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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 May 2020 (22.05.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/102499 A3

(51) International Patent Classification:
A61K 9/127 (2006.01) *C12N 15/88* (2006.01)

(21) International Application Number:
PCT/US2019/061424

(22) International Filing Date:
14 November 2019 (14.11.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/767,358 14 November 2018 (14.11.2018) US
62/900,064 13 September 2019 (13.09.2019) US

(71) Applicant: **FLAGSHIP PIONEERING INNOVATIONS V, INC.** [US/US]; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US).

(72) Inventors: **VON MALTZAHN, Geoffrey, A.**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **RUBENS, Jacob, Rosenblum**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **SHAH, Jagesh, Vijaykumar**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **RUZO MATIAS, Albert**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **PUCCI, Ferdinando**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **MILWID, John, Miles**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **MEE, Michael, Travis**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US). **GORDON, Neal, Francis**; 55 Cambridge Parkway, 8th Floor, Cambridge, MA 02142 (US).

(74) Agent: **KOYFMAN, Hannah, R.** et al.; Lando & Anastasi, LLP, 60 State Street, 23rd Floor, Boston, MA 02109 (US).

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

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

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

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

(88) Date of publication of the international search report:
23 July 2020 (23.07.2020)

(54) Title: FUSOSOME COMPOSITIONS FOR CNS DELIVERY

(57) Abstract: The present disclosure provides, at least in part, methods and compositions for in vivo fusosome delivery. In some embodiments, the fusosome comprises a combination of elements that promote specificity for target cells, e.g., one or more of a fusogen, a positive target cell-specific regulatory element, and a non-target cell-specific regulatory element. In some embodiments, the fusosome comprises one or more modifications that decrease an immune response against the fusosome.

WO 2020/102499 A3

FUSOSOME COMPOSITIONS FOR CNS DELIVERY

Cross-Reference to Related Applications

This application claims priority to U.S. provisional applications: 62/767,358, filed November 14, 2018, entitled "FUSOSOME COMPOSITIONS FOR CNS CELL DELIVERY"; and 62/900,064, filed September 13, 2019, entitled "FUSOSOME COMPOSITIONS FOR CNS CELL DELIVERY", the contents of which are incorporated by reference in their entirety for all purposes.

Incorporation by Reference of Sequence Listing

The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled 186152003340SeqList.TXT, created November 14, 2019, which is 819 kilobytes in size. The information in the electronic format of the Sequence Listing is incorporated by reference in its entirety.

BACKGROUND

Complex biologics are promising therapeutic candidates for a variety of diseases. However, it is difficult to deliver large biologic agents into a cell because the plasma membrane acts as a barrier between the cell and the extracellular space. There is a need in the art for new methods of delivering complex biologics into cells in a subject.

SUMMARY

The present disclosure provides, at least in part, fusosome methods and compositions for in vivo delivery. In some embodiments, the fusosome comprises a combination of elements that promote specificity for target cells, e.g., one or more of a fusogen, a positive target cell-specific regulatory element, and a non-target cell-specific regulatory element. In some embodiments, the fusosome comprises one or more modifications that decrease an immune response against the fusosome.

Enumerated embodiments

1. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogenous agent, e.g. a payload gene encoding an exogneous agent of Table 5 or Table 6, optionally wherein the exogenous agent is set forth in any one of SEQ ID NOS: 134-154, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 134-154; and

(ii) a positive target cell-specific regulatory element (e.g., a target cell-specific promoter) operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a CNS cell.

2. The fusosome of embodiment 1, wherein the nucleic acid further comprises a non-target cell-specific regulatory element (NTCSRE) (e.g., a non-target cell-specific miRNA recognition sequence), operatively linked to the payload gene, wherein the NTCSRE decreases expression of the payload gene in a non-target cell relative to an otherwise similar fusosome lacking the NTCSRE, optionally wherein the target cell is a first type of CNS cell and the non-target cell is a second, different type of CNS cell or a non-CNS cell, optionally wherein:

the target cell is a neuron and the non-target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglia cell), or

the target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglia cell) and the non-target cell is a neuron.

3. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogenous agent, e.g., an exogenous agent of Table 5 or Table 6, optionally wherein the exogenous agent is set forth in any one of SEQ ID NOS: 134-154, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 134-154; and

(ii) a promoter operatively linked to the payload gene, wherein the promoter is chosen from a SYN, NSE, CaMKII, aTubulin, PDGF, fSST, fNPY, GAD67, DLX5/6, VGLUT1, Dock10, ChAT, VAChT, Drd1a, TPH-2, GFAP, EAAT1, GS, CX3CR1, TMEM119, MBP, CNP, or CRFR2β promoter, e.g., according to a sequence of a promoter in Table 3, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

4. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogenous agent, e.g. a payload gene encoding an exogenous agent of Table 5 or Table 6, optionally wherein the exogenous agent is set forth in any one of SEQ ID NOS: 134-154, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 134-154; and

(ii) a non-target cell-specific regulatory element (NTCSRE) (e.g., a non-target cell-specific miRNA recognition sequence), operatively linked to the payload gene, wherein the NTCSRE decreases expression of the payload gene in a non-target cell or tissue relative to an otherwise similar fusosome lacking the NTCSRE.

5. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogneous agent, e.g. a payload gene encoding an exogenous agent of Table 5 or Table 6, optionally wherein the exogenous agent is

set forth in any one of SEQ ID NOS: 134-154, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 134-154; and

(ii) a negative target cell-specific regulatory element (negative TCSRE) (e.g., a tissue-specific miRNA recognition sequence), operatively linked to the payload gene, wherein the negative TCSRE decreases expression of the exogenous agent in a non-target cell or tissue relative to an otherwise similar nucleic acid lacking the negative TCSRE.

6. The fusosome of either embodiment 4 or 5, wherein the nucleic acid further comprises a positive target cell-specific regulatory element (e.g., a target cell-specific promoter) operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a first type of CNS cell, optionally wherein the non-target cell is a second, different type of CNS cell or a non-CNS cell, optionally wherein:

the target cell is a neuron and the non-target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglia cell), or

the target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglia cell) and the non-target cell is a neuron.

7. A fusosome comprising:

a) a lipid bilayer comprising a fusogen;

b) a nucleic acid that comprises a payload gene encoding an exogenous agent, e.g. a payload gene encoding an exogenous agent of Table 5 or Table 6, optionally wherein the exogenous agent is set forth in any one of SEQ ID NOS: 134-154, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 134-154; and

c) one or both of:

(i) a first exogenous or overexpressed immunosuppressive protein on the lipid bilayer; or

(ii) a first immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell.

8. The fusosome of any of the preceding embodiments, wherein one or more of:

i) the fusosome fuses at a higher rate with a target cell than with a non-target cell, e.g., by at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold;

ii) the fusosome fuses at a higher rate with a target cell than with another fusosome, e.g., by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold;

iii) the fusosome fuses with target cells at a rate such that an agent in the fusosome is delivered to at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, of target cells after 24, 48, or 72 hours;

iv) the fusosome delivers the nucleic acid, e.g., retroviral nucleic acid, to a target cell at a higher rate than to a non-target cell, e.g., by at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold;

v) the fusosome delivers the nucleic acid, e.g., retroviral nucleic acid, to a target cell at a higher rate than to another fusosome, e.g., by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold; or

vi) the fusosome delivers the nucleic acid, e.g., retroviral nucleic acid, to a target cell at a rate such that an agent in the fusosome is delivered to at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, of target cells after 24, 48, or 72 hours.

9. The fusosome of any of the preceding embodiments, wherein one or more of (e.g., 2 or all 3 of) the following apply: the fusosome is a retroviral vector, the lipid bilayer is comprised by an envelope, e.g., a viral envelope, and the nucleic acid is a retroviral nucleic acid.

10. The fusosome of any of the preceding embodiments, wherein the nucleic acid comprises one or more of (e.g., all of) the following nucleic acid sequences: 5' LTR (e.g., comprising U5 and lacking a functional U3 domain), Psi packaging element (Psi), Central polypurine tract (cPPT) Promoter operatively linked to the payload gene, payload gene (optionally comprising an intron before the open reading frame), Poly A tail sequence, WPRE, and 3' LTR (e.g., comprising U5 and lacking a functional U3).

11. The fusosome of any of the preceding embodiments, which comprises one or more of (e.g., all of) a polymerase (e.g., a reverse transcriptase, e.g., pol or a portion thereof), an integrase (e.g., pol or a portion thereof, e.g., a functional or non-functional variant), a matrix protein (e.g., gag or a portion thereof), a capsid protein (e.g., gag or a portion thereof), a nucleocaspid protein (e.g., gag or a portion thereof), and a protease (e.g., pro).

12. The fusosome of embodiment 7, which comprises (i) and (ii).

13. The fusosome of any of embodiments 7-12, which further comprises a second exogenous or overexpressed immunosuppressive protein on the lipid bilayer.

14. The fusosome of any of embodiments 7-13, which further comprises a second immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell.

15. The fusosome of any of embodiments 7-14, wherein the nucleic acid, e.g., retroviral vector, further comprises a positive target cell-specific regulatory element (e.g., a target cell-specific promoter) operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a CNS cell.

16. The fusosome of any of embodiments 7-15, wherein the nucleic acid, e.g., retroviral nucleic acid, further comprises a non-target cell-specific regulatory element (NTCSRE) (e.g., a non-target cell-specific miRNA recognition sequence), operatively linked to the payload gene, wherein the NTCSRE decreases expression of the payload gene in a non-target cell or tissue relative to an otherwise similar fusosome lacking the NTCSRE, optionally wherein the target cell is a first type of CNS cell and the non-target cell is a second, different type of CNS cell or a non-CNS cell, optionally wherein:

the target cell is a neuron and the non-target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglia cell), or

the target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglia cell) and the non-target cell is a neuron.

17. The fusosome of any of embodiments 7-15, wherein the nucleic acid, e.g., retroviral nucleic acid, further comprises a negative target cell-specific regulatory element (negative TCSRE) (e.g., a tissue-specific miRNA recognition sequence), operatively linked to the payload gene, wherein the negative TCSRE decreases expression of the exogenous agent in a non-target cell or tissue relative to an otherwise similar nucleic acid, e.g., retroviral nucleic acid, lacking the negative TCSRE.

18. The fusosome of any of embodiments 7-17, wherein, when administered to a subject (e.g., a human subject or a mouse), one or more of:

i) the fusosome does not produce a detectable antibody response (e.g., after a single administration or a plurality of administrations), or antibodies against the fusosome are present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level, e.g., by a FACS antibody detection assay, e.g., an assay of Example 13 or Example 14);

ii) the fusosome does not produce a detectable cellular immune response (e.g., T cell response, NK cell response, or macrophage response), or a cellular immune response against the fusosome is present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level, e.g., by a PBMC lysis assay (e.g., an assay of Example 5), by an NK cell lysis assay (e.g., an assay of Example 6), by a CD8 killer T cell lysis assay (e.g., an assay of Example 7), or by a macrophage phagocytosis assay (e.g., an assay of Example 8);

iii) the fusosome does not produce a detectable innate immune response, e.g., complement activation (e.g., after a single administration or a plurality of administrations), or the innate immune response against the fusosome is present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level, e.g., by a complement activity assay (e.g., an assay of Example 9);

iv) less than 10%, 5%, 4%, 3%, 2%, or 1% of fusosomes are inactivated by serum, e.g., by a serum inactivation assay, e.g., an assay of Example 11 or Example 12;

v) a target cell that has received the exogenous agent from the fusosome does not produce a detectable antibody response (e.g., after a single administration or a plurality of administrations), or antibodies against the target cell are present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level, e.g., by a FACS antibody detection assay, e.g., an assay of Example 15; or

vi) a target cell that has received the exogenous agent from the fusosome does not produce a detectable cellular immune response (e.g., T cell response, NK cell response, or macrophage response), or a cellular response against the target cell is present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level, e.g., by a macrophage phagocytosis assay (e.g., an assay of Example 16), by a PBMC lysis assay (e.g., an assay of Example 17), by an NK cell lysis assay (e.g., an assay of Example 18), or by a CD8 killer T cell lysis assay (e.g., an assay of Example 19).

19. The fusosome of embodiment 18, wherein the background level is the corresponding level in the same subject prior to administration of the fusosome.

20. The fusosome of any of embodiments 7-19, wherein the immunosuppressive protein (e.g., first immunosuppressive protein or second immunosuppressive protein) is a complement regulatory protein or CD47.

21. The fusosome of any of embodiments 7-20, wherein the immunostimulatory protein (e.g., first immunostimulatory protein or second immunostimulatory protein) is an MHC I (e.g., HLA-A, HLA-B, HLA-C, HLA-E, or HLA-G) or MHC II (e.g., HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ, or HLA-DR) protein.

22. The fusosome of any of the preceding embodiments, wherein the exogenous agent is chosen from: SYNE1, SETX, FMR1, SLC6A8, UBE3A, SOD1, TDP43, C9orf72, FXN, MECP2, ASPA, or ALDH7A1; or the exogenous agent is chosen from: TPP1, FUCA1, GALC, HEXA, HEXB, MANBA, ARSA, GNPTAB, or MCOLN1.

23. The fusosome of any of the preceding embodiments, wherein the fusogen comprises VSV-G.

24. The fusosome of any embodiments 1, 2, 6, 15, 22, or 23, wherein the positive target cell-specific regulatory element comprises a CNS cell-specific promoter, a CNS cell-specific enhancer, a CNS cell-specific splice site, a CNS cell-specific site extending half-life of an RNA or protein, a CNS cell-specific mRNA nuclear export promoting site, a CNS cell-specific translational enhancing site, or a CNS cell-specific post-translational modification site.

25. The fusosome of any embodiments 1, 2, 6, 15, or 22-24, wherein the positive target cell-specific regulatory element comprises a CNS cell-specific promoter.

26. The fusosome of embodiment 25, wherein the CNS cell-specific promoter comprises a motif of Table 3.

27. The fusosome of embodiment 25 or 26, wherein the positive CNS cell-specific regulatory element comprises a promoter chosen from a SYN, NSE, CaMKII, aTubulin, PDGF, fSST, fNPY, GAD67, DLX5/6, VGLUT1, Dock10, ChAT, VAChT, Drd1a, TPH-2, GFAP, EAAT1, GS, CX3CR1, TMEM119, MBP, CNP, or CRFR2β promoter.

28. The fusosome of any of embodiments 4-6, or 16-21, wherein the negative TCSRE or NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site.

29. The fusosome of any of embodiments 4-6, 16-21, or 28, wherein the negative TCSRE or NTCSRE comprises a tissue-specific miRNA recognition sequence, tissue-specific protease recognition site, tissue-specific ubiquitin ligase site, tissue-specific transcriptional repression site, or tissue-specific epigenetic repression site.

30. The fusosome of any of embodiments 4-6, 16-21, 28, or 29, wherein the negative TCSRE or NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site.

31. The fusosome of any of embodiments 4-6, 16-21, or 28-30, wherein the negative TCSRE or NTCSRE comprises a non-target cell-specific miRNA recognition sequence bound by a miRNA of Table 4, e.g., by one or more of (e.g., two or more of) miR-338-3p, miR-9, miR-125b-5p, miR-342-3p, or miR-124 optionally wherein the miRNA is or comprises the sequence set forth in any one of SEQ ID NOS: 156-162.

32. The fusosome of any of embodiments 28-31, wherein the negative TCSRE or NTCSRE is situated or encoded within a transcribed region (e.g., the transcribed region encoding the exogenous agent), e.g., such that an RNA produced by the transcribed region comprises the miRNA recognition sequence within a UTR or coding region.

33. The fusosome of any of the preceding embodiments, wherein the nucleic acid, e.g., retroviral nucleic acid, comprises one or more insulator elements.

34. The fusosome of embodiment 33, wherein the nucleic acid, e.g., retroviral nucleic acid, comprises two insulator elements, e.g., a first insulator element upstream of the payload gene and a second insulator element downstream of the payload gene, e.g., wherein the first insulator element and second insulator element comprise the same or different sequences.

35. The fusosome of any of the preceding embodiments, which is not genotoxic or does not increase the rate of tumor formation in target cells.

36. The fusosome of any of the preceding embodiments, wherein the nucleic acid, e.g., retroviral nucleic acid, is capable of integrating into the genome of a target cell.

37. The fusosome of embodiment 36, wherein the nucleic acid, e.g., retroviral nucleic acid, is an integration-competent lentivirus or an integration-deficient lentivirus.

38. The fusosome of any of the preceding embodiments, wherein the target cell is chosen from a CNS cell, a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, an astrocyte, a microglia, an oligodendrocytes, or a choroid plexus cell.

39. The fusosome of any of embodiments 4-6 and 9-38, wherein one or more of:

i) less than 10%, 5%, 4%, 3%, 2%, or 1% of the exogenous agent detectably present in the subject is in non-target cells;

ii) at least 90%, 95%, 96%, 97%, 98%, or 99% of the cells of the subject that detectably comprise the exogenous agent, are target cells (e.g., cells of a single cell type);

iii) less than 1,000,000, 500,000, 200,000, 100,000, 50,000, 20,000, or 10,000 cells of the cells of the subject that detectably comprise the exogenous agent are non-target cells;

iv) average levels of the exogenous agent in all target cells in the subject are at least 100-fold, 200-fold, 500-fold, or 1,000-fold higher than average levels of the exogenous agent in all non-target cells in the subject; or

v) the exogenous agent is not detectable in any non-target cell in the subject.

40. The fusosome of any of the preceding embodiments, wherein the nucleic acid, e.g., retroviral nucleic acid, encodes a positive TCSRE and/or a NTCSRE or negative TCSRE.

41. The fusosome of any of the preceding embodiments, wherein the nucleic acid, e.g., retroviral nucleic acid, comprises the complement of a positive TCSRE and/or a NTCSRE or negative TCSRE.

42. The fusosome of either embodiment 40 or 41, wherein the positive TCSRE comprises a target cell-specific promoter that is at least 10%, 25%, 50%, 75%, 100%, 150%, 200%, 250%, 300%, 400%, 500%, 750%, 1000% or more active in a target cell than a non-target cell.

43. The fusosome of any of embodiments 40-42, wherein the negative TCSRE or NTCSRE comprises a miRNA recognition sequence that decreases gene expression by at least 10%, 25%, 50%, 75%, or 100% in a non-target cell compared to a target cell.

44. The fusosome of any of the preceding embodiments, which does not deliver nucleic acid, e.g., retroviral nucleic acid, to a non-target cell, e.g., a neuron, a glial cell, an antigen presenting cell, an MHC class II+ cell, a professional antigen presenting cell, an atypical antigen presenting cell, a macrophage, a dendritic cell, a myeloid dendritic cell, a plasmacyteoid dendritic cell, a CD11c+ cell, a CD11b+ cell, a splenocyte, a B cell, a hepatocyte, a endothelial cell, or a non-cancerous cell.

45. The fusosome of any of the preceding embodiments, wherein less than 10%, 5%, 2.5%, 1%, 0.5%, 0.1%, 0.01%, 0.001%, 0.0001%, 0.00001%, or 0.000001% of a non-target cell type (e.g., one or more of a neuron, a glial cell, an antigen presenting cell, an MHC class II+ cell, a professional antigen presenting cell, an atypical antigen presenting cell, a macrophage, a dendritic cell, a myeloid dendritic cell, a plasmacyteoid dendritic cell, a CD11c+ cell, a CD11b+ cell, a splenocyte, a B cell, a hepatocyte, a endothelial cell, or a non-cancerous cell) comprise the nucleic acid, e.g., retroviral nucleic acid, e.g., using quantitative PCR, e.g., using an assay of Example 1.

46. The fusosome of any of the preceding embodiments, wherein the target cells comprise 0.00001-10, .0001-10, .001-10, .01-10, .1-10, .5 – 5, 1-4, 1-3, or 1-2 copies of the nucleic acid, e.g., retroviral nucleic acid, or a portion thereof, per host cell genome, e.g., wherien

copy number of the nucleic acid, e.g., retroviral nucleic acid, is assessed after administration in vivo.

47. The fusosome of any of the preceding embodiments, wherein:

less than 10%, 5%, 2.5%, 1%, 0.5%, 0.1%, 0.01% of the non-target cells (e.g., a neuron, a glial cell, an antigen presenting cell, an MHC class II+ cell, a professional antigen presenting cell, an atypical antigen presenting cell, a macrophage, a dendritic cell, a myeloid dendritic cell, a plasmacyteoid dendritic cell, a CD11c+ cell, a CD11b+ cell, a splenocyte, a B cell, a hepatocyte, a endothelial cell, or a non-cancerous cell) comprise the exogenous agent; or

the exogenous agent (e.g., protein) is not detectably present in a non-target cell, e.g., a neuron, a glial cell, an antigen presenting cell, an MHC class II+ cell, a professional antigen presenting cell, an atypical antigen presenting cell, a macrophage, a dendritic cell, a myeloid dendritic cell, a plasmacyteoid dendritic cell, a CD11c+ cell, a CD11b+ cell, a splenocyte, a B cell, a hepatocyte, a endothelial cell, or a non-cancerous cell.

48. The fusosome of any of the preceding embodiments, wherein the fusosome delivers the nucleic acid, e.g., retroviral nucleic, acid to a target cell, e.g., a CNS cell, a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell.

49. The fusosome of any of the preceding embodiments, wherein at least 0.00001%, 0.0001%, 0.001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of target cells (e.g., one or more of a CNS cell, a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell) comprise the nucleic acid, e.g., retroviral nucleic acid, e.g., using quantitative PCR, e.g., using an assay of Example 3.

50. The fusosome of any of the preceding embodiments, wherein at least 0.00001%, 0.0001%, 0.001%, 0.001%, 0.01%, 0.1%, 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%,

80%, or 90% of target cells (e.g., a CNS cell, a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell) comprise the exogenous agent.

51. The fusosome of any of the preceding embodiments, wherein, upon administration, the ratio of target cells comprising the nucleic acid, e.g., retroviral nucleic acid, to non-target cells comprising the nucleic acid, e.g., retroviral nucleic acid, is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a quantitative PCR assay, e.g., using assays of Example 1 and Example 3.

52. The fusosome of any of the preceding embodiments, wherein the ratio of the average copy number of nucleic acid, e.g., retroviral nucleic acid, or a portion thereof in target cells to the average copy number of nucleic acid, e.g., retroviral nucleic acid, or a portion thereof in non-target cells is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a quantitative PCR assay, e.g., using assays of Example 1 and Example 3.

53. The fusosome of any of the preceding embodiments, wherein the ratio of the median copy number of of nucleic acid, e.g., retroviral nucleic acid, or a portion thereof in target cells to the median copy number of nucleic acid, e.g., retroviral nucleic acid, or a portion thereof in non-target cells is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a quantitative PCR assay, e.g., using assays of Example 1 and Example 3.

54. The fusosome of any of the preceding embodiments, wherein the ratio of target cells comprising the exogenous RNA agent to non-target cells comprising the exogenous RNA agent is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a reverse transcription quantitative PCR assay.

55. The fusosome of any of the preceding embodiments, wherein the ratio of the average exogenous RNA agent level of target cells to the average exogenous RNA agent level of

non-target cells is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a reverse transcription quantitative PCR assay.

56. The fusosome of any of the preceding embodiments, wherein the ratio of the median exogenous RNA agent level of target cells to the median exogenous RNA agent level of non-target cells is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a reverse transcription quantitative PCR assay.

57. The fusosome of any of the preceding embodiments, wherein the ratio of target cells comprising the exogenous protein agent to non-target cells comprising the exogenous protein agent is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a FACS assay, e.g., using assays of Example 2 and Example 4.

58. The fusosome of any of the preceding embodiments, wherein the ratio of the average exogenous protein agent level of target cells to the average exogenous protein agent level of non-target cells is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a FACS assay, e.g., using assays of Example 2 and Example 4.

59. The fusosome of any of the preceding embodiments, wherein the ratio of the median exogenous protein agent level of target cells to the median exogenous protein agent level of non-target cells is at least 1.5, 2, 3, 4, 5, 10, 25, 50, 100, 500, 1000, 5000, 10,000, e.g., according to a FACS assay, e.g., using assays of Example 2 and Example 4.

60. The fusosome of any of the preceding embodiments, which comprises one or both of:

i) an exogenous or overexpressed immunosuppressive protein on the lipid bilayer, e.g., envelope; and

ii) an immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell.

61. The fusosome of any of the preceding embodiments, which comprises one or more of:

i) a first exogenous or overexpressed immunosuppressive protein on the lipid bilayer, e.g., envelope, and a second exogenous or overexpressed immunosuppressive protein on the lipid bilayer, e.g., envelope;

ii) a first exogenous or overexpressed immunosuppressive protein on the lipid bilayer, e.g., envelope, and a second immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell; or

iii) a first immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell and a second immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell.

62. The fusosome of any of the preceding embodiments, wherein the fusosome is in circulation at least 0.5, 1, 2, 3, 4, 6, 12, 18, 24, 36, or 48 hours after administration to the subject.

63. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 30 minutes after administration.

64. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 1 hour after administration.

65. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 2 hours after administration.

66. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 4 hours after administration.

67. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 8 hours after administration.

68. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 12 hours after administration.

69. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 18 hours after administration.

70. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 24 hours after administration.

71. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 36 hours after administration.

72. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are in circulation 48 hours after administration.

73. The fusosome of any of the preceding embodiments, which has a reduction in immunogenicity as measured by a reduction in humoral response following one or more administration of the fusosome to an appropriate animal model, e.g., an animal model described

herein, compared to reference fusosome, e.g., an unmodified fusosome otherwise similar to the fusosome.

74. The fusosome of embodiment 73, wherein the reduction in humoral response is measured in a serum sample by an anti-cell antibody titre, e.g., anti-retroviral antibody titre, e.g., by ELISA.

75. The fusosome of any of the preceding embodiments, wherein a serum sample from animals administered the fusosome has a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of an anti-fusosome antibody titer compared to the serum sample from a subject administered an unmodified cell.

76. The fusosome of any of the preceding embodiments, wherein a serum sample from a subject administered the fusosome has an increased anti-cell antibody titre, e.g., increased by 1%, 2%, 5%, 10%, 20%, 30%, or 40% from baseline, e.g., wherein baseline refers to serum sample from the same subject before administration of the fusosome.

77. The fusosome of any of the preceding embodiments, wherein:

the subject to be administered the fusosome or a pharmaceutical composition comprising the fusosome has, or is known to have, or is tested for, a pre-existing antibody (e.g., IgG or IgM) reactive with the fusosome;

the subject to be administered the fusosome does not have detectable levels of a pre-existing antibody reactive with the fusosome;

a subject that has received the fusosome or a pharmaceutical composition comprising the fusosome has, or is known to have, or is tested for, an antibody (e.g., IgG or IgM) reactive with the fusosome;

the subject that received the fusosome or a pharmaceutical composition comprising the fusosome (e.g., at least once, twice, three times, four times, five times, or more) does not have detectable levels of antibody reactive with the fusosome; or

levels of antibody do not rise more than 1%, 2%, 5%, 10%, 20%, or 50% between two timepoints, the first timepoint being before the first administration of the fusosome, and the second timepoint being after one or more administrations of the fusosome.

78. The fusosome of any of the preceding embodiments, wherein the fusosome is produced by the methods of Example 5, 6, or 7, e.g., from cells transfected with HLA-G or HLA-E cDNA.

79. The fusosome of any of the preceding embodiments, wherein fusosomes generated from NMC-HLA-G cells have a decreased percentage of lysis, e.g., PBMC mediated lysis, NK cell mediated lysis, and/or CD8+ T cell mediated lysis, at specific timepoints as compared to fusosomes generated from NMCs or NMC-empty vector.

80. The fusosome of any of the preceding embodiments, wherein the modified fusosome evades phagocytosis by macrophages.

81. The fusosome of any of the preceding embodiments, wherein the fusosome is produced by the methods of Example 8, e.g., from cells transfected with CD47 cDNA.

82. The fusosome of any of the preceding embodiments, wherein the phagocytic index is reduced when macrophages are incubated with fusosomes derived from NMC-CD47, versus those derived from NMC, or NMC-empty vector.

83. The fusosome of any of the preceding embodiments, which has a reduction in macrophage phagocytosis, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in macrophage phagocytosis compared to a reference fusosome, e.g., an unmodified fusosome otherwise similar to the fusosome, wherein the reduction in macrophage phagocytosis is determined by assaying the phagocytosis index *in vitro*, e.g., as described in Example 8.

84. The fusosome of any of the preceding embodiments, wherein the fusosome composition has a phagocytosis index of 0, 1, 10, 100, or more, e.g., as measured by an assay of Example 8, when incubated with macrophages in an *in vitro* assay of macrophage phagocytosis.

85. The fusosome of any of the preceding embodiments, which is modified and has reduced complement activity compared to an unmodified fusosome.

86. The fusosome of any of the preceding embodiments, which is produced by the methods of Example 9, e.g., from cells transfected with a cDNA coding for a complement regulatory protein, e.g., DAF.

87. The fusosome of any of the preceding embodiments, wherein the dose of fusosome at which 200 pg/ml of C3a is present is greater for the modified fusosome (e.g., HEK293-DAF) incubated with corresponding mouse sera (e.g., HEK-293 DAF mouse sera) than for the reference fusosome (e.g., HEK293 retroviral vector) incubated with corresponding mouse sera (e.g., HEK293 mouse sera).

88. The fusosome of any of the preceding embodiments, wherein the dose of fusosome at which 200 pg/ml of C3a is present is greater for for the modified fusosome (e.g., HEK293-DAF) incubated with naive mouse sera than for the reference fusosome (e.g., HEK293 retroviral vector) incubated with naive mouse sera.

89. The fusosome of any of the preceding embodiments, wherein the fusosome is resistant to complement mediated inactivation in patient serum 30 minutes after administration according to an assay of Example 9.

90. The fusosome of any of the preceding embodiments, wherein at least 0.001%, 0.01%, 0.1%, 1%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of fusosomes are resistant to complement mediated inactivation.

91. The fusosome of any of embodiments 86-90, wherein the complement regulatory protein comprises one or more of proteins that bind decay-accelerating factor (DAF, CD55), e.g. factor H (FH)-like protein-1 (FHL-1), e.g. C4b-binding protein (C4BP), e.g. complement receptor 1 (CD35), e.g. Membrane cofactor protein (MCP, CD46), eg. Protectin (CD59), e.g. proteins that inhibit the classical and alternative complement pathway CD/C5 convertase enzymes, e.g. proteins that regulate MAC assembly.

92. The fusosome of any of the preceding embodiments, which is produced by the methods of Example 10, e.g., from cells transfected with a DNA coding for an shRNA targeting MHC class I, e.g., wherein retroviral vectors derived from NMC- shMHC class I has lower expression of MHC class I compared to NMCs and NMC-vector control.

93. The fusosome of any of the preceding embodiments, wherein a measure of immunogenicity for fusosomes is serum inactivation, e.g., serum inactivation measured as described herein, e.g., as described in Example 11.

94. The fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is not different between fusosome samples that have been incubated with serum and heat-inactivated serum from fusosome naïve mice.

95. The fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is not different between fusosome samples that have been incubated with serum from fusosome naïve mice and no-serum control incubations.

96. fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is less in fusosome samples that have been incubated with positive control serum than in fusosome samples that have been incubated with serum from fusosome naïve mice.

97. The fusosome of any of the preceding embodiments, wherein a modified fusosome, e.g., modified by a method described herein, has a reduced (e.g., reduced compared to

administration of an unmodified fusosome) serum inactivation following multiple (e.g., more than one, e.g., 2 or more), administrations of the modified fusosome.

98. The fusosome of any of the preceding embodiments, wherein a fusosome described herein is not inactivated by serum following multiple administrations.

99. The fusosome of any of the preceding embodiments, wherein a measure of immunogenicity for the fusosome is serum inactivation, e.g., after multiple administrations, e.g., serum inactivation after multiple administrations measured as described herein, e.g., as described in Example 12.

100. The fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is not different between fusosome samples that have been incubated with serum and heat-inactivated serum from mice treated with modified (e.g., HEK293-HLA-G) fusosomes.

101. The fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is not different between fusosome samples that have been incubated from mice treated 1, 2, 3, 5 or 10 times with modified (e.g., HEK293-HLA-G) fusosomes.

102. The fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is not different between fusosome samples that have been incubated with serum from mice treated with vehicle and from mice treated with modified (e.g., HEK293-HLA-G) fusosomes.

103. The fusosome of any of the preceding embodiments, wherein the percent of cells which receive the exogenous agent is less for fusosomes derived from a reference cell (e.g., HEK293) than for modified (e.g., HEK293-HLA-G) fusosomes.

104. The fusosome of any of the preceding embodiments, wherein a measure of immunogenicity for a fusosome is antibody response.

105. The fusosome of any of the preceding embodiments, wherein a subject that receives a fusosome described herein has pre-existing antibodies which bind to and recognize fusosome, e.g., measured as described herein, e.g., as described in Example 13.

106. The fusosome of any of the preceding embodiments, wherein serum from fusosome -naïve mice shows more signal (e.g., fluorescence) than the negative control, e.g., serum from a mouse depleted of IgM and IgG, e.g., indicating that in immunogenicity has occurred.

107. The fusosome of any of the preceding embodiments, wherein serum from fusosome -naïve mice shows similar signal (e.g., fluorescence) compared to the negative control, e.g., indicating that immunogenicity did not detectably occur.

108. The fusosome of any of the preceding embodiments, which is a modified fusosome, e.g., modified by a method described herein, and which has a reduced (e.g., reduced compared to administration of an unmodified fusosome) humoral response following multiple (e.g., more than one, e.g., 2 or more), administrations of the modified fusosome, e.g., measured as described herein, e.g., as described in Example 14.

109. The fusosome of any of the preceding embodiments, wherein the fusosome is produced by the methods of Example 5, 6, 7, or 14, e.g., from cells transfected with HLA-G or HLA-E cDNA.

110. The fusosome of any of the preceding embodiments, wherein humoral response is assessed by determining a value for the level of anti-fusosome antibodies (e.g., IgM, IgG1, and/or IgG2 antibodies).

111. The fusosome of any of the preceding embodiments, wherein modified (e.g., NMC-HLA-G) fusosomes have decreased anti-viral IgM or IgG1/2 antibody titers (e.g., as measured by fluorescence intensity on FACS) after injections, as compared to a control, e.g., NMC fusosomes or NMC-empty fusosomes.

112. The fusosome of any of the preceding embodiments, wherein recipient cells are not targeted by an antibody response, or an antibody response will be below a reference level, e.g., measured as described herein, e.g., as described in Example 15.

113. The fusosome of any of the preceding embodiments, signal (e.g., mean fluorescence intensity) is similar for recipient cells from mice treated with fusosomes and mice treated with PBS.

114. The fusosome of any of the preceding embodiments, wherein a measure of the immunogenicity of recipient cells is the macrophage response.

115. The fusosome of any of the preceding embodiments, wherein recipient cells are not targeted by macrophages, or are targeted below a reference level.

116. The fusosome of any of the preceding embodiments, wherein the phagocytic index, e.g., measured as described herein, e.g., as described in Example 16, is similar for recipient cells derived from mice treated with fusosomes and mice treated with PBS.

117. The fusosome of any of the preceding embodiments, wherein a measure of the immunogenicity of recipient cells is the PBMC response.

118. The fusosome of any of the preceding embodiments, wherein recipient cells do not elicit a PBMC response.

119. The fusosome of any of the preceding embodiments, wherein the percent of CD3+/CMG+ cells is similar for recipient cells derived from mice treated with fusosome and mice treated with PBS, e.g., as measured as described herein, e.g., as described in Example 17.

120. The fusosome of any of the preceding embodiments, wherein a measure of the immunogenicity of recipient cells is the natural killer cell response.

121. The fusosome of any of the preceding embodiments, wherein recipient cells do not elicit a natural killer cell response or elicit a lower natural killer cell response, e.g., lower than a reference value.

122. The fusosome of any of the preceding embodiments, wherein the percent of CD3+/CMG+ cells is similar for recipient cells derived from mice treated with fusosome and mice treated with PBS, e.g., as measured as described herein, e.g., as described in Example 18.

123. The fusosome of any of the preceding embodiments, wherein a measure of the immunogenicity of recipient cells is the CD8+ T cell response.

124. The fusosome of any of the preceding embodiments, wherein recipient cells do not elicit a CD8+ T cell response or elicit a lower CD8+ T cell response, e.g., lower than a reference value.

125. The fusosome of any of the preceding embodiments, wherein the percent of CD3+/CMG+ cells is similar for recipient cells derived from mice treated with fusosome and mice treated with PBS, e.g., as measured as described herein, e.g., as described in Example 19.

126. The fusosome of any of the preceding embodiments, wherein the fusogen is a re-targeted fusogen.

127. The fusosome of any of the preceding embodiments, which comprises a nucleic acid, e.g., retroviral nucleic acid, that encodes one or both of: (i) a positive target cell-specific

regulatory element operatively linked to a nucleic acid encoding an exogenous agent, or (ii) a non-target cell-specific regulatory element or negative TCSRE operatively linked to the nucleic acid encoding the exogenous agent.

128. A pharmaceutical composition comprising the fusosome of any of the preceding embodiments, and a pharmaceutically acceptable carrier, diluent, or excipient.

129. A method of delivering an exogenous agent to a subject (e.g., a human subject) comprising administering to the subject a fusosome of any of embodiments 1-127 or pharmaceutical composition of embodiment 128, thereby delivering the exogenous agent to the subject.

130. A method of modulating a function, in a subject (e.g., a human subject), target tissue or target cell (e.g., a CNS cell, e.g., a neuron or a glial cell), comprising contacting, e.g., administering to, the subject, the target tissue or the target cell a fusosome of any of embodiments 1-127, or the pharmaceutical compositon of embodiment 128.

131. The method of embodiment 130, wherein the target tissue or the target cell is present in a subject.

132. A method of treating a genetic deficiency in a subject (e.g., a human subject) comprising administering to the subject a fusosome of any of embodiments 1-127 or the pharmaceutical composition of claim 128.

133. The method of embodiment 132, wherein the genetic deficiency is a genetic deficiency of Table 5 or Table 6.

134. The method of embodiment 132 or 133, wherein the genetic deficiency is a genetic deficiency able to be treated by the payload gene encoding the exogenous agent.

135. The method of any of embodiments 132-134, wherien the genetic deficiency is associated with a CNS disease or disorder or a lysosomal disease or disorder, wherein the method treats the CNS disease or disorder or a lysosomal disease or disorder.

136. The method of embodiment 135, wherein the CNS disease or disorder or a lysosomal disease or disorder Spinocerebellar Ataxia; Autosomal Recessive, Type 1; Ataxia with Oculomotor Apraxia, Type 2; Fragile X Syndrome; Cerebral Creatine Deficiency Syndrome 1; Angelman Syndrome; Amyotrophic Lateral Sclerosis; Friedreich's Ataxia; Rett Syndrome; Canavan Disease; Pyridoxine-Dependent Epilepsy; Batten Disease, Fucosidosis; Krabbe Disease; Tay Sachs Disease; Sandhoff Disease; Beta-mannosidosis; Metachromatic Leukodystrophy; Mucolipidosis Type IIIa; Mucolipidosis Type IIIb; or Mucolipidosis Type IV.

137. A fusosome of any of embodiments 1-127 or pharmaceutical composition of embodiment 128 for use in treating a subject (e.g. a human subject) with a genetic deficiency.

138. Use of a fusosome of any of embodiments 1-127 or pharmaceutical composition of embodiment 128 for manufacture of a medicament for use in treating a subject (e.g. a human subject) with a genetic deficiency.

139. The fusosome or pharmaceutical composition for use of embodiment 137 or the use of embodiments 138, wherein the fusosome comprises a payload gene encoding an exogenous agent for treating the genetic deficiency.

140. The fusosome or pharmaceutical composition for use of embodimetn 137 or 139 or the use of embodiment 138 or 139, wherien the genetic deficiency is associated with a CNS disease or disorder or a lysosomal disease or disorder, wherein the method treats the CNS disease or disorder or a lysosomal disease or disorder.

141. The fusosome or pharmaceutical composition for use of embodimetn 137, 139 or 140, or the use of embodiment 138, 139 or 140, wherein the CNS disease or disorder or a lysosomal disease or disorder Spinocerebellar Ataxia; Autosomal Recessive, Type 1; Ataxia with

Oculomotor Apraxia, Type 2; Fragile X Syndrome; Cerebral Creatine Deficiency Syndrome 1; Angelman Syndrome; Amyotrophic Lateral Sclerosis; Friedreich's Ataxia; Rett Syndrome; Canavan Disease; Pyridoxine-Dependent Epilepsy; Batten Disease, Fucosidosis; Krabbe Disease; Tay Sachs Disease; Sandhoff Disease; Beta-mannosidosis; Metachromatic Leukodystrophy; Mucolipidosis Type IIIa; Mucolipidosis Type IIIb; or Mucolipidosis Type IV.

142. A method of making a fusosome of any of embodiments 1-127, comprising:

a) providing a cell that comprises the nucleic acid, e.g., retroviral nucleic acid, and the fusogen;

b) culturing the cell under conditions that allow for production of the fusosome, and

c) separating, enriching, or purifying the fusosome from the cell, thereby making the fusosome.

Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. For example, all GenBank, Unigene, and Entrez sequences referred to herein, e.g., in any Table herein, are incorporated by reference. Unless otherwise specified, the sequence accession numbers specified herein, including in any Table herein, refer to the database entries current as of May 15, 2018. When one gene or protein references a plurality of sequence accession numbers, all of the sequence variants are encompassed. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

BRIEF DESCRIPTION OF THE DRAWINGS

The following detailed description of the invention will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there are shown in the drawings described herein certain embodiments, which are presently

exemplified. It should be understood, however, that the invention is not limited to the precise arrangement and instrumentalities of the embodiments shown in the drawings.

FIG. 1 quantifies staining of fusosomes with a dye for F-actin.

FIG. 2 is a graph showing the capacity for fusosomes and parent cells to polymerase actin over a period of 3, 5, and 24 hours.

FIG. 3 is a table showing size distribution statistics of fusosomes and parental cells as measured by NTA and microscopy.

FIG. 4 is a table showing the average size and volume of fusosomes and parental cells.

FIG. 5 is a series of diagrams showing the soluble:insoluble ratio observed for fusosomes or a cell preparation.

FIG. 6 is a series of diagrams showing MvH(CD8)+F fusosome fusion to target or non-target cells and absolute amount of targeted fusion.

FIG. 7 is a diagram showing 2-NBDG mean fluorescence intensity in VSV-G fusosomes.

FIG. 8 is a diagram showing esterase activity in the cytosol of VSV-G fusosomes.

FIGS. 9A-9B are a series of diagrams showing Cre recombinase delivery by fusosomes as detected by biolumniscent imaging in mice. (A) Ventral image and luminescent signal overlay of exposed liver and spleen of IV fusosome treated mice (1x and 3x concentration). Lower portion is luminescent signal alone. (B) Total flux signal of fusosome targeted spleen and liver; y-scale is on log10 scale. Mice treated with a concentration of 3x fusosome treatment had a significantly greater signal in the spleen ($p=0.0004$) than background 72 hours post-treatment.

FIGS. 10A-10B are a series of diagrams showing Cre recombinase to murine liver and spleen by fusosomes as detected by bioluminescent imaging. (A) From left to right; dorsal image and luminescent signal overlay of excised liver, heart, lungs, kidney, small intestines, pancreas, and spleen collected and imaged within 5 minutes of euthanasia. Lower portion is luminescent signal alone. (B) Total flux signal of fusosome targeted spleen and liver and other tissues; y-scale is on log10 scale. Mice treated with a concentration of 3x fusosome treatment had a significantly greater signal in the spleen($p<0.0001$) as compared to the tissue with the lowest signal (heart).

FIG. 11 is a table showing delivery of Cre cargo by NivG+F fusosomes via a non-endocytic pathway.

FIG. 12 is a graph showing GAPDH: Total protein ratios measured by bicinchoninic acid assay in fusosomes and parental cells.

FIG. 13 is a graph showing lipid: protein ratios measured by bicinchoninic acid assay in fusosomes and parental cells.

FIG. 14 is a graph showing protein: DNA ratios measured by bicinchoninic acid assay in fusosomes and parental cells.

FIG. 15 is a graph showing lipids: DNA ratios measured by bicinchoninic acid assay in fusosomes and parental cells.

FIG. 16 is a graph showing protein levels of the exosome marker CD63 in exosomes and fusosomes.

FIG. 17 is a graph showing the intensity of calnexin signal detected in fusosomes and parental cells.

FIG. 18 is a graph showing lipid:DNA ratios determined for fusosomes and parental cells.

FIGS. 19A-19B are a series of graphs showing the proportion of lipid species as a percentage of total lipids in parental cells, exosomes, and fusosomes.

FIG. 20 is a series of graphs showing the protein content of parental cells, exosomes, and fusosomes with respect to proteins associated with specific compartments, as indicated.

FIG. 21 is a series of graphs showing the level of ARRDC1 (left panel) or TSG101 (right panel) as a percentage of total protein content in parental cells, exosomes, and fusosomes.

DETAILED DESCRIPTION

The present disclosure provides, at least in part, fusosome methods and compositions for in vivo delivery. In some embodiments, the fusosome comprises a combination of elements that promote specificity for target cells, e.g., one or more of a re-targeted fusogen, a positive target cell-specific regulatory element, and a non-target cell-specific regulatory element. In some embodiments, the fusosome comprises one or more modifications that decrease an immune response against the fusosome.

I. Definitions

Terms used in the claims and specification are defined as set forth below unless otherwise specified.

As used herein, "detectably present", when used in the context of an exogenous agent being detectably present, means that the exogenous agent itself is detectably present. For instance, if the exogenous agent is a protein, the exogenous protein agent can be detectably present regardless of whether a nucleic acid that encodes it is detectably present or not.

As used herein, "fusosome" refers to a bilayer of amphipathic lipids enclosing a lumen or cavity and a fusogen that interacts with the amphipathic lipid bilayer. In embodiments, the fusosome comprises a nucleic acid. In some embodiments, the fusosome is a membrane enclosed preparation. In some embodiments, the fusosome is derived from a source cell.

As used herein, "fusosome composition" refers to a composition comprising one or more fusosomes.

As used herein, "fusogen" refers to an agent or molecule that creates an interaction between two membrane enclosed lumens. In embodiments, the fusogen facilitates fusion of the membranes. In other embodiments, the fusogen creates a connection, e.g., a pore, between two lumens (e.g., a lumen of a retroviral vector and a cytoplasm of a target cell). In some embodiments, the fusogen comprises a complex of two or more proteins, e.g., wherein neither protein has fusogenic activity alone. In some embodiments, the fusogen comprises a targeting domain.

As used herein, an "insulator element" refers to a nucleotide sequence that blocks enhancers or prevents heterochromatin spreading. An insulator element can be wild-type or mutant.

The term "effective amount" as used herein means an amount of a pharmaceutical composition which is sufficient enough to significantly and positively modify the symptoms and/or conditions to be treated (e.g., provide a positive clinical response). The effective amount of an active ingredient for use in a pharmaceutical composition will vary with the particular condition being treated, the severity of the condition, the duration of treatment, the nature of concurrent therapy, the particular active ingredient(s) being employed, the particular pharmaceutically-acceptable excipient(s) and/or carrier(s) utilized, and like factors with the knowledge and expertise of the attending physician.

An "exogenous agent" as used herein with reference to a virus, VLP or fusosome, refers to an agent that is neither comprised by nor encoded in the corresponding wild-type virus or fusogen made from a corresponding wild-type source cell. In some embodiments, the exogenous agent does not naturally exist, such as a protein or nucleic acid that has a sequence that is altered (e.g., by insertion, deletion, or substitution) relative to a naturally occurring protein. In some embodiments, the exogenous agent does not naturally exist in the source cell. In some embodiments, the exogenous agent exists naturally in the source cell but is exogenous to the virus. In some embodiments, the exogenous agent does not naturally exist in the recipient cell. In some embodiments, the exogenous agent exists naturally in the recipient cell, but is not present at a desired level or at a desired time. In some embodiments, the exogenous agent comprises RNA or protein.

The term "pharmaceutically acceptable" as used herein, refers to excipients, compositions and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

As used herein, a "promoter" refers to a cis- regulatory DNA sequence that, when operably linked to a gene coding sequence, drives transcription of the gene. The promoter may comprise a transcription factor binding sites. In some embodiments, a promoter works in concert with one or more enhancers which are distal to the gene.

As used herein, a "positive target cell-specific regulatory element" (or positive TCSRE) refers to a nucleic acid sequence that increases the level of an exogenous agent in a target cell compared to in a non-target cell, wherein the nucleic acid encoding the exogenous agent is operably linked to the positive TCSRE. In some embodiments, the positive TCSRE is a functional nucleic acid sequence, e.g., the positive TCSRE can comprise a promoter or enhancer. In some embodiments, the positive TCSRE encodes a functional RNA sequence, e.g., the positive TCSRE can encode a splice site that promotes correct splicing of the RNA in the target cell. In some embodiments, the positive TCSRE encodes a functional protein sequence, or the positive TCSRE can encode a protein sequence that promotes correct post-translational modification of the protein. In some embodiments, the positive TCSRE decreases the level or activity of a downregulator or inhibitor of the exogenous agent.

As used herein, a "negative target cell-specific regulatory element" (or negative TCSRE) refers to a nucleic acid sequence that decreases the level of an exogenous agent in a non-target cell compared to in a target cell, wherein the nucleic acid encoding the exogenous agent is operably linked to the negative TCSRE. In some embodiments, the negative TCSRE is a functional nucleic acid sequence, e.g., a miRNA recognition site that causes degradation or inhibition of the retroviral nucleic acid in a non-target cell. In some embodiments, the nucleic acid sequence encodes a functional RNA sequence, e.g., the nucleic acid encodes an miRNA sequence present in an mRNA encoding an exogenous protein agent, such that the mRNA is degraded or inhibited in a non-target cell. In some embodiments, the negative TCSRE increases the level or activity of a downregulator or inhibitor of the exogenous agent.

As used herein, a "non-target cell-specific regulatory element" (or NTCSRE) refers to a nucleic acid sequence that decreases the level of an exogenous agent in a non-target cell compared to in a target cell, wherein the nucleic acid encoding the exogenous agent is operably linked to the NTCSRE. In some embodiments, the NTCSRE is a functional nucleic acid sequence, e.g., a miRNA recognition site that causes degradation or inhibition of the retroviral nucleic acid in a non-target cell. In some embodiments, the nucleic acid sequence encodes a functional RNA sequence, e.g., the nucleic acid encodes an miRNA sequence present in an mRNA encoding an exogenous protein agent, such that the mRNA is degraded or inhibited in a non-target cell. In some embodiments, the NTCSRE increases the level or activity of a downregulator or inhibitor of the exogenous agent. The terms "negative TCSRE" and "NTCSRE" are used interchangeably herein.

As used herein, a "non-CNS cell specific regulatory element" refers to a non-target cell-specific regulatory element (NTCSRE), wherein the target cell is a CNS cell. Thus, a non-CNS cell specific regulatory element refers to a nucleic acid sequence that decreases the level of an exogenous agent in a non-CNS cell compared to in a CNS cell, wherein the nucleic acid encoding the exogenous agent is operably linked to the non-CNS cell-specific regulatory element.

As used herein, a "re-targeted fusogen" refers to a fusogen that comprises a targeting moiety having a sequence that is not part of the naturally-occurring form of the fusogen. In embodiments, the fusogen comprises a different targeting moiety relative to the targeting moiety in the naturally-occurring form of the fusogen. In embodiments, the naturally-occurring form of

the fusogen lacks a targeting domain, and the re-targeted fusogen comprises a targeting moiety that is absent from the naturally-occurring form of the fusogen. In embodiments, the fusogen is modified to comprise a targeting moiety. In embodiments, the fusogen comprises one or more sequence alterations outside of the targeting moiety relative to the naturally-occurring form of the fusogen, e.g., in a transmembrane domain, fusogenically active domain, or cytoplasmic domain.

As used herein, a "retroviral nucleic acid" refers to a nucleic acid containing at least the minimal sequence requirements for packaging into a retrovirus or retroviral vector, alone or in combination with a helper cell, helper virus, or helper plasmid. In some embodiments, the retroviral nucleic acid further comprises or encodes an exogenous agent, a positive target cell-specific regulatory element, a non-target cell-specific regulatory element, or a negative TCSRE. In some embodiments, the retroviral nucleic acid comprises one or more of (e.g., all of) a 5' LTR (e.g., to promote integration), U3 (e.g., to activate viral genomic RNA transcription), R (e.g., a Tat-binding region), U5, a 3' LTR (e.g., to promote integration), a packaging site (e.g., psi (Ψ)), RRE (e.g., to bind to Rev and promote nuclear export). The retroviral nucleic acid can comprise RNA (e.g., when part of a virion) or DNA (e.g., when being introduced into a source cell or after reverse transcription in a recipient cell). In some embodiments, the retroviral nucleic acid is packaged using a helper cell, helper virus, or helper plasmid which comprises one or more of (e.g., all of) gag, pol, and env.

As used herein, a "target cell" refers to a cell of a type to which it is desired that a fusosome (e.g., lentiviral vector) deliver an exogenous agent. In embodiments, a target cell is a cell of a specific tissue type or class, e.g., a CNS cell, e.g., a neuron or a glial cell. In some embodiments, a target cell is a diseased cell, e.g., a cancer cell. In some embodiments, the fusogen, e.g., re-targeted fusogen (alone or in combination with the positive TCSRE, NTCSRE, negative TCSRE, or any combination thereof) leads to preferential delivery of the exogenous agent to a target cell compared to a non-target cell.

As used herein a "non-target cell" refers to a cell of a type to which it is not desired that a lentiviral vector delivers an exogenous agent. In some embodiments, a non-target cell is a cell of a specific tissue type or class. In some embodiments, a non-target cell is a non-diseased cell, e.g., a non-cancerous cell. In some embodiments, the fusogen, e.g., re-targeted fusogen (alone or in combination with the positive TCSRE, NTCSRE, negative TCSRE or any combination

thereof) leads to lower delivery of the exogenous agent to a non-target cell compared to a target cell.

As used herein, the terms "treat," "treating," or "treatment" refer to ameliorating a disease or disorder, e.g., slowing or arresting or reducing the development of the disease or disorder, e.g., a root cause of the disorder or at least one of the clinical symptoms thereof.

As used herein, "cytobiologic" refers to a portion of a cell that comprises a lumen and a cell membrane, or a cell having partial or complete nuclear inactivation. In some embodiments, the cytobiologic comprises one or more of a cytoskeleton component, an organelle, and a ribosome. In embodiments, the cytobiologic is an enucleated cell, a microvesicle, or a cell ghost.

II. Fusosomes, e.g. , cell-derived fusosomes

Fusosomes can take various forms. For example, in some embodiments, a fusosome described herein is derived from a source cell. A fusosome may be or comprise, e.g., an extracellular vesicle, a microvesicle, a nanovesicle, an exosome, an apoptotic body (from apoptotic cells), a microparticle (which may be derived from, e.g., platelets), an ectosome (derivable from, e.g., neutrophiles and monocytes in serum), a prostatosome (obtainable from prostate cancer cells), a cardiosome (derivable from cardiac cells), or any combination thereof. In some embodiments, a fusosome is released naturally from a source cell, and in some embodiments, the source cell is treated to enhance formation of fusosomes. In some embodiments, the fusosome is between about 10-10,000 nm in diameter, e.g., about 30-100 nm in diameter. In some embodiments, the fusosome comprises one or more synthetic lipids.

In some embodiments, the fusosome is or comprises a virus, e.g., a retrovirus, e.g., a lentivirus. For instance, in some embodiments, the fusosome's bilayer of amphipathic lipids is or comprises the viral envelope. The viral envelope may comprise a fusogen, e.g., a fusogen that is endogenous to the virus or a pseudotyped fusogen. In some embodiments, the fusosome's lumen or cavity comprises a viral nucleic acid, e.g., a retroviral nucleic acid, e.g., a lentiviral nucleic acid. The viral nucleic acid may be a viral genome. In some embodiments, the fusosome further comprises one or more viral non-structural proteins, e.g., in its cavity or lumen.

Fusosomes may have various properties that facilitate delivery of a payload to a target cell. For instance, in some embodiments, the fusosome and the source cell together comprise nucleic acid(s) sufficient to make a particle that can fuse with a target cell. In embodiments,

these nucleic acid(s) encode proteins having one or more of (e.g., all of) the following activities: gag polyprotein activity, polymerase activity, integrase activity, protease activity, and fusogen activity.

Fusosomes may also comprise various structures that facilitate delivery of a payload to a target cell. For instance, in some embodiments, the fusosome and the source cell together comprise nucleic acid(s) sufficient to make a particle that can fuse with a target cell. In embodiments, these nucleic acid(s) encode proteins having one or more of (e.g., all of) the following activities: gag polyprotein activity, polymerase activity, integrase activity, protease activity, and fusogen activity.

Fusosomes may also comprise various structures that facilitate delivery of a payload to a target cell. For instance, in some embodiments, the fusosome (e.g., virus, e.g., retrovirus, e.g., lentivirus) comprises one or more of (e.g., all of) the following proteins: gag polyprotein, polymerase (e.g., pol), integrase (e.g., a functional or non-functional variant), protease, and a fusogen. In some embodiments, the fusosome further comprises rev. In some embodiments, one or more of the aforesaid proteins are encoded in the retroviral genome, and in some embodiments, one or more of the aforesaid proteins are provided in trans, e.g., by a helper cell, helper virus, or helper plasmid. In some embodiments, the fusosome nucleic acid (e.g., retroviral nucleic acid) comprises one or more of (e.g., all of) the following nucleic acid sequences: 5' LTR (e.g., comprising U5 and lacking a functional U3 domain), Psi packaging element (Psi), Central polypurine tract (cPPT) Promoter operatively linked to the payload gene, payload gene (optionally comprising an intron before the open reading frame), Poly A tail sequence, WPRE, and 3' LTR (e.g., comprising U5 and lacking a functional U3). In some embodiments the fusosome nucleic acid (e.g., retroviral nucleic acid) further comprises one or more insulator element. In some embodiments the fusosome nucleic acid (e.g., retroviral nucleic acid) further comprises one or more miRNA recognition sites. In some embodiments, one or more of the miRNA recognition sites are situated downstream of the poly A tail sequence, e.g., between the poly A tail sequence and the WPRE.

In some embodiments, a fusosome provided herein is administered to a subject, e.g., a mammal, e.g., a human. In such embodiments, the subject may be at risk of, may have a symptom of, or may be diagnosed with or identified as having, a particular disease or condition (e.g., a disease or condition described herein). In one embodiment, the subject has a genetic

deficiency, such as any listed in Table 5 or Table 6. In some embodiments, the fusosome contains nucleic acid sequences encoding an exogenous agent for treating the disease or condition, such as for treating the genetic deficiency.

A. Fusosomes generated from viruses.

For instance, in some embodiments, the fusosome (e.g., virus, e.g., retrovirus, e.g., lentivirus) comprises one or more of (e.g., all of) the following proteins: gag polyprotein, polymerase (e.g., pol), integrase (e.g., a functional or non-functional variant), protease, and a fusogen. In some embodiments, the fusosome further comprises rev. In some embodiments, one or more of the aforesaid proteins are encoded in the retroviral genome, and in some embodiments, one or more of the aforesaid proteins are provided in trans, e.g., by a helper cell, helper virus, or helper plasmid. In some embodiments, the fusosome nucleic acid (e.g., retroviral nucleic acid) comprises one or more of (e.g., all of) the following nucleic acid sequences: 5' LTR (e.g., comprising U5 and lacking a functional U3 domain), Psi packaging element (Psi), Central polypurine tract (cPPT) Promoter operatively linked to the payload gene, payload gene (optionally comprising an intron before the open reading frame), Poly A tail sequence, WPRE, and 3' LTR (e.g., comprising U5 and lacking a functional U3). In some embodiments the fusosome nucleic acid (e.g., retroviral nucleic acid) further comprises one or more insulator element. In some embodiments the fusosome nucleic acid (e.g., retroviral nucleic acid) further comprises one or more miRNA recognition sites. In some embodiments, one or more of the miRNA recognition sites are situated downstream of the poly A tail sequence, e.g., between the poly A tail sequence and the WPRE.

i) Lentiviral components and helper cells

In some embodiments, the retroviral nucleic acid comprises one or more of (e.g., all of): a 5' promoter (e.g., to control expression of the entire packaged RNA), a 5' LTR (e.g., that includes R (polyadenylation tail signal) and/or U5 which includes a primer activation signal), a primer binding site, a psi packaging signal, a RRE element for nuclear export, a promoter directly upstream of the transgene to control transgene expression, a transgene (or other exogenous agent element), a polypurine tract, and a 3' LTR (e.g., that includes a mutated U3, a R, and U5). In some embodiments, the retroviral nucleic acid further comprises one or more of a cPPT, a WPRE, and/or an insulator element.

A retrovirus typically replicates by reverse transcription of its genomic RNA into a linear double-stranded DNA copy and subsequently covalently integrates its genomic DNA into a host genome. Illustrative retroviruses suitable for use in particular embodiments, include, but are not limited to: Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), feline leukemia virus (FLV), spumavirus, Friend murine leukemia virus, Murine Stem Cell Virus (MSCV) and Rous Sarcoma Virus (RSV)) and lentivirus.

In some embodiments the retrovirus is a Gammretrovirus. In some embodiments the retrovirus is an Epsilonretrovirus. In some embodiments the retrovirus is an Alpharetrovirus. In some embodiments the retrovirus is a Betaretrovirus. In some embodiments the retrovirus is a Deltaretrovirus. In some embodiments the retrovirus is a Lentivirus. In some embodiments the retrovirus is a Spumaretrovirus. In some embodiments the retrovirus is an endogenous retrovirus.

Illustrative lentiviruses include, but are not limited to: HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2); visna-maedi virus (VMV) virus; the caprine arthritis-encephalitis virus (CAEV); equine infectious anemia virus (EIAV); feline immunodeficiency virus (FIV); bovine immune deficiency virus (BIV); and simian immunodeficiency virus (SIV). In some embodiments, HIV based vector backbones (i.e., HIV cis-acting sequence elements) are used.

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

A viral vector can comprise, e.g., a nucleic acid molecule (e.g., a transfer plasmid) that includes virus-derived nucleic acid elements that typically facilitate transfer of the nucleic acid molecule or integration into the genome of a cell or to a viral particle that mediates nucleic acid transfer. Viral particles will typically include various viral components and sometimes also host cell components in addition to nucleic acid(s). A viral vector can comprise, e.g., a virus or viral

particle capable of transferring a nucleic acid into a cell, or to the transferred nucleic acid (e.g., as naked DNA). Viral vectors and transfer plasmids can comprise structural and/or functional genetic elements that are primarily derived from a virus. A retroviral vector can comprise a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from a retrovirus. A lentiviral vector can comprise a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, including LTRs that are primarily derived from a lentivirus.

In embodiments, a lentiviral vector (e.g., lentiviral expression vector) may comprise a lentiviral transfer plasmid (e.g., as naked DNA) or an infectious lentiviral particle. With respect to elements such as cloning sites, promoters, regulatory elements, heterologous nucleic acids, etc., it is to be understood that the sequences of these elements can be present in RNA form in lentiviral particles and can be present in DNA form in DNA plasmids.

In some vectors described herein, at least part of one or more protein coding regions that contribute to or are essential for replication may be absent compared to the corresponding wild-type virus. This makes the viral vector replication-defective. In some embodiments, the vector is capable of transducing a target non-dividing host cell and/or integrating its genome into a host genome.

The structure of a wild-type retrovirus genome often comprises a 5' long terminal repeat (LTR) and a 3' LTR, between or within which are located a packaging signal to enable the genome to be packaged, a primer binding site, integration sites to enable integration into a host cell genome and gag, pol and env genes encoding the packaging components which promote the assembly of viral particles. More complex retroviruses have additional features, such as rev and RRE sequences in HIV, which enable the efficient export of RNA transcripts of the integrated provirus from the nucleus to the cytoplasm of an infected target cell. In the provirus, the viral genes are flanked at both ends by regions called long terminal repeats (LTRs). The LTRs are involved in proviral integration and transcription. LTRs also serve as enhancer-promoter sequences and can control the expression of the viral genes. Encapsidation of the retroviral RNAs occurs by virtue of a psi sequence located at the 5' end of the viral genome.

The LTRs themselves are typically similar (e.g., identical) sequences that can be divided into three elements, which are called U3, R and U5. U3 is derived from the sequence unique to the 3' end of the RNA. R is derived from a sequence repeated at both ends of the RNA and U5 is

derived from the sequence unique to the 5' end of the RNA. The sizes of the three elements can vary considerably among different retroviruses.

For the viral genome, the site of transcription initiation is typically at the boundary between U3 and R in one LTR and the site of poly (A) addition (termination) is at the boundary between R and U5 in the other LTR. U3 contains most of the transcriptional control elements of the provirus, which include the promoter and multiple enhancer sequences responsive to cellular and in some cases, viral transcriptional activator proteins. Some retroviruses comprise any one or more of the following genes that code for proteins that are involved in the regulation of gene expression: tot, rev, tax and rex.

With regard to the structural genes gag, pol and env themselves, gag encodes the internal structural protein of the virus. Gag protein is proteolytically processed into the mature proteins MA (matrix), CA (capsid) and NC (nucleocapsid). The pol gene encodes the reverse transcriptase (RT), which contains DNA polymerase, associated RNase H and integrase (IN), which mediate replication of the genome. The env gene encodes the surface (SU) glycoprotein and the transmembrane (TM) protein of the virion, which form a complex that interacts specifically with cellular receptor proteins. This interaction promotes infection, e.g., by fusion of the viral membrane with the cell membrane.

In a replication-defective retroviral vector genome gag, pol and env may be absent or not functional. The R regions at both ends of the RNA are typically repeated sequences. U5 and U3 represent unique sequences at the 5' and 3' ends of the RNA genome respectively.

Retroviruses may also contain additional genes which code for proteins other than gag, pol and env. Examples of additional genes include (in HIV), one or more of vif, vpr, vpx, vpu, tat, rev and nef. EIAV has (amongst others) the additional gene S2. Proteins encoded by additional genes serve various functions, some of which may be duplicative of a function provided by a cellular protein. In EIAV, for example, tat acts as a transcriptional activator of the viral LTR (Derse and Newbold 1993 Virology 194:530-6; Maury et al. 1994 Virology 200:632-42). It binds to a stable, stem-loop RNA secondary structure referred to as TAR. Rev regulates and co-ordinates the expression of viral genes through rev-response elements (RRE) (Martarano et al. 1994 J. Virol. 68:3102-11). The mechanisms of action of these two proteins are thought to be broadly similar to the analogous mechanisms in the primate viruses. In addition, an EIAV

protein, Ttm, has been identified that is encoded by the first exon of tat spliced to the env coding sequence at the start of the transmembrane protein.

In addition to protease, reverse transcriptase and integrase, non-primate lentiviruses contain a fourth pol gene product which codes for a dUTPase. This may play a role in the ability of these lentiviruses to infect certain non-dividing or slowly dividing cell types.

In embodiments, a recombinant lentiviral vector (RLV) is a vector with sufficient retroviral genetic information to allow packaging of an RNA genome, in the presence of packaging components, into a viral particle capable of infecting a target cell. Infection of the target cell can comprise reverse transcription and integration into the target cell genome. The RLV typically carries non-viral coding sequences which are to be delivered by the vector to the target cell. In embodiments, an RLV is incapable of independent replication to produce infectious retroviral particles within the target cell. Usually the RLV lacks a functional gag-pol and/or env gene and/or other genes involved in replication. The vector may be configured as a split-intron vector, e.g., as described in PCT patent application WO 99/15683, which is herein incorporated by reference in its entirety.

In some embodiments, the lentiviral vector comprises a minimal viral genome, e.g., the viral vector has been manipulated so as to remove the non-essential elements and to retain the essential elements in order to provide the required functionality to infect, transduce and deliver a nucleotide sequence of interest to a target host cell, e.g., as described in WO 98/17815, which is herein incorporated by reference in its entirety.

A minimal lentiviral genome may comprise, e.g., (5')R-U5-one or more first nucleotide sequences-U3-R(3'). However, the plasmid vector used to produce the lentiviral genome within a source cell can also include transcriptional regulatory control sequences operably linked to the lentiviral genome to direct transcription of the genome in a source cell. These regulatory sequences may comprise the natural sequences associated with the transcribed retroviral sequence, e.g., the 5' U3 region, or they may comprise a heterologous promoter such as another viral promoter, for example the CMV promoter. Some lentiviral genomes comprise additional sequences to promote efficient virus production. For example, in the case of HIV, rev and RRE sequences may be included. Alternatively or combination, codon optimization may be used, e.g., the gene encoding the exogenous agent may be codon optimized, e.g., as described in WO 01/79518, which is herein incorporated by reference in its entirety. Alternative sequences which

perform a similar or the same function as the rev/RRE system may also be used. For example, a functional analogue of the rev/RRE system is found in the Mason Pfizer monkey virus. This is known as CTE and comprises an RRE-type sequence in the genome which is believed to interact with a factor in the infected cell. The cellular factor can be thought of as a rev analogue. Thus, CTE may be used as an alternative to the rev/RRE system. In addition, the Rex protein of HTLV-I can functionally replace the Rev protein of HIV-I . Rev and Rex have similar effects to IRE-BP.

In some embodiments, a retroviral nucleic acid (e.g., a lentiviral nucleic acid, e.g., a primate or non-primate lentiviral nucleic acid) (1) comprises a deleted gag gene wherein the deletion in gag removes one or more nucleotides downstream of about nucleotide 350 or 354 of the gag coding sequence; (2) has one or more accessory genes absent from the retroviral nucleic acid; (3) lacks the tat gene but includes the leader sequence between the end of the 5' LTR and the ATG of gag; and (4) combinations of (1), (2) and (3). In an embodiment the lentiviral vector comprises all of features (1) and (2) and (3). This strategy is described in more detail in WO 99/32646, which is herein incorporated by reference in its entirety.

In some embodiments, a primate lentivirus minimal system requires none of the HIV/SIV additional genes vif, vpr, vpx, vpu, tat, rev and nef for either vector production or for transduction of dividing and non-dividing cells. In some embodiments, an EIAV minimal vector system does not require S2 for either vector production or for transduction of dividing and non-dividing cells.

The deletion of additional genes may permit vectors to be produced without the genes associated with disease in lentiviral (e.g. HIV) infections. In particular, tat is associated with disease. Secondly, the deletion of additional genes permits the vector to package more heterologous DNA. Thirdly, genes whose function is unknown, such as S2, may be omitted, thus reducing the risk of causing undesired effects. Examples of minimal lentiviral vectors are disclosed in WO 99/32646 and in WO 98/17815.

In some embodiments, the retroviral nucleic acid is devoid of at least tat and S2 (if it is an EIAV vector system), and possibly also vif, vpr, vpx, vpu and nef. In some embodiments, the retroviral nucleic acid is also devoid of rev, RRE, or both.

In some embodiments the retroviral nucleic acid comprises vpx. The Vpx polypeptide binds to and induces the degradation of the SAMHD1 restriction factor, which degrades free

dNTPs in the cytoplasm. Thus, the concentration of free dNTPs in the cytoplasm increases as Vpx degrades SAMHD1 and reverse transcription activity is increased, thus facilitating reverse transcription of the retroviral genome and integration into the target cell genome.

Different cells differ in their usage of particular codons. This codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. By altering the codons in the sequence so that they are tailored to match with the relative abundance of corresponding tRNAs, it is possible to increase expression. By the same token, it is possible to decrease expression by deliberately choosing codons for which the corresponding tRNAs are known to be rare in the particular cell type. Thus, an additional degree of translational control is available. An additional description of codon optimization is found, e.g., in WO 99/41397, which is herein incorporated by reference in its entirety.

Many viruses, including HIV and other lentiviruses, use a large number of rare codons and by changing these to correspond to commonly used mammalian codons, increased expression of the packaging components in mammalian producer cells can be achieved.

Codon optimization has a number of other advantages. By virtue of alterations in their sequences, the nucleotide sequences encoding the packaging components may have RNA instability sequences (INS) reduced or eliminated from them. At the same time, the amino acid sequence coding sequence for the packaging components is retained so that the viral components encoded by the sequences remain the same, or at least sufficiently similar that the function of the packaging components is not compromised. In some embodiments, codon optimization also overcomes the Rev/RRE requirement for export, rendering optimized sequences Rev independent. In some embodiments, codon optimization also reduces homologous recombination between different constructs within the vector system (for example between the regions of overlap in the gag-pol and env open reading frames). In some embodiments, codon optimization leads to an increase in viral titer and/or improved safety.

In some embodiments, only codons relating to INS are codon optimized. In other embodiments, the sequences are codon optimized in their entirety, with the exception of the sequence encompassing the frameshift site of gag-pol.

The gag-pol gene comprises two overlapping reading frames encoding the gag-pol proteins. The expression of both proteins depends on a frameshift during translation. This frameshift occurs as a result of ribosome "slippage" during translation. This slippage is thought

to be caused at least in part by ribosome-stalling RNA secondary structures. Such secondary structures exist downstream of the frameshift site in the gag-pol gene. For HIV, the region of overlap extends from nucleotide 1222 downstream of the beginning of gag (wherein nucleotide 1 is the A of the gag ATG) to the end of gag (nt 1503). Consequently, a 281 bp fragment spanning the frameshift site and the overlapping region of the two reading frames is preferably not codon optimized. In some embodiments, retaining this fragment will enable more efficient expression of the gag-pol proteins. For EIAV, the beginning of the overlap is at nt 1262 (where nucleotide 1 is the A of the gag ATG). The end of the overlap is at nt 1461. In order to ensure that the frameshift site and the gag-pol overlap are preserved, the wild type sequence may be retained from nt 1156 to 1465.

Derivations from optimal codon usage may be made, for example, in order to accommodate convenient restriction sites, and conservative amino acid changes may be introduced into the gag-pol proteins.

In some embodiments, codon optimization is based on codons with poor codon usage in mammalian systems. The third and sometimes the second and third base may be changed.

Due to the degenerate nature of the genetic code, it will be appreciated that numerous gag-pol sequences can be achieved by a skilled worker. Also, there are many retroviral variants described which can be used as a starting point for generating a codon optimized gag-pol sequence. Lentiviral genomes can be quite variable. For example there are many quasi-species of HIV-I which are still functional. This is also the case for EIAV. These variants may be used to enhance particular parts of the transduction process. Examples of HIV-I variants may be found in the HIV databases maintained by Los Alamos National Laboratory. Details of EIAV clones may be found at the NCBI database maintained by the National Institutes of Health.

The strategy for codon optimized gag-pol sequences can be used in relation to any retrovirus, e.g., EIAV, FIV, BIV, CAEV, VMR, SIV, HIV-I and HIV -2. In addition this method could be used to increase expression of genes from HTLV-I, HTLV-2, HFV, HSRV and human endogenous retroviruses (HERV), MLV and other retroviruses.

As described above, the packaging components for a retroviral vector can include expression products of gag, pol and env genes. In addition, packaging can utilize a short sequence of 4 stem loops followed by a partial sequence from gag and env as a packaging signal. Thus, inclusion of a deleted gag sequence in the retroviral vector genome (in addition to the full

gag sequence on the packaging construct) can be used. In embodiments, the retroviral vector comprises a packaging signal that comprises from 255 to 360 nucleotides of gag in vectors that still retain env sequences, or about 40 nucleotides of gag in a particular combination of splice donor mutation, gag and env deletions. In some embodiments, the retroviral vector includes a gag sequence which comprises one or more deletions, e.g., the gag sequence comprises about 360 nucleotides derivable from the N-terminus.

The retroviral vector, helper cell, helper virus, or helper plasmid may comprise retroviral structural and accessory proteins, for example gag, pol, env, tat, rev, vif, vpr, vpu, vpx, or nef proteins or other retroviral proteins. In some embodiments the retroviral proteins are derived from the same retrovirus. In some embodiments the retroviral proteins are derived from more than one retrovirus, e.g. 2, 3, 4, or more retroviruses.

The gag and pol coding sequences are generally organized as the Gag-Pol Precursor in native lentivirus. The gag sequence codes for a 55-kD Gag precursor protein, also called p55. The p55 is cleaved by the virally encoded protease4 (a product of the pol gene) during the process of maturation into four smaller proteins designated MA (matrix [p17]), CA (capsid [p24]), NC (nucleocapsid [p9]), and p6. The pol precursor protein is cleaved away from Gag by a virally encoded protease, and further digested to separate the protease (p10), RT (p50), RNase H (p15), and integrase (p31) activities.

Native Gag-Pol sequences can be utilized in a helper vector (e.g., helper plasmid or helper virus), or modifications can be made. These modifications include, chimeric Gag-Pol, where the Gag and Pol sequences are obtained from different viruses (e.g., different species, subspecies, strains, clades, etc.), and/or where the sequences have been modified to improve transcription and/or translation, and/or reduce recombination.

In various examples, the retroviral nucleic acid includes a polynucleotide encoding a 150-250 (e.g., 168) nucleotide portion of a gag protein that (i) includes a mutated INS1 inhibitory sequence that reduces restriction of nuclear export of RNA relative to wild-type INS1, (ii) contains two nucleotide insertion that results in frame shift and premature termination, and/or (iii) does not include INS2, INS3, and INS4 inhibitory sequences of gag.

In some embodiments, a vector described herein is a hybrid vector that comprises both retroviral (e.g., lentiviral) sequences and non-lentiviral viral sequences. In some embodiments, a

hybrid vector comprises retroviral e.g., lentiviral, sequences for reverse transcription, replication, integration and/or packaging.

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

At each end of the provirus, long terminal repeats (LTRs) are typically found. An LTR typically comprises a domain located at the ends of retroviral nucleic acid which, in their natural sequence context, are direct repeats and contain U3, R and U5 regions. LTRs generally promote the expression of retroviral genes (e.g., promotion, initiation and polyadenylation of gene transcripts) and viral replication. The LTR can comprise numerous regulatory signals including transcriptional control elements, polyadenylation signals and sequences for replication and integration of the viral genome. The viral LTR is typically divided into three regions called U3, R and U5. The U3 region typically contains the enhancer and promoter elements. The U5 region is typically the sequence between the primer binding site and the R region and can contain the polyadenylation sequence. The R (repeat) region can be flanked by the U3 and U5 regions. The LTR is typically composed of U3, R and U5 regions and can appear at both the 5′ and 3′ ends of the viral genome. In some embodiments, adjacent to the 5′ LTR are sequences for reverse transcription of the genome (the tRNA primer binding site) and for efficient packaging of viral RNA into particles (the Psi site).

A packaging signal can comprise a sequence located within the retroviral genome which mediate insertion of the viral RNA into the viral capsid or particle, see e.g., Clever et al., 1995. J. of Virology, Vol. 69, No. 4; pp. 2101-2109. Several retroviral vectors use a minimal packaging signal (a psi [Ψ] sequence) for encapsidation of the viral genome.

In various embodiments, retroviral nucleic acids comprise modified 5′ LTR and/or 3′ LTRs. Either or both of the LTR may comprise one or more modifications including, but not limited to, one or more deletions, insertions, or substitutions. Modifications of the 3′ LTR are

often made to improve the safety of lentiviral or retroviral systems by rendering viruses replication-defective, e.g., virus that is not capable of complete, effective replication such that infective virions are not produced (e.g., replication-defective lentiviral progeny).

In some embodiments, a vector is a self-inactivating (SIN) vector, e.g., replication-defective vector, e.g., retroviral or lentiviral vector, in which the right (3′) LTR enhancer-promoter region, known as the U3 region, has been modified (e.g., by deletion or substitution) to prevent viral transcription beyond the first round of viral replication. This is because the right (3′) LTR U3 region can be used as a template for the left (5′) LTR U3 region during viral replication and, thus, absence of the U3 enhancer-promoter inhibits viral replication. In embodiments, the 3′ LTR is modified such that the U5 region is removed, altered, or replaced, for example, with an exogenous poly(A) sequence The 3′ LTR, the 5′ LTR, or both 3′ and 5′ LTRs, may be modified LTRs.

In some embodiments, the U3 region of the 5′ LTR is replaced with a heterologous promoter to drive transcription of the viral genome during production of viral particles. Examples of heterologous promoters which can be used include, for example, viral simian virus 40 (SV40) (e.g., early or late), cytomegalovirus (CMV) (e.g., immediate early), Moloney murine leukemia virus (MoMLV), Rous sarcoma virus (RSV), and herpes simplex virus (HSV) (thymidine kinase) promoters. In some embodiments, promoters are able to drive high levels of transcription in a Tat-independent manner. In certain embodiments, the heterologous promoter has additional advantages in controlling the manner in which the viral genome is transcribed. For example, the heterologous promoter can be inducible, such that transcription of all or part of the viral genome will occur only when the induction factors are present. Induction factors include, but are not limited to, one or more chemical compounds or the physiological conditions such as temperature or pH, in which the host cells are cultured.

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

The R region, e.g., the region within retroviral LTRs beginning at the start of the capping group (i.e., the start of transcription) and ending immediately prior to the start of the poly A tract

can be flanked by the U3 and U5 regions. The R region plays a role during reverse transcription in the transfer of nascent DNA from one end of the genome to the other.

The retroviral nucleic acid can also comprise a FLAP element, e.g., a nucleic acid whose sequence includes the central polypurine tract and central termination sequences (cPPT and CTS) of a retrovirus, e.g., HIV-1 or HIV-2. Suitable FLAP elements are described in U.S. Pat. No. 6,682,907 and in Zennou, et al., 2000, Cell, 101:173, which are herein incorporated by reference in their entireties. During HIV-1 reverse transcription, central initiation of the plus-strand DNA at the central polypurine tract (cPPT) and central termination at the central termination sequence (CTS) can lead to the formation of a three-stranded DNA structure: the HIV-1 central DNA flap. In some embodiments, the retroviral or lentiviral vector backbones comprise one or more FLAP elements upstream or downstream of the gene encoding the exogenous agent. For example, in some embodiments a transfer plasmid includes a FLAP element, e.g., a FLAP element derived or isolated from HIV-1.

In embodiments, a retroviral or lentiviral nucleic acid comprises one or more export elements, e.g., a cis-acting post-transcriptional regulatory element which regulates the transport of an RNA transcript from the nucleus to the cytoplasm of a cell. Examples of RNA export elements include, but are not limited to, the human immunodeficiency virus (HIV) rev response element (RRE) (see e.g., Cullen et al., 1991. J. Virol. 65: 1053; and Cullen et al., 1991. Cell 58: 423), and the hepatitis B virus post-transcriptional regulatory element (HPRE), which are herein incorporated by reference in their entireties. Generally, the RNA export element is placed within the 3′ UTR of a gene, and can be inserted as one or multiple copies.

In some embodiments, expression of heterologous sequences in viral vectors is increased by incorporating one or more of, e.g., all of, posttranscriptional regulatory elements, polyadenylation sites, and transcription termination signals into the vectors. A variety of posttranscriptional regulatory elements can increase expression of a heterologous nucleic acid at the protein, e.g., woodchuck hepatitis virus posttranscriptional regulatory element (WPRE; Zufferey et al., 1999, J. Virol., 73:2886); the posttranscriptional regulatory element present in hepatitis B virus (HPRE) (Huang et al., Mol. Cell. Biol., 5:3864); and the like (Liu et al., 1995, Genes Dev., 9:1766), each of which is herein incorporated by reference in its entirety. In some embodiments, a retroviral nucleic acid described herein comprises a posttranscriptional regulatory element such as a WPRE or HPRE

In some embodiments, a retroviral nucleic acid described herein lacks or does not comprise a posttranscriptional regulatory element such as a WPRE or HPRE.

Elements directing the termination and polyadenylation of the heterologous nucleic acid transcripts may be included, e.g., to increases expression of the exogenous agent. Transcription termination signals may be found downstream of the polyadenylation signal. In some embodiments, vectors comprise a polyadenylation sequence 3′ of a polynucleotide encoding the exogenous agent. A polyA site may comprise a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript by RNA polymerase II. Polyadenylation sequences can promote mRNA stability by addition of a polyA tail to the 3′ end of the coding sequence and thus, contribute to increased translational efficiency. Illustrative examples of polyA signals that can be used in a retroviral nucleic acid, include AATAAA, ATTAAA, AGTAAA, a bovine growth hormone polyA sequence (BGHpA), a rabbit β-globin polyA sequence (rβgpA), or another suitable heterologous or endogenous polyA sequence.

In some embodiments, a retroviral or lentiviral vector further comprises one or more insulator elements, e.g., an insulator element described herein.

In various embodiments, the vectors comprise a promoter operably linked to a polynucleotide encoding an exogenous agent. The vectors may have one or more LTRs, wherein either LTR comprises one or more modifications, such as one or more nucleotide substitutions, additions, or deletions. The vectors may further comprise one of more accessory elements to increase transduction efficiency (e.g., a cPPT/FLAP), viral packaging (e.g., a Psi (Ψ) packaging signal, RRE), and/or other elements that increase exogenous gene expression (e.g., poly (A) sequences), and may optionally comprise a WPRE or HPRE.

In some embodiments, a lentiviral nucleic acid comprises one or more of, e.g., all of, e.g., from 5’ to 3’, a promoter (e.g., CMV), an R sequence (e.g., comprising TAR), a U5 sequence (e.g., for integration), a PBS sequence (e.g., for reverse transcription), a DIS sequence (e.g., for genome dimerization), a psi packaging signal, a partial gag sequence, an RRE sequence (e.g., for nuclear export), a cPPT sequence (e.g., for nuclear import), a promoter to drive expression of the exogenous agent, a gene encoding the exogenous agent, a WPRE sequence (e.g., for efficient transgene expression), a PPT sequence (e.g., for reverse transcription), an R sequence (e.g., for polyadenylation and termination), and a U5 signal (e.g., for integration).

ii) Vectors engineered to remove splice sites

Some lentiviral vectors integrate inside active genes and possess strong splicing and polyadenylation signals that could lead to the formation of aberrant and possibly truncated transcripts.

Mechanisms of proto-oncogene activation may involve the generation of chimeric transcripts originating from the interaction of promoter elements or splice sites contained in the genome of the insertional mutagen with the cellular transcriptional unit targeted by integration (Gabriel et al. 2009. Nat Med 15: 1431 -1436; Bokhoven, et al. J Virol 83:283-29). Chimeric fusion transcripts comprising vector sequences and cellular mRNAs can be generated either by read- through transcription starting from vector sequences and proceeding into the flanking cellular genes, or vice versa.

In some embodiments, a lentiviral nucleic acid described herein comprises a lentiviral backbone in which at least two of the splice sites have been eliminated, e.g., to improve the safety profile of the lentiviral vector. Species of such splice sites and methods of identification are described in WO2012156839A2, all of which is included by reference.

iii) Retroviral production methods

Large scale viral particle production is often useful to achieve a desired viral titer. Viral particles can be produced by transfecting a transfer vector into a packaging cell line that comprises viral structural and/or accessory genes, e.g., gag, pol, env, tat, rev, vif, vpr, vpu, vpx, or nef genes or other retroviral genes.

In embodiments, the packaging vector is an expression vector or viral vector that lacks a packaging signal and comprises a polynucleotide encoding one, two, three, four or more viral structural and/or accessory genes. Typically, the packaging vectors are included in a packaging cell, and are introduced into the cell via transfection, transduction or infection. A retroviral, e.g., lentiviral, transfer vector can be introduced into a packaging cell line, via transfection, transduction or infection, to generate a source cell or cell line. The packaging vectors can be introduced into human cells or cell lines by standard methods including, e.g., calcium phosphate transfection, lipofection or electroporation. In some embodiments, the packaging vectors are introduced into the cells together with a dominant selectable marker, such as neomycin, hygromycin, puromycin, blastocidin, zeocin, thymidine kinase, DHFR, Gln synthetase or ADA,

followed by selection in the presence of the appropriate drug and isolation of clones. A selectable marker gene can be linked physically to genes encoding by the packaging vector, e.g., by IRES or self cleaving viral peptides.

Packaging cell lines include cell lines that do not contain a packaging signal, but do stably or transiently express viral structural proteins and replication enzymes (e.g., gag, pol and env) which can package viral particles. Any suitable cell line can be employed, e.g., mammalian cells, e.g., human cells. Suitable cell lines which can be used include, for example, CHO cells, BHK cells, MDCK cells, C3H 10T1/2 cells, FLY cells, Psi-2 cells, BOSC 23 cells, PA317 cells, WEHI cells, COS cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, W138 cells, MRC5 cells, A549 cells, HT1080 cells, 293 cells, 293T cells, B-50 cells, 3T3 cells, NIH3T3 cells, HepG2 cells, Saos-2 cells, Huh7 cells, HeLa cells, W163 cells, 211 cells, and 211A cells. In embodiments, the packaging cells are 293 cells, 293T cells, or A549 cells.

A source cell line includes a cell line which is capable of producing recombinant retroviral particles, comprising a packaging cell line and a transfer vector construct comprising a packaging signal. Methods of preparing viral stock solutions are illustrated by, e.g., Y. Soneoka et al. (1995) *Nucl. Acids Res.* 23:628-633, and N. R. Landau et al. (1992) *J. Virol.* 66:5110-5113, which are incorporated herein by reference. Infectious virus particles may be collected from the packaging cells, e.g., by cell lysis, or collection of the supernatant of the cell culture. Optionally, the collected virus particles may be enriched or purified.

iv) Packaging plasmids and cell lines

In some embodiments, the source cell comprises one or more plasmids coding for viral structural proteins and replication enzymes (e.g., gag, pol and env) which can package viral particles. In some embodiments, the sequences coding for at least two of the gag, pol, and env precursors are on the same plasmid. In some embodiments, the sequences coding for the gag, pol, and env precursors are on different plasmids. In some embodiments, the sequences coding for the gag, pol, and env precursors have the same expression signal, e.g., promoter. In some embodiments, the sequences coding for the gag, pol, and env precursors have a different expression signal, e.g., different promoters. In some embodiments, expression of the gag, pol, and env precursors is inducible. In some embodiments, the plasmids coding for viral structural proteins and replication enzymes are transfected at the same time or at different times. In some

embodiments, the plasmids coding for viral structural proteins and replication enzymes are transfected at the same time or at a different time from the packaging vector.

In some embodiments, the source cell line comprises one or more stably integrated viral structural genes. In some embodiments expression of the stably integrated viral structural genes is inducible.

In some embodiments, expression of the viral structural genes is regulated at the transcriptional level. In some embodiments, expression of the viral structural genes is regulated at the translational level. In some embodiments, expression of the viral structural genes is regulated at the post-translational level.

In some embodiments, expression of the viral structural genes is regulated by a tetracycline (Tet)-dependent system, in which a Tet-regulated transcriptional repressor (Tet-R) binds to DNA sequences included in a promoter and represses transcription by steric hindrance (Yao et al, 1998; Jones et al, 2005). Upon addition of doxycycline (dox), Tet-R is released, allowing transcription. Multiple other suitable transcriptional regulatory promoters, transcription factors, and small molecule inducers are suitable to regulate transcription of viral structural genes.

In some embodiments, the third-generation lentivirus components, human immunodeficiency virus type 1 (HIV) Rev, Gag/Pol, and an envelope under the control of Tet-regulated promoters and coupled with antibiotic resistance cassettes are separately integrated into the source cell genome. In some embodiments the source cell only has one copy of each of Rev, Gag/Pol, and an envelope protein integrated into the genome.

In some embodiments a nucleic acid encoding the exogenous agent (e.g., a retroviral nucleic acid encoding the exogenous agent) is also integrated into the source cell genome. In some embodiments a nucleic acid encoding the exogenous agent is maintained episomally. In some embodiments a nucleic acid encoding the exogenous agent is transfected into the source cell that has stably integrated Rev, Gag/Pol, and an envelope protein in the genome. See, e.g., Milani et al. *EMBO Molecular Medicine*, 2017, which is herein incorporated by reference in its entirety.

In some embodiments, a retroviral nucleic acid described herein is unable to undergo reverse transcription. Such a nucleic acid, in embodiments, is able to transiently express an exogenous agent. The retrovirus or VLP, may comprise a disabled reverse transcriptase protein,

or may not comprise a reverse transcriptase protein. In embodiments, the retroviral nucleic acid comprises a disabled primer binding site (PBS) and/or att site. In embodiments, one or more viral accessory genes, including rev, tat, vif, nef, vpr, vpu, vpx and S2 or functional equivalents thereof, are disabled or absent from the retroviral nucleic acid. In embodiments, one or more accessory genes selected from S2, rev and tat are disabled or absent from the retroviral nucleic acid.

v) Strategies for packaging a retroviral nucleic acid

Typically, modern retroviral vector systems consist of viral genomes bearing cis-acting vector sequences for transcription, reverse-transcription, integration, translation and packaging of viral RNA into the viral particles, and (2) producer cells lines which express the trans-acting retroviral gene sequences (e.g., gag, pol and env) needed for production of virus particles. By separating the cis-and trans-acting vector sequences completely, the virus is unable to maintain replication for more than one cycle of infection. Generation of live virus can be avoided by a number of strategies, e.g., by minimizing the overlap between the cis-and trans-acting sequences to avoid recombination.

A viral vector particle which comprises a sequence that is devoid of or lacking viral RNA may be the result of removing or eliminating the viral RNA from the sequence. In one embodiment this may be achieved by using an endogenous packaging signal binding site on gag. Alternatively, the endogenous packaging signal binding site is on pol. In this embodiment, the RNA which is to be delivered will contain a cognate packaging signal. In another embodiment, a heterologous binding domain (which is heterologous to gag) located on the RNA to be delivered, and a cognate binding site located on gag or pol, can be used to ensure packaging of the RNA to be delivered. The heterologous sequence could be non-viral or it could be viral, in which case it may be derived from a different virus. The vector particles could be used to deliver therapeutic RNA, in which case functional integrase and/or reverse transcriptase is not required. These vector particles could also be used to deliver a therapeutic gene of interest, in which case pol is typically included.

In an embodiment, gag-pol are altered, and the packaging signal is replaced with a corresponding packaging signal. In this embodiment, the particle can package the RNA with the

new packaging signal. The advantage of this approach is that it is possible to package an RNA sequence which is devoid of viral sequence for example, RNAi.

An alternative approach is to rely on over-expression of the RNA to be packaged. In one embodiment the RNA to be packaged is over-expressed in the absence of any RNA containing a packaging signal. This may result in a significant level of therapeutic RNA being packaged, and that this amount is sufficient to transduce a cell and have a biological effect.

In some embodiments, a polynucleotide comprises a nucleotide sequence encoding a viral gag protein or retroviral gag and pol proteins, wherein the gag protein or pol protein comprises a heterologous RNA binding domain capable of recognising a corresponding sequence in an RNA sequence to facilitate packaging of the RNA sequence into a viral vector particle.

In some embodiments, the heterologous RNA binding domain comprises an RNA binding domain derived from a bacteriophage coat protein, a Rev protein, a protein of the U1 small nuclear ribonucleoprotein particle, a Nova protein, a TF111A protein, a TIS11 protein, a trp RNA-binding attenuation protein (TRAP) or a pseudouridine synthase.

In some embodiments, a method herein comprises detecting or confirming the absence of replication competent retrovirus. The methods may include assessing RNA levels of one or more target genes, such as viral genes, e.g. structural or packaging genes, from which gene products are expressed in certain cells infected with a replication-competent retrovirus, such as a gammaretrovirus or lentivirus, but not present in a viral vector used to transduce cells with a heterologous nucleic acid and not, or not expected to be, present and/or expressed in cells not containing replication-competent retrovirus. Replication competent retrovirus may be determined to be present if RNA levels of the one or more target genes is higher than a reference value, which can be measured directly or indirectly, e.g. from a positive control sample containing the target gene. For further disclosure, see WO2018023094A1.

vi) Repression of a gene encoding an exogenous agent in a source cell

(Over-)expressed protein in the source cell may have an indirect or direct effect on vector virion assembly and/or infectivity. Incorporation of the exogenous agent into vector virions may also impact downstream processing of vector particles.

In some embodiments, a tissue-specific promoter is used to limit expression of the exogenous agent in source cells. In some embodiments, a heterologous translation control system

is used in eukaryotic cell cultures to repress the translation of the exogenous agent in source cells. More specifically, the retroviral nucleic acid may comprise a binding site operably linked to the gene encoding the exogenous agent, wherein the binding site is capable of interacting with an RNA-binding protein such that translation of the exogenous agent is repressed or prevented in the source cell.

In some embodiments, the RNA-binding protein is tryptophan RNA-binding attenuation protein (TRAP), for example bacterial tryptophan RNA-binding attenuation protein. The use of an RNA-binding protein (e.g. the bacterial trp operon regulator protein, tryptophan RNA-binding attenuation protein, TRAP), and RNA targets to which it binds, will repress or prevent transgene translation within a source cell. This system is referred to as the Transgene Repression In vector Production cell system or TRIP system.

In embodiments, the placement of a binding site for an RNA binding protein (e.g., a TRAP-binding sequence, tbs) upstream of the NOI translation initiation codon allows specific repression of translation of mRNA derived from the internal expression cassette, while having no detrimental effect on production or stability of vector RNA. The number of nucleotides between the tbs and translation initiation codon of the gene encoding the exogenous agent may be varied from 0 to 12 nucleotides. The tbs may be placed downstream of an internal ribosome entry site (IRES) to repress translation of the gene encoding the exogenous agent in a multicistronic mRNA.

vii) Kill switch systems and amplification

In some embodiments, a polynucleotide or cell harboring the gene encoding the exogenous agent utilizes a suicide gene, e.g., an inducible suicide gene, to reduce the risk of direct toxicity and/or uncontrolled proliferation. In specific aspects, the suicide gene is not immunogenic to the host cell harboring the exogenous agent. Examples of suicide genes include caspase-9, caspase-8, or cytosine deaminase. Caspase-9 can be activated using a specific chemical inducer of dimerization (CID).

In certain embodiments, vectors comprise gene segments that cause target cells, e.g., immune effector cells, e.g., T cells, to be susceptible to negative selection in vivo. For instance, the transduced cell can be eliminated as a result of a change in the in vivo condition of the individual. The negative selectable phenotype may result from the insertion of a gene that

confers sensitivity to an administered agent, for example, a compound. Negative selectable genes are known in the art, and include, inter alia the following: the Herpes simplex virus type I thymidine kinase (HSV-I TK) gene (Wigler et al., Cell 11:223, 1977) which confers ganciclovir sensitivity; the cellular hypoxanthine phosphribosyltransferase (HPRT) gene, the cellular adenine phosphoribosyltransferase (APRT) gene, and bacterial cytosine deaminase, (Mullen et al., Proc. Natl. Acad. Sci. USA. 89:33 (1992)).

In some embodiments, transduced cells, e.g., immune effector cells, such as T cells, comprise a polynucleotide further comprising a positive marker that enables the selection of cells of the negative selectable phenotype in vitro. The positive selectable marker may be a gene which, upon being introduced into the target cell, expresses a dominant phenotype permitting positive selection of cells carrying the gene. Genes of this type include, inter alia, hygromycin-B phosphotransferase gene (hph) which confers resistance to hygromycin B, the amino glycoside phosphotransferase gene (neo or aph) from Tn5 which codes for resistance to the antibiotic G418, the dihydrofolate reductase (DHFR) gene, the adenosine deaminase gene (ADA), and the multi-drug resistance (MDR) gene.

In some embodiments, the positive selectable marker and the negative selectable element are linked such that loss of the negative selectable element necessarily also is accompanied by loss of the positive selectable marker. For instance, the positive and negative selectable markers can be fused so that loss of one obligatorily leads to loss of the other. An example of a fused polynucleotide that yields as an expression product a polypeptide that confers both the desired positive and negative selection features described above is a hygromycin phosphotransferase thymidine kinase fusion gene (HyTK). Expression of this gene yields a polypeptide that confers hygromycin B resistance for positive selection in vitro, and ganciclovir sensitivity for negative selection in vivo. See Lupton S. D., et al, Mol. and Cell. Biology 1 1:3374-3378, 1991. In addition, in embodiments, the polynucleotides encoding the chimeric receptors are in retroviral vectors containing the fused gene, particularly those that confer hygromycin B resistance for positive selection in vitro, and ganciclovir sensitivity for negative selection in vivo, for example the HyTK retroviral vector described in Lupton, S. D. et al. (1991), supra. See also the publications of PCT U591/08442 and PCT/U594/05601, describing the use of bifunctional selectable fusion genes derived from fusing dominant positive selectable markers with negative selectable markers.

Suitable positive selectable markers can be derived from genes selected from the group consisting of hph, nco, and gpt, and suitable negative selectable markers can be derived from genes selected from the group consisting of cytosine deaminase, HSV-I TK, VZV TK, HPRT, APRT and gpt. Other suitable markers are bifunctional selectable fusion genes wherein the positive selectable marker is derived from hph or neo, and the negative selectable marker is derived from cytosine deaminase or a TK gene or selectable marker.

viii) Strategies for regulating lentiviral integration

Retroviral and lentiviral nucleic acids are disclosed which are lacking or disabled in key proteins/sequences so as to prevent integration of the retroviral or lentiviral genome into the target cell genome. For instance, viral nucleic acids lacking each of the amino acids making up the highly conserved DDE motif (Engelman and Craigie (1992) *J. Virol.* 66:6361-6369; Johnson et al. (1986) *Proc. Natl. Acad. Sci. USA* 83:7648-7652; Khan et al. (1991) *Nucleic Acids Res.* 19:851-860) of retroviral integrase enables the production of integration defective retroviral nucleic acids.

For instance, in some embodiments, a retroviral nucleic acid herein comprises a lentiviral integrase comprising a mutation that causes said integrase to be unable to catalyze the integration of the viral genome into a cell genome. In some embodiments, said mutations are type I mutations which affect directly the integration, or type II mutations which trigger pleiotropic defects affecting virion morphogenesis and/or reverse transcription. Illustrative non-limitative examples of type I mutations are those mutations affecting any of the three residues that participate in the catalytic core domain of the integrase: $DX_{39-58}DX_{35}E$ (D64, D116 and E152 residues of the integrase of the HIV-1). In a particular embodiment, the mutation that causes said integrase to be unable to catalyze the integration of the viral genome into a cell genome is the substitution of one or more amino acid residues of the DDE motif of the catalytic core domain of the integrase, preferably the substitution of the first aspartic residue of said DEE motif by an asparagine residue. In some embodiment the retroviral vector does not comprise an integrase protein.

In some embodiments the retrovirus integrates into active transcription units. In some embodiments the retrovirus does not integrate near transcriptional start sites, the 5' end of genes, or DNAse1 cleavage sites. In some embodiments the retrovirus integration does not active proto-

oncogenes or inactive tumor suppressor genes. In some embodiments the retrovirus is not genotoxic. In some embodiments the lentivirus integrates into introns.

In some embodiments, the retroviral nucleic acid integrates into the genome of a target cell with a particular copy number. The average copy number may be determined from single cells, a population of cells, or individual cell colonies. Exemplary methods for determining copy number include polymerase chain reaction (PCR) and flow cytometry.

In some embodiments DNA encoding the exogenous agent is integrated into the genome. In some embodiments DNA encoding the exogenous agent is maintained episomally. In some embodiments the ratio of integrated to episomal DNA encoding the exogenous agent is at least 0.01, 0.1, 0.5, 1.0, 2, 5, 10, 100.

In some embodiments DNA encoding the exogenous agent is linear. In some embodiments DNA encoding the exogenous agent is circular. In some embodiments the ratio of linear to circular copies of DNA encoding the exogenous agent is at least 0.01, 0.1, 0.5, 1.0, 2, 5, 10, 100.

In embodiments the DNA encoding the exogenous agent is circular with 1 LTR. In some embodiments the DNA encoding the exogenous agent is circular with 2 LTRs. In some embodiments the ratio of circular, 1 LTR-comprising DNA encoding the exogenous agent to circular, 2 LTR-comprising DNA encoding the exogenous agent is at least 0.1, 0.5, 1.0, 2, 5, 10, 20, 50, 100.

ix) Maintenance of an episomal virus

In retroviruses deficient in integration, circular cDNA off-products of the retrotranscription (e.g., 1-LTR and 2-LTR) can accumulate in the cell nucleus without integrating into the host genome (see Yáñez-Muñoz R J et al., Nat. Med. 2006, 12: 348-353). Like other exogenous DNA those intermediates can then integrate in the cellular DNA at equal frequencies (e.g., 10^3 to 10^5/cell).

In some embodiments, episomal retroviral nucleic acid does not replicate. Episomal virus DNA can be modified to be maintained in replicating cells through the inclusion of eukaryotic origin of replication and a scaffold/matrix attachment region (S/MAR) for association with the nuclear matrix.

Thus, in some embodiments, a retroviral nucleic acid described herein comprises a eukaryotic origin of replication or a variant thereof. Examples of eukaryotic origins of replication of interest are the origin of replication of the β-globin gene as have been described by Aladjem et al (Science, 1995, 270: 815-819), a consensus sequence from autonomously replicating sequences associated with alpha-satellite sequences isolated previously from monkey CV-1 cells and human skin fibroblasts as has been described by Price et al Journal of Biological Chemistry, 2003, 278 (22): 19649-59, the origin of replication of the human c-myc promoter region has have been described by McWinney and Leffak (McWinney C. and Leffak M., Nucleic Acid Research 1990, 18(5): 1233-42). In embodiments, the variant substantially maintains the ability to initiate the replication in eukaryotes. The ability of a particular sequence of initiating replication can be determined by any suitable method, for example, the autonomous replication assay based on bromodeoxyuridine incorporation and density shift (Araujo F. D. et al., supra; Frappier L. et al., supra).

In some embodiments, the retroviral nucleic acid comprises a scaffold/matrix attachment region (S/MAR) or variant thereof, e.g., a non-consensus-like AT-rich DNA element several hundred base pairs in length, which organizes the nuclear DNA of the eukaryotic genome into chromatin domains, by periodic attachment to the protein scaffold or matrix of the cell nucleus. They are typically found in non-coding regions such as flanking regions, chromatin border regions, and introns. Examples of S/MAR regions are 1.8 kbp S/MAR of the human IFN-γ gene (hIFN-γ^{large}) as described by Bode et al (Bode J. et al., Science, 1992, 255: 195-7), the 0.7 Kbp minimal region of the S/MAR of the human IFN-γ gene (hIFN-γ^{short}) as has have been described by Ramezani (Ramezani A. et al., Blood 2003, 101: 4717-24), the 0.2 Kbp minimal region of the S/MAR of the human dehydrofolate reductase gene (hDHFR) as has been described by Mesner L. D. et al., Proc Natl Acad Sci USA, 2003, 100: 3281-86). In embodiments, the functionally equivalent variant of the S/MAR is a sequence selected based on the set six rules that together or alone have been suggested to contribute to S/MAR function (Kramer et al (1996) *Genomics* 33, 305; Singh et al (1997) *Nucl. Acids Res* 25, 1419). These rules have been merged into the MAR-Wiz computer program freely available at genomecluster.secs.oakland.edu/MAR-Wiz. In embodiments, the variant substantially maintains the same functions of the S/MAR from which it derives, in particular, the ability to specifically bind to the nuclear the matrix. The skilled person can determine if a particular variant is able to specifically bind to the nuclear matrix, for example

by the in vitro or in vivo MAR assays described by Mesner et al. (Mesner L. D. et al, supra). In some embodiments, a specific sequence is a variant of a S/MAR if the particular variant shows propensity for DNA strand separation. This property can be determined using a specific program based on methods from equilibrium statistical mechanics. The stress-induced duplex destabilization (SIDD) analysis technique "[. . .] calculates the extent to which the imposed level of superhelical stress decreases the free energy needed to open the duplex at each position along a DNA sequence. The results are displayed as an SIDD profile, in which sites of strong destabilization appear as deep minima [. . .]" as defined in Bode et al (2005) *J. Mol. Biol.* 358,597. The SIDD algorithm and the mathematical basis (Bi and Benham (2004) Bioinformatics 20, 1477) and the analysis of the SIDD profile can be performed using the freely available internet resource at WebSIDD (www.genomecenter.ucdavis.edu/benham). Accordingly, in some embodiment, the polynucleotide is considered a variant of the S/MAR sequence if it shows a similar SIDD profile as the S/MAR.

B. Cell-derived fusosomes

Compositions of fusosomes may be generated from cells in culture, for example cultured mammalian cells, e.g., cultured human cells. The cells may be progenitor cells or non-progenitor (e.g., differentiated) cells. The cells may be primary cells or cell lines (e.g., a mammalian, e.g., human, cell line described herein). In embodiments, the cultured cells are progenitor cells, e.g., bone marrow stromal cells, marrow derived adult progenitor cells (MAPCs), endothelial progenitor cells (EPC), blast cells, intermediate progenitor cells formed in the subventricular zone, neural stem cells, muscle stem cells, satellite cells, liver stem cells, hematopoietic stem cells, bone marrow stromal cells, epidermal stem cells, embryonic stem cells, mesenchymal stem cells, umbilical cord stem cells, precursor cells, muscle precursor cells, myoblast, cardiomyoblast, neural precursor cells, glial precursor cells, neuronal precursor cells, hepatoblasts.

In some embodiments, the source cell is an endothelial cell, a fibroblast, a blood cell (e.g., a macrophage, a neutrophil, a granulocyte, a leukocyte), a stem cell (e.g., a mesenchymal stem cell, an umbilical cord stem cell, bone marrow stem cell, a hematopoietic stem cell, an induced pluripotent stem cell e.g., an induced pluripotent stem cell derived from a subject's cells), an embryonic stem cell (e.g., a stem cell from embryonic yolk sac, placenta, umbilical cord, fetal skin, adolescent skin, blood, bone marrow, adipose tissue, erythropoietic tissue,

hematopoietic tissue), a myoblast, a parenchymal cell (e.g., hepatocyte), an alveolar cell, a neuron (e.g., a retinal neuronal cell) a precursor cell (e.g., a retinal precursor cell, a myeloblast, myeloid precursor cells, a thymocyte, a meiocyte, a megakaryoblast, a promegakaryoblast, a melanoblast, a lymphoblast, a bone marrow precursor cell, a normoblast, or an angioblast), a progenitor cell (e.g., a cardiac progenitor cell, a satellite cell, a radial glial cell, a bone marrow stromal cell, a pancreatic progenitor cell, an endothelial progenitor cell, a blast cell), or an immortalized cell (e.g., HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cell).

The cultured cells may be from epithelial, connective, muscular, or nervous tissue or cells, and combinations thereof. Fusosome can be generated from cultured cells from any eukaryotic (e.g., mammalian) organ system, for example, from the cardiovascular system (heart, vasculature); digestive system (esophagus, stomach, liver, gallbladder, pancreas, intestines, colon, rectum and anus); endocrine system (hypothalamus, pituitary gland, pineal body or pineal gland, thyroid, parathyroids, adrenal glands); excretory system (kidneys, ureters, bladder); lymphatic system (lymph, lymph nodes, lymph vessels, tonsils, adenoids, thymus, spleen); integumentary system (skin, hair, nails); muscular system (e.g., skeletal muscle); nervous system (brain, spinal cord, nerves); reproductive system (ovaries, uterus, mammary glands, testes, vas deferens, seminal vesicles, prostate); respiratory system (pharynx, larynx, trachea, bronchi, lungs, diaphragm); skeletal system (bone, cartilage), and combinations thereof. In embodiments, the cells are from a highly mitotic tissue (e.g., a highly mitotic healthy tissue, such as epithelium, embryonic tissue, bone marrow, intestinal crypts). In embodiments, the tissue sample is a highly metabolic tissue (e.g., skeletal tissue, neural tissue, cardiomyocytes).

In some embodiments, the cells are from a young donor, e.g., a donor 25 years, 20 years, 18 years, 16 years, 12 years, 10 years, 8 years of age, 5 years of age, 1 year of age, or less. In some embodiments, the cells are from fetal tissue.

In some embodiments, the cells are derived from a subject and administered to the same subject or a subject with a similar genetic signature (e.g., MHC-matched).

In certain embodiments, the cells have telomeres of average size greater than 3000, 4000, 5000, 6000, 7000, 8000, 9000, or 10000 nucleotides in length (e.g., between 4,000-10,000 nucleotides in length, between 6,000-10,000 nucleotides in length).

In some embodiments, fusosomes are generated from a cell clone identified, chosen, or

selected based on a desirable phenotype or genotype for use as a source for fusosome composition described herein. For example, a cell clone is identified, chosen, or selected based on low mitochondrial mutation load, long telomere length, differentiation state, or a particular genetic signature (e.g., a genetic signature to match a recipient).

A fusosome composition described herein may be comprised of fusosomes from one cellular or tissue source, or from a combination of sources. For example, a fusosome composition may comprise fusosomes from xenogeneic sources (e.g., animals, tissue culture of the aforementioned species' cells), allogeneic, autologous, from specific tissues resulting in different protein concentrations and distributions (liver, skeletal, neural, adipose, etc.), from cells of different metabolic states (e.g., glycolytic, respiring). A composition may also comprise fusosomes in different metabolic states, e.g. coupled or uncoupled, as described elsewhere herein.

In some embodiments, fusosomes are generated from source cells expressing a fusogen, e.g., a fusogen described herein. In some embodiments, the fusogen is disposed in a membrane of the source cell, e.g., a lipid bilayer membrane, e.g., a cell surface membrane, or a subcellular membrane (e.g., lysosomal membrane). In some embodiments, fusosomes are generated from source cells with a fusogen disposed in a cell surface membrane.

In some embodiments, fusosomes are generated by inducing budding of an exosome, microvesicle, membrane vesicle, extracellular membrane vesicle, plasma membrane vesicle, giant plasma membrane vesicle, apoptotic body, mitoparticle, pyrenocyte, lysosome, or other membrane enclosed vesicle.

In some embodiments, fusosomes are generated by inducing cell enucleation. Enucleation may be performed using assays such as genetic, chemical (e.g., using Actinomycin D, see Bayona-Bafaluyet al., "A chemical enucleation method for the transfer of mitochondrial DNA to ρ° cells" Nucleic Acids Res. 2003 Aug 15; 31(16): e98), mechanical methods (e.g., squeezing or aspiration, see Lee et al., "A comparative study on the efficiency of two enucleation methods in pig somatic cell nuclear transfer: effects of the squeezing and the aspiration methods." Anim Biotechnol. 2008;19(2):71-9), or combinations thereof. Enucleation refers not only to a complete removal of the nucleus but also the displacement of the nucleus from its typical location such that the cell contains the nucleus but it is non-functional.

In embodiments, making a fusosome comprises producing cell ghosts, giant plasma membrane vesicle, or apoptotic bodies. In embodiments, a fusosome composition comprises one or more of cell ghosts, giant plasma membrane vesicle, and apoptotic bodies.

In some embodiments, fusosomes are generated by inducing cell fragmentation. In some embodiments, cell fragmentation can be performed using the following methods, including, but not limited to: chemical methods, mechanical methods (e.g., centrifugation (e.g., ultracentrifugation, or density centrifugation), freeze-thaw, or sonication), or combinations thereof.

In some embodiments, a fusosome can be generated from a source cell expressing a fusogen, e.g., as described herein, by any one, all of, or a combination of the following methods:

i) inducing budding of a mitoparticle, exosome, or other membrane enclosed vesicle;

ii) inducing nuclear inactivation, e.g., enucleation, by any of the following methods or a combination thereof:

a) a genetic method;

b) a chemical method, e.g., using Actinomycin D; or

c) a mechanical method, e.g., squeezing or aspiration; or

iii) inducing cell fragmentation, e.g., by any of the following methods or a combination thereof:

a) a chemical method;

b) a mechanical method, e.g., centrifugation (e.g., ultracentrifugation or density centrifugation); freeze thaw; or sonication.

i) Modifications to Cells Prior to Fusosome Generation

In some aspects, a modification is made to a cell, such as modification of a subject, tissue or cell, prior to fusosome generation. Such modifications can be effective to, e.g., improve fusion, fusogen expression or activity, structure or function of the cargo, or structure or function of the target cell.

a) Physical Modifications

In some embodiments, a cell is physically modified prior to generating the fusosome. For example, as described elsewhere herein, a fusogen may be linked to the surface of the cell.

In some embodiments, a cell is treated with a chemical agent prior to generating the fusosome. For example, the cell may be treated with a chemical or lipid fusogen, such that the chemical or lipid fusogen non-covalently or covalently interacts with the surface of the cell or embeds within the surface of the cell. In some embodiments, the cell is treated with an agent to enhance fusogenic properties of the lipids in the cell membrane.

In some embodiments, the cell is physically modified prior to generating the fusosome with one or more covalent or non-covalent attachment sites for synthetic or endogenous small molecules or lipids on the cell surface that enhance targeting of the fusosome to an organ, tissues, or cell-type.

In embodiments, a fusosome comprises increased or decreased levels of an endogenous molecule. For instance, the fusosome may comprise an endogenous molecule that also naturally occurs in the naturally occurring source cell but at a higher or lower level than in the fusosome. In some embodiments, the polypeptide is expressed from an exogenous nucleic acid in the source cell or fusosome. In some embodiments, the polypeptide is isolated from a source and loaded into or conjugated to a source cell or fusosome.

In some embodiments, a cell is treated with a chemical agent, e.g., small molecule, prior to generating the fusosome to increase the expression or activity of an endogenous fusogen in the cell (e.g., in some embodiments, endogenous relative to the source cell, and in some embodiments, endogenous relative to the target cell). In some embodiments, a small molecule may increase expression or activity of a transcriptional activator of the endogenous fusogen. In some embodiments, a small molecule may decrease expression or activity of a transcriptional repressor of the endogenous fusogen. In some embodiments, a small molecule is an epigenetic modifier that increases expression of the endogenous fusogen.

In some embodiments, fusosomes are generated from cells treated with fusion arresting compounds, e.g., lysophosphatidylcholine. In some embodiments, fusosomes are generated from cells treated with dissociation reagents that do not cleave fusogens, e.g., Accutase.

In some embodiments, a source cell is physically modified with, e.g., CRISPR activators, prior to generating a fusosome to add or increase the concentration of fusogens.

In some embodiments, the cell is physically modified to increase or decrease the quantity, or enhance the structure or function of organelles, e.g., mitochondria, Golgi apparatus, endoplasmic reticulum, intracellular vesicles (such as lysosomes, autophagosomes).

b) Genetic Modifications

In some embodiments, a cell is genetically modified prior to generating the fusosome to increase the expression of an endogenous fusogen in the cell (e.g., in some embodiments, endogenous relative to the source cell, and in some embodiments, endogenous relative to the target cell) . In some embodiments, a genetic modification may increase expression or activity of a transcriptional activator of the endogenous fusogen. In some embodiments, a genetic modification may decrease expression or activity of a transcriptional repressor of the endogenous fusogen. In some embodiments the activator or repressor is a nuclease-inactive cas9 (dCas9) linked to a transcriptional activator or repressor that is targeted to the endogenous fusogen by a guide RNA. In some embodiments, a genetic modification epigenetically modifies an endogenous fusogen gene to increase its expression. In some embodiments the epigenetic activator a nuclease-inactive cas9 (dCas9) linked to an epigenetic modifier that is targeted to the endogenous fusogen by a guide RNA.

In some embodiments, a cell is genetically modified prior to generating the fusosome to increase the expression of an exogenous fusogen in the cell, e.g., delivery of a transgene. In some embodiments, a nucleic acid, e.g., DNA, mRNA or siRNA, is transferred to the cell prior to generating the fusosome, e.g., to increase or decrease the expression of a cell surface molecule (protein, glycan, lipid or low molecular weight molecule) used for organ, tissue, or cell targeting. In some embodiments, the nucleic acid targets a repressor of a fusogen, e.g., an shRNA, siRNA construct. In some embodiments, the nucleic acid encodes an inhibitor of a fusogen repressor.

In some embodiments, the method comprises introducing a nucleic acid , that is exogenous relative to the source cell encoding a fusogen into a source cell. The exogenous nucleic acid may be, e.g., DNA or RNA. In some embodiments the exogenous nucleic acid may be e.g., a DNA, a gDNA, a cDNA, an RNA, a pre-mRNA, an mRNA, an miRNA, an siRNA, etc. In some embodiments, the exogenous DNA may be linear DNA, circular DNA, or an artificial

chromosome. In some embodiments the DNA is maintained episomally. In some embodiments the DNA is integrated into the genome. The exogenous RNA may be chemically modified RNA, e.g., may comprise one or more backbone modification, sugar modifications, noncanonical bases, or caps. Backbone modifications include, e.g., phosphorothioate, N3′ phosphoramidite, boranophosphate, phosphonoacetate, thio-PACE, morpholino phosphoramidites, or PNA. Sugar modifications include, e.g., 2′-O-Me, 2′F, 2′F-ANA, LNA, UNA, and 2′-O-MOE. Noncanonical bases include, e.g., 5-bromo-U, and 5-iodo-U, 2,6-diaminopurine, C-5 propynyl pyrimidine, difluorotoluene, difluorobenzene, dichlorobenzene, 2-thiouridine, pseudouridine, and dihydrouridine. Caps include, e.g., ARCA. Additional modifications are discussed, e.g., in Deleavey et al., "Designing Chemically Modified Oligonucleotides for Targeted Gene Silencing" Chemistry & Biology Volume 19, Issue 8, 24 August 2012, Pages 937-954, which is herein incorporated by reference in its entirety.

In some embodiments, a cell is treated with a chemical agent, e.g. a small molecule, prior to generating the fusosome to increase the expression or activity of a fusogen that is exogenous relative to the source cell in the cell. In some embodiments, a small molecule may increase expression or activity of a transcriptional activator of the exogenous fusogen. In some embodiments, a small molecule may decrease expression or activity of a transcriptional repressor of the exogenous fusogen. In some embodiments, a small molecule is an epigenetic modifier that increases expression of the exogenous fusogen.

In some embodiments, the nucleic acid encodes a modified fusogen. For example, a fusogen that has regulatable fusogenic activity, e.g., specific cell-type, tissue-type or local microenvironment activity. Such regulatable fusogenic activity may include, activation and/or initiation of fusogenic activity by low pH, high pH, heat, infrared light, extracellular enzyme activity (eukaryotic or prokaryotic), or exposure of a small molecule, a protein, or a lipid. In some embodiments, the small molecule, protein, or lipid is displayed on a target cell.

In some embodiments, a cell is genetically modified prior to generating the fusosome to alter (i.e., upregulate or downregulate) the expression of signaling pathways (e.g., the Wnt/Beta-catenin pathway). In some embodiments, a cell is genetically modified prior to generating the fusosome to alter (e.g., upregulate or downregulate) the expression of a gene or genes of interest. In some embodiments, a cell is genetically modified prior to generating the fusosome to alter

(e.g., upregulate or downregulate) the expression of a nucleic acid (e.g. a miRNA or mRNA) or nucleic acids of interest. In some embodiments, nucleic acids, e.g., DNA, mRNA or siRNA, are transferred to the cell prior to generating the fusosome, e.g., to increase or decrease the expression of signaling pathways, genes, or nucleic acids. In some embodiments, the nucleic acid targets a repressor of a signaling pathway, gene, or nucleic acid, or represses a signaling pathway, gene, or nucleic acid. In some embodiments, the nucleic acid encodes a transcription factor that upregulates or downregulates a signaling pathway, gene, or nucleic acid. In some embodiments the activator or repressor is a nuclease-inactive cas9 (dCas9) linked to a transcriptional activator or repressor that is targeted to the signaling pathway, gene, or nucleic acid by a guide RNA. In some embodiments, a genetic modification epigenetically modifies an endogenous signaling pathway, gene, or nucleic acid to its expression. In some embodiments the epigenetic activator a nuclease-inactive cas9 (dCas9) linked to a epigenetic modifier that is targeted to the signaling pathway, gene, or nucleic acid by a guide RNA. In some embodiments, a cell's DNA is edited prior to generating the fusosome to alter (e.g., upregulate or downregulate) the expression of signaling pathways (e.g. the Wnt/Beta-catenin pathway), gene, or nucleic acid. In some embodiments, the DNA is edited using a guide RNA and CRISPR-Cas9/Cpf1 or other gene editing technology.

A cell may be genetically modified using recombinant methods. A nucleic acid sequence coding for a desired gene can be obtained using recombinant methods, such as, for example by screening libraries from cells expressing the gene, by deriving the gene from a vector known to include the same, or by isolating directly from cells and tissues containing the same, using standard techniques. Alternatively, a gene of interest can be produced synthetically, rather than cloned.

Expression of natural or synthetic nucleic acids is typically achieved by operably linking a nucleic acid encoding the gene of interest to a promoter, and incorporating the construct into an expression vector. The vectors can be suitable for replication and integration in eukaryotes. Typical cloning vectors contain transcription and translation terminators, initiation sequences, and promoters useful for expression of the desired nucleic acid sequence.

In some embodiments, a cell may be genetically modified with one or more expression regions, e.g., a gene. In some embodiments, the cell may be genetically modified with an

exogenous gene (e.g., capable of expressing an exogenous gene product such as an RNA or a polypeptide product) and/or an exogenous regulatory nucleic acid. In some embodiments, the cell may be genetically modified with an exogenous sequence encoding a gene product that is endogenous to a target cell and/or an exogenous regulatory nucleic acid capable of modulating expression of an endogenous gene. In some embodiments, the cell may be genetically modified with an exogenous gene and/or a regulatory nucleic acid that modulates expression of an exogenous gene. In some embodiments, the cell may be genetically modified with an exogenous gene and/or a regulatory nucleic acid that modulates expression of an endogenous gene. It will be understood by one of skill in the art that the cell described herein may be genetically modified to express a variety of exogenous genes that encode proteins or regulatory molecules, which may, e.g., act on a gene product of the endogenous or exogenous genome of a target cell. In some embodiments, such genes confer characteristics to the fusosome, e.g., modulate fusion with a target cell. In some embodiments, the cell may be genetically modified to express an endogenous gene and/or regulatory nucleic acid. In some embodiments, the endogenous gene or regulatory nucleic acid modulates the expression of other endogenous genes. In some embodiments, the cell may be genetically modified to express an endogenous gene and/or regulatory nucleic acid which is expressed differently (e.g., inducibly, tissue-specifically, constitutively, or at a higher or lower level) than a version of the endogenous gene and/or regulatory nucleic acid on other chromosomes.

The promoter elements, e.g., enhancers, regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In the thymidine kinase (tk) promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either cooperatively or independently to activate transcription.

One example of a suitable promoter is the immediate early cytomegalovirus (CMV) promoter sequence. This promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto. Another example of a suitable promoter is Elongation Growth Factor-1α (EF-1α). However,

other constitutive promoter sequences may also be used, including, but not limited to the simian virus 40 (SV40) early promoter, mouse mammary tumor virus (MMTV), human immunodeficiency virus (HIV) long terminal repeat (LTR) promoter, MoMuLV promoter, an avian leukemia virus promoter, an Epstein-Barr virus immediate early promoter, a Rous sarcoma virus promoter, as well as human gene promoters such as, but not limited to, the actin promoter, the myosin promoter, the hemoglobin promoter, and the creatine kinase promoter.

Further, the invention should not be limited to the use of constitutive promoters. Inducible promoters are also contemplated as part of the invention. The use of an inducible promoter provides a molecular switch capable of turning on expression of the polynucleotide sequence which it is operatively linked when such expression is desired, or turning off the expression when expression is not desired. Examples of inducible promoters include, but are not limited to a tissue-specific promoter, metallothionine promoter, a glucocorticoid promoter, a progesterone promoter, and a tetracycline promoter. In some embodiments, expression of a fusogen is upregulated before fusosomes are generated, e.g., 3, 6, 9, 12, 24, 26, 48, 60, or 72 hours before fusosomes are generated.

The expression vector to be introduced into the source can also contain either a selectable marker gene or a reporter gene or both to facilitate identification and selection of expressing cells from the population of cells sought to be transfected or infected through viral vectors. In other aspects, the selectable marker may be carried on a separate piece of DNA and used in a co-transfection procedure. Both selectable markers and reporter genes may be flanked with appropriate regulatory sequences to enable expression in the host cells. Useful selectable markers include, for example, antibiotic-resistance genes, such as neo and the like.

Reporter genes may be used for identifying potentially transfected cells and for evaluating the functionality of regulatory sequences. In general, a reporter gene is a gene that is not present in or expressed by the recipient source and that encodes a polypeptide whose expression is manifested by some easily detectable property, e.g., enzymatic activity. Expression of the reporter gene is assayed at a suitable time after the DNA has been introduced into the recipient cells. Suitable reporter genes may include genes encoding luciferase, beta-galactosidase, chloramphenicol acetyl transferase, secreted alkaline phosphatase, or the green fluorescent protein gene (e.g., Ui-Tei et al., 2000 FEBS Letters 479: 79-82). Suitable expression

systems are well known and may be prepared using known techniques or obtained commercially. In general, the construct with the minimal 5' flanking region showing the highest level of expression of reporter gene is identified as the promoter. Such promoter regions may be linked to a reporter gene and used to evaluate agents for the ability to modulate promoter-driven transcription.

In some embodiments, a cell may be genetically modified to alter expression of one or more proteins. Expression of the one or more proteins may be modified for a specific time, e.g., development or differentiation state of the source. In some embodiments, fusosomes are generated from a source of cells genetically modified to alter expression of one or more proteins, e.g., fusogen proteins or non-fusogen proteins that affect fusion activity, structure or function. Expression of the one or more proteins may be restricted to a specific location(s) or widespread throughout the source.

In some embodiments, the expression of a fusogen protein is modified. In some embodiments, fusosomes are generated from cells with modified expression of a fusogen protein, e.g., an increase or a decrease in expression of a fusogen by at least 10%, 15%, 20%, 30%, 40%, 50%, 60%, 75%, 80%, 90% or more.

In some embodiments, cells may be engineered to express a cytosolic enzyme (e.g., proteases, phosphatases, kinases, demethylases, methyltransferases, acetylases) that targets a fusogen protein. In some embodiments, the cytosolic enzyme affects one or more fusogens by altering post-translational modifications. Post-translational protein modifications of proteins may affect responsiveness to nutrient availability and redox conditions, and protein–protein interactions. In some embodiments, a fusosome comprises fusogens with altered post-translational modifications, e.g., an increase or a decrease in post-translational modifications by at least 10%, 15%, 20%, 30%, 40%, 50%, 60%, 75%, 80%, 90% or more.

Methods of introducing a modification into a cell include physical, biological and chemical methods. See, for example, Geng. & Lu, Microfluidic electroporation for cellular analysis and delivery. Lab on a Chip. 13(19):3803-21. 2013; Sharei, A. et al. A vector-free microfluidic platform for intracellular delivery. PNAS vol. 110 no. 6. 2013; Yin, H. et al., Non-viral vectors for gene-based therapy. Nature Reviews Genetics. 15: 541–555. 2014. Suitable methods for modifying a cell for use in generating the fusosomes described herein include, for

example, diffusion, osmosis, osmotic pulsing, osmotic shock, hypotonic lysis, hypotonic dialysis, ionophoresis, electroporation, sonication, microinjection, calcium precipitation, membrane intercalation, lipid mediated transfection, detergent treatment, viral infection, receptor mediated endocytosis, use of protein transduction domains, particle firing, membrane fusion, freeze-thawing, mechanical disruption, and filtration.

Confirming the presence of a genetic modification includes a variety of assays. Such assays include, for example, molecular biological assays, such as Southern and Northern blotting, RT-PCR and PCR; biochemical assays, such as detecting the presence or absence of a particular peptide, e.g., by immunological means (ELISAs and Western blots) or by assays described herein.

The present disclosure provides, in some aspects, a fusosome comprising: (a) a lipid bilayer, (b) a lumen (e.g., comprising cytosol) surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen, e.g., wherein the fusogen is disposed in the lipid bilayer, wherein the fusosome is derived from a source cell; and wherein the fusosome has partial or complete nuclear inactivation (e.g., nuclear removal).

The present disclosure provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a nucleic acid, e.g., a nucleic acid comprising a payload gene; and wherein the fusosome does not comprise a nucleus; wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein; wherein: (i) when the plurality of fusosomes are contacted with a cell population comprising target cells and non-target cells, the cargo is present in at least 10-fold more target cells than non-target cells or reference cells, or (ii) the fusosomes of the plurality fuse at a higher rate with a target cell than with a non-target cell or reference cell by at least at least 50%; wherein the target cell is chosen from a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell.

The present disclosure provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by

the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a nucleic acid comprising a payload gene encoding an exogenous agent of Table 5 or Table 6, wherein the fusosome does not comprise a nucleus; and wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein.

The present disclosure provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a nucleic acid comprising a payload gene, wherein the nucleic acid comprises a NTCSRE operably linked to the payload gene, wherein the NTCSRE comprises a non- target cell-specific miRNA recognition sequence, e.g., a non-target cell-specific miRNA recognition sequence bound by a miRNA present in a non-target cell at a higher level than in a target cell, e.g., a non-target cell-specific miRNA recognition sequence bound by a miRNA of Table 4, wherein the target cell is a first type of CNS cell, optionally wherein the non-target cell is a second, different type of CNS cell or a non-CNS cell; and wherein the fusosome does not comprise a nucleus; and wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein.

In some embodiments, the miRNA is present in a non-target cell (e.g., a non-target cell described herein) at a level at least 10, 100, 1,000, or 10,000 times higher than the level of the miRNA present in the target cell (e.g., a CNS cell). In some embodiments, the miRNA is not detectably present in a target cell (e.g., a CNS cell, e.g., a CNS cell described herein). In some embodiments, the miRNA is not present in the target cell (e.g., a CNS cell, e.g., a CNS cell described herein).

The present disclosure provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a nucleic acid comprising a payload gene, wherein the nucleic acid comprises a promoter operably linked to the payload gene, wherein the promoter is a CNS cell-specific promoter, e.g., is a promoter specific for a CNS cell, a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell; wherein the fusosome

does not comprise a nucleus; and wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein.

The present disclosure provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a nucleic acid comprising a payload gene, wherein the nucleic acid comprises a promoter having sequence of a promoter in Table 3, or a sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto; wherein the fusosome does not comprise a nucleus; and wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein;

The present disclosure provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a nucleic acid comprising: (i) a payload gene; (ii) a NTCSRE operably linked to the payload gene, e.g., wherein the NTCSRE comprises a non-target cell-specific miRNA recognition sequence, e.g., a non-target cell-specific miRNA recognition sequence bound by a miRNA of Table 4, and

(iii) optionally, a positive target cell-specific regulatory element, e.g., a positive target cell-specific regulatory element (e.g., a target cell-specific promoter) operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a first type of CNS cell; optionally wherein the non-target cell is a second, different type of CNS cell or a non-CNS cell, optionally wherein: the target cell is a neuron and the non-target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglial cell), or the target cell is a glial cell (e.g., an oligodendrocyte, an astrocyte, or a microglial cell) and the non-target cell is a neuron; wherein the fusosome does not comprise a nucleus; and wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein.

In some embodiments, one or more of the following is present: i) the fusosome comprises or is comprised by a cytobiologic; ii) the fusosome comprises an enucleated cell; iii) the

fusosome comprises an inactivated nucleus; iv) the fusosome fuses at a higher rate with a target cell than with a non-target cell, e.g., by at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold, e.g., in an assay of Example 42; v) the fusosome fuses at a higher rate with a target cell than with other fusosomes, e.g., by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold, e.g., in an assay of Example 42; vi) the fusosome fuses with target cells at a rate such that an agent in the fusosome is delivered to at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, of target cells after 24, 48, or 72 hours, e.g., in an assay of Example 42; vii) the fusogen is present at a copy number of at least, or no more than, 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies, e.g., as measured by an assay of Example 26; viii) the fusosome comprises a therapeutic agent at a copy number of at least, or no more than, 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies, e.g., as measured by an assay of Example 88; ix) the ratio of the copy number of the fusogen to the copy number of the therapeutic agent is between 1,000,000:1 and 100,000:1, 100,000:1 and 10,000:1, 10,000:1 and 1,000:1, 1,000:1 and 100:1, 100:1 and 50:1, 50:1 and 20:1, 20:1 and 10:1, 10:1 and 5:1, 5:1 and 2:1, 2:1 and 1:1, 1:1 and 1:2, 1:2 and 1:5, 1:5 and 1:10, 1:10 and 1:20, 1:20 and 1:50, 1:50 and 1:100, 1:100 and 1:1,000, 1:1,000 and 1:10,000, 1:10,000 and 1:100,000, or 1:100,000 and 1:1,000,000; x) the fusosome comprises a lipid composition substantially similar to that of the source cell or wherein one or more of CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM and TAG is within 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, or 75% of the corresponding lipid level in the source cell; xi) the fusosome comprises a proteomic composition similar to that of the source cell, e.g., using an assay of Example 87; xii) the fusosome comprises a ratio of lipids to proteins that is within 10%, 20%, 30%, 40%, or 50% of the corresponding ratio in the source cell, e.g., as measured using an assay of Example 40; xiii) the fusosome comprises a ratio of proteins to nucleic acids (e.g., DNA) that is within 10%, 20%, 30%, 40%, or 50% of the corresponding ratio in the source cell, e.g., as measured using an assay of Example 41; xiv) the fusosome comprises a ratio of lipids to nucleic acids (e.g., DNA) that is within 10%, 20%, 30%, 40%, or 50% of the

corresponding ratio in the source cell, e.g., as measured using an assay of Example 91; xv) the fusosome has a half-life in a subject, e.g., in a mouse, that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% of the half life of a reference cell, e.g., the source cell, e.g., by an assay of Example 60; xvi) the fusosome transports glucose (e.g., labeled glucose, e.g., 2-NBDG) across a membrane, e.g., by at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% more (e.g., about 11.6% more) than a negative control, e.g., an otherwise similar fusosome in the absence of glucose, e.g., as measured using an assay of Example 50; xvii) the fusosome comprises esterase activity in the lumen that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of that of the esterase activity in a reference cell, e.g., the source cell or a mouse embryonic fibroblast, e.g., using an assay of Example 51; xviii) the fusosome comprises a metabolic activity level that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the citrate synthase activity in a reference cell, e.g., the source cell, e.g., as described in Example 53; xix) the fusosome comprises a respiration level (e.g., oxygen consumption rate) that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the respiration level in a reference cell, e.g., the source cell, e.g., as described in Example 54; xx) the fusosome comprises an Annexin-V staining level of at most 18,000, 17,000, 16,000, 15,000, 14,000, 13,000, 12,000, 11,000, or 10,000 MFI, e.g., using an assay of Example 55, or wherein the fusosome comprises an Annexin-V staining level at least 5%, 10%, 20%, 30%, 40%, or 50% lower than the Annexin-V staining level of an otherwise similar fusosome treated with menadione in the assay of Example 55, or wherein the fusosome comprises an Annexin-V staining level at least 5%, 10%, 20%, 30%, 40%, or 50% lower than the Annexin-V staining level of a macrophage treated with menadione in the assay of Example 55, xxi) the fusosome has a miRNA content level of at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater than that of the source cell, e.g., by an assay of Example 33; xxii) the fusosome has a soluble : non-soluble protein ratio is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater than that of the source cell, e.g., within 1%-2%, 2%-3%, 3%-4%, 4%-5%, 5%-10%, 10%-20%, 20%-30%, 30%-40%, 40%-50%, 50%-60%, 60%-70%, 70%-80%, or 80%-90% of that of the source cell, e.g., by an assay of Example 38; xxiii) the fusosome has an LPS level less than 5%, 1%, 0.5%, 0.01%, 0.005%, 0.0001%, 0.00001% or less of the LPS content of the source cell, e.g., as measured by mass spectrometry,

e.g., in an assay of Example 39; xxiv) the fusosome is capable of signal transduction, e.g., transmitting an extracellular signal, e.g., AKT phosphorylation in response to insulin, or glucose (e.g., labeled glucose, e.g., 2-NBDG) uptake in response to insulin, e.g., by at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% more than a negative control, e.g., an otherwise similar fusosome in the absence of insulin, e.g., using an assay of Example 49; xxv) the fusosome targets a tissue, e.g., liver, lungs, heart, spleen, pancreas, gastrointestinal tract, kidney, testes, ovaries, brain, reproductive organs, central nervous system, peripheral nervous system, skeletal muscle, endothelium, inner ear, or eye, when administered to a subject, e.g., a mouse, e.g., wherein at least 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the fusosomes in a population of administered fusosomes are present in the target tissue after 24, 48, or 72 hours, e.g., by an assay of Example 64; xxvi) the fusosome has juxtacrine-signaling level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% greater than the level of juxtacrine signaling induced by a reference cell, e.g., the source cell or a bone marrow stromal cell (BMSC), e.g., by an assay of Example 56; xxvii) the fusosome has paracrine-signaling level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% greater than the level of paracrine signaling induced by a reference cell, e.g., the source cell or a macrophage, e.g., by an assay of Example 57; xxviii) the fusosome polymerizes actin at a level within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% compared to the level of polymerized actin in a reference cell, e.g., the source cell or a C2C12 cell, e.g., by the assay of Example 58; xxix) the fusosome has a membrane potential within about 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% of the membrane potential of a reference cell, e.g., the source cell or a C2C12 cell, e.g., by an assay of Example 59, or wherein the fusosome has a membrane potential of about -20 to -150mV, -20 to -50mV, -50 to -100mV, or -100 to -150mV; xxx) the fusosome is capable of extravasation from blood vessels, e.g., at a rate at least 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% the rate of extravasation of the source cell or of a cell of the same type as the source cell, e.g., using an assay of Example 44, e.g., wherein the source cell is a neutrophil, lymphocyte, B cell, macrophage, or NK cell; xxxi) the fusosome is capable of crossing a cell membrane, e.g., an endothelial cell membrane or the blood brain barrier; xxxii) the fusosome is capable of secreting a protein, e.g., at a rate at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%,

80%, 90%, or 100% greater than a reference cell, e.g., a mouse embryonic fibroblast, e.g., using an assay of Example 48; xxxiii) the fusosome meets a pharmaceutical or good manufacturing practices (GMP) standard; xxxiv) the fusosome was made according to good manufacturing practices (GMP); xxxv) the fusosome has a pathogen level below a predetermined reference value, e.g., is substantially free of pathogens; xxxiv) the fusosome has a contaminant level below a predetermined reference value, e.g., is substantially free of contaminants; xxxvii) the fusosome has low immunogenicity, e.g., as described herein; xxxviii) the source cell is selected from a neutrophil, a granulocyte, a mesenchymal stem cell, a bone marrow stem cell, an induced pluripotent stem cell, an embryonic stem cell, a myeloblast, a myoblast, a hepatocyte, or a neuron e.g., retinal neuronal cell; or xxxix) the source cell is other than a 293 cell, HEK cell, human endothelial cell, or a human epithelial cell, monocyte, macrophage, dendritic cell, or stem cell.

The present disclosure also provides, in some aspects, a fusosome comprising: a) a lipid bilayer and a lumen that is miscible with an aqueous solution, e.g., water, wherein the fusosome is derived from a source cell, b) an exogenous or overexpressed fusogen disposed in the lipid bilayer, and c) an organelle, e.g., a therapeutically effective number of organelles, disposed in the lumen.

In some embodiments, one or more of the following is present: i) the source cell is selected from an endothelial cell, a macrophage, a neutrophil, a granulocyte, a leukocyte, a stem cell (e.g., a mesenchymal stem cell, a bone marrow stem cell, an induced pluripotent stem cell, an embryonic stem cell), a myeloblast, a myoblast, a hepatocyte, or a neuron e.g., retinal neuronal cell; ii) the organelle is selected from a Golgi apparatus, lysosome, endoplasmic reticulum, mitochondria, vacuole, endosome, acrosome, autophagosome, centriole, glycosome, glyoxysome, hydrogenosome, melanosome, mitosome, cnidocyst, peroxisome, proteasome, vesicle, and stress granule; iii) the fusosome has a size of greater than 5 um, 10 um, 20 um, 50 um, or 100 um; iv) the fusosome, or a composition or preparation comprising a plurality of the fusosomes, has a density of other than between 1.08 g/ml and 1.12 g/ml, e.g., the fusosome has a density of 1.25 g/ml +/- 0.05, e.g., as measured by an assay of Example 30; v) the fusosome is not captured by the scavenger system in circulation or by Kupffer cells in the sinus of the liver; vi) the source cell is other than a 293 cell; vii) the source cell is not transformed or immortalized; viii) the source cell is transformed, or immortalized using a method other than adenovirus-

mediated immortalization, e.g., immortalized by spontaneous mutation, or telomerase expression; ix) the fusogen is other than VSVG, a SNARE protein, or a secretory granule protein; x) the fusosome does not comprise Cre or GFP, e.g., EGFP; xi) the fusosome further comprises an exogenous protein other than Cre or GFP, e.g., EGFP xii) the fusosome further comprises an exogenous nucleic acid (e.g., RNA, e.g., mRNA, miRNA, or siRNA) or an exogenous protein (e.g., an antibody, e.g., an antibody), e.g., in the lumen; or xiii) the fusosome does not comprise mitochondria.

The present disclosure also provides, in some aspects, a fusosome comprising: (a) a lipid bilayer, (b) a lumen (e.g., comprising cytosol) surrounded by the lipid bilayer, (c) an exogenous or overexpressed fusogen, e.g., wherein the fusogen is disposed in the lipid bilayer, and (d) a functional nucleus, wherein the fusosome is derived from a source cell.

In some embodiments, one or more of the following is present: i) the source cell is other than a dendritic cell or tumor cell, e.g., the source cell is selected from an endothelial cell, a macrophage, a neutrophil, a granulocyte, a leukocyte, a stem cell (e.g., a mesenchymal stem cell, a bone marrow stem cell, an induced pluripotent stem cell, an embryonic stem cell), a myeloblast, a myoblast, a hepatocyte, or a neuron e.g., retinal neuronal cell; ii) the fusogen is other than a fusogenic glycoprotein; iii) the fusogen is a mammalian protein other than fertilin-beta, iv) the fusosome has low immunogenicity, e.g., as described herein; v) the fusosome meets a pharmaceutical or good manufacturing practices (GMP) standard; vi) the fusosome was made according to good manufacturing practices (GMP); vii) the fusosome has a pathogen level below a predetermined reference value, e.g., is substantially free of pathogens; or viii) the fusosome has a contaminant level below a predetermined reference value, e.g., is substantially free of contaminants.

The present disclosure also provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise:(a) a lipid bilayer,(b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a cargo; and wherein the fusosome does not comprise a nucleus; wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein; wherein the plurality of fusosomes, when contacted with a target cell population in the presence of an inhibitor of endocytosis, and when contacted with a reference target cell population not treated

with the inhibitor of endocytosis, delivers the cargo to at least 30% of the number of cells in the target cell population compared to the reference target cell population.

The present disclosure also provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, and wherein the fusosomes of the plurality comprise:(a) a lipid bilayer,(b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed re-targeted fusogen disposed in the lipid bilayer;(d) a cargo; and wherein the fusosome does not comprise a nucleus; wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein; wherein:(i) when the plurality of fusosomes are contacted with a cell population comprising target cells and non-target cells, the cargo is present in at least 2-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold more target cells than non-target cells, or (ii) the fusosomes of the plurality fuse at a higher rate with a target cell than with a non-target cell by at least at least 50%.

The present disclosure also provides, in some aspects, a fusosome composition comprising a plurality of fusosomes derived from a source cell, and wherein the fusosomes of the plurality comprise: (a) a lipid bilayer, (b) a lumen surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen, wherein the fusogen is disposed in the lipid bilayer; and (d) a cargo; wherein the fusosome does not comprise a nucleus; and wherein one or more of (e.g., at least 2, 3, 4, or 5 of): i) the fusogen is present at a copy number of at least 1,000 copies; ii) the fusosome comprises a therapeutic agent at a copy number of at least 1,000 copies; iii) the fusosome comprises a lipid wherein one or more of CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM and TAG is within 75% of the corresponding lipid level in the source cell; iv) the fusosome comprises a proteomic composition similar to that of the source cell; v) the fusosome is capable of signal transduction, e.g., transmitting an extracellular signal, e.g., AKT phosphorylation in response to insulin, or glucose (e.g., labeled glucose, e.g., 2-NBDG) uptake in response to insulin, e.g., by at least 10% more than a negative control, e.g., an otherwise similar fusosome in the absence of insulin; vi) the fusosome targets a tissue, e.g., liver, lungs, heart, spleen, pancreas, gastrointestinal tract, kidney, testes, ovaries, brain, reproductive organs, central nervous system, peripheral nervous system, skeletal muscle, endothelium, inner ear, or eye, when administered to a subject, e.g., a mouse, e.g., wherein at least 0.1%, or 10%, of the fusosomes in a population of administered fusosomes are present in the target tissue after 24 hours; or the source cell is selected from a neutrophil, a granulocyte, a

mesenchymal stem cell, a bone marrow stem cell, an induced pluripotent stem cell, an embryonic stem cell, a myeloblast, a myoblast, a hepatocyte, or a neuron e.g., retinal neuronal cell.

In embodiments, one or more of: i) the source cell is other than a 293 cell; ii) the source cell is not transformed or immortalized; iii) the source cell is transformed or immortalized using a method other than adenovirus-mediated immortalization, e.g., immortalized by spontaneous mutation or telomerase expression; iv) the fusogen is other than VSVG, a SNARE protein, or a secretory granule protein; v) the therapeutic agent is other than Cre or EGFP; vi) the therapeutic agent is a nucleic acid (e.g., RNA, e.g., mRNA, miRNA, or siRNA) or an exogenous protein (e.g., an antibody, e.g., an antibody), e.g., in the lumen; or vii) the fusosome does not comprise mitochondria.

In embodiments, one or more of: i) the source cell is other than a 293 or HEK cell; ii) the source cell is not transformed or immortalized; iii) the source cell is transformed or immortalized using a method other than adenovirus-mediated immortalization, e.g., immortalized by spontaneous mutation or telomerase expression; iv) the fusogen is not a viral fusogen; or v) the fusosome has a size of other than between 40 and 150 nm, e.g., greater than 150 nm, 200 nm, 300 nm, 400 nm, or 500 nm.

In embodiments, one or more of: i) the therapeutic agent is a soluble protein expressed by the source cell; ii) the fusogen is other than TAT, TAT-HA2, HA-2, gp41, Alzheimer's beta-amyloid peptide, a Sendai virus protein, or amphipathic net-negative peptide (WAE 11); iii) the fusogen is a mammalian fusogen; iv) the fusosome comprises in its lumen a polypeptide selected from an enzyme, antibody, or anti-viral polypeptide; v) the fusosome does not comprise an exogenous therapeutic transmembrane protein; or vi) the fusosome does not comprise CD63 or GLUT4, or the fusosome comprises less than or equal to 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, or 10% CD63 (e.g., about 0.048% or less), e.g., as determined according to the method described in Example 89.

In embodiments, the fusosome: i) does not comprise a virus, is not infectious, or does not propagate in a host cell; ii) is not a viral vector iii) is not a VLP (virus like particle); iv) does not comprise a viral structural protein, e.g., a protein derived from gag, e.g. a viral capsid protein, e.g. a viral capsule protein, e.g., a viral nucleocapsid protein, or wherein the amount of viral

capsid protein is less than 10%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.2%, or 0.1% of total protein, e.g., by mass spectrometry, e.g. using an assay of Example 93; v) does not comprise a viral matrix protein; vi) does not comprise a viral non-structural protein; e.g. pol or a fragment or variant thereof, a viral reverse transcriptase protein, a viral integrase protein, or a viral protease protein. vii) does not comprise viral nucleic acid; e.g. viral RNA or viral DNA; viii) comprises less than 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies per vesicle of a viral structural protein; or ix) the fusosome is not a virosome.

In some embodiments, the fusosome comprises (or is identified as comprising) less than about 0.01%, 0.05%, 0.1%, 0.5%, 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% viral capsid protein (e.g., about 0.05% viral capsid protein). In embodiments, the viral capsid protein is Complex of Rabbit Endogenous Lentivirus (RELIK) Capsid with Cyclophilin A. In embodiments, the viral capsid protein: total protein ratio is (or is identified as being) about 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, or 0.1.

In some embodiments, the fusosome does not comprise (or is identified as not comprising) a gag protein or a fragment or variant thereof, or the amount of gag protein or fragment or variant thereof is less than 10%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.2%, or 0.1% of total protein, e.g., by an assay of Example 93.

In embodiments, the ratio of the copy number of the fusogen to the copy number of viral structural protein on the fusosome is at least 1,000,000:1, 100,000:1, 10,000:1, 1,000:1, 100:1, 50:1, 20:1, 10:1, 5:1, or 1:1; or is between 100:1 and 50:1, 50:1 and 20:1, 20:1 and 10:1, 10:1 and 5:1 or 1:1. In embodiments, the ratio of the copy number of the fusogen to the copy number of viral matrix protein on the fusosome is at least 1,000,000:1, 100.000:1, 10,000:1, 1,000:1, 100:1, 50:1, 20:1, 10:1, 5:1, or 1:1.

In embodiments, one or more of: i) the fusosome does not comprise a water-immiscible droplet; ii) the fusosome comprises an aqueous lumen and a hydrophilic exterior; iii) the fusogen is a protein fusogen; or iv) the organelle is selected from a mitochondrion, Golgi apparatus, lysosome, endoplasmic reticulum, vacuole, endosome, acrosome, autophagosome, centriole, glycosome, glyoxysome, hydrogenosome, melanosome, mitosome, cnidocyst, peroxisome, proteasome, vesicle, and stress granule.

In embodiments, one or more of: i) the fusogen is a mammalian fusogen or a viral fusogen; ii) the fusosome was not made by loading the fusosome with a therapeutic or diagnostic substance;iii) the source cell was not loaded with a therapeutic or diagnostic substance; iv) the fusosome does not comprise doxorubicin, dexamethasone, cyclodextrin; polyethylene glycol, a micro RNA e.g., miR125, VEGF receptor, ICAM-1, E-selectin, iron oxide, a fluorescent protein e.g., GFP or RFP, a nanoparticle, or an RNase, or does not comprise an exogenous form of any of the foregoing; or v) the fusosome further comprises an exogenous therapeutic agent having one or more post-translational modifications, e.g., glycosylation.

In embodiments, the fusosome is unilamellar or multilamellar.

In embodiments, one or more of: i) the fusosome is not an exosome; ii) the fusosome is a microvesicle; iii) the fusosome comprises a non-mammalian fusogen; iv) the fusosome has been engineered to incorporate a fusogen; v) the fusosome comprises an exogenous fusogen; vi) the fusosome has a size of at least 80 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1200 nm, 1400 nm, or 1500 nm, or a population of fusosomes has an average size of at least 80 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1200 nm, 1400 nm, or 1500 nm; vii) the fusosome comprises one or more organelles, e.g., a mitochondrion, Golgi apparatus, lysosome, endoplasmic reticulum, vacuole, endosome, acrosome, autophagosome, centriole, glycosome, glyoxysome, hydrogenosome, melanosome, mitosome, cnidocyst, peroxisome, proteasome, vesicle, and stress granule; viii) the fusosome comprises a cytoskeleton or a component thereof, e.g., actin, Arp2/3, formin, coronin, dystrophin, keratin, myosin, or tubulin; ix) the fusosome, or a composition or preparation comprising a plurality of the fusosomes, does not have a flotation density of 1.08–1.22 g/ml, or has a density of at least 1.18–1.25 g/ml, or 1.05–1.12 g/ml, e.g., in a sucrose gradient centrifugation assay, e.g., as described in Théry et al., "Isolation and characterization of exosomes from cell culture supernatants and biological fluids." Curr Protoc Cell Biol. 2006 Apr; Chapter 3:Unit 3.22; x)the lipid bilayer is enriched for ceramides or sphingomyelins or a combination thereof compared to the source cell, or the lipid bilayer is not enriched (e.g., is depleted) for glycolipids, free fatty acids, or phosphatidylserine, or a combination thereof, compared to the source cell; xi) the fusosome comprises Phosphatidyl serine (PS) or CD40 ligand or both of PS and CD40 ligand, e.g., when measured in an assay of Example 92; xii) the fusosome is enriched for PS compared to the source cell, e.g., in a population of fusosomes at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% are positive for PS, e.g., by an

assay of Kanada M, et al. (2015) Differential fates of biomolecules delivered to target cells via extracellular vesicles. Proc Natl Acad Sci USA 112:E1433–E1442; xiii) the fusosome is substantially free of acetylcholinesterase (AChE), or contains less than 0.001, 0.002, 0.005, 0.01,0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500, or 1000 AChE activity units/ug of protein , e.g., by an assay of Example 52; xiv) the fusosome is substantially free of a Tetraspanin family protein (e.g., CD63, CD9, or CD81), an ESCRT-related protein (e.g., TSG101, CHMP4A-B, or VPS4B), Alix, TSG101, MHCI, MHCII, GP96, actinin-4, mitofilin, syntenin-1, TSG101, ADAM10, EHD4, syntenin-1, TSG101, EHD1, flotillin-1, heat-shock 70-kDa proteins (HSC70/HSP73, HSP70/HSP72), or any combination thereof, or contains less than 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 5%, or 10% of any individual exosomal marker protein and/or less than 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, or 25% of total exosomal marker proteins of any of said proteins, or is de-enriched for any one or more of these proteins compared to the source cell, or is not enriched for any one or more of these proteins, e.g., by an assay of Example 89; xv) the fusosome comprises a level of Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) that is below 500, 250, 100, 50, 20, 10, 5, or 1 ng GAPDH/ug total protein or below the level of GAPDH in the source cell, e.g., less than 1%, 2.5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, less than the level of GAPDH per total protein in ng/ug in the source cell, e.g., using an assay of Example 36; xvi) the fusosome is enriched for one or more endoplasmic reticulum proteins (e.g., calnexin), one or more proteasome proteins, or one or more mitochondrial proteins, or any combination thereof, e.g., wherein the amount of calnexin is less than 500, 250, 100, 50, 20, 10, 5, or 1 ng Calnexin / ug total protein, or wherein the fusosome comprises less Calnexin per total protein in ng/ug compared to the source cell by 1%, 2.5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, e.g., using an assay of Example 37 or 90, or wherein the average fractional content of Calnexin in the fusosome is less than about $1x10^{-4}$, $1.5x10^{-4}$, $2x10^{-4}$, $2.1x10^{-4}$, $2.2x10^{-4}$, $2.3x10^{-4}$,; $2.4x10^{-4}$, $2.43x10^{-4}$, $2.5x10^{-4}$, $2.6x10^{-4}$, $2.7x10^{-4}$, $2.8x10^{-4}$, $2.9x10^{-4}$, $3x10^{-4}$, $3.5x10^{-4}$, or $4x10^{-4}$, or wherein the fusosome comprises an amount of Calnexin per total protein that is lower than that of the parental cell by about 70%, 75%, 80%, 85%, 88%, 90%, 95%, 99%, or more; xvii) the fusosome comprises an exogenous agent (e.g., an exogenous protein, mRNA, or siRNA) e.g., as measured using an assay of Example 34; or xviii) the fusosome can be immobilized on a mica surface by atomic force microscopy for at least 30 min, e.g., by an assay

of Kanada M, et al. (2015) Differential fates of biomolecules delivered to target cells via extracellular vesicles. Proc Natl Acad Sci USA 112:E1433–E1442.

In embodiments, one or more of: i) the fusosome is an exosome; ii) the fusosome is not a microvesicle; iii) the fusosome has a size of less than 80 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1200 nm, 1400 nm, or 1500 nm, or a population of fusosomes has an average size of less than 80 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1200 nm, 1400 nm, or 1500 nm; iv) the fusosome does not comprise an organelle; v) the fusosome does not comprise a cytoskeleton or a component thereof, e.g., actin, Arp2/3, formin, coronin, dystrophin, keratin, myosin, or tubulin; vi) the fusosome, or a composition or preparation comprising a plurality of the fusosomes, has flotation density of 1.08–1.22 g/ml, e.g., in a sucrose gradient centrifugation assay, e.g., as described in Théry et al., "Isolation and characterization of exosomes from cell culture supernatants and biological fluids." Curr Protoc Cell Biol. 2006 Apr; Chapter 3:Unit 3.22; vii) the lipid bilayer is not enriched (e.g., is depleted) for ceramides or sphingomyelins or a combination thereof compared to the source cell, or the lipid bilayer is enriched for glycolipids, free fatty acids, or phosphatidylserine, or a combination thereof, compared to the source cell; viii) the fusosome does not comprise, or is depleted for relative to the source cell, Phosphatidyl serine (PS) or CD40 ligand or both of PS and CD40 ligand, e.g., when measured in an assay of Example 92; ix) the fusosome is not enriched (e.g., is depleted) for PS compared to the source cell, e.g., in a population of fusosomes less than 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% are positive for PS, e.g., by an assay of Kanada M, et al. (2015) Differential fates of biomolecules delivered to target cells via extracellular vesicles. Proc Natl Acad Sci USA 112:E1433–E1442; x) the fusosome comprises acetylcholinesterase (AChE), e.g. at least 0.001, 0.002, 0.005, 0.01,0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500, or 1000 AChE activity units/ug of protein , e.g., by an assay of Example 52; xi) the fusosome comprises a Tetraspanin family protein (e.g., CD63, CD9, or CD81), an ESCRT-related protein (e.g., TSG101, CHMP4A-B, or VPS4B), Alix, TSG101, MHCI, MHCII, GP96, actinin-4, mitofilin, syntenin-1, TSG101, ADAM10, EHD4, syntenin-1, TSG101, EHD1, flotillin-1, heat-shock 70-kDa proteins (HSC70/HSP73, HSP70/HSP72), or any combination thereof, e.g., contains more than 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 5%, or 10% of any individual exosomal marker protein and/or less than 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, or 25% of total exosomal marker proteins of any of said proteins, or is enriched for any one or more of these proteins compared to

the source cell, e.g., by an assay of Example 89; xii) the fusosome comprises a level of Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) that is above 500, 250, 100, 50, 20, 10, 5, or 1 ng GAPDH/ug total protein or below the level of GAPDH in the source cell, e.g., at least 1%, 2.5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, greater than the level of GAPDH per total protein in ng/ug in the source cell, e.g., using an assay of Example 36; xiii) the fusosome is not enriched for (e.g., is depleted for) one or more endoplasmic reticulum proteins (e.g., calnexin), one or more proteasome proteins, or one or more mitochondrial proteins, or any combination thereof, e.g., wherein the amount of calnexin is less than 500, 250, 100, 50, 20, 10, 5, or 1 ng Calnexin / ug total protein, or wherein the fusosome comprises less Calnexin per total protein in ng/ug compared to the source cell by 1%, 2.5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, e.g., using an assay of Example 90, or wherein the average fractional content of Calnexin in the fusosome is less than about $1x10^{-4}$, $1.5x10^{-4}$, $2x10^{-4}$, $2.1x10^{-4}$, $2.2x10^{-4}$, $2.3x10^{-4}$,; $2.4x10^{-4}$, $2.43x10^{-4}$, $2.5x10^{-4}$, $2.6x10^{-4}$, $2.7x10^{-4}$, $2.8x10^{-4}$, $2.9x10^{-4}$, $3x10^{-4}$, $3.5x10^{-4}$, or $4x10^{-4}$, or wherein the fusosome comprises an amount of Calnexin per total protein that is lower than that of the parental cell by about 70%, 75%, 80%, 85%, 88%, 90%, 95%, 99%, or more; or xiv) the fusosome can not be immobilized on a mica surface by atomic force microscopy for at least 30 min, e.g., by an assay of Kanada M, et al. (2015) Differential fates of biomolecules delivered to target cells via extracellular vesicles. Proc Natl Acad Sci USA 112:E1433–E1442.

In embodiments, one or more of: i) the fusosome does not comprise a VLP; ii) the fusosome does not comprise a virus; iii) the fusosome does not comprise a replication-competent virus; iv) the fusosome does not comprise a viral protein, e.g., a viral structural protein, e.g., a capsid protein or a viral matrix protein; v) the fusosome does not comprise a capsid protein from an enveloped virus; vi) the fusosome does not comprise a nucleocapsid protein; or vii) the fusogen is not a viral fusogen.

In embodiments, the fusosome comprises cytosol.

In embodiments, one or more of: i) the fusosome or the source cell does not form a teratoma when implanted into subject, e.g., by an assay of Example 65; ii) the fusosome is capable of chemotaxis, e.g., of within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than a reference cell, e.g., a macrophage, e.g., using an assay of Example 45; iii) the fusosome is capable of homing, e.g., at the site of an injury, wherein the

fusosome or cytobiologic is from a human cell, e.g., using an assay of Example 46, e.g., wherein the source cell is a neutrophil; or iv) the fusosome is capable of phagocytosis, e.g., wherein phagocytosis by the fusosome is detectable within 0.5, 1, 2, 3, 4, 5, or 6 hours in using an assay of Example 47, e.g., wherein the source cell is a macrophage.

In embodiments, the fusosome or fusosome composition retains one, two, three, four, five, six or more of any of the characteristics for 5 days or less, e.g., 4 days or less, 3 days or less, 2 days or less, 1 day or less, e.g., about 12-72 hours, after administration into a subject, e.g., a human subject.

In embodiments, the fusosome has one or more of the following characteristics: a) comprises one or more endogenous proteins from a source cell, e.g., membrane proteins or cytosolic proteins; b) comprises at least 10, 20, 50, 100, 200, 500, 1000, 2000, or 5000 different proteins; c) comprises at least 1, 2, 5, 10, 20, 50, or 100 different glycoproteins; d) at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% by mass of the proteins in the fusosome are naturally-occurring proteins; e) comprises at least 10, 20, 50, 100, 200, 500, 1000, 2000, or 5000 different RNAs; or f) comprises at least 2, 3, 4, 5, 10, or 20 different lipids, e.g., selected from CL, Cer, DAG, HexCer, LPA, LPC, LPE, LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM and TAG.

In embodiments, the fusosome has been manipulated to have, or the fusosome is not a naturally occurring cell and has, or wherein the nucleus does not naturally have one, two, three, four, five or more of the following properties: a) the partial nuclear inactivation results in a reduction of at least 50%, 60%, 70%, 80%, 90% or more in nuclear function, e.g., a reduction in transcription or DNA replication, or both, e.g., wherein transcription is measured by an assay of Example 24 and DNA replication is measured by an assay of Example 25; b) the fusosome is not capable of transcription or has transcriptional activity of less than 1%, 2.5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of that of the transcriptional activity of a reference cell, e.g., the source cell, e.g., using an assay of Example 24; c) the fusosome is not capable of nuclear DNA replication or has nuclear DNA replication of less than 1%, 2.5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the nuclear DNA replication of a reference cell, e.g., the source cell, e.g., using an assay of Example 25; d) the fusosome lacks chromatin or has a chromatin content of less than 1%, 2.5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the of the chromatin content of a reference cell, e.g., the source cell, e.g., using an assay

of Example 32; e) the fusosome lacks a nuclear membrane or has less than 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, or 1% the amount of nuclear membrane of a reference cell, e.g., the source cell or a Jurkat cell, e.g., by an assay of Example 31; f) the fusosome lacks functional nuclear pore complexes or has reduced nuclear import or export activity, e.g., by at least 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, or 1% by an assay of Example 31, or the fusosome lacks on or more of a nuclear pore protein, e.g., NUP98 or Importin 7; g) the fusosome does not comprise histones or has histone levels less than 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the histone level of the source cell (e.g., of H1, H2a, H2b, H3, or H4), e.g., by an assay of Example 32; h) the fusosome comprises less than 20, 10, 5, 4, 3, 2, or 1 chromosome; i) nuclear function is eliminated; j) the fusosome is an enucleated mammalian cell; k) the nucleus is removed or inactivated, e.g., extruded by mechanical force, by radiation or by chemical ablation; or l) the fusosome is from a mammalian cell having DNA that is completely or partially removed, e.g., during interphase or mitosis.

In embodiments, the fusosome comprises mtDNA or vector DNA. In embodiments, the fusosome does not comprise DNA.

In embodiments, the source cell is a primary cell, immortalized cell or a cell line (e.g., myelobast cell line, e.g., C2C12). In embodiments, the fusosome is from a source cell having a modified genome, e.g., having reduced immunogenicity (e.g., by genome editing, e.g., to remove an MHC protein or MHC complexes). In embodiments, the source cell is from a cell culture treated with an anti-inflammatory signal. In embodiments, the source cell is from a cell culture treated with an immunosuppressive agent. In embodiments, the source cell is substantially non-immunogenic, e.g., using an assay described herein. In embodiments, the source cell comprises an exogenous agent, e.g., a therapeutic agent. In embodiments, the source cell is a recombinant cell.

In embodiments, the fusosome further comprises an exogenous agent, e.g., a therapeutic agent, e.g., a protein or a nucleic acid (e.g., a DNA, a chromosome (e.g. a human artificial chromosome), an RNA, e.g., an mRNA or miRNA). In embodiments, the exogenous agent is present at at least, or no more than, 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies, e.g., comprised by the fusosome, or is present at an average level of at least, or no more than, 10, 20, 50, 100, 200, 500, 1,000, 2,000,

5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000 or 1,000,000 copies per fusosome. In embodiments, the fusosome has an altered, e.g., increased or decreased level of one or more endogenous molecules, e.g., protein or nucleic acid, e.g., due to treatment of the mammalian cell with a siRNA or gene editing enzyme. In embodiments, the endogenous molecule is present at, e.g. an average level, of at least, or no more than, 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies (e.g., copies comprised by the fusosome), or is present at an average level of at least, or no more than, 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000 or 1,000,000 copies per fusosome. In embodiments, the endogenous molecule (e.g., an RNA or protein) is present at a concentration of at least 1, 2, 3, 4, 5, 10, 20, 50, 100, 500, 10^3, 5.0 x 10^3, 10^4, 5.0 x 10^4, 10^5, 5.0 x 10^5, 10^6, 5.0 x 10^6, 1.0 x 10^7, 5.0 x 10^7, or 1.0 x 10^8, greater than its concentration in the source cell.

In embodiments, the active agent is selected from a protein, protein complex (e.g., comprising at least 2, 3, 4, 5, 10, 20, or 50 proteins, e.g., at least at least 2, 3, 4, 5, 10, 20, or 50 different proteins) polypeptide, nucleic acid (e.g., DNA, chromosome, or RNA, e.g., mRNA, siRNA, or miRNA) or small molecule. In embodiments, the exogenous agent comprises a site-specific nuclease, e.g., Cas9 molecule, TALEN, or ZFN.

In embodiments, the fusogen is a viral fusogen, e.g., HA, HIV-1 ENV, HHV-4, gp120, or VSV-G. In embodiments, the fusogen is a mammalian fusogen, e.g., a SNARE, a Syncytin, myomaker, myomixer, myomerger, or FGFRL1. In embodiments, the fusogen is active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10. In embodiments, the fusogen is not active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10. In embodiments, the fusosome fuses to a target cell at the surface of the target cell. In embodiments, the fusogen promotes fusion in a lysosome-independent manner. In embodiments, the fusogen is a protein fusogen. In embodiments, the fusogen is a lipid fusogen, e.g., oleic acid, glycerol mono-oleate, a glyceride, diacylglycerol, or a modified unsaturated fatty acid. In embodiments, the fusogen is a chemical fusogen, e.g., PEG. In embodiments, the fusogen is a small molecule fusogen, e.g., halothane, an NSAID such as meloxicam, piroxicam, tenoxicam, and chlorpromazine. In embodiments, the fusogen is recombinant. In embodiments, the fusogen is biochemically incorporated, e.g., the fusogen is provided as a purified protein and contacted with a lipid bilayer under conditions that allow for associate of the fusogen with the

lipid bilayer. In embodiments, the fusogen is biosynthetically incorporated, e.g. expressed in a source cell under conditions that allow the fusogen to associate with the lipid bilayer.

In embodiments, the fusosome binds a target cell. In embodiments, the target cell is other than a HeLa cell, or the target cell is not transformed or immortalized.

In some embodiments involving fusosome compositions, the plurality of fusosomes are the same. In some embodiments, the plurality of fusosomes are different. In some embodiments the plurality of fusosomes are from one or more source cells. In some embodiments at least 50%, 60%, 70%, 80%, 90%, 95%, or 99% of fusosomes in the plurality have a diameter within 10%, 20%, 30%, 40%, or 50% of the mean diameter of the fusosomes in the fusosome composition. In some embodiments at least 50%, 60%, 70%, 80%, 90%, 95%, or 99% of fusosomes in the plurality have a volume within 10%, 20%, 30%, 40%, or 50% of the mean volume of the fusosomes in the fusosome composition. In some embodiments, the fusosome composition has less than about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, variability in size distribution within 10%, 50%, or 90% of the source cell population variability in size distribution, e.g., based on Example 28. In some embodiments, at least 50%, 60%, 70%, 80%, 90%, 95%, or 99% of fusosomes in the plurality have a copy number of the fusogen within 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the mean fusogen copy number in the fusosomes in the fusosome composition. In some embodiments, at least 50%, 60%, 70%, 80%, 90%, 95%, or 99% of fusosomes in the plurality have a copy number of the therapeutic agent within 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the mean therapeutic agent copy number in the fusosomes in the fusosome composition. In some embodiments, the fusosome composition comprises at least 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, or 10^{15} or more fusosomes. In some embodiments, the fusosome composition is in a volume of at least 1 ul, 2 ul, 5 ul, 10 ul, 20 ul, 50 ul, 100 ul, 200 ul, 500 ul, 1 ml, 2 ml, 5 ml, or 10 ml.

In some embodiments, the fusosome composition delivers the cargo to at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% of the number of cells in the target cell population compared to the reference target cell population.

In some embodiments, the fusosome composition delivers at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo to the target cell population compared to the reference target cell population or to a non-target cell population. In some embodiments, the fusosome composition delivers at least 40%, 45%, 50%,

55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% more of the cargo to the target cell population compared to the reference target cell population or to a non-target cell population.

In some embodiments, less than 10% of cargo enters the cell by endocytosis.

In some embodiments, the inhibitor of endocytosis is an inhibitor of lysosomal acidification, e.g., bafilomycin A1. In some embodiments, the inhibitor of endocytosis is a dynamin inhibitor, e.g., Dynasore.

In some embodiments, the target cell population is at a physiological pH (e.g., between 7.3-7.5, e.g., between 7.38-7.42).

In some embodiments, the cargo delivered is determined using an endocytosis inhibition assay, e.g., an assay of Example 80.

In some embodiments, cargo enters the cell through a dynamin-independent pathway or a lysosomal acidification-independent pathway, a macropinocytosis-independent pathway (e.g., wherein the inhibitor of endocytosis is an inhibitor of macropinocytosis, e.g., 5-(N-ethyl-N-isopropyl)amiloride (EIPA), e.g., at a concentration of 25 μM), or an actin-independent pathway (e.g., wherein the inhibitor of endocytosis is an inhibitor of actin polymerization is, e.g., Latrunculin B, e.g., at a concentration of 6 μM).

In some embodiments, the fusosomes of the plurality further comprise a targeting moiety. In embodiments, the targeting moiety is comprised by the fusogen or is comprised by a separate molecule.

In some embodiments, when the plurality of fusosomes are contacted with a cell population comprising target cells and non-target cells, the cargo is present in at least 10-fold more target cells than non-target cells.

In some embodiments, when the plurality of fusosomes are contacted with a cell population comprising target cells and non-target cells, the cargo is present at least 2-fold, 5-fold, 10-fold, 20-fold, or 50-fold higher in target cells than non-target cells and/or the cargo is present at least 2-fold, 5-fold, 10-fold, 20-fold, or 50-fold higher in target cells than reference cells.

In some embodiments, the fusosomes of the plurality fuse at a higher rate with a target cell than with a non-target cell by at least 50%.

In some embodiments, the fusosome, when contacted with a target cell population, delivers cargo to a target cell location other than an endosome or lysosome, e.g., to the cytosol.

In embodiments, less 50%, 40%, 30%, 20%, or 10% of the cargo is delivered to an endosome or lysosome.

In some embodiments, the fusosomes of the plurality comprise exosomes, microvesicles, or a combination thereof.

In some embodiments, the plurality of fusosomes has an average size of at least 50 nm, 100 nm, 200 nm, 500 nm, 1000 nm, 1200 nm, 1400 nm, or 1500 nm. In other embodiments, the plurality of fusosomes has an average size of less than 100 nm, 80 nm, 60 nm, 40 nm, or 30 nm.

In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a mammalian fusogen. In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a viral fusogen. In some embodiments, the fusogen (e.g., re-targeted fusogen) is a protein fusogen. In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a sequence chosen from a Nipah virus protein F, a measles virus F protein, a tupaia paramyxovirus F protein, a paramyxovirus F protein, a Hendra virus F protein, a Henipavirus F protein, a Morbilivirus F protein, a respirovirus F protein, a Sendai virus F protein, a rubulavirus F protein, or an avulavirus F protein, or a derivative thereof.

In some embodiments, the fusogen (e.g., re-targeted fusogen) is active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10. In some embodiments, the fusogen (e.g., re-targeted fusogen) is not active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10.

In some embodiments, the fusogen is present at a copy number of at least 1, 2, 5, or 10 copies per fusosome.

In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a Nipah virus protein G, a measles protein H, a tupaia paramyxovirus H protein, a paramyxovirus G protein, a paramyxovirus H protein, a paramyxovirus HN protein, a Morbilivirus H protein, a respirovirus HN protein, a sendai HN protein, a rubulavirus HN protein, an avulavirus HN protein, or a derivative thereof. In some embodimetns, the fusogen (e.g., re-targeted fusogen) comprises a sequence chosen from Nipah virus F and G proteins, measles virus F and H proteins, tupaia paramyxovirus F and H proteins, paramyxovirus F and G proteins or F and H proteins or F and HN proteins, Hendra virus F and G proteins, Henipavirus F and G proteins, Morbilivirus F and H proteins, respirovirus F and HN protein, a Sendai virus F and HN protein, rubulavirus F and HN proteins, or avulavirus F and HN proteins, or a derivative thereof, or any combination thereof.

In some embodiments, the cargo comprises an exogenous protein or an exogenous nucleic acid. In some embodiments, the cargo comprises or encodes a cytosolic protein. In some embodiments the cargo comprises or encodes a membrane protein. In some embodiments, the cargo comprises a therapeutic agent. In some embodiments, the cargo is present at a copy number of at least 1, 2, 5, 10, 20, 50, 100, or 200 copies per fusosome (e.g., up to about 1,000 copies per fusosome). In some embodiments, the ratio of the copy number of the fusogen (e.g., re-targeted fusogen) to the copy number of the cargo is between 1000:1 and 1:1, or between 500:1 and 1:1 or between 250:1 and 1:1, or between 150:1 and 1:1, or between 100:1 and 1:1, or between 75:1 and 1:1 or between 50:1 and 1:1 or between 25:1 and 1:1 or between 20:1 and 1:1 or between 15:1 and 1:1 or between 10:1 and 1:1 or between 5:1 and 1:1 or between 2:1 and 1:1 or between 1:1 and 1:2.

In some embodiments, the fusosome composition comprises a viral capsid protein or a DNA integration polypeptide. In some embodiments, the cargo comprises a viral genome.

In some embodiments, the fusosome composition is capable of delivering a nucleic acid to a target cell, e.g., to stably modify the genome of the target cell, e.g., for gene therapy.

In some embodiments, the fusosome composition does not comprise a viral nucleocapsid protein, or the amount of viral nucleocapside protein is less than 10%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.2%, or 0.1% of total protein, e.g., by mass spectrometry, e.g. using an assay of Example 93.

In embodiments, the fusosome composition comprises at least 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, 10^{14}, or 10^{15} fusosomes. In embodiments, the fusosome composition comprises at least 10 ml, 20 ml, 50 ml, 100 ml, 200 ml, 500 ml, 1 L, 2 L, 5 L, 10 L, 20 L, or 50 L.

In embodiments, the fusosome is from a mammalian cell having a modified genome, e.g., to reduce immunogenicity (e.g., by genome editing, e.g., to remove an MHC protein or MHC complexes). In embodiments, the source cell is from a cell culture treated with an anti-inflammatory signal. In embodiments, the method further comprises contacting the source cell of step a) with an immunosuppressive agent or anti-inflammatory signal, e.g., before or after inactivating the nucleus, e.g., enucleating the cell.

In one aspect, provided herein is a fusosome composition comprising a plurality of fusosomes derived from a source cell, wherein the fusosomes of the plurality comprise: (a) a

lipid bilayer, (b) a lumen comprising cytosol, wherein the lumen is surrounded by the lipid bilayer; (c) an exogenous or overexpressed fusogen disposed in the lipid bilayer, (d) a cargo; and wherein the fusosome does not comprise a nucleus; wherein the amount of viral capsid protein in the fusosome composition is less than 1% of total protein; wherein the plurality of fusosomes, when contacted with a target cell population in the presence of an inhibitor of endocytosis, and when contacted with a reference target cell population not treated with the inhibitor of endocytosis, delivers the cargo to at least 30% of the number of cells in the target cell population compared to the reference target cell population.

In embodiments, the fusosome composition delivers the cargo to at least 40%, 50%, 60%, 70%, or 80% of the number of cells in the target cell population compared to the reference target cell population or to a non-target cell population; or delivers the cargo, e.g., at least 40%, 50%, 60%, 70%, or 80% of the cargo, to the target cell population compared to the reference target cell population or to a non-target cell population. In embodiments, less than 10% of cargo enters the cell by endocytosis. In embodiments, the inhibitor of endocytosis is an inhibitor of lysosomal acidification, e.g., bafilomycin A1. In embodiments, cargo delivered is determined using an endocytosis inhibition assay, e.g., an assay of Example 80. In embodiments, cargo enters the cell through a dynamin-independent pathway or a lysosomal acidification-independent pathway, a macropinocytosis-independent pathway (e.g., wherein the inhibitor of endocytosis is an inhibitor of macropinocytosis, e.g., 5-(N-ethyl-N-isopropyl)amiloride (EIPA), e.g., at a concentration of 25 μM), or an actin-independent pathway (e.g., wherein the inhibitor of endocytosis is an inhibitor of actin polymerization is, e.g., Latrunculin B, e.g., at a concentration of 6 μM).

C. Fusogens and pseudotyping

In some embodiments, the fusosome described herein (e.g., comprising a vesicle or a portion of a cell) includes one or more fusogens, e.g., to facilitate the fusion of the fusosome to a membrane, e.g., a cell membrane. Also these compositions may include surface modifications made during or after synthesis to include one or more fusogens. The surface modification may comprise a modification to the membrane, e.g., insertion of a lipid or protein into the membrane.

In some embodiments, the fusosomes comprise one or more fusogens on their exterior surface (e.g., integrated into the cell membrane) to target a specific cell or tissue type (e.g.,CNS

cell). In some embodiments, the specific cell type targeted by the one or more fusogens is a CNS cell, a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell. Fusosomes may comprise a targeting domain. Fusogens include without limitation protein based, lipid based, and chemical based fusogens. The fusogen may bind a partner, e.g., a feature on a target cells' surface. In some embodiments the partner on a target cells' surface is a target cell moiety. In particular embodiments, a fusogen is a fusogen or a re-targeted fusogen that binds to a target cell from among a CNS cell a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell. In some embodiments, the fusosome comprising the fusogen will integrate the membrane into a lipid bilayer of a target cell. In some embodiments, one or more of the fusogens described herein may be included in the fusosome.

The fusosomes (e.g., retroviral vectors) described herein can comprise a fusogen, e.g., an endogenous fusogen or a pseudotyped fusogen.

i) Protein fusogens

In some embodiments, the fusogen comprises a protein (e.g., glycoprotein), lipid, or small molecule. A fusogen can be, for instance, a mammalian fusogen or a viral fusogen. In some embodiments, the fusogen is a protein fusogen, e.g., a mammalian protein or a homologue of a mammalian protein (e.g., having 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or greater identity), a non-mammalian protein such as a viral protein or a homologue of a viral protein (e.g., having 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or greater identity), a native protein or a derivative of a native protein, a synthetic protein, a fragment thereof, a variant thereof, a protein fusion comprising one or more of the fusogens or fragments, and any combination thereof. In some embodiments, a viral fusogen is a Class I viral membrane fusion protein, a Class II viral membrane protein, a Class III viral membrane fusion protein, a viral membrane glycoprotein, or other viral fusion proteins, or a homologue thereof, a fragment thereof, a variant thereof, or a protein fusion comprising one or more proteins or fragments thereof.

In embodiments, the fusogen is a viral fusogen, e.g., HA, HIV-1 ENV, HHV-4, gp120, or VSV-G. In embodiments, the fusogen is a mammalian fusogen, e.g., a SNARE, a Syncytin,

myomaker, myomixer, myomerger, or FGFRL1. In embodiments, the fusogen is active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10. In embodiments, the fusogen is not active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10. In embodiments, the fusosome fuses to a target cell at the surface of the target cell. In embodiments, the fusogen promotes fusion in a lysosome-independent manner. In embodiments, the fusogen is a protein fusogen. In embodiments, the fusogen is a lipid fusogen, e.g., oleic acid, glycerol mono-oleate, a glyceride, diacylglycerol, or a modified unsaturated fatty acid. In embodiments, the fusogen is a chemical fusogen, e.g., PEG. In embodiments, the fusogen is a small molecule fusogen, e.g., halothane, an NSAID such as meloxicam, piroxicam, tenoxicam, and chlorpromazine. In embodiments, the fusogen is recombinant. In embodiments, the fusogen is biochemically incorporated, e.g., the fusogen is provided as a purified protein and contacted with a lipid bilayer under conditions that allow for associate of the fusogen with the lipid bilayer. In embodiments, the fusogen is biosynthetically incorporated, e.g. expressed in a source cell under conditions that allow the fusogen to associate with the lipid bilayer.

In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a mammalian fusogen. In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a viral fusogen. In some embodiments, the fusogen (e.g., re-targeted fusogen) is a protein fusogen. In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a sequence chosen from a Nipah virus protein F, a measles virus F protein, a tupaia paramyxovirus F protein, a paramyxovirus F protein, a Hendra virus F protein, a Henipavirus F protein, a Morbilivirus F protein, a respirovirus F protein, a Sendai virus F protein, a rubulavirus F protein, or an avulavirus F protein, or a derivative thereof.

In some embodiments, the fusogen (e.g., re-targeted fusogen) is active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10. In some embodiments, the fusogen (e.g., re-targeted fusogen) is not active at a pH of 4-5, 5-6, 6-7, 7-8, 8-9, or 9-10.

In some embodiments, the fusogen is present at a copy number of at least 1, 2, 5, or 10 copies per fusosome.

In some embodiments, the fusogen (e.g., re-targeted fusogen) comprises a Nipah virus protein G, a measles protein H, a tupaia paramyxovirus H protein, a paramyxovirus G protein, a paramyxovirus H protein, a paramyxovirus HN protein, a Morbilivirus H protein, a respirovirus HN protein, a sendai HN protein, a rubulavirus HN protein, an avulavirus HN protein, or a derivative thereof. In some embodimetns, the fusogen (e.g., re-targeted fusogen) comprises a

sequence chosen from Nipah virus F and G proteins, measles virus F and H proteins, tupaia paramyxovirus F and H proteins, paramyxovirus F and G proteins or F and H proteins or F and HN proteins, Hendra virus F and G proteins, Henipavirus F and G proteins, Morbilivirus F and H proteins, respirovirus F and HN protein, a Sendai virus F and HN protein, rubulavirus F and HN proteins, or avulavirus F and HN proteins, or a derivative thereof, or any combination thereof.

Non-mammalian fusogens include viral fusogens, homologues thereof, fragments thereof, and fusion proteins comprising one or more proteins or fragments thereof. Viral fusogens include class I fusogens, class II fusogens, class III fusogens, and class IV fusogens. In embodiments, class I fusogens such as human immunodeficiency virus (HIV) gp41, have a characteristic postfusion conformation with a signature trimer of α-helical hairpins with a central coiled-coil structure. Class I viral fusion proteins include proteins having a central postfusion six-helix bundle. Class I viral fusion proteins include influenza HA, parainfluenza F, HIV Env, Ebola GP, hemagglutinins from orthomyxoviruses, F proteins from paramyxoviruses (e.g. Measles, (Katoh et al. BMC Biotechnology 2010, 10:37)), ENV proteins from retroviruses, and fusogens of filoviruses and coronaviruses. In embodiments, class II viral fusogens such as dengue E glycoprotein, have a structural signature of β- sheets forming an elongated ectodomain that refolds to result in a trimer of hairpins. In embodiments, the class II viral fusogen lacks the central coiled coil. Class II viral fusogen can be found in alphaviruses (e.g., E1 protein) and flaviviruses (e.g., E glycoproteins). Class II viral fusogens include fusogens from Semliki Forest virus, Sinbis, rubella virus, and dengue virus. In embodiments, class III viral fusogens such as the vesicular stomatitis virus G glycoprotein, combine structural signatures found in classes I and II. In embodiments, a class III viral fusogen comprises α helices (e.g., forming a six-helix bundle to fold back the protein as with class I viral fusogens), and β sheets with an amphiphilic fusion peptide at its end, reminiscent of class II viral fusogens. Class III viral fusogens can be found in rhabdoviruses and herpesviruses. In embodiments, class IV viral fusogens are fusion-associated small transmembrane (FAST) proteins (doi:10.1038/sj.emboj.7600767, Nesbitt, Rae L., "Targeted Intracellular Therapeutic Delivery Using Liposomes Formulated with Multifunctional FAST proteins" (2012). Electronic Thesis and Dissertation Repository. Paper 388), which are encoded by nonenveloped reoviruses. In embodiments, the class IV viral fusogens are sufficiently small that they do not form hairpins (doi: 10.1146/annurev-cellbio-101512-122422, doi:10.1016/j.devcel.2007.12.008).

Fusogens, which include viral envelope proteins (env), generally determine the range of host cells which can be infected and transformed by fusosomes. In the case of lentiviruses, such as HIV-1, HIV-2, SIV, FIV and EIV, the native env proteins include gp41 and gp120. In some embodiments, the viral env proteins expressed by source cells described herein are encoded on a separate vector from the viral gag and pol genes, as has been previously described.

Illustrative examples of retroviral-derived env genes which can be employed include, but are not limited to: MLV envelopes, 10A1 envelope, BAEV, FeLV-B, RD114, SSAV, Ebola, Sendai, FPV (Fowl plague virus), and influenza virus envelopes. Similarly, genes encoding envelopes from RNA viruses (e.g., RNA virus families of Picornaviridae, Calciviridae, Astroviridae, Togaviridae, Flaviviridae, Coronaviridae, Paramyxoviridae, Rhabdoviridae, Filoviridae, Orthomyxoviridae, Bunyaviridae, Arenaviridae, Reoviridae, Birnaviridae, Retroviridae) as well as from the DNA viruses (families of Hepadnaviridae, Circoviridae, Parvoviridae, Papovaviridae, Adenoviridae, Herpesviridae, Poxyiridae, and Iridoviridae) may be utilized. Representative examples include, FeLV, VEE, HFVW, WDSV, SFV, Rabies, ALV, BIV, BLV, EBV, CAEV, SNV, ChTLV, STLV, MPMV, SMRV, RAV, FuSV, MH2, AEV, AMV, CT10, and EIAV.

In some embodiments, envelope proteins for display on a fusosome include, but are not limited to any of the following sources: Influenza A such as H1N1, H1N2, H3N2 and H5N1 (bird flu), Influenza B, Influenza C virus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Hepatitis D virus, Hepatitis E virus, Rotavirus, any virus of the Norwalk virus group, enteric adenoviruses, parvovirus, Dengue fever virus, Monkey pox, Mononegavirales, Lyssavirus such as rabies virus, Lagos bat virus, Mokola virus, Duvenhage virus, European bat virus 1 & 2 and Australian bat virus, Ephemerovirus, Vesiculovirus, Vesicular Stomatitis Virus (VSV), Herpesviruses such as Herpes simplex virus types 1 and 2, varicella zoster, cytomegalovirus, Epstein-Bar virus (EBV), human herpesviruses (HHV), human herpesvirus type 6 and 8, Human immunodeficiency virus (HIV), papilloma virus, murine gammaherpesvirus, Arenaviruses such as Argentine hemorrhagic fever virus, Bolivian hemorrhagic fever virus, Sabia-associated hemorrhagic fever virus, Venezuelan hemorrhagic fever virus, Lassa fever virus, Machupo virus, Lymphocytic choriomeningitis virus (LCMV), Bunyaviridiae such as Crimean-Congo hemorrhagic fever virus, Hantavirus, hemorrhagic fever with renal syndrome causing virus, Rift Valley fever virus, Filoviridae (filovirus) including Ebola hemorrhagic fever and Marburg

hemorrhagic fever, Flaviviridae including Kaysanur Forest disease virus, Omsk hemorrhagic fever virus, Tick-borne encephalitis causing virus and Paramyxoviridae such as Hendra virus and Nipah virus, variola major and variola minor (smallpox), alphaviruses such as Venezuelan equine encephalitis virus, eastern equine encephalitis virus, western equine encephalitis virus, SARS-associated coronavirus (SARS-CoV), West Nile virus, any encephaliltis causing virus.

In some embodiments, a source cell described herein produces a fusosome, e.g., recombinant retrovirus, e.g., lentivirus, pseudotyped with the VSV-G glycoprotein.

A fusosome or pseudotyped virus generally has a modification to one or more of its envelope proteins, e.g., an envelope protein is substituted with an envelope protein from another virus. For example, HIV can be pseudotyped with a fusion protein from rhabdovirus, e.g., vesicular stomatitis virus G-protein (VSV-G) envelope proteins, which allows HIV to infect a wider range of cells because HIV envelope proteins (encoded by the env gene) normally target the virus to CD4+ presenting cells. In some embodiments, lentiviral envelope proteins are pseudotyped with VSV-G. In one embodiment, source cells produce recombinant retrovirus, e.g., lentivirus, pseudotyped with the VSV-G envelope glycoprotein.

Furthermore, a fusogen or viral envelope protein can be modified or engineered to contain polypeptide sequences that allow the transduction vector to target and infect host cells outside its normal range or more specifically limit transduction to a cell or tissue type. For example, the fusogen or envelope protein can be joined in-frame with targeting sequences, such as receptor ligands, antibodies (using an antigen-binding portion of an antibody or a recombinant antibody-type molecule, such as a single chain antibody), and polypeptide moieties or modifications thereof (e.g., where a glycosylation site is present in the targeting sequence) that, when displayed on the transduction vector coat, facilitate directed delivery of the virion particle to a target cell of interest. Furthermore, envelope proteins can further comprise sequences that modulate cell function. Modulating cell function with a transducing vector may increase or decrease transduction efficiency for certain cell types in a mixed population of cells. For example, stem cells could be transduced more specifically with envelope sequences containing ligands or binding partners that bind specifically to stem cells, rather than other cell types that are found in the blood or bone marrow. Non-limiting examples are stem cell factor (SCF) and Flt-3 ligand. Other examples include, e.g., antibodies (e.g., single-chain antibodies that are

specific for a cell-type), and essentially any antigen (including receptors) that binds tissues as lung, liver, pancreas, heart, endothelial, smooth, breast, prostate, epithelial, vascular cancer, etc.

Protein fusogens or viral envelope protein may be re-targeted by mutating amino acid residues in a fusion protein or a targeting protein (e.g. the hemagglutinin protein). In some embodiments the fusogen is randomly mutated. In some embodiments the fusogen is rationally mutated. In some embodiments the fusogen is subjected to directed evolution. In some embodiments the fusogen is truncated and only a subset of the peptide is used in the retroviral vector or VLP. For example, amino acid residues in the measles hemagglutinin protein may be mutated to alter the binding properties of the protein, redirecting fusion (doi:10.1038/nbt942, Molecular Therapy vol. 16 no. 8, 1427–1436 Aug. 2008, doi:10.1038/nbt1060, DOI: 10.1128/JVI.76.7.3558–3563.2002, DOI: 10.1128/JVI.75.17.8016–8020.2001, doi: 10.1073pnas.0604993103).

In some embodiments, the protein fusogen or viral envelope protein is re-targeted by i) mutating amino acid resides in the natural fusogen protein sequence or viral envelope protein sequence and/or ii) engineering the fusogen protein or viral envelope protein to contain polypeptide sequences that allow the fusogen or viral envelope protein to target and fuse or infect host cells outside its normal range.

In some embodiments, the fusosomes comprise one or more fusogens on their exterior surface (e.g., integrated into the cell membrane) to target a specific cell or tissue type. Fusogens include without limitation protein based, lipid based, and chemical based fusogens. The fusogen may bind a partner on a target cells' surface. In some embodiments, the fusosome comprising the fusogen will integrate the membrane into a lipid bilayer of a target cell.

In some embodiments the fusogen is a paramyxovirus fusogen. In some embodiments the fusogen is a Nipah virus protein F, a measles virus F protein, a tupaia paramyxovirus F protein, a paramyxovirus F protein, a Hendra virus F protein, a Henipavirus F protein, a Morbilivirus F protein, a respirovirus F protein, a Sendai virus F protein, a rubulavirus F protein, or an avulavirus F protein.

In some embodiments, the fusogen is a poxviridae fusogen.

Additional exemplary fusogens are disclosed in US 9,695,446, US 2004/0028687, US 6,416,997, US 7,329,807, US 2017/0112773, US 2009/0202622, WO 2006/027202, and US 2004/0009604, the entire contents of all of which are hereby incorporated by reference.

In some embodiments, a fusogen described herein comprises an amino acid sequence of Table 1, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 100, 200, 300, 400, 500, or 600 amino acids in length. For instance, in some embodiments, a fusogen described herein comprises an amino acid sequence having at least 80% identity to any amino acid sequence of Table 1. In some embodiments, a nucleic acid sequence described herein encodes an amino acid sequence of Table 1, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 40, 50, 60, 80, 100, 200, 300, 400, 500, or 600 amino acids in length.

In some embodiments, a fusogen described herein comprises an amino acid sequence set forth in any one of SEQ ID NOS: 1-57, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 100, 200, 300, 400, 500, or 600 amino acids in length. For instance, in some embodiments, a fusogen described herein comprises an amino acid sequence having at least 80% identity to an amino acid sequence set forth in any one of SEQ ID NOS: 1-57. In some embodiments, a nucleic acid sequence described herein encodes an amino acid sequence set forth in any one of SEQ ID NOS: 1-57, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 40, 50, 60, 80, 100, 200, 300, 400, 500, or 600 amino acids in length.

Table 1. **Paramyxovirus F sequence clusters. Column 1, Genbank ID includes the Genbank ID of the whole genome sequence of the virus that is the centroid sequence of the cluster. Column 2, Nucleotides of CDS provides the nucleotides corresponding to the CDS of the gene in the whole genome. Column 3, Full Gene Name, provides the full name of the gene including Genbank ID, virus species, strain, and protein name. Column 4, Sequence, provides the amino acid sequence of the gene. Column 5, #Sequences/Cluster, provides the number of sequences that cluster with this centroid sequence.**

Genbank ID	Nucleotides of CDS	Full Gene Name	Sequence	#Sequences /Cluster	SEQ ID NO

KP317927	5630-7399	gb:KP317927:5630-7399|Organism:Human respiratory syncytial virus|Strain Name:Kilifi_9465_7_RSVB_2011|Protein Name:fusion glycoprotein|Gene Symbol:F	MIPQARTELNLGQITMELLIHRSSAIFL TLAINALYLTSSQNITEEFYQSTCSAVS RGYLSALRTGWYTSVITIELSNIKETK CNGTDTKVKLIKQELDKYKNAVTELQ LLMQNTPAANNRARREAPQYMNYTI NTTGSLNVSISKKRKRRFLGFLLGVGS AIASGIAVSKVLHLEGEVNKIKNALLS TNKAVVSLSNGVSVLTSKVLDLKNYI NNQLLPIVNQQSCRISNIETVIEFQQKN SRLLEITREFSVNAGVTTPLSTYMLTN SELLSLINDMPITNDQKKLMSSNVQIV RQQSYSIMSIIKEEVLAYVVQLPIYGVI DTPCWKLHTSPLCTTNIKEGSNICLTR TDRGWYCDNAGSVSFFPQADTCKVQ SNRVFCDTMNSLTLPSEVSLCNTDIFN SKYDCKIMTSKTDISSSVITSLGAIVSC YGKTKCTASNKNRGIIKTFSNGCDYV SNKGVDTVSVGNTLYYVNKLEGKNL YVKGEPIINYYDPLVFPSDEFDASISQV NEKINQSLAFIRRSDELLHNVNTGKST TNIMITAIIIVIIVVLLSLIAIGLLLYCKA KNTPVTLSKDQLSGINNIAFSK	993	1
AB524405	4556-6217	gb:AB524405:4556-6217|Organism:Newcastle disease virus|Strain Name:Goose/Alaska/415/91|Protein Name:fusion protein|Gene Symbol:F	MDPKPSTSYLHAFPLIFVAISLVFMAG RASALDGRPLAAAGIVVTGDKAVNIY TSSQTGTIIIKLLPNMPKDKEQCAKSPL DAYNRTLTTLLAPLGDSIRRIQESVTT SGGERQERLVGAIIGGVALGVATAAQ ITAASALIQANQNAANILKLKESIAAT NEAVHEVTSGLSQLAVAVGKMQQFV NDQFNKTAQEIDCIKITQQVGVELNLY LTELTTVFGPQITSPALTQLTIQALYNL AGGNMDYMLTKLGVGNNQLSSLISS GLISGNPILYDSQTQLLGIQVTLPSVGN LNNMRATYLETLSVSTNKGFASALVP KVVTQVGSVIEELDTSYCIETDLDLYC TRIVTFPMSPGIFSCLGGNTSACMYSK TEGALTTPYMTLKGSVIANCKMTTCR CADPPGIISQNYGEAVSLIDKKVCNILT LDGITLRLSGEFDATYQKNISIQDSQV VITGNLDISTELGNVNNSISNALDKLE ESNSKLDKVNVRLTSTSALITYIVLTTI ALICGIVSLVLACYIMYKQKAQQKTL LWLGNNTLDQMRATTKM	418	2
AF266286	4875-7247	gb:AF266286:4875-7247|Organism:Measles virus strain AIK-C|Strain Name:Measles virus strain Edmonston (AIK-C	MSIMGLKVNVSAIFMAVLLTLQTPTG QIHWGNLSKIGVVGIGSASYKVMTRS SHQSLVIKLMPNITLLNNCTRVEIAEY RRLLRTVLEPIRDALNAMTQNIRPVQS VASSRRHKRFAGVVLAGAALGVATA AQITAGIALHQSMLNSQAIDNLRASLE TTNQAIEAIRQAGQEMILAVQGVQDY INNELIPSMNQLSCDLIGQKLGLKLLR YYTEILSLFGPSLRDPISAEISIQALSYA LGGDINKVLEKLGYSGGDLLGILESRG IKARITHVDTESYFIVLSIAYPTLSEIKG VIVHRLEGVSYNIGSQEWYTTVPKYV ATQGYLISNFDESSCTFMPEGTVCSQN	128	3

		vaccine)|Protein Name:fusion protein|Gene Symbol:F	ALYPMSPLLQECLRGYTKSCARTLVS GSFGNRFILSQGNLIANCASILCKCYTT GTIINQDPDKILTYIAADNCPVVEVNG VTIQVGSRRYPDAVYLHRIDLGPPILL ERLDVGTNLGNAIAKLEDAKELLESS DQILRSMKGLSSTCIVYILIAVCLGGLI GIPALICCCRGRCNKKGEQVGMSRPG LKPDLTGTSKSYVRSL		
AB503857	3068-4687	gb:AB503857:3068-4687|Organism:Human metapneumovirus|Strain Name:Jpn03-1|Protein Name:fusion glycoprotein precursor|Gene Symbol:F	MSWKVVIIFSLLITPQHGLKESYLEES CSTITEGYLSVLRTGWYTNVFTLEVG DVENLTCSDGPSLIKTELDLTKSALRE LKTVSADQLAREEQIEKPRQSRFVLG AIALGVATAAAVTAGVAIAKTIRLESE VTAIKNALKTTNEAVSTLGNGVRVLA TAVRELKDFVSKNLTRAINKNKCDID DLKMAVSFSQFNRRFLNVVRQFSDNA GITPAISLDLMTDAELARAVSNMPTSA GQIKLMLENRAMVRRKGFGILIGVYG SSVIYMVQLPIFGVIDTPCWIVKAAPS CSEKKGNYACLLREDQGWYCQNAGS TVYYPNEKDCETRGDHVFCDTAAGIN VAEQSKECNINISTTNYPCKVSTGRHP ISMVALSPLGALVACYKGVSCSIGSNR VGIIKQLNKGCSYITNQDADTVTIDNT VYQLSKVEGEQHVIKGRPVSSSFDPIK FPEDQFNVALDQVFENIENSQALVDQ SNRILSSAEKGNTGFIIVIILIAVLGSSM ILVSIFIIIKKTKKPTGAPPELSGVTNNG FIPHS	125	4
EU277658	5078-6700	gb:EU277658:5078-6700|Organism:Bovine parainfluenza virus 3|Strain Name:Q592|Protein Name:fusion protein|Gene Symbol:F	MIIIVITMILSLTPSSLCQIDITKLQSVG VLVNSPKGIKISQNFETRYLILSLIPKIE DSHSCGNQQIDQYKKLLDRLIIPLYDG LKLQKDVIVVNHESHNNTNLRTKRFF GEIIGTIAIGIATSAQITAAVALVEAKQ ARSDIDKLKEAIKDTNKAVQSIQSSVG NLIVAVKSVQDYVNNEIVPSITRLGCE AAGLQLGIALTQHYSELTNIFGDNIGT LGEKGVKLQGIASLYRTNITEVFTTST VDQYDIYDLLFTESIKMRVIDVDLSDY SITLQVRLPLLTKVSNTQIYKVDSISYN IQGKEWYIPLPHHIMTKGAFLGGADIK ECIESFSNYICPSDPGFILNHEMENCLS GNITQCPKTIVTSDIVPRYAFVDGGVI ANCIPTTCTCNGIDNRINQSPDQGIKIIT YKECQIVGINGMLFKTNQEGTLAKYT FDNIKLNNSVALNPIDISLELNKAKSD LEESKRWIEKSNQKLDSIGSWHQSSVT IIIIIVMIVVLLIINAIIIMIMIRYLRDRNR HLNNKDSEPYVLTNRQ	93	5
AB040874	4546-6162	gb:AB040874:4546-6162|Organism:Mumps virus|Strain	MKVFLVTCLGFAVFSSSVCVNINILQQ IGYIKQQVRQLSYYSQSSSSYIVVKLL PNIQPTDNSCEFKSVTQYNKTLSNLLL PIAENINNIASPSSGSRRHKRFAGIAIGI AALGVATAAQVTAAVSLVQAQTNAR AIAAMKNSIQATNRAVFEVKEGTQRL AIAVQAIQDHINTIMNTQLNNMSCQIL	89	6

		Name:Miyahara|Protein Name:fusion protein|Gene Symbol:F	DNQLATSLGLYLTELTTVFQPQLINPA LSPISIQALRSLLGSMTPAVVQATLSTS ISAAEILSAGLMEGQIVSVLLDEMQMI VKINIPTIVTQSNALVIDFYSISSFINNQ ESIIQLPDRILEIGNEQWSYPAKNCKLT RHHIFCQYNEAERLSLESKLCLAGNIS ACVFSPIAGSYMRRFVALDGTIVANC RSLTCLCKSPSYPIYQPDHHAVTTIDL TACQTLSLDGLDFSIVSLSNITYAENLT ISLSQTINTQPIDISTELSKVNASLQNA VKYIKESNHQLQSVNVNSKIGAIIVAA LVLSILSIIISLLFCCWAYVATKEIRRIN FKTNHINTISSSVDDLIRY		
AB475097	4908-6923	gb:AB475097:4908-6923|Organism:Canine distemper virus|Strain Name:M25CR|Protein Name:fusion protein|Gene Symbol:F	MNPHEQTIPMHEKIPKRSKTQTHTQQ DLPQQHSTKSAESKTSRARHSITSAQR STHYDPRTADWPDYYIMKRTRSCKQ ASYRSDNIPAHGDHDGIIHHTPESVSQ GAKSRLKMGQSNAVKSGSQCTWLVL WCIGVASLFLCSKAQIHWNNLSTIGII GTDSVHYKIMTRPSHQYLVIKLMPNV SLIDNCTKAELDEYEKLLSSILEPINQA LTLMTKNVKPLQSVGSGRRQRRFAG VVLAGAALGVATAAQITAGIALHQSN LNAQAIQSLRTSLEQSNKAIEEIREATQ ETVIAVQGVQDYVNNELVPAMQHMS CELVGQRLGLKLLRYYTELLSIFGPSL RDPISAEISIQALSYALGGEIHKILEKL GYSGNDMIAILESRGIKTKITHVDLPG KFIILSVSYPTLSEVKGVIVHRLEAVSY NIGSQEWYTTVPRYVATNGYLISNFD ESSCVFVSESAICSQNSLYPMSPLLQQ CIRGDTSSCARTLVSGTMGNKFILSKG NIVANCASILCKCYSTSTIINQSPDKLL TFIASDTCPLVEIDGVTIQVGSRQYPD MVYESKVALGPAISLERLDVGTNLGN ALKKLDDAKVLIDSSNQILETVRRSSF NFGSLLSVPILSCTALALLLLICCCKRR YQQTHKQNTKVDPTFKPDLTGTSRSY VRSL	46	7
AJ849636	5526-7166	gb:AJ849636:5526-7166|Organism:Peste-des-petits-ruminants virus|Strain Name:Turkey 2000|Protein Name:fusion protein|Ge	MTRVAILTFLFLFPNAVACQIHWGNL SKIGIVGTGSASYKVMTRPSHQTLVIK LMPNITAIDNCTKSEIAEYKRLLITVLK PVEDALSVITKNVRPIQTLTPGRRTRR FAGAVLAGVALGVATAAQITAGVAL HQSLMNSQAIESLKTSLEKSNQAIEEIR LANKETILAVQGVQDYINNELVPSVH RMSCELVGHKLGLKLLRYYTEILSIFG PSLRDPIAAEISIQALSYALGGDINRIL DKLGYSGGDFLAILESKGIKARVTYV DTRDYFIILSIAYPTLSEIKGVIVHKIEA ITYNIGAQEWYTTIPKYVATQGYLISN FDETSCVFTPDGTVCSQNALYPMSPLL QECFQGSTKSCARTLVSGTISNRFILSK GNLIANCASVLCKCYTTETVISQDPDK LLTVVASDKCPVVEVDGVTIQVGSRE YPDSVYLHKIDLGPAISLEKLDVGTNL	34	8

		ne Symbol:F	GNAVTRLENAKELLDASDQILKTVKG VPFGGNMYIALAACIGVSLGLVTLICC CKGRCKNKEVPISKINPGLKPDLTGTS KSYVRSL		
AF017149	6618-8258	gb:AF017149|Organism:Hendra virus|Strain Name:UNKNOWN-AF017149|Protein Name:fusion|Gene Symbol:F	MATQEVRLKCLLCGIIVLVLSLEGLGI LHYEKLSKIGLVKGITRKYKIKSNPLT KDIVIKMIPNVSNVSKCTGTVMENYK SRLTGILSPIKGAIELYNNNTHDLVGD VKLAGVVMAGIAIGIATAAQITAGVA LYEAMKNADNINKLKSSIESTNEAVV KLQETAEKTVYVLTALQDYINTNLVP TIDQISCKQTELALDLALSKYLSDLLF VFGPNLQDPVSNSMTIQAISQAFGGN YETLLRTLGYATEDFDDLLESDSIAGQ IVYVDLSSYYIIVRVYFPILTEIQQAYV QELLPVSFNNDNSEWISIVPNFVLIRNT LISNIEVKYCLITKKSVICNQDYATPM TASVRECLTGSTDKCPRELVVSSHVPR FALSGGVLFANCISVTCQCQTTGRAIS QSGEQTLLMIDNTTCTTVVLGNIIISLG KYLGSINYNSESIAVGPPVYTDKVDIS SQISSMNQSLQQSKDYIKEAQKILDTV NPSLISMLSMIILYVLSIAALCIGLITFIS FVIVEKKRGNYSRLDDRQVRPVSNGD LYYIGT	29	9
AB005795	4866-6563	gb:AB005795:4866-6563|Organism:Sendai virus|Strain Name:Ohita|Protein Name:fusion protein|Gene Symbol:F	MATYIQRVQCISALLSVVLTTLVSCQI PRDRLSNIGVIVDEGKSLKIAGSHESR YIVLSLVPGIDLENGCGTAQVIQYKSL LNRLLIPLRDALDLQEALITVTNDTMT GADVPQSRFFGAVIGTIALGVATSAQI TAGIALAEAREAKRDIALIKESMTKTH KSIELLQNAVGEQILALKTLQDFVNDE IKPAISELGCETAALRLGIKLTQHYSEL LTAFGSNFGTIGEKSLTLQALSSLYSA NITEIMTTIRTGQSNIYDVIYTEQIKGT VIDVDLERYMVTLSVKIPILSEVPGVLI HKASSISYNIDGEEWYVTVPSHILSRA SFLGGANIADCVESRLTYICPRDPAQLI PDSQQKCILGDTTRCPVTKVVDNIIPK FAFVNGGVVANCIASTCTCGTGRRPIS QDRSKGVVFLTHDNCGLIGVNGIELY ANRKGHDATWGVQNLTVGPAIAIRPV DISLNLAAATDFLQDSRAELEKARKIL SEVGRWYNSGATLITIIVVMIVVLVVII VIVIVLYRLRRSMLMSNPAGRISRDTY TLEPKIRHMYTNGGFDAMTEKR	23	10
AF457102	5088-6755	gb:AF457102|Organism:Human parainfluenza virus 1 strain Washington/1964|Strain	MQKSEILFLVYSSLLLSSSLCQIPVEKL SNVGVIINEGKLLKIAGSYESRYIVLSL VPSIDLQDGCGTTQIIQYKNLLNRLLIP LKDALDLQESLITITNDTTVTNDNPQT RFFGAVIGTIALGVATAAQITAGIALA EAREARKDIALIKDSIVKTHNSVELIQ RGIGEQIIALKTLQDFVNDEIRPAIGEL RCETTALKLGIKLTQHYSELATAFSSN LGTIGEKSLTLQALSSLYSANITEILST TKKDKSDIYDIIYTEQVKGTVIDVDLE	21	11

		Name:Washington 1964|Protein Name:F glycoprotein|Gene Symbol:F	KYMVTLLVKIPILSEIPGVLIYRASSISY NIEGEEWHVAIPNYIINKASSLGGADV TNCIESKLAYICPRDPTQLIPDNQQKCI LGDVSKCPVTKVINNLVPKFAFINGG VVANCIASTCTCGTNRIPVNQDRSRG VTFLTYTNCGLIGINGIELYANKRGRD TTWGNQIIKVGPAVSIRPVDISLNLAS ATNFLEESKTELMKARAIISAVGGWH NTESTQIIMIIIVCILIIIICGILYYLYRVR RLLVMINSTHNSPVNAYTLESRMRNP YMGNNSN		
AB910309	4951-6582	gb:AB910309:4951-6582|Organism:Feline morbillivirus|Strain Name:SS1|Protein Name:fusion protein|Gene Symbol:F	MGKIRVIIISSLLLSNITTAQVGWDNLT SIGVISTKQYDYKITTLNTNQLMVIKM VPNISSIINCTKPELMKYRELVLGVIRP INESLELMNSYINMRAGSERFIGAVIA GVALGVATAAQITSGIALHNSIMNKR QIQELRKALSTTNKAIDEIRIAGERTLI AVQGVQDYINNIIIPMQDKLQCDILSS QLAIALLRYYTNILTVFGPSIRDPVTSII SIQALSQAFNGNLQALLDGLGYTGRD LRDLLESRSITGQIIHADMTDLFLVLRI NYPSITEMQGVTIYELNSITYHIGPEE WYTIMPNFIAVQGFLTSNFDERKCSIT KSSILCQQNSIYPMSTEMQRCIKGEIRF CPRSKAVGTLVNRFILTKGNLMANCL GVICRCYSSGQIITQDPSKLITIISQEEC KEVGVDGIRIMVGPRKLPDVIFNARLE VGVPISLSKLDVGTDLAIASAKLNNSK ALLEQSDKILDSMSKLDSINSRITGLIL AIMAIFIITVTIIWIIYKRCRNKDNKFST SIEPLYIPPSYNSPHSVVKSI	12	12
KT071755	4310-6070	gb:KT071755:4310-6070|Organism:Avian paramyxovirus 2|Strain Name:APMV-2/Procarduelis nipalensis/China/Suiling/53/2013|Protein Name:fusion protein|Gene Symbol:F	MIAALFISLFATCGALDNSVLAPVGIA SAQEWQLAAYTNTLSGTIAVRFVPVL PGNLSTCAQATLAEYNKTVTNILGPL KENLETLLSEPTKTAARFVGAIIGTVA LGVATSAQITAAVALNQAQENARNIW RLKESIRKTNEAVLELKDGLASTAIAL DKVQKFINEDIIPQIKEIDCQVVANKL GVYLSLYLTELTTIFGAQITNPALTPLS YQALYNLCGGDMGKLTELIGVKAKDI NSLYEANLITGQVIGYDSESQIILIQVS YPSVSEVTGVRATELVTVSVTTPKGE GRAIAPKYVAQSRVVTEELDTSTCRFS KTTLYCRSIITRPLPPLIANCLNGLYQD CQYTTEIGALSSRFITVNGGIIANCRAT ICKCVNPPKIIVQSDASSLTVIDSAICK DVVLDNVQLRLEGKLSAQYFTNITIDL SQITTSGSLDISSEIGSINNTVNKVEELI AESNAWLQAVNPHLVNNTSIIVLCVL AAIFVVWLVALTGCLAYYIKKSSATR MVGIGSSPAGNPYVAQSATKM	12	13
AY029299	4598-6265	gb:AY029299|Organism:Avian paramyxo	MGARLGPLAMAPGRYVIIFNLILLHRV VSLDNSRLLQQGIMSATEREIKVYTNS ITGSIAVRLIPNLPQEVLKCSAGQIKSY NDTLNRIFTPIKANLERLLATPSMLED	11	14

		virus 6|Strain Name:APMV-6/duck/Taiwan/Y1/98|Protein Name:fusion protein|Gene Symbol:F	NQNPAPEPRLIGAIIGTAALGLATAAQVTAALALNQAQDNAKAILNLKESITKTNEAVLELKDATGQIAIALDKTQRFINDNILPAINNLTCEVAGAKVGVELSLYLTELSTVFGSQITNPALSTLSIQALMSLCGNDFNYLLNLMGAKHSDLGALYEANLINGRIIQYDQASQIMVIQVSVPSISSISGLRLTELFTLSIETPVGEGKAVVPQFVVESGQLLEEIDTQACTLTDTTAYCTIVRTKPLPELVAQCLRGDESRCQYTTGIGMLESRFGVFDGLVIANCKATICRCLAPEMIITQNKGLPLTVISQETCKRILIDGVTLQIEAQVSGSYSRNITVGNSQIAPSGPLDISSELGKVNQSLSNVEDLIDQSNQLLNRVNPNIVNNTAIIVTIVLLVLLVLWCLALTISILYVSKHAVRMIKTVPNPYVMQAKSPGSATQF		
AY141760	5028-6665	gb:AY141760|Organism:Fer-de-Lance paramyxovirus|Strain Name:ATCC VR-895|Protein Name:fusion protein F|Gene Symbol:F	MTRITILQIILTLTLPVMCQVSFDNLEQVGVMFDKPKFLKITGPASTATMIIKLIPTLGTMESCGTSAVNEYKKTLDTILVPLRDTINKLSTDITVVEGTSNISNKREKRFVGIAIAVGAVALATSAQITAGIALSNTIKNAEAIESIKSSIQASNQAIQKVIDAQGRTVTVINGIQDHINSVINPALNQLGCDVAKNTLAISLTQYFSKLSLLFGPNLRNPVEQPLSVQAIAGLMDGDINAVVSQLGYTQSDLLDLLSTESIVGTVTAIDMVNYMIQIEMSFPQYITIPDTKVLEGHKITFNDKGSEWQTQVPSTIAVRDILIAGVDPDGCSITSTSYICKNDPTYAMSEVLTNCFRGNTQECPRARITSTFATRFAIARSTVIANCVAAVCLCGDPGIPVVQKAEVTLTAMTLDQCSLITVDGLQIKPSKSIANVTANFGNITLGPVVSVGDLDLSAELTKVQSDLKEAQDKLDESNAILQGINNKILTAPTSIALIVVSVVVILLIIGMISWLVWLTKAVRRSNTRSERVTPSAYNNLGFIK	8	15
EU877976	4330-6410	gb:EU877976:4330-6410|Organism:Avian paramyxovirus 4|Strain Name:APMV-4/KR/YJ/06|Protein Name:fusion protein|Gene Symbol:F	MRLSRTILTLILGTLTGYLMGAHSTNVNEGPKSEGIRGDLIPGAGIFVTQVRQLQIYQQSGYHDLVIRLLPLLPAELNDCQREVVTEYNNTVSQLLQPIKTNLDTLLADGGTRDADIQPRFIGAIIATGALAVATVAEVTAAQALSQSKTNAQNILKLRDSIQATNQAVFEISQGLEATATVLSKLQTELNENIIPSLNNLSCAAMGNRLGVSLSLYLTLMTTLFGDQITNPVLTPISYSTLSAMAGGHIGPVMSKILAGSVTSQLGAEQLIASGLIQSQVVGYDSQYQLLVIRVNLVRIQEVQNTRVVSLRTLAVNRDGGLYRAQVPPEVVERSGIAERFYADDCVLTTTDYICSSIRSSRLNPELVKCLSGALDSCTFERESALLSTPFFVYNKAVVANCKAATCRCNKPPSIIAQYSASALVTITTDTCADLEIEGYRFNIQTESNSWVAPNFTVSTSQIVSVDPIDISSDIAKINSSIEAA	8	16

			REQLELSNQILSRINPRIVNDESLIAIIV TIVVLSLLVIGLIVVLGVMYKNLKKV QRAQAAMMMQQMSSSQPVTTKLGTP F		
AB176531	4793-6448	gb:AB176531:4793-6448|Organism:Human parainfluenza virus 2|Strain Name:Nishio|Protein Name:fusion protein|Gene Symbol:F	MHHLHPMIVCIFVMYTGIVGSDAIAG DQLLNIGVIQSKIRSLMYYTDGGASFI VVKLLPNLPPSNGTCNITSLDAYNVTL FKLLTPLIENLSKISTVTDTKTRQKRFA GVVVGLAALGVATAAQITAAVAIVK ANANAAAINNLASSIQSTNKAVSDVID ASRTIATAVQAIQDRINGAIVNGITSAS CRAHDALIGSILNLYLTELTTIFHNQIT NPALTPLSIQALRILLGSTLPIVIESKLN TNFNTAELLSSGLLTGQIISISPMYMQ MLIQINVPTFIMQPGAKVIDLIAISANH KLQEVVVQVPNRILEYANELQNYPAN DCVVTPNSVFCRYNEGSPIPESQYQCL RGNLNSCTFTPIIGNFLKRFAFANGVL YANCKSLLCRCADPPHVVSQDDTQGI SIIDIKRCSEMMLDTFSFRITSTFNATY VTDFSMINANIVHLSPLDLSNQINSINK SLKSAEDWIADSNFFANQARTAKTLY SLSAIALILSVITLVVVGLLIAYIIKLVS QIHQFRSLAATTMFHRENPAFFSKNN HGNIYGIS	7	17
BK005918	4677-6302	gb:BK005918|Organism:Porcine rubulavirus|Strain Name:UNKNOWN-BK005918|Protein Name:fusion protein|Gene Symbol:F	MPQQQVAHTCVMLWGIISTVSGINTE ALSQYGVVVTNVRQLTYYTQAGSTY LAVRLLPSLASPDQSCALHSIINYNAT LQAILSPIAENLNLISTALREQHRKKRF AGVAIGLTALGVATAAQATAAVALV RANKNAEKVEQLSQALGETNAAISDL IDATKNLGFAVQAIQNQINTAILPQIH NLSCQVIDAQLGNILSLYLTELTTVFQ PQLTNPALSPLTIQALRAVLGTTLPAL LSEKLKSNIPLGDLMSSGLLKGQLVGL NLQNMLMIIELYIPTLSTHSTAKVLDL VTISSHVNGREVEIQVPNRVLELGSEV LGYGGSECALTMSHILCPFNDARVLS TDMKYCLQGNITHCIFSPVVGSFLRRF ALVNGVVIANCADMSCVCFDPQEIIY QNFQEPTTVIDIKKCGKVQLDTLTFTIS TFANRTYGPPAYVPPDNIIQSEPLDISG NLIAVNNSLSSALNHLATSEILRNEQI WTSSLGISTIVALVIIGILIICLVVTWAA LWALLKEVRGLNSAVNSQLSSYVMG DKFIRY	7	18
KC237063	4530-6185	gb:KC237063:4530-6185|Organism:Parainfluenza virus 5|Strain Name:08-1990|Protein	MGTRIQFLMVSCLLAGTGSLDPAALM QIGVIPTNVRQLMYYTEASSAFIVVKL MPTIDSPISGCNITSISSYNATMTKLLQ PIGENLETIRYQLIPTRRRRRFVGVVIG LAALGVATAAQVTAAVALVKANKN AAAILNLKNAIQKTNAAVADVVQAT QSLGTAVQAVQDHINSVVSPAITAAN CKAQDAIIGSILNLYLTELTTIFHNQIT NPALSPITIQALRILLGSTLPTVVRKSF NTQISAAELLSSGLLTGQIVGLDLTYM	7	19

		Name:fusion protein|Gene Symbol:F|Segment:4	QMVIKIELPTLTVQPATQIIDLVTISAFINNREVMAQLPTRIIVTGSLIQAYPASQCTITPNTVYCRYNDAQVLSDDTMACLQGNLTRCTFSPVVGSFLTRFVLFDGIVYANCRSMLCKCMQPAAVILQPSSSPVTVIDMHKCVSLQLDNLRFTITQLANITYNSTIKLETSQILPIDPLDISQNLAAVNKSLSDALQHLAQSDTYLSAITSATTTSVLSIIAICLGSLGLILIILISVVVWKLLTIVAANRNRMENFVYHNSAFHHSRSDLSEKNQPATLGTR		
AY729016	5862-7523	gb:AY729016:5862-7523|Organism:Murine pneumonia virus|Strain Name:15; ATCC VR-25|Protein Name:fusion glycoprotein precursor|Gene Symbol:F	MIPGRIFLVLLVIFNTKPIHPNTLTEKFYESTCSVETAGYKSALRTGWHMTVMSIKLSQINIESCKSSNSLLAHELAIYSSAVDELRTLSSNALKSKRKKRFLGLILGLGAAVTAGVALAKTVQLESEIALIRDAVRNTNEAVVSLTNGMSVLAKVVDDLKNFISKELLPKINRVSCDVHDITAVIRFQQLNKRLLEVSREFSSNAGLTHTVSSFMLTDRELTSIVGGMAVSAGQKEIMLSSKAIMRRNGLAILSSVNADTLVYVIQLPLFGVMDTDCWVIRSSIDCHNIADKYACLARADNGWYCHNAGSLSYFPSPTDCEIHNGYAFCDTLKSLTVPVTSRECNSNMYTTNYDCKISTSKTYVSTAVLTTMGCLVSCYGHNSCTVINNDKGIIRTLPDGCHYISNKGVDRVQVGNTVYYLSKEVGKSIVVRGEPLVLKYDPLSFPDDKFDVAIRDVEHSINQTRTFLKASDQLLDLSENRENKNLNKSYILTTLLFVVMLIIIMAVIGFILYKVLKMIRDNKLKSKSTPGLTVLS	6	20
AB543336	5174-6805	gb:AB543336:5174-6805|Organism:Human parainfluenza virus 4a|Strain Name:M-25|Protein Name:fusion protein|Gene Symbol:F	MGVKGLSLIMIGLLISPITNLDITHLMNLGTVPTAIRSLVYYTYTKPSYLTVDLIPNLKNLDQNCNYSSLNYYNKTALSLIQPIADNINRLTKPITSSEIQSRFFGAVIGTIALGVATAAQVTAAIGLAKAQENAKLILTLKKAATETNEAVRDLANSNKIVVKMISAIQNQINTIIQPAIDQINCQIKDLQVANILNLYLTEITTVFHNQLTNPALESISIQALKSLLGPTLPEVLSKLDLNNISAASVMASGLIKGQIIAVDIPTMTLVLMVQIPSISPLRQAKIIDLTSITIHTNSQEVQAVVPARFLEIGSEILGFDGSVCQITKDTIFCPYNDAYELPIQQKRCLQGQTRDCVFTPVAGTFPRRFLTTYGTIVANCRDLVCSCLRPPQIIYQPDENPVTIIDKDLCTTLTLDSITIEIQKSINSTFRREVVLESTQVRSLTPLDLSTDLNQYNQLLKSAEDHIQRSTDYLNSINPSIVNNNAIIILILCILLILTVTICIIWLKYLTKEVKNVARNQRLNRDADLFYKIPSQIPVPR	5	21
AF298895	4834-6450	gb:AF298895|Organism:Tioman	MRIALTAVIVSIHFDLAFPMNKNSLLSVGLVHKSVKNLYFYSQGSPSYIVVKLVPTLGNVPGNCTLNSLVRYKSTVSSLLSPLAENLEYLQKTLTVSRGGRRRRF	5	22

		virus|Strain Name:UNKNOWN-AF298895|Protein Name:fusion protein|Gene Symbol:F	AGVAIGLAALGVAAAAQATAAVALV EARQNAAQIQSLSEAIQNTNLAVNEL KTAIGASATAIQAIQTQINEVINPAINR LSCEILDAQLASMLNLYLIHLTTVFQN QLTNPALTPLSIQSLQSLLQGTSSVLTN ITSSSKLALNDALVTGLITGQVVGLN MTSLQIVIAAYVPSVAKLSNAVVHNFI RITTSVNGTEVIIQSPTIIMEQNEVMYD LKTGHCTESDLNIYCPYVDAQLLSPG MTNCINGRLNDCTFSKVVGSFPTRFA AVEGAILANCKYLQCNCLTPPYIITPL NGEMISMIDLSKCQRLDLGTIVFDINN PVNVTFNGNYRADVGQMIVTNPLDIS AELNQINTSLSNAQGFLSKSDAWLHV SQWVTNSGTIFIILIIGLIVGIVYMIINT YVVVQIIKEINRMRTSDRAHLLKGSIS SIST		
FJ215863	4499-6130	gb:FJ215863:4499-6130|Organism:Avian paramyxovirus 8|Strain Name:goose/Delaware/1053/76|Protein Name:fusion protein|Gene Symbol:F	MGQISVYLINSVLLLLVYPVNSIDNTLI APIGVASANEWQLAAYTTSLSGTIAV RFLPVLPDNMTTCLRETITTYNNTVN NILGPLKSNLDALLSSETYPQTRLIGA VIGSIALGVATSAQITAAVALKQAQD NARNILALKEALSKTNEAVKELSSGL QQTAIALGKIQSFVNEEILPSINQLSCE VTANKLGVYLSLYLTELTTIFGAQLTN PALTSLSYQALYNLCGGNMAMLTQKI GIKQQDVNSLYEAGLITGQVIGYDSQ YQLLVIQVNYPSISEVTGVRATELVTV SVTTDKGEGKAIVPQFVAESRVTIEEL DVASCKFSSTTLYCRQVNTRALPPLV ASCLRGNYDDCQYTTEIGALSSRYITL DGGVLVNCKSIVCRCLNPSKIISQNTN AAVTYVDATICKTIQLDDIQLQLEGSL SSVYARNISIEISQVTTSGSLDISSEIGNI NNTVNRVEDLIHQSEEWLAKVNPHIV NNTTLIVLCVLSALAVIWLAVLTAIIIY LRTKLKTISALAVTNTIQSNPYVNQTK RESKF	5	23
JN689227	4689-6521	gb:JN689227:4689-6521|Organism:Tailam virus|Strain Name:TL8K|Protein Name:fusion protein|Gene Symbol:F	MKLSVVYTTLLVSTFYSDLARSQLAL SELTKIGVIPGRSYDLKISTQASYQYM VVKLIPNLTGLNNCTNGTIEAYKKML NRLLSPIDAALRKMKDAVNDKPPESV GNVKFWGAVIGGVALGVATSAQITA GVALHNSIQNANAILALKDSIRQSNKA IQELQTAMSTTVVVLNALQDQINNQL VPAINSLGCQVVANTLGLKLNQYFSEI SLVFGPNLRDPTSETLSIQALSRAFNG DFDSMLSKLKYDDSDFLDLLESDSIRG RIIDVSLSDYLITIQIEYPALLSIKDAVI QTFNLISYNTRGTEWISIFPKQLLVRGT YISNIDISQCVIAATSIICKSDTSTPISSA TWSCATGNITNCARTRVVNAHVPRFA LYGGVVFANCAPVVCKCQDPLYSINQ EPKVTNVMVDVDACKEMYLDGLYIT LGKTQISRAMYAEDVSLGGPISVDPID LGNEINSINSAINRSEEHLNHANELLD	5	24

			KVNPRIVNVKTFGVMIGLLVLVVLWC VITLVWLICLTKQLARTAYAGSMGSR ASTVNSLSGFVG		
JX857409	4831-6615	gb:JX857409:4831-6615|Organism:Porcine parainfluenza virus 1|Strain Name:S206N|Protein Name:fusion protein|Gene Symbol:F	MQVTTLRPAIILSIALLVTGQVPRDKL ANLGIIIKDSKALKIAGSYENRYIVLSL VPTIDNVNGCGSIQIAKYKEMLERLLI PIKDALDLQESLIVIDNETVNNNYSPQ YRFVGAIIGTIALGVATAAQVTAGVA LMEAREAKRDISMLKEAIEKTQNSIEK LQNSAGEQILALKMLQDYVNGEIKPA IEELGCETAALKLGIALTQHYTELTNA FGSNLGSIGEKSLTLQALSSLYKTNITN ILTATNLGKTDIYDIIYAEQVKGRVID VDLKRYMVTISVKIPILSEIPGVLIYEV SSISYNIDGAEWYAAVPDHILSKSAYI GGADISDCIESRLTYICPQDPAQIIADN QQQCFFGHLDKCPITKVIDNLVPKFAF INGGVVANCIASTCTCGEERIQVSQDR NKGVTFLTHNNCGLIGINGIEFHANKK GSDATWNVSPIGVGPAVSLRPVDISLQ IVAATNFLNSSRKDLMKAKEILNQVG NLKDLTTITIINIVIIILLICVIGLGILYH QLRSALGMRDKMSVLNNSSYSLEPRT AQVQVIKPTSFMG	5	25
AY640317	2932-4571	gb:AY640317:2932-4571|Organism:Avian metapneumovirus|Strain Name:LAH A|Protein Name:F|Gene Symbol:F	MDVRICLLLFLISNPSSCIQETYNEESC STVTRGYKSVLRTGWYTNVFNLEIGN VENITCNDGPSLIDTELVLTKNALREL KTVSADQVAKESRLSSPRRRRFVLGAI ALGVATAAAVTAGVALAKTIRLEGEV KAIKNALRNTNEAVSTLGNGVRVLAT AVNDLKEFISKKLTPAINQNKCNIADI KMAISFGQNNRRFLNVVRQFSDSAGI TSAVSLDLMTDDELVRAINRMPTSSG QISLMLNNRAMVRRKGFGILIGVYDG TVVYMVQLPIFGVIETPCWRVVAAPL CRKRRGNYACILREDQGWYCTNAGS TAYYPNKDDCEVRDDYVFCDTAAGI NVALEVDQCNYNISTSKYPCKVSTGR HPVSMVALTPLGGLVSCYESVSCSIGS NKVGIIKQLGKGCTHIPNNEADTITID NTVYQLSKVVGEQRTIKGAPVVNNFN PILFPVDQFNVALDQVFESIDRSQDLID KSNDLLGADAKSKAGIAIAIVVLVILG IFFLLAVIYYCSRVRKTKPKHDYPATT GHSSMAYVS	4	26
KU646513	4641-6498	gb:KU646513:4641-6498|Organism:Avian paramyxovirus 13 goose/Kazakhstan/5751/2013|Strain	MARFSWEIFRLSTILLIAQTCQGSIDGR LTLAAGIVPVGDRPISIYTSSQTGIIVV KLIPNLPDNKKDCAKQSLQSYNETLS RILTPLATAMSAIRGNSTTQVRENRLV GAIIGSVALGVATAAQITAATALIQAN QNAANIARLANSIAKTNEAVTDLTEG LGTLAIGVGKLQDYVNEQFNNTAVAI DCLTLESRLGIQLSLYLTELMGVFGNQ LTSPALTPITIQALYNLAGGNLNALLS RLGASETQLGSLINSGLIKGMPIMYDD ANKLLAVQVELPSIGKLNGARSTLLET	4	27

		Name:APMV-13/white fronted goose/Northern Kazakhstan/5751/2013|Protein Name:fusion protein|Gene Symbol:F	LAVDTTRGPSSPIIPSAVIEIGGAMEELDLSPCITTDLDMFCTKIISYPLSQSTLSCLNGNLSDCVFSRSEGVLSTPYMTIKGKIVANCKQVICRCMDPPQILSQNYGEALLLIDENTCRSLELSGVILKLAGTYESEYTRNLTVDPSQVIITGPLDISAELSKVNQSIDSAKENIAESNKFLSQVNVKLLSSSAMITYIVATVVCLIIAITGCVIGIYTLTKLKSQQKTLLWLGNNAEMHGSRSKTSF		
AF326114	4818-6482	gb:AF326114|Organism:Menangle virus|Strain Name:UNKNOWN-AF326114|Protein Name:fusion protein|Gene Symbol:F	MMPRVLGMIVLYLTHSQILCINRNTLYQIGLIHRSVKKVNFYSQGSPSYIVVKLVPTLAAIPPNCSIKSLQRYKETVTSLVQPISDNLGYLQDKLVTGQSRRRRRFAGVAIGLAALGVAAAAQATAAVALVETRENAGKIQALSESIQNTNQAVHSLKTALGFSATAIQAIQNQVNEVINPAINKLSCEVLDSQLASMLNLYLIHLTTVFQTQLTNPALTPLSIQALTSVLQGTSGVLMNSTNSTLTQPIDLLATGLITGQIISVNMTSLQLIIATFMPSIAELPNAVLHSFFRITTSVNLTEVMIQSPEFIMEQNGVFYDFNTAHCQLGDNNVYCPYIDAARLSSMMTNCINGNLGECVFSRVIGSFPSRFVSLNGAILANCKFMRCNCLSPEKIITPLDGEMISLIDLRVCQKLTLGTITFEISQPVNVSFQGGFVANAGQIIVTNPFDISAELGQINNSLNDAQGFLDQSNNWLKVSGWINNSGSLFIAGIVVIGLIVLCIVIIIYINVQIIREVNRLRSFIYRDYVLDHDKAPYSPESSSPHRKSLKTVS	3	28
GU206351	5441-7468	gb:GU206351:5441-7468|Organism:Avian paramyxovirus 5|Strain Name:budgerigar/Kunitachi/74|Protein Name:fusion protein|Gene Symbol:F	MLQLPLTILLSILSAHQSLCLDNSKLIHAGIMSTTEREVNVYAQSITGSIVVRLIPNIPSNHKSCATSQIKLYNDTLTRLLTPIKANLEGLISAVSQDQSQNSGKRKKRFVGAVIGAAALGLATAAQVTATVALNQAQENARNILRLKNSIQKTNEAVMELKDAVGQTAVAIDKTQAFINNQILPAISNLSCEVLGNKIGVQLSLYLTELTTVFGNQLTNPALTTLSLQALYNLCGDDFNYLINLLNAKNRNLASLYEANLIQGRITQYDSMNQLLIIQVQIPSISTVSGMRVTELFTLSVDTPIGEGKALVPKYVLSSGRIMEEVDLSSCAITSTSVFCSSIISRPLPLETINCLNGNVTQCQFTANTGTLESRYAVIGGLVIANCKAIVCRCLNPPGVIAQNLGLPITIISSNTCQRINLEQITLSLGNSILSTYSANLSQVEMNLAPSNPLDISVELNRVNTSLSKVESLIKESNSILDSVNPQILNVKTVIILAVIIGLIVVWCFILTCLIVRGFMLLVKQQKFKGLSVQNNPYVSNNSH	3	29

JQ001776	6129-8166	gb:JQ001776:6129-8166|Organism:Cedar virus|Strain Name:CG1a|Protein Name:fusion glycoprotein|Gene Symbol:F	MSNKRTTVLIIISYTLFYLNNAAIVGFD FDKLNKIGVVQGRVLNYKIKGDPMTK DLVLKFIPNIVNITECVREPLSRYNETV RRLLLPIHNMLGLYLNNTNAKMTGL MIAGVIMGGIAIGIATAAQITAGFALY EAKKNTENIQKLTDSIMKTQDSIDKLT DSVGTSILILNKLQTYINNQLVPNLELL SCRQNKIEFDLMLTKYLVDLMTVIGP NINNPVNKDMTIQSLSLLFDGNYDIM MSELGYTPQDFLDLIESKSITGQIIYVD MENLYVVIRTYLPTLIEVPDAQIYEFN KITMSSNGGEYLSTIPNFILIRGNYMSN IDVATCYMTKASVICNQDYSLPMSQN LRSCYQGETEYCPVEAVIASHSPRFAL TNGVIFANCINTICRCQDNGKTITQNIN QFVSMIDNSTCNDVMVDKFTIKVGKY MGRKDINNINIQIGPQIIIDKVDLSNEIN KMNQSLKDSIFYLREAKRILDSVNISLI SPSVQLFLIIISVLSFIILLIIIVYLYCKSK HSYKYNKFIDDPDYYNDYKRERINGK ASKSNNIYYVGD	3	30
LC168749	4869-7235	gb:LC168749:4869-7235|Organism:Rinderpest morbillivirus|Strain Name:Lv|Protein Name:F protein|Gene Symbol:F	MGILFAALLAMTNPHLATGQIHWGNL SKIGVVGTGSASYKVMTQSSHQSLVI KLMPNVTAIDNCTKTEIMEYKRLLGT VLKPIREALNAITKNIKPIQSSTTSRRH KRFAGVVLAGAALGVATAAQITAGIA LHQSMMNSQAIESLKASLETTNQAIEE IRQAGQEMVLAVQGVQDYINNELVP AMGQLSCEIVGQKLGLKLLRYYTEILS LFGPSLRDPVSAELSIQALSYALGGDI NKILEKLGYSGSDLLAILESKGIKAKIT YVDIESYFIVLSIAYPSLSEIKGVIVHRL ESVSYNIGSQEWYTTVPRYVATQGYL ISNFDDTPCAFTPEGTICSQNALYPMSP LLQECFRGSTRSCARTLVSGSIGNRFIL SKGNLIANCASILCKCYTTGSIISQDPD KILTYIAADQCPVVEVGGVTIQVGSRE YSDAVYLHEIDLGPPISLEKLDVGTNL WNAVTKLEKAKDLLDSSDLILENIKG VSVTNTGYILVGVGLIAVVGILIITCCC KKRRSDNKVSTMVLNPGLRPDLTGTS KSYVRSL	2	31
LC187310	6250-7860	gb:LC187310:6250-7860|Organism:Avian paramyxovirus 10|Strain Name:rAPMV-10-FI324/YmHA|Protein Name:fusi	MTRTRLLFLLTCYIPGAVSLDNSILAP AGIISASERQIAIYTQTLQGTIALRFIPV LPQNLSSCAKDTLESYNSTVSNLLLPI AENLNALLKDADKPSQRIIGAIIGSVA LGVATTAQVTAALAMTQAQQNARNI WKLKESIKNTNQAVLELKDGLQQSAI ALDKVQSFINSEILPQINQLGCEVAAN KLGIFLSLYLTEITTVFKNQITNPALST LSYQALYNLCGGNMAALTKQIGIKDT EINSLYEAELITGQVIGYDSADQILLIQ VSYPSVSRVQGVRAVELLTVSVATPK GEGKAIAPSFIAQSNIIAEELDTQPCKF SKTTLYCRQVNTRTLPVRVANCLKGK YNDCQYTTEIGALASRYVTITNGVVA	2	32

		on protein|Gene Symbol:F	NCRSIICRCLDPEGIVAQNSDAAITVID RSTCKLIQLGDITLRLEGKLSSSYSKNI TIDISQVTTSGSLDISSELGSINNTITKV EDLISKSNDWLSKVNPTLISNDTIIALC VIAGIVVIWLVIITILSYYILIKLKNVAL LSTMPKKDLNPYVNNTKF		
NC_005283	5277-6935	gb:NC_005283:5277-6935|Organism:Dolphin morbillivirus|Strain Name:UNKNOWN-NC_005283|Protein Name:fusion protein|Gene Symbol:F	MAASNGGVMYQSFLTIIILVIMTEGQI HWGNLSKIGIVGTGSASYKVMTRPNH QYLVIKLMPNVTMIDNCTRTEVTEYR KLLKTVLEPVKNALTVITKNIKPIQSLT TSRRSKRFAGVVLAGVALGVATAAQI TAGVALHQSIMNSQSIDNLRTSLEKSN QAIEEIRQASQETVLAVQGVQDFINNE LIPSMHQLSCEMLGQKLGLKLLRYYT EILSIFGPSLRDPVSAEISIQALSYALGG DINKILEKLGYSGADLLAILESRGIKA KVTHVDLEGYFIVLSIAYPTLSEVKGV IVHKLEAVSYNLGSQEWYTTLPKYVA TNGYLISNFDESSCAFMSEVTICSQNA LYPMSPLLQQCLRGSTASCARSLVSG TIGNRFILSKGNLIANCASVLCKCYST GTIISQDPDKLLTFVAADKCPLVEVDG ITIQVGSREYPDSVYVSRIDLGPAISLE KLDVGTNLGSALTKLDNAKDLLDSSN QILENVRRSSFGGAMYIGILVCAGALV ILCVLVYCCRRHCRKRVQTPPKATPG LKPDLTGTTKSYVRSL	2	33
NC_005339	5374-7602	gb:NC_005339:5374-7602|Organism:Mossman virus|Strain Name:UNKNOWN-NC_005339|Protein Name:fusion protein|Gene Symbol:F	MSNYFPARVIIIVSLITAVSCQISFQNLS TIGVFKFKEYDYRVSGDYNEQFLAIK MVPNVTGVENCTASLIDEYRHVIYNL LQPINTTLTASTSNVDPYAGNKKFFGA VIAGVALGVATAAQVTAGVALYEAR QNAAAIAEIKESLHYTHKAIESLQISQ KQTVVAIQGIQDQINTNIIPQINALTCEI ANQRLRLMLLQYYTEMLSSFGPIIQDP LSGHITVQALSQAAGGNITGLMRELG YSSKDLRYILSVNGISANIIDADPEIGSI ILRIRYPSMIKIPDVAVMELSYLAYHA AGGDWLTVGPRFILKRGYSLSNLDITS CTIGEDFLLCSKDVSSPMSLATQSCLR GDTQMCSRTAVQDREAPRFLLLQGNL IVNCMSVNCKCEDPEETITQDPAYPL MVLGSDTCKIHYIDGIRIKLGKVQLPPI TVLNTLSLGPIVVLNPIDVSNQLSLVE TTVKESEDHLKNAIGALRSQSRVGGV GIVAIVGLIIATVSLVVLVISGCCLVKY FSRTATLESSLTTIEHGPTLAPKSGPIIP TYINPVYRHD	2	34
NC_007454	4635-6384	gb:NC_007454:4635-6384|Organism:J-virus|Strain Name:UN	MKPVALIYLTILAFTVKVRSQLALSDL TKIGIIPAKSYELKISTQAAQQLMVIKL IPNVNGLTNCTIPVMDSYKKMLDRIL KPIDDALNHVKNAIQDKQGDGVPGV RFWGAIIGGVALGVATSAQITAGVAL HNSIQNANAILQLKESIRNSNKAIEELQ AGLQSTVLVINALQDQINSQLVPAINT LGCSVIANTLGLRLNQYFSEISLVFGP	2	35

		KNOWN-NC_007454|Protein Name:fusion protein|Gene Symbol:F	NLRDPTSQTLSIQAIAKAFNGDFDSM MKKMHYTDSDFLDLLESDSIRGRIISV SLEDYLIIIQIDYPGLTTIPNSVVQTFNL ITYNYKGTEWESIFPRELLIRGSYISNI DISQCVGTSKSMICKSDTSTTISPATW ACATGNLTSCARTRVVNSHSTRFALS GGVLFANCAPIACRCQDPQYSINQEPK TTNVMVTSEDCKELYIDGFYLTLGKK MLDRAMYAEDVALGGSVSVDPIDIGN ELNSINESINKSHEYLDKANELLEQVN PNIVNVSSFSFILVISILLIIWFIVTLVWL IYLTKHMNFIVGKVAMGSRSSTVNSL SGFVG		
NC_009489	4620-6500	gb:NC_009489:4620-6500|Organism:Mapuera virus|Strain Name:BeAnn 370284|Protein Name:fusion protein|Gene Symbol:F	MRSSLFLVLTLLVPFAHSIDSITLEQYG TVITSVRSLAYFLETNPTYISVRLMPAI QTDSSHCSYHSIENYNLTLTKLLLPLQ ENLHQITDSLSSRRRKKRFAGVAVGL AALGVATAAQVTAAIAVVKAKENSA KIAQLTSAISETNRAVQDLIEGSKQLA VAVQAIQDQINNVIQPQLTNLSCQVA DAQVGTILNMYLTELTTVFHPQITNSA LTPITIQALRSLLGSTLPQVVTSTIKTD VPLQDLLTSGLLKGQIVYLDLQSMIM VVSVSVPTIALHSMAKVYTLKAISAH VNNAEVQMQVPSRVMELGSEIMGYD IDQCEETSRYLFCPYNGGSILSATMKM CLNGNISQCVFTPIYGSFLQRFVLVDG VIVANCRDMTCACKSPSKIITQPDSLP VTIIDSTSCSNLVLDTLELPIISINNATY RPVQYVGPNQIIFSQPLDLLSQLGKINS SLSDAIEHLAKSDEILEQIQWDSPQGY TLIALTSVLAFVVVAIVGLLISTRYLIF EIRRINTTLTQQLSSYVLSNKIIQY	2	36
NC_017937	4534-6330	gb:NC_017937:4534-6330|Organism:Nariva virus|Strain Name:UNKNOWN-NC_017937|Protein Name:fusion protein|Gene Symbol:F	MAEQEKTPLRYKILLIIIVINHYNITNV FGQIHLANLSSIGVFVTKTLDYRTTSD PTEQLLVINMLPNISNIQDCAQGVVNE YKHLISSLLTPINDTLDLITSNINPYSGR NKLFGEIIAGAALTVATSAQITAGVAL YEARQNAKDIAAIKESLGYAYKAIDK LTTATREITVVINELQDQINNRLIPRIN DLACEVWATRLQAMLLQYYAEIFSVI GPNLQDPLSGKISIQALARAAGGNIKL MVDELNYSGQDLSRLVKVGAIKGQII DADPSLGVVIIKMRYPNIIKIPNVAISE LSYVSYSSDGQDWITTGPNYIVTRGYS IANIQTSSCSVGDDFVLCDRDMTYPM SQVTQDCLRGNIALCSRMVVRDREAP RYLILQGNMVANCMSITCRCEEPESEI YQSPDQPLTLLTRDTCDTHVVDGIRIR LGVRKLPTISVINNITLGPIITTDPIDVS NQLNAVVSTIDQSAELLHQAQRVLSE RARGARDHILATAAIVICVVLAVLILV LLIGLVYLYRTQNEILVKTTMLEQVPT FAPKSFPMESQIYSGKTNKGYDPAE	2	37
NC_025256	6865-8853	gb:NC_025256:6865	MKKKTDNPTISKRGHNHSRGIKSRAL LRETDNYSNGLIVENLVRNCHHPSKN	2	38

		-8853|Organism:Bat Paramyxovirus Eid_hel/GH-M74a/GHA/2009|Strain Name:Bat PV/Eid_hel/GH-M74a/GHA/2009|Protein Name:fusion protein|Gene Symbol:F	NLNYTKTQKRDSTIPYRVEERKGHYP KIKHLIDKSYKHIKRGKRRNGHNGNII TIILLLILILKTQMSEGAIHYETLSKIGLI KGITREYKVKGTPSSKDIVIKLIPNVTG LNKCTNISMENYKEQLDKILIPINNIIE LYANSTKSAPGNARFAGVIIAGVALG VAAAAQITAGIALHEARQNAERINLL KDSISATNNAVAELQEATGGIVNVITG MQDYINTNLVPQIDKLQCSQIKTALDI SLSQYYSEILTVFGPNLQNPVTTSMSI QAISQSFGGNIDLLLNLLGYTANDLLD LLESKSITGQITYINLEHYFMVIRVYYP IMTTISNAYVQELIKISFNVDGSEWVS LVPSYILIRNSYLSNIDISECLITKNSVI CRHDFAMPMSYTLKECLTGDTEKCPR EAVVTSYVPRFAISGGVIYANCLSTTC QCYQTGKVIAQDGSQTLMMIDNQTCS IVRIEEILISTGKYLGSQEYNTMHVSV GNPVFTDKLDITSQISNINQSIEQSKFY LDKSKAILDKINLNLIGSVPISILFIIAIL SLILSIITFVIVMIIVRRYNKYTPLINSDP SSRRSTIQDVYIIPNPGEHSIRSAARSID RDRD		
NC_025347	4471-6386	gb:NC_025347:4471-6386|Organism:Avian paramyxovirus 7|Strain Name:APMV-7/dove/Tennessee/4/75|Protein Name:fusion protein|Gene Symbol:F	MRVRPLIIILVLLVLLWLNILPVIGLDN SKIAQAGIISAQEYAVNVYSQSNEAYI ALRTVPYIPPHNLSCFQDLINTYNTTIQ NIFSPIQDQITSITSASTLPSSRFAGLVV GAIALGVATSAQITAAVALTKAQQNA QEIIRLRDSIQNTINAVNDITVGLSSIGV ALSKVQNYLNDVINPALQNLSCQVSA LNLGIQLNLYLTEITTIFGPQITNPSLTP LSIQALYTLAGDNLMQFLTRYGYGET SVSSILESGLISAQIVSFDKQTGIAILYV TLPSIATLSGSRVTKLMSVSVQTGVGE GSAIVPSYVIQQGTVIEEFIPDSCIFTRS DVYCTQLYSKLLPDSILQCLQGSMAD CQFTRSLGSFANRFMTVAGGVIANCQ TVLCRCYNPVMIIPQNNGIAVTLIDGS LCKELELEGIRLTMADPVFASYSRDLII NGNQFAPSDALDISSELGQLNNSISSA TDNLQKAQESLNKSIIPAATSSWLIILL FVLVSISLVIGCISIYFIYKHSTTNRSRN LSSDIISNPYIQKAN	2	39
NC_025348	4790-6570	gb:NC_025348:4790-6570|Organism:Tuhoko virus 2|Strain Name:UNKNOWN-NC_025348|Protein Name:fusion	MAPCVLFLSSLLLISTISPSHGINQPAL RRIGAIVSSVKQLKFYSKTKPNYIIVKL LPTINLSKSNCNLTSINRYKESVIEIIKP LADNIDNLNQKLLPKNRRKRMAGVAI GLAALGVAAAAQATAAVALVEARKN TQMIQSLADSIQDTNAAVQAVNIGLQ NSAVAIQAIQNQINNVINPALDRLNCE VLDAQIASILNLYLIKSVTIFQNQLTNP ALQQLSIQMLSIVMQDTAKILGNFTIG DKFDQHDLLGSGLITGQVVGVNLTNL QLIIAAFIPSIAPLPQAYIIDLISITISVND TEAVIQIPERIMEHGSSIYQFGGKQCV YGQFSAYCPFSDAVLMTQDLQLCMK	2	40

		protein|Gene Symbol:F	GNIEHCIFSSVLGSFPNRFASVDGVFY ANCKYMSCACSDPLQVIHQDDSVNL MVIDSSVCRSLTLGHVTFPIIAFSNVSY QMKTNISIEQMIVTSPLDLSTELKQINN SVNIANTFLDSSNRALKTSIFGTSSQIIL IVLLIFTCLLILYVIFLTYIIKILIKEVKR LRDGNSRTGSKLSFINPDV		
NC_025350	4663-6428	gb:NC_025350:4663-6428|Organism:Tuhoko virus 3|Strain Name:UNKNOWN-NC_025350|Protein Name:fusion protein|Gene Symbol:F	MLWLTILIALVGNHESTCMNINFLQSL GQINSQKRFLNFYTQQPPSYMVIRLVP TLQLSANNCTLGSIVRYRNAIKELIQP MDENLRWLSSNLIPQRRGKRFAGVAV GLAALGVAVAAQATAAVALVEARAN AEKIASMSQSIQETNKAVTSLSQAVSA SGIAIQAIQNEINNVIHPILNQVQCDVL DARVGNILNLYLIKVTTIFQNQLTNPA LQRLSTQALSMLMQSTSSYLRNLSSSE SAINADLSMTNLIEAQIVGINMTNLQL VLAVFIPSIARLNGALLYDFISITISSNQ TEVMLQIPHRVLEIGNSLYTFEGTQCE MTKLNAYCLYSDAIPVTESLRDCMNG LFSQCGFVRIIGSFANRFASVNGVIYA NCKHLTCSCLQPDEIITQDTNVPLTIID TKRCTKISLGHLTFTIREYANVTYSLR TEIANSQITVVSPLDLSSQLTTINNSLA DATNHIMNSDRILDRLNSGLYSKWVII FLICASIVSLIGLVFLGFLIRGLILELRS KHRSNLNKASTYSIDSSIGLT	2	41
NC_025352	5950-8712	gb:NC_025352:5950-8712|Organism:Mojiang virus|Strain Name:Tongguan1|Protein Name:fusion protein|Gene Symbol:F	MALNKNMFSSLFLGYLLVYATTVQSS IHYDSLSKVGVIKGLTYNYKIKGSPST KLMVVKLIPNIDSVKNCTQKQYDEYK NLVRKALEPVKMAIDTMLNNVKSGN NKYRFAGAIMAGVALGVATAATVTA GIALHRSNENAQAIANMKSAIQNTNE AVKQLQLANKQTLAVIDTIRGEINNNI IPVINQLSCDTIGLSVGIRLTQYYSEIIT AFGPALQNPVNTRITIQAISSVFNGNF DELLKIMGYTSGDLYEILHSELIRGNII DVDVDAGYIALEIEFPNLTLVPNAVV QELMPISYNIDGDEWVTLVPRFVLTRT TLLSNIDTSRCTITDSSVICDNDYALPM SHELIGCLQGDTSKCAREKVVSSYVP KFALSDGLVYANCLNTICRCMDTDTP ISQSLGATVSLLDNKRCSVYQVGDVLI SVGSYLGDGEYNADNVELGPPIVIDKI DIGNQLAGINQTLQEAEDYIEKSEEFL KGVNPSIITLGSMVVLYIFMILIAIVSVI ALVLSIKLTVKGNVVRQQFTYTQHVP SMENINYVSH	2	42
NC_025363	4622-6262	gb:NC_025363:4622-6262|Organism:Avian paramyxovirus	MAIPVPSSTALMIFNILVSLAPASALD GRLLLGAGIVPTGDRQVNVYTSSQTGI IALKLLPNLPKDKENCAEVSIRSYNET LTRILTPLAQSMAAIRGNSTVSTRGRE PRLVGAIIGGVALGVATAAQITAATAL IQANQNAENIARLAKGLAATNEAVTD LTKGVGSLAIGVGKLQDYVNEQFNRT GEAIECLTIESRVGVQLSLYLTEVIGVF	2	43

		12|Strain Name:Wigeon/Italy/3920_1/2005|Protein Name:fusion protein|Gene Symbol:F	GDQITSPALSDISIQALYNLAGGNLNV LLQKMGIEGTQLGSLINSGLIKGRPIM YDDGNKILGIQVTLPSVGRINGARATL LEAIAVATPKGNASPLIPRAVISVGSLV EELDMTPCVLTPTDIFCTRILSYPLSDS LTTCLKGNLSSCVFSRTEGALSTPYVS VHGKIVANCKSVVCRCVEPQQIISQN YGEALSLIDESLCRILELNGVILKMDG QFTSEYTKNITIDPVQVIISGPIDISSELS QVNQSLDSALENIKESNSYLSKVNVK LISSSAMITYIVITVICLILTFVALVLGI YSYTKIRSQQKTLIWMGNNIARSKEG NRF		
NC_025373	4617-6582	gb:NC_025373:4617-6582|Organism:Avian paramyxovirus 3|Strain Name:turkey/Wisconsin/68|Protein Name:fusion protein|Gene Symbol:F	MASPMVPLLIITVVPALISSQSANIDKL IQAGIIMGSGKELHIYQESGSLDLYLR LLPVIPSNLSHCQSEVITQYNSTVTRLL SPIAKNLNHLLQPRPSGRLFGAVIGSIA LGVATSAQISAAIALVRAQQNANDIL ALKAAIQSSNEAIKQLTYGQEKQLLAI SKIQKAVNEQVIPALTALDCAVLGNK LAAQLNLYLIEMTTIFGDQINNPVLTPI PLSYLLRLTGSELNDVLLQQTRSSLSLI HLVSKGLLSGQIIGYDPSVQGIIIRIGLI RTQRIDRSLVFXPYVLPITISSNIATPIIP DCVVKKGVIIEGMLKSNCIELERDIIC KTINTYQITKETRACLQGNITMCKYQ QSRTQLSTPFITYNGVVIANCDLVSCR CIRPPMIITQVKGYPLTIINRNLCTELS VDNLILNIETNHNFSLNPTIIDSQSRLIA TSPLEIDALIQDAQHHAAAALLKVEES NAHLLRVTGLGSSSWHIILILTLLVCTI AWLIGLSIYVCRIKNDDSTDKEPTTQS SNRGIGVGSIQYMT	2	44
NC_025386	5548-7206	gb:NC_025386:5548-7206|Organism:Salem virus|Strain Name:UNKNOWN-NC_025386|Protein Name:fusion protein|Gene Symbol:F	MNPLNQTLIAKVLGFLLLSSSFTVGQI GFENLTRIGVHQVKQYGYKLAHYNS HQLLLIRMIPTVNGTHNCTHQVITRYR EMVREIITPIKGALDIMKKAVSPDLVG ARIFGAIVAGAALGIATSAQITAGVAL HRTKLNGQEISKLKEAVSLTNEAVEQ LQYSQGKSILAIQGIQDFINFNVVPLLE EHTCGIAKLHLEMALMEYFQKLILVF GPNLRDPIGSTIGIQALATLFQNNMFE VSLRLGYAGDDLEDVLQSNSIRANIIE AEPDSGFIVLAIRYPTLTLVEDQVITEL AHITFNDGPQEWVATIPQFVTYRGLV LANIDVSTCTFTERNVICARDQTYPMII DLQLCMRGNIAKCGRTRVTGSTASRF LLKDGNMYANCIATMCRCMSSSSIIN QEPSHLTTLIVKETCSEVMIDTIRITLG ERKHPPIDYQTTITLGQPIALAPLDVGT ELANAVSYLNKSKVLLEHSNEVLSSV STAHTSLTATIVLGIVVGGLAILIVVMF LFLEAQVIKVQRAMMLCPITNHGYLP NEDLLTRGHSIPTIG	2	45
NC_025390	4805-6460	gb:NC_025390:4805	MGYFHLLLLILTAIAISAHLCYTTTLDG RKLLGAGIVITEEKQVRVYTAAQSGTI	2	46

		-6460|Organism:Avian paramyxovirus 9|Strain Name:duck/New York/22/1978|Protein Name:fusion protein|Gene Symbol:F	VLRSFRVVSLDRYSCMESTIESYNKTV YNILAPLGDAIRRIQASGVSVERIREGR IFGAILGGVALGVATAAQITAAIALIQ ANENAKNILRIKDSITKTNEAVRDVTN GVSQLTIAVGKLQDFVNKEFNKTTEAI NCVQAAQQLGVELSLYLTEITTVFGP QITSPALSKLTIQALYNLAGVSLDVLL GRLGADNSQLSSLVSSGLITGQPILYD SESQILALQVSLPSISDLRGVRATYLDT LAVNTAAGLASAMIPKVVIQSNNIVEE LDTTACIAAEADLYCTRITTFPIASAVS ACILGDVSQCLYSKTNGVLTTPYVAV KGKIVANCKHVTCRCVDPTSIISQNYG EAATLIDDQLCKVINLDGVSIQLSGTF ESTYVRNVSISANKVIVSSSIDISNELE NVNSSLSSALEKLDESDAALSKVNVH LTSTSAMATYIVLTVIALILGFVGLGL GCFAMIKVKSQAKTLLWLGAHADRS YILQSKPAQSST		
NC_025403	4826-6649	gb:NC_025403:4826-6649|Organism:Achimota virus 1|Strain Name:UNKNOWN-NC_025403|Protein Name:fusion protein|Gene Symbol:F	MWIMIILSLFQIIPGVTPINSKVLTQLG VITKHTRQLKFYSHSTPSYLVVKLVPT INTESTVCNFTSLSRYKDSVRELITPLA KNIDNLNSILTIPKRRKRMAGVVIGLA ALGVAAAAQATAAVALIEAKKNTEQI QALSESIQNTNKAVSSIEKGLSSAAIA VQAIQNQINNVINPALTALDCGVTDA QLGNILNLYLIKTLTVFQKQITNPALQ PLSIQALNIIMQETSSVLRNFTKTDEIE HTDLLTSGLITGQVVGVNLTNLQLIIA AFIPSIAPLNQAYILDFIRITVNINNSES MIQIPERIMEHGISLYQFGGDQCTFSD WSAYCPYSDATLMAPGLQNCFRGQA ADCVFSTVMGSFPNRFVSVQGVFYVN CKFIRCACTQPQRLITQDDSLSLTQIDA KTCRMLTLGFVQFSINEYANVTYSFK NNVTAGQLIMTNPIDLSTEIKQMNDS VDEAARYIEKSNAALNKLMYGGRSDI VTTVLLVGFILLVVYVIFVTYILKILM KEVARLRNSNHPDLIKPYNYPM	2	47
NC_025404	4772-6647	gb:NC_025404:4772-6647|Organism:Achimota virus 2|Strain Name:UNKNOWN-NC_025404|Protein Name:fusion protein|Gene Symbol:F	MLNSFYQIICLAVCLTTYTVISIDQHNL LKAGVIVKSIKGLNFYSRGQANYIIVK LIPNVNVTDTDCDIGSIKRYNETVYSLI KPLADNIDYLRTQFAPTKRKKRFAGV AIGLTALGVATAAQVTAAVALVKAQ ENARKLDALADSIQATNEAVQDLSTG LQAGAIAIQAIQSEINHVINPALERLSC EIIDTRVASILNLYLIRLTTVFHRQLVN PALTPLSIQALNHLLQGETEGLVKNES KMTDSKIDLLMSGLITGQVVGVNIKH MQLMIAVFVPTTAQLPNAYVINLLTIT ANINNSEVLVQLPNQILERSGIIYQFRG KDCVSSPNHMYCPYSDASILSPELQLC LQGRLEMCLFTQVVGSFPTRFASDKGI VYANCRHLQCACSEPEGIIYQDDTSAI TQIDASKCSTLKLDMLTFKLSTYANK TFDASFSVGKDQMLVTNLLDLSAELK	2	48

			TMNASVAHANKLIDKSNLLIQSNALIG HSNTIFIVVIVILAVMVLYLIIVTYIIKVI MVEVSRLKRMNIYSIDK		
NC_025410	4958-6751	gb:NC_025410:4958-6751|Organism:Tuhoko virus 1|Strain Name:UNKNOWN-NC_025410|Protein Name:fusion protein|Gene Symbol:F	MVTIIKPLILLVTVILQISGHIDTTALTS IGAVIASSKEIMYYAQSTPNYIVIKLIP NLPNIPSQCNFSSIAYYNKTLLDLFTPI SDNINMLHQRLSNTGRNRRFAGVAIG LAALGVATAAQVTAAFALVEAKSNT AKIAQIGQAIQNTNAAINSLNAGIGGA VTAIQAIQTQINGIITDQINAATCTALD AQIGTLLNMYLLQLTTTFQPQIQNPAL QPLSIQALHRIMQGTSIVLSNLTDSSK YGLNDALSAGLITGQIVSVDLRLMQIT IAANVPTLSRLENAIAHDIMRITTNVN NTEVIVQLPETIMEHAGRLYQFNKDH CLSSTQRFFCPYSDAKLLTSKISSCLSG IRGDCIFSPVVGNFATRFISVKGVIIAN CKFIRCTCLQPEGIISQLDDHTLTVIDL KLCNKLDLGLIQFDLQVLSNISYEMTL NTSQNQLILTDPLDLSSELQTMNQSIN NAANFIEKSNSLLNSSTYEFNRSVALL VALILLSLTILYVIVLTCVVKLLVHEVS KNRRHIQDLESHHK	2	49
NC_028249	4850-7055	gb:NC_028249:4850-7055|Organism:Phocine distemper virus|Strain Name:PDV/Wadden_Sea.NLD/1988|Protein Name:fusion protein|Gene Symbol:F	MTRVKKLPVPTNPPMHHSLDSPFLNP EHATGKISITDDTSSQLTNFLYHKYHK TTINHLSRTISGTDPPSAKLNKFGSPILS TYQIRSALWWIAMVILVHCVMGQIH WTNLSTIGIIGTDSSHYKIMTRSSHQY LVLKLMPNVSIIDNCTKAELDEYEKLL NSVLEPINQALTLMTKNVKSLQSLGS GRRQRRFAGVVIAGAALGVATAAQIT AGVALYQSNLNAQAIQSLRASLEQSN KAIDEVRQASQNIIIAVQGVQDYVNN EIVPALQHMSCELIGQRLGLKLLRYYT ELLSVFGPSLRDPVSAEISIQALSYALG GEIHKILEKLGYSGNDMVAILETKGIR AKITHVDLSGKFIVLSISYPTLSEVKGV VVHRLEAVSYNIGSQEWYTTVPRYVA TNGYLISNFDESSCVFVSESAICSQNSL YPMSPILQQCLRGETASCARTLVSGTL GNKFILSKGNIIANCASILCKCHSTSKII NQSPDKLLTFIASDTCSLVEIDGVTIQV GSRQYPDVVYASKVILGPAISLERLDV GTNLGSALKKLNDAKVLIESSDQILDT VKNSYLSLGTLIALPVSIGLGLILLLLIC CCKKRYQHLFSQSTKVAPVFKPDLTG TSKSYVRSL	2	50
NC_028362	5217-6842	gb:NC_028362:5217-6842|Organism:Caprine parainfluenza virus 3|Strain	MIKKIICIFSMPILLSFCQVDIIKLQRVG ILVSKPKSIKISQNFETRYLVLNLIPNIE NAQSCGDQQIKQYKKLLDRLIIPLYDG LRLQQDIIVVDNNLKNNTNHRAKRFF GEIIGTIALGVATSAQITAAVALVEAK QARSDIERVKNAVRDTNKAVQSIQGS VGNLIVAVKSVQDYVNNEIVPSIKRLG CEAAGLQLGIALTQHYSELTNIFGDNI GTLKEKGIKLQGIASLYHTNITEIFTTS	2	51

		Name:JS2013|Protein Name:fusion protein|Gene Symbol:F	TVDQYDIYDLLFTESIKMRVIDVDLND YSITLQVRLPLLTKISDAQIYNVDSVS YNIGGTEWYIPLPRNIMTKGAFLGGA NLQDCIESFSDYICPSDPGFILNRDIEN CLSGNITQCPKTLVISDIVPRYAFVDG GVIANCLSTTCTCNGIDNRINQAPDQG IKIITYKDCQTIGINGMLFKTNQEGTLA AYTPVDITLNNSVNLDPIDLSIELNRA RSDLAESKEWIKRSEAKLDSVGSWYQ SSTTEIIQIVMIIVLFIINIIVLIVLIKYSRS QNQSMNNHMNEPYILTNKVQ		
AF079780	5919-7580	gb:AF079780|Organism:Tupaia paramyxovirus|Strain Name:UNKNOWN-AF079780|Protein Name:fusion protein|Gene Symbol:F	MASLLKTICYIYLITYAKLEPTPKSQL DLDSLASIGVVDAGKYNYKLMTTGSE KLMVIKLVPNITYATNCNLTAHTAYT KMIERLLTPINQSLYEMRSVITERDGG TIFWGAIIAGAALGVATAAAITAGVAL HRAEQNARNIAALKDALRNSNEAIQH LKDAQGHTVLAIQGLQEQINNNIIPKL KESHCLGVNNQLGLLLNQYYSEILTV FGPNLQNPVSASLTIQAIAKAFNGDFN SLMTNLNYDPTDLLDILESNSINGRIID VNLNEKYIALSIEIPNFITLTDAKIQTFN RITYGYGSNEWLTLIPDNILEYGNLISN VDLTSCVKTKSSYICNQDTSYPISSELT RCLRGDTSSCPRTPVVNSRAPTFALSG GHIYANCAKAACRCEKPPMAIVQPAT STLTFLTEKECQEVVIDQINIQLAPNRL NKTIITDGIDLGPEVIINPIDVSAELGNI ELEMDKTQKALDRSNKILDSMITEVT PDKLLIAMIVVFGILLLWLFGVSYYAF KIWSKLHFLDSYVYSLRNPSHHRSNG HQNHSFSTDISG	1	52
EU403085	4664-6585	gb:EU403085:4664-6585|Organism:Avian paramyxovirus 3|Strain Name:APMV3/PKT/Netherland/449/75|Protein Name:fusion protein|Gene Symbol:F	MQPGSALHLPHLYIIIALVSDGTLGQT AKIDRLIQAGIVLGSGKELHISQDSGTL DLFVRLLPVLPSNLSHCQLEAITQYNK TVTRLLAPIGKNLEQVLQARPRGRLF GPIIGSIALGVATSAQITAAIALVRAQQ NANDILALKNALQSSNEAIRQLTYGQ DKQLLAISKIQKAVNEQILPALDQLDC AVLGTKLAVQLNLYLIEMTTIFGEQIN NPVLATIPLSYILRLTGAELNNVLMKQ ARSSLSLVQLVSKGLLSGQVIGYDPSV QGLIIRVNLMRTQKIDRALVYQPYVLP ITLNSNIVTPIAPECVIQKGTIIEGMSRK DCTELEQDIICRTVTTYTLARDTRLCL QGNISSCRYQQSGTQLHTPFITYNGAV IANCDLVSCRCLRPPMIITQVKGYPLTI ITRSVCQELSVDNLVLNIETHHNFSLN PTIIDPLTRVIATTPLEIDSLIQEAQDHA NAALAKVEESDKYLRAVTGGNYSNW YIVLVIVLLFGNLGWSLLLTVLLCRSR KQQRRYQQDDSVGSERGVGVGTIQY MS	1	53
KX258200	4443-6068	gb:KX258200:4443-6068|Orga	MEKGTVLFLAALTLYNVKALDNTKL LGAGIASGKEHELKIYQSSVNGYIAVK LIPFLPSTKRECYNEQLKNYNATINRL	1	54

		nism:Avian paramyxovirus 14|Strain Name:APMV14/duck/Japan/11OG0352/2011|Protein Name:fusion protein|Gene Symbol:F	MGPINDNIKLVLSGVKTRTREGKLIGAIIGTAALGLATAAQVTAAIALEQAQDNARAILTLKESIRNTNNAVSELKTGLSEVSIALSKTQDYINTQIMPALSNLSCEIVGLKIGIQLSQYLTEVTAVFGNQITNPALQPLSMQALYQLCGGDFSLLLDKIGADRNELESLYEANLVTGRIVQYDTADQLVIIQVSIPSVSTLSGYRVTELQSISVDMDHGEGKAVIPRYIVTSGRVIEEMDISPCVLTATAVYCNRLLTTSLPESVLKCLDGDHSSCTYTSNSGVLETRYIAFDGMLIANCRSIVCKCLDPPYIIPQNKGKPLTIISKEVCKKVTLDGITLLIDAEFTGEYGLNITIGPDQFAPSGALDISTELGKLNNSINKAEDYIDKSNELLNRVNVDIVNDTAVIVLCVMSALVVVWCIGLTVGLIYVSKNTLRAVAIKGTSIENPYVSSGKHAKNSS
KY511044	4592-6247	gb:KY511044:4592-6247|Organism:Avian paramyxovirus UPO216|Strain Name:APMV-15/WB/Kr/UPO216/2014|Protein Name:fusion protein|Gene Symbol:F	MIFTMYHVTVLLLLSLLTLPLGIQLARASIDGRQLAAAGIVVTGEKAINLYTSSQTGTIVVKLLPNVPQGREACMRDPLTSYNKTLTSLLSPLGEAIRRIHESTTETAGLVQARLVGAIIGSVALGVATSAQITAAAALIQANKNAENILKLKQSIAATNEAVHEVTDGLSQLAVAVGKMQDFINTQFNNTAQEIDCIRISQQLGVELNLYLTELTTVFGPQITSPALSPLSIQALYNLAGGNLDVLLSKIGVGNNQLSALISSGLISGSPILYDSQTQLLGIQVTLPSVSSLNNMRAIFLETLSVSTDKGFAAALIPKVVTTVGTVTEELDTSYCIETDIDLFCTRIVTFPMSPGIYACLNGNTSECMYSKTQGALTTPYMSVKGSIVANCKMTTCRCADPASIISQNYGEAVSLIDSSVCRVITLDGVTLRLSGSFDSTYQKNITIRDSQVIITGSLDISTELGNVNNSINNALDKIEESNQILESVNVSLTSTNALIVYIICTALALICGITGLILSCYIMYKMRSQQKTLMWLGNNTLDQMRAQTKM
NC_025360	6104-8123	gb:NC_025360:6104-8123|Organism:Atlantic salmon paramyxovirus|Strain Name:ASPV/Yrkje371/95|Protein Name:fusion protein|Ge	MDGPKFRFVLLILLTAPARGQVDYDKLLKVGIFEKGTANLKISVSSQQRYMVIKMMPNLGPMNQCGIKEVNLYKESILRLITPISTTLNYIKSEIQVEREVALQPNGTIVRFFGLIVAAGALTLATSAQITAGIALHNSLENAKAIKGLTDAIKESNLAIQKIQDATAGTVIALNALQDQVNTNIIPAINTLGCTAAGNTLGIALTRYYSELIMIFGPSLGNPVEAPLTIQALAGAFNGDLHGMIREYGYTPSDIEDILRTNSVTGRVIDVDLVGMNIVLEINLPTLYTLRDTKIVNLGKITYNVDGSEWQTLVPEWLAIRNTLMGGVDLSRCVVSSRDLICKQDPVFSLDTSIISCLNGNTESCPRNRVVNSVAPRYAVIRGNILANCISTTCLCGDPGVPIIQKGDNTLTAMSINDCKLVGVDGYVFRPGPKAVNVTFNLPHLNLGPEVNVNP

		ne Symbol:F	VDISGALGKVEQDLASSRDHLAKSEKI LSGINPNIINTEMVLVAVILSLVCAMV VIGIVCWLSILTKWVRSCRADCRRPN KGPDLGPIMSSQDNLSF		
UniProt ID: Q9IH63		FUS_NIPAV Fusion glycoprotein F0 OS=Nipah virus	MVVILDKRCYCNLLILILMISECSVGIL HYEKLSKIGLVKGVTRKYKIKSNPLT KDIVIKMIPNVSNMSQCTGSVMENYK TRLNGILTPIKGALEIYKNNTHDLVGD VRLAGVIMAGVAIGIATAAQITAGVA LYEAMKNADNINKLKSSIESTNEAVV KLQETAEKTVYVLTALQDYINTNLVP TIDKISCKQTELSLDLALSKYLSDLLFV FGPNLQDPVSNSMTIQAISQAFGGNYE TLLRTLGYATEDFDDLLESDSITGQIIY VDLSSYYIIVRVYFPILTEIQQAYIQEL LPVSFNNDNSEWISIVPNFILVRNTLIS NIEIGFCLITKRSVICNQDYATPMTNN MRECLTGSTEKCPRELVVSSHVPRFA LSNGVLFANCISVTCQCQTTGRAISQS GEQTLLMIDNTTCPTAVLGNVIISLGK YLGSVNYNSEGIAIGPPVFTDKVDISS QISSMNQSLQQSKDYIKEAQRLLDTV NPSLISMLSMIILYVLSIASLCIGLITFIS FIIVEKKRNTYSRLEDRRVRPTSSGDL YYIGT		57

In some embodiments, a fusogen described herein comprises an amino acid sequence of Table 2, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 100, 200, 300, 400, 500, or 600 amino acids in length. For instance, in some embodiments, a fusogen described herein comprises an amino acid sequence having at least 80% identity to any amino acid sequence of Table 2. In some embodiments, a nucleic acid sequence described herein encodes an amino acid sequence of Table 2, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 40, 50, 60, 80, 100, 200, 300, 400, 500, or 600 amino acids in length.

In some embodiments, a fusogen described herein comprises an amino acid sequence set forth in any one of SEQ ID NOS: 58-133, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the

sequence, e.g., a portion of 100, 200, 300, 400, 500, or 600 amino acids in length. For instance, in some embodiments, a fusogen described herein comprises an amino acid sequence having at least 80% identity to an amino acid sequence set forth in any one of SEQ ID NOS: 58-133. In some embodiments, a nucleic acid sequence described herein encodes an amino acid sequence set forth in any one of SEQ ID NOS: 58-133, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity thereto, or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a portion of the sequence, e.g., a portion of 40, 50, 60, 80, 100, 200, 300, 400, 500, or 600 amino acids in length.

Table 2. Paramyxovirus protein G, H, and HN sequence clusters. Column 1, Genbank ID includes the Genbank ID of the whole genome sequence of the virus that is the centroid sequence of the cluster. Column 2, nucleotides of CDS provides the nucleotides corresponding to the CDS of the gene in the whole genome. Column 3, Full Gene Name, provides the full name of the gene including Genbank ID, virus species, strain, and protein name. Column 4, Sequence, provides the amino acid sequence of the gene. Column 5, #Sequences/Cluster, provides the number of sequences that cluster with this centroid sequence.

Genbank ID	**Nucleotides of CDS**	**Full sequence ID**	**Sequence**	**#Sequences /Cluster**	**SEQ ID NO**
KU950686	4643-5638	gb:KU950686:4643-5638|Organism:Human respiratory syncytial virus|Strain Name:RSVA/Homo sapiens/USA/TH_10506/2014|Protein Name:attachment glycoprotein|Gene Symbol:G	MSKTKDQRTAKTLERTWDTLNHLLFISSC LYKLNLKSIAQITLSILAMIISTSLIIAAIIFIA SANHKVTLTTAIIQDATNQIKNTTPTYLTQ NPQLGISFSNLSGTTLQSTTILASTTPSAEST PQSTTVKIINTTTTQILPSKPTTKQRQNKPQ NKPNNDFHFEVFNFVPCSICSNNPTCWAIC KRIPNKKPGKKTTTKPTKKPTLKTTKKDP KPQTTKPKEALTTKPTGKPTINTTKTNIRTT LLTSNTKGNPEHTSQEETLHSTTSEGYLSP SQVYTTSGQEETLHSTTSEGYLSPSQVYTT SEYLSQSLSSSNTTK	706	58
AB524405	6424-8274	gb:AB524405:6424-8274|Organism:Newcastle disease virus|Strain Name:Goose/Alaska/415/91|Protein Name:hemaggl	MERGVSQVALENDEREAKNTWRLVFRVT VLFLTIVTLAISAAALAFSMNASTPQDLEGI PVAISKVEDKITSALGASQDVMDRIYKQV ALESPLALLNTESTIMNALTSLSYQINGAA NASGCGAPVPDPDYIGGIGKELIVDDTSDV TSFYPSAFQEHLNFIPAPTTGSGCTRIPSFD MSATHYCYTHNVILSGCRDHSHSHQYLAL GVLRTSATGRVFFSTLRSINLDDTQNRKSC SVSATPLGCDMLCSKVTETEEEDYQSTDP TLMVHGRLGFDGQYHERDLDVHTLFGDW	418	59

		utinin-neuraminidase protein|Gene Symbol:HN	VANYPGVGGGSFINNRVWFPVYGGLKPGS PTDKRQEGQYAIYKRYNDTCPDDQEYQV RMAKSAYKPNRFGGKRVQQAILSIGVSTT LADDPVLTVTSNTITLMGAEGRVMTVGTS HYLYQRGSSYYSPAILYPLTIANKTATLQD PYKFNAFTRPGSVPCQASARCPNSCVTGV YTDPYPIVFHKNHTLRGVFGTMLDDEQAR LNPVSAVFDSIARSRVTRVSSSSTKAAYTT STCFKVVKTGKVYCLSIAEISNTLFGEFRIV PLLVEILRDEGRSEARSALTTQGHPGWND EVVDPIFCAVTNQTDHRQKLEEYAQSWP		
JQ582844	4686-5636	gb:JQ582844:4686-5636|Organism:Human respiratory syncytial virus|Strain Name:NH1067|Protein Name:receptor-binding glycoprotein|Gene Symbol:G	MSKNKNQRTARTLEKTWDTLNHLIVISSC LYKLNLKSIAQIALSVLAMIISTSLIIAAIIFII SANHKVTLTTVTVQTIKNHTEKNITTYLTQ VSPERVSPSKQPTTTPPIHTNSATISPNTKSE THHTTAQTKGRTTTPTQNNKPSTKPRPKN PPKKPKDDYHFEVFNFVPCSICGNNQLCKS ICKTIPNNKPKKKPTTKPTNKPPTKTTNKR DPKTPAKTLKKETTINPTTKKPTPKTTERD TSTPQSTVLDTTTSKHTERDTSTPQSTVLD TTTSKHTIQQQSLHSITPENTPNSTQTPTAS EPSTSNSTQKL	278	60
AB254456	7271-9136	gb:AB254456:7271-9136|Organism:Measles virus|Strain Name:SSPE-Kobe-1|Protein Name:Hemagglutinin|Gene Symbol:H	MSPHRDRINAFYRDNPHPKGSRIVINREHL MIDRPYVLLAVLFVMFLSLIGLLAIAGIRL HRAAIYTAEIHKSLSTNLDVTNSIEHQVKD VLTPLFKIIGDEVGLRTPQRFTDLVKFISDK IKFLNPDREYDFRDLTWCINPPERIKLDYD QYCADVAAEELMNALVNSTLLEARATNQ FLAVSKGNCSGPTTIRGQFSNMSLSLLDLY LSRGYNVSSIVTMTSQGMYGGTYLVGKPN LSSKGSELSQLSMHRVFEVGVIRNPGLGAP VFHMTNYFEQPVSNDFSNCMVALGELRFA ALCHREDSVTVPYQGSGKGVSFQLVKLGV WKSPTDMQSWVPLSTDDPVIDRLYLSSHR GVIADNQAKWAVPTTRTDDKLRMETCFQ QACKGKNQALCENPEWAPLKDNRIPSYG VLSVNLSLTVELKIKIASGFGPLITHGSGM DLYKTNHDNVYWLTIPPMKNLALGVINTL EWIPRFKVSPNLFTVPIKEAGEDCHAPTYL PAEVDGDVKLSSNLVILPGQDLQYVLATY DTSRVEHAVVYYVYSPSRSFSYFYPFRLPI KGVPIELQVECFTWDQKLWCRHFCVLADS ESGGHITHSGMVGMGVSCTVTREDGTNR RQGCQ	128	61
AB040874	6614-8362	gb:AB040874:6614-8362|Organism:Mumps virus|Strain Name:Miyahara|Protein Name:hemagglutinin-neuraminidase|	MEPSKLFTMSDNATFAPGPVINAADKKTF RTCFRILVLSVQAVTLILVIVTLGELVRMIN DQGLSNQLSSIADKIRESATMIASAVGVM NQVIHGVTVSLPLQIEGNQNQLLSTLATIC TGKKQVSNCSTNIPLVNDLRFINGINKFIIE DYATHDFSIGHPLNMPSFIPTATSPNGCTRI PSFSLGKTHWCYTHNVINANCKDHTSSNQ YISMGILVQTASGYPMFKTLKIQYLSDGLN RKSCSIATVPDGCAMYCYVSTQLETDDYA GSSPPTQKLTLLFYNDTVTERTISPTGLEGN	87	62

		Gene Symbol:HN	WATLVPGVGSGIYFENKLIFPAYGGVLPNS SLGVKSAREFFRPVNPYNPCSGPQQDLDQ RALRSYFPSYFSNRRVQSAFLVCAWNQIL VTNCELVVPSNNQTLMGAEGRVLLINNRL LYYQRSTSWWPYELLYEISFTFTNSGQSSV NMSWIPIYSFTRPGSGNCSGENVCPTACVS GVYLDPWPLTPYSHQSGINRNFYFTGALL NSSTTRVNPTLYVSALNNLKVLAPYGNQG LFASYTTTTCFQDTGDASVYCVYIMELAS NIVGEFQILPVLTRLTIT		
AB7361 66	6709-8427	gb:AB736166:6709-8427|Organism:Human respirovirus 3|Strain Name:ZMLS/2011|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	MEYWKHTNHGKDAGNELETATATHGNR LTNKITYILWTITLVLLSIVFIIVLINSIKSEK AHESLLQDINNEFMEVTEKIQVASDNTND LIQSGVNTRLLTIQSHVQNYIPISLTQQISDL RKFISEITIRNDNQEVPPQRITHDVGIKPLNP DDFWRCTSGLPSLMRTPKIRLMPGPGLLA MPTTVDGCVRTPSLVINDLIYAYTSNLITR GCQDIGKSYQVLQIGIITVNSDLVPDLNPRI SHTFNINDNRKSCSLALLNTDVYQLCSTPK VDERSDYASSGIEDIVLDIVNYDGSISTTRF KNNNISFDQPYAALYPSVGPGIYYKGKIIFL GYGGLEHPINENAICNTTGCPGKTQRDCN QASHSPWFSDRRMVNSIIVVDKGLNSVPK LKVWTISMRQNYWGSEGRLLLLGNKIYIY TRSTSWHSKLQLGIIDITDYSDIRIKWTWH NVLSRPGNNECPWGHSCPDGCITGVYTDA YPLNPTGSIVSSVILDSQKSRVNPVITYSTA TERVNELAIRNKTLSAGYTTTSCITHYNKG YCFHIVEINHKSLNTFQPMLFKTEIPKSCS	78	63
KJ627396	6166-6885	gb:KJ627396:6166-6885|Organism:Human metapneumovirus|Strain Name:HMPV/Homo sapiens/PER/FLI1305/2010/A|Protein Name:attachment glycoprotein G|Gene Symbol:G	MEVKVENIRAIDMLKARVKNRVARSKCF KNASLILIGITTLSIALNIYLIINYTIQKTTSE SEHHTSSPPTESNKETSTIPIDNPDITPNSQH PTQQSTESLTLYPASSMSPSETEPASTPGIT NRLSLADRSTTQPSESRTKTNSTVHKKNK KNISSTISRTQSPPRTTAKAVSRTTALRMSS TGERPTTTSVQSDSSTTAQNHEETGPANPQ ASVSTM	71	64
AB475097	7079-8902	gb:AB475097:7079-8902|Organism:Canine distemper virus|Strain Name:M25CR|Protein Name:hemagglutinin|Gene Symbol:H	MLSYQDKVGAFYKDNARANSSKLSLVTE EQGGRRPPYLLFVLLILLVGILALLAIAGVR FRQVSTSNVEFGRLLKDDLEKSEAVHHQV MDVLTPLFKIIGDEIGLRLPQKLNEIKQFIL QKTNFFNPNREFDFRDLHWCINPPSKIKVN FTNYCDAIGVRKSIASAANPILLSALSGGR GDIFPPYRCSGATTSVGRVFPLSVSLSMSLI SKTSEIISMLTAISDGVYGKTYLLVPDYIER EFDTQKIRVFEIGFIKRWLNDMPLLQTTNY MVLPENSKAKVCTIAVGELTLASLCVDES TVLLYHDSNGSQDSILVVTLGIFGATPMNQ VEEVIPVAHPSVERIHITNHRGFIKDSVAT	45	65

			WMVPALVSEQQEGQKNCLESACQRKSYP MCNQTSWEPFGGVQLPSYGRLTLPLDASI DLQLNISFTYGPVILNGDGMDYYENPLLDS GWLTIPPKNGTILGLINKASRGDQFTVTPH VLTFAPRESSGNCYLPIQTSQIMDKDVLTE SNLVVLPTQNFRYVVATYDISRENHAIVY YVYDPIRTISYTYPFRLTTKGRPDFLRIECF VWDDDLWCHQFYRFESDITNSTTSVEDLV RIRFSCNRSKP		
AJ849636	7326-9155	gb:AJ849636:7326-9155|Organism:Peste-des-petits-ruminants virus|Strain Name:Turkey 2000|Protein Name:haemagglutinin|Gene Symbol:H	MSAQRERINAFYKDNPHNKNHRVILDRER LVIERPYILLGVLLVMFLSLIGLLAIAGIRL HRATVGTSEIQSRLNTNIELTESIDHQTKD VLTPLFKIIGDEVGIRIPQKFSDLVKFISDKI KFLNPDREYDFRDLRWCMNPPERVKINFD QFCEYKAAVKSIEHIFESPLNKSKKLQSLT LGPGTGCLGRTVTRAHFSELTLTLMDLDL EMKHNVSSVFTVVEEGLFGRTYTVWRSD ARDPSTDLGIGHFLRVFEIGLVRDLGLGPP VFHMTNYLTVNMSDDYRRCLLAVGELKL TALCSSSETVTLGERGVPKREPLVVVILNL AGPTLGGELYSVLPTSDLMVEKLYLSSHR GIIKDDEANWVVPSTDVRDLQNKGECLVE ACKTRPPSFCNGTGSGPWSEGRIPAYGVIR VSLDLASDPGVVITSVFGPLIPHLSGMDLY NNPFSRAVWLAVPPYEQSFLGMINTIGFPN RAEVMPHILTTEIRGPRGRCHVPIELSRRV DDDIKIGSNMVILPTIDLRYITATYDVSRSE HAIVYYIYDTGRSSSYFYPVRLNFKGNPLS LRIECFPWRHKVWCYHDCLIYNTITDEEV HTRGLTGIEVTCNPV	34	66
AB005795	6693-8420	gb:AB005795:6693-8420|Organism:Sendai virus|Strain Name:Ohita|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MDGDRSKRDSYWSTSPGGSTTKLVSDSER SGKVDTWLLILAFTQWALSIATVIICIVIAA RQGYSMERYSMTVEALNTSNKEVKESLTS LIRQEVITRAANIQSSVQTGIPVLLNKNSRD VIRLIEKSCNRQELTQLCDSTIAVHHAEGIA PLEPHSFWRCPAGEPYLSSDPEVSLLPGPSL LSGSTTISGCVRLPSLSIGEAIYAYSSNLITQ GCADIGKSYQVLQLGYISLNSDMFPDLNP VVSHTYDINDNRKSCSVVATGTRGYQLCS MPIVDERTDYSSDGIEDLVLDILDLKGRTK SHRYSNSEIDLDHPFSALYPSVGSGIATEGS LIFLGYGGLTTPLQGDTKCRIQGCQQVSQD TCNEALKITWLGGKQVVSVLIQVNDYLSE RPRIRVTTIPITQNYLGAEGRLLKLGDQVYI YTRSSGWHSQLQIGVLDVSHPLTISWTPHE ALSRPGNEDCNWYNTCPKECISGVYTDAY PLSPDAANVATVTLYANTSRVNPTIMYSN TTNIINMLRIKDVQLEAAYTTTSCITHFGK GYCFHIIEINQKSLNTLQPMLFKTSIPKLCK AES	23	67
AF457102	6903-8630	gb:AF457102|Organism:Human parainfluenza virus 1 strain Washington/19	MAEKGKTNSSYWSTTRNDNSTVNTHINTP AGRTHIWLLIATTMHTVLSFIIMILCIDLIIK QDTCMKTNIMTVSSMNESAKIIKETITELIR QEVISRTINIQSSVQSGIPILLNKQSRDLTQL IEKSCNRQELAQICENTIAIHHADGISPLDP HDFWRCPVGEPLLSNNPNISLLPGPSLLSG	21	68

		64|Strain Name:Washington 1964|Protein Name:HN glycoprotein|Gene Symbol:HN	STTISGCVRLPSLSIGDAIYAYSSNLITQGC ADIGKSYQVLQLGYISLNSDMYPDLNPVIS HTYDINDNRKSCSVIAAGTRGYQLCSLPTV NETTDYSSEGIEDLVFDILDLKGKTKSHRY KNEDITFDHPFSAMYPSVGSGIKIENTLIFL GYGGLTTPLQGDTKCVINRCTNVNQSVCN DALKITWLKKRQVVNVLIRINNYLSDRPKI VVETIPITQNYLGAEGRLLKLGKKIYIYTRS SGWHSNLQIGSLDINNPMTIKWAPHEVLS RPGNQDCNWYNRCPRECISGVYTDAYPLS PDAVNVATTTLYANTSRVNPTIMYSNTSEI INMLRLKNVQLEAAYTTTSCITHFGKGYC FHIVEINQASLNTLQPMLFKTSIPKICKITS		
KJ627397	6146-6888	gb:KJ627397:6146-6888|Organism:Human metapneumovirus|Strain Name:HMPV/Homo sapiens/PER/FPP00098/2010/B|Protein Name:attachment glycoprotein G|Gene Symbol:G	MEVRVENIRAIDMFKAKMKNRIRSSKCYR NATLILIGLTALSMALNIFLIIDYATLKNMT KVEHