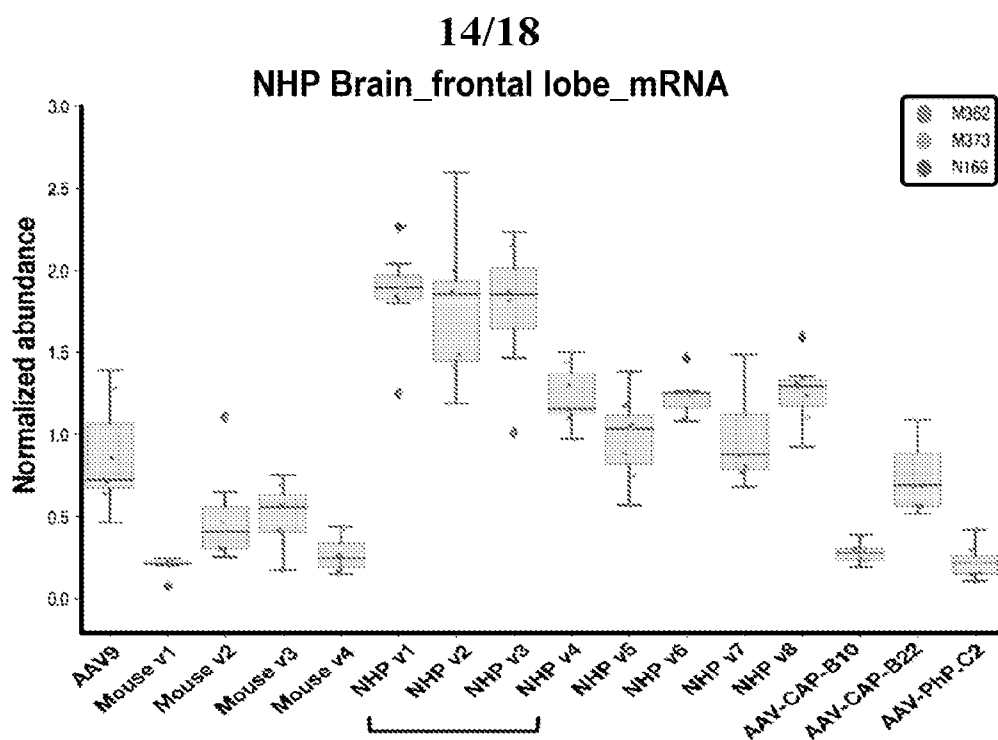


FIG. 11G

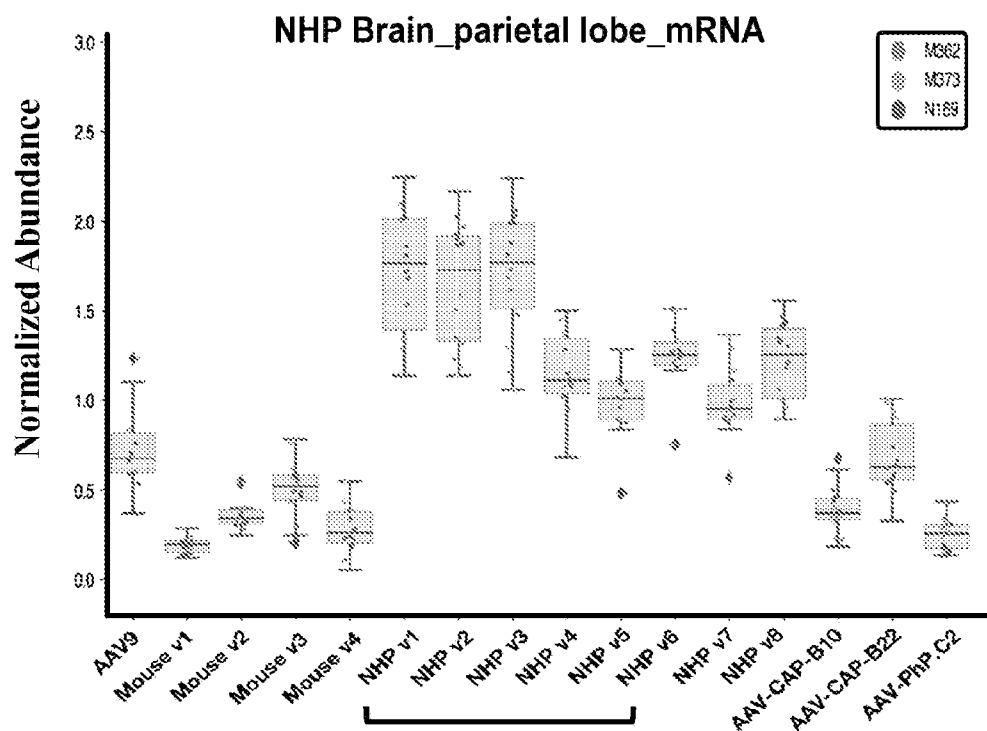
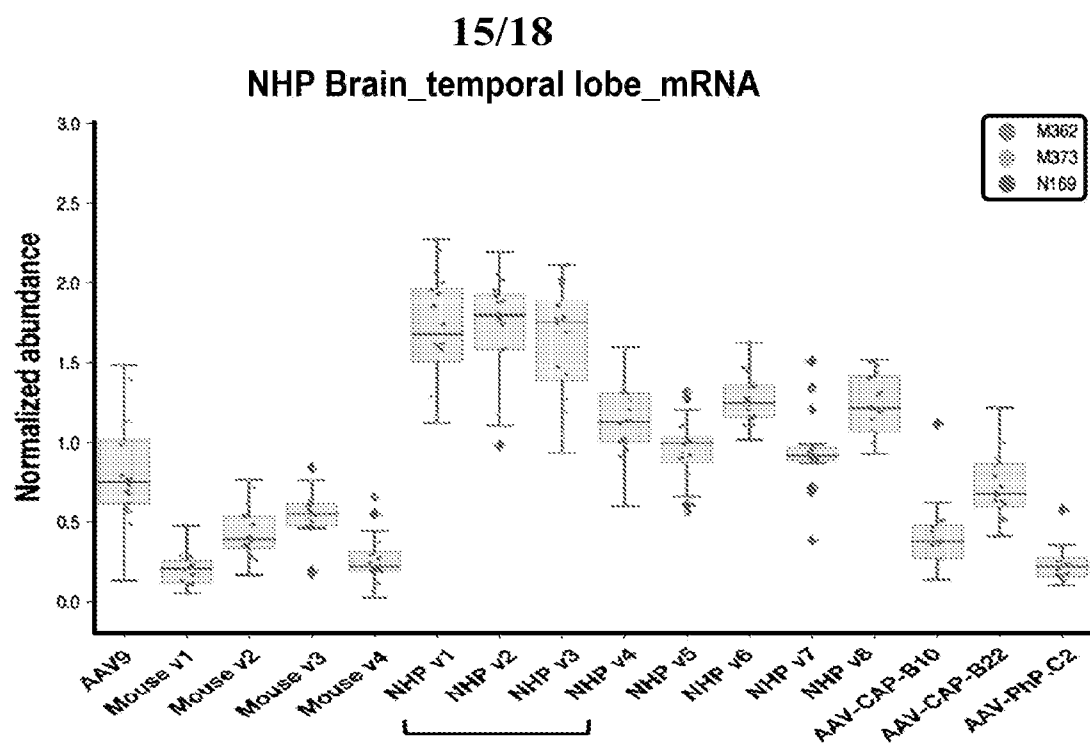
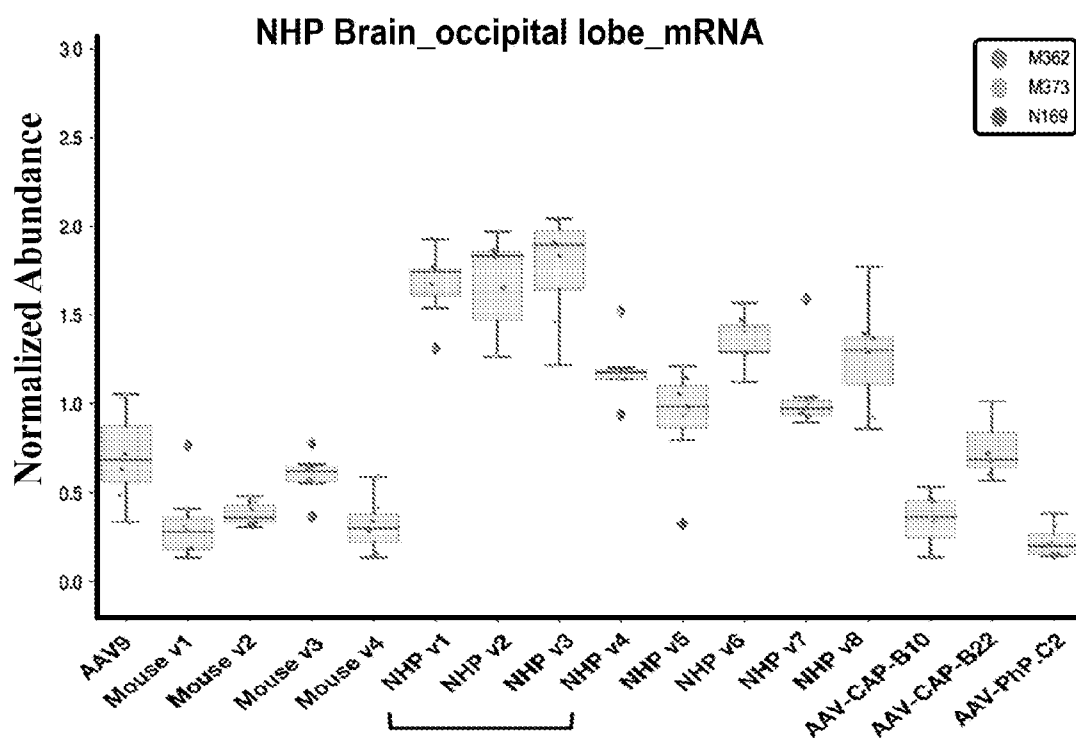
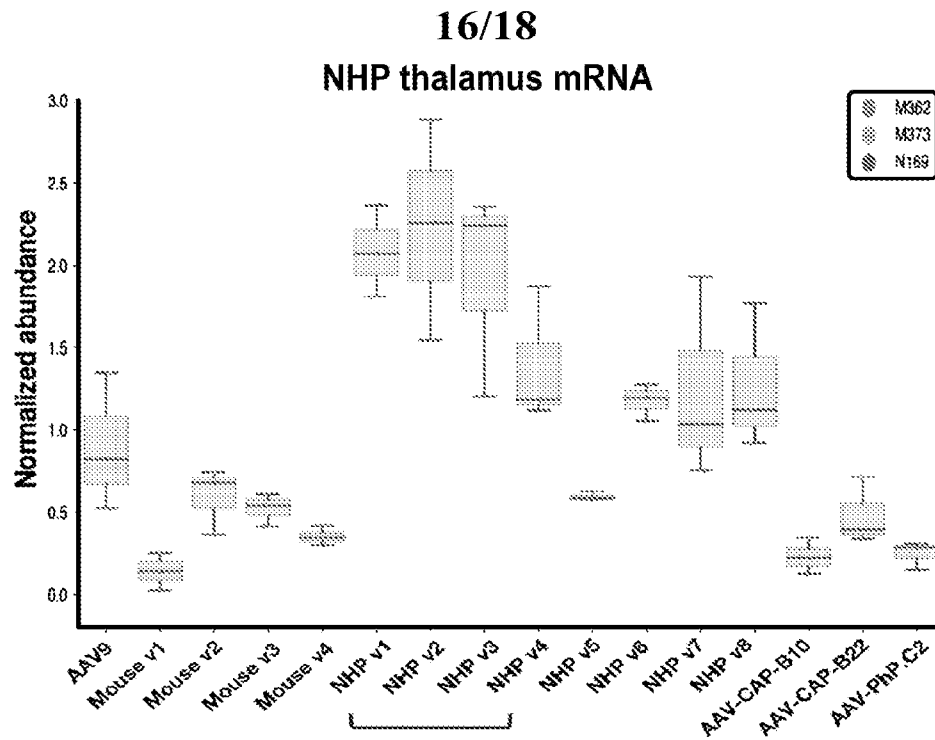
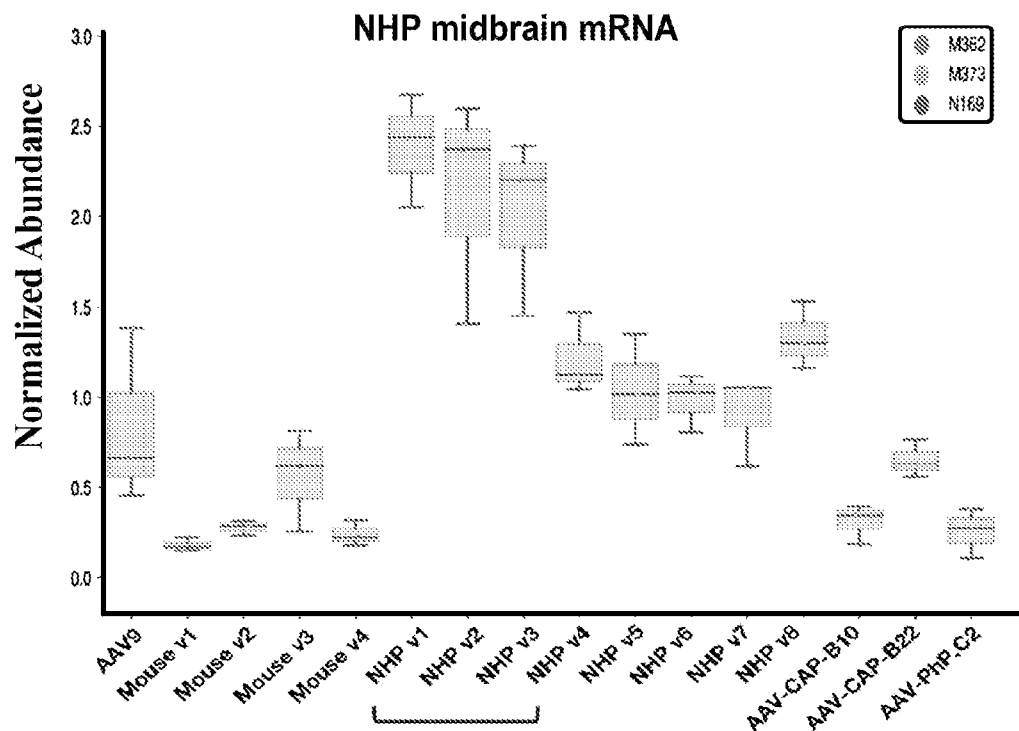
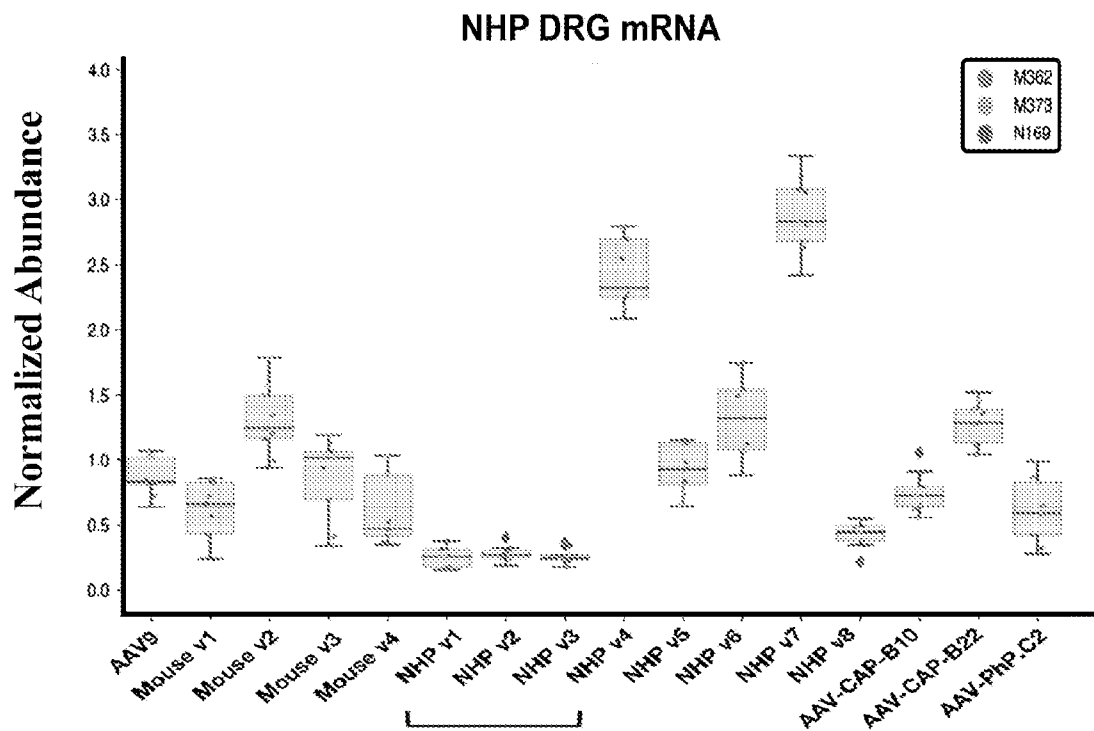
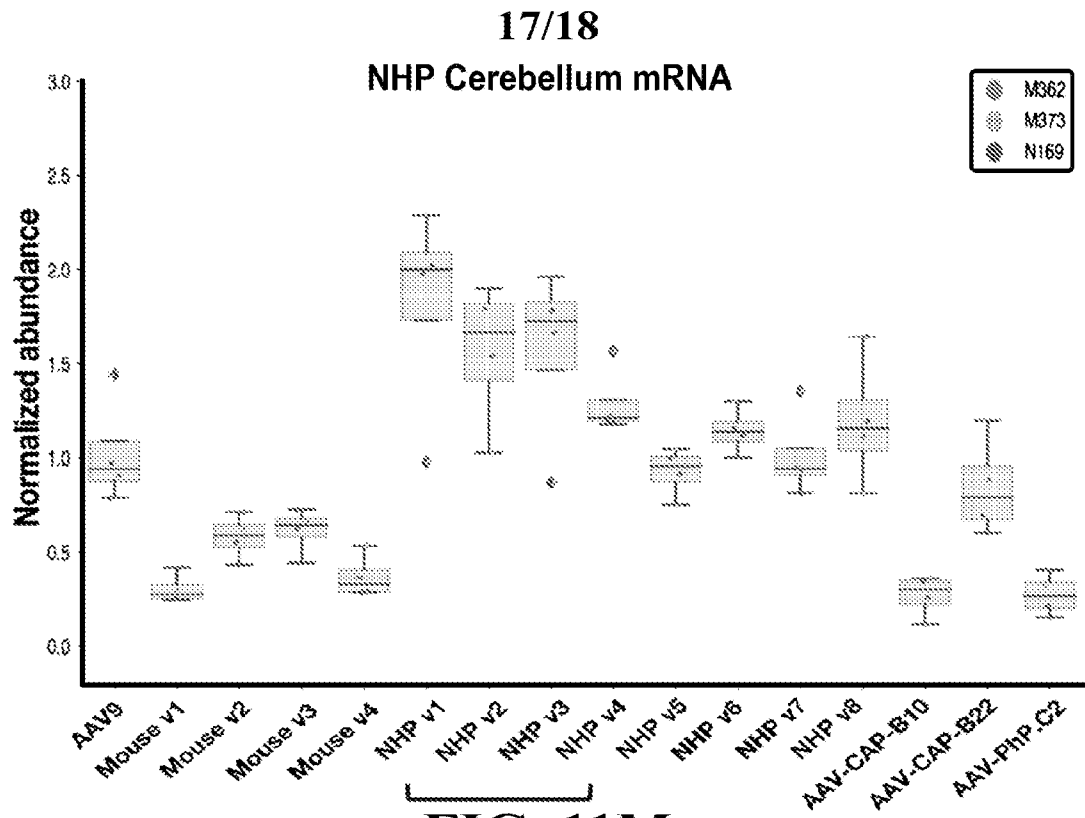


FIG. 11H

**FIG. 11I****FIG. 11J**

**FIG. 11K****FIG. 11L**



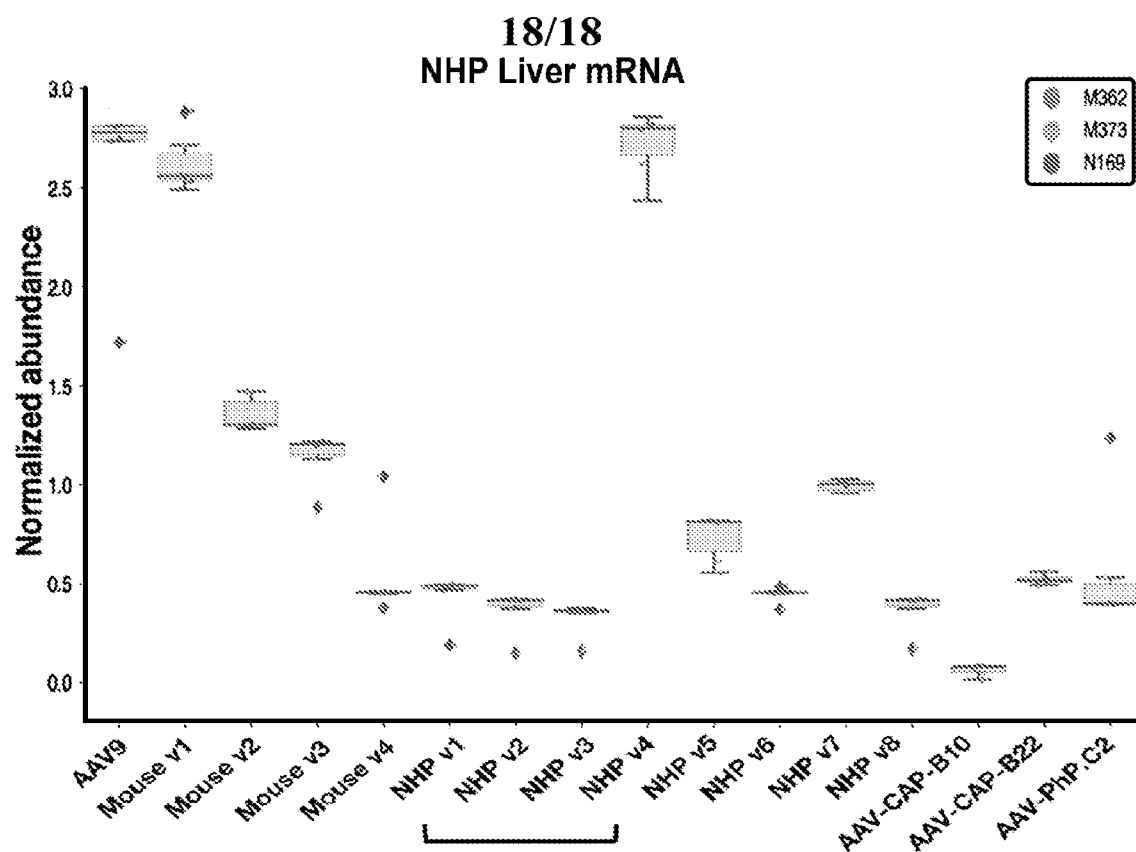


FIG. 11O
NHP Liver DNA

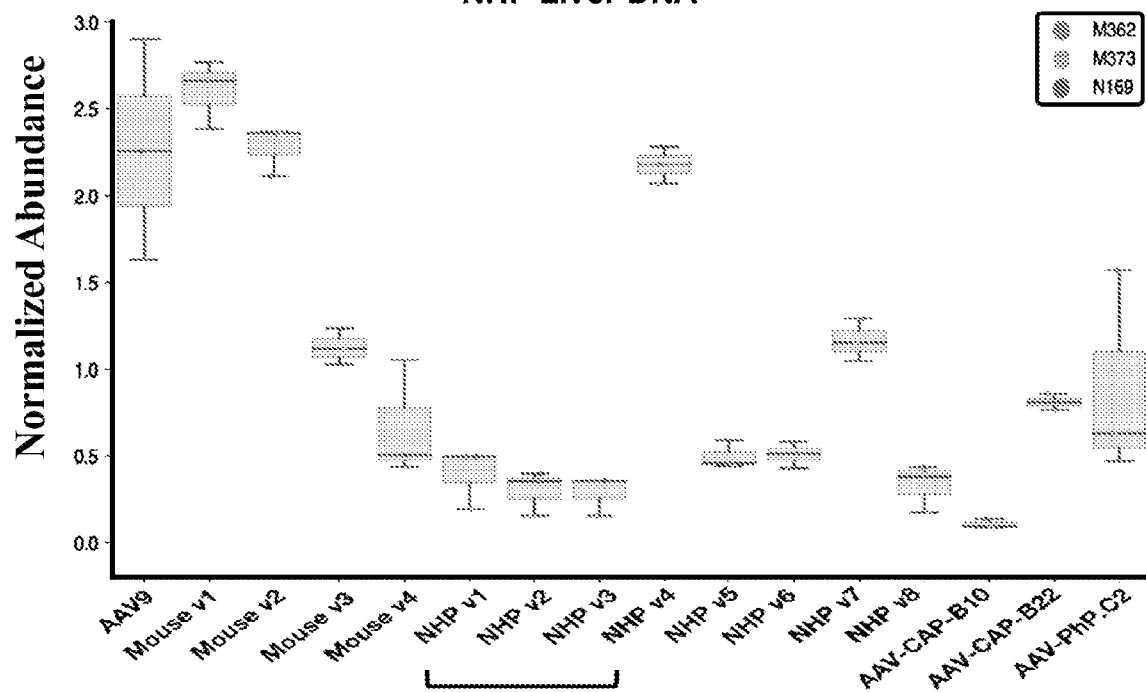


FIG. 11P



AAV capsid
virus library



C57BL/6J
mouse



BALB/cJ
mouse



Macaque

First round of capsid variant selection

Second round of selection

Benchmarking against state-of-the-art



Capsid variant for potent CNS-directed
therapeutic gene delivery

FIG. 6A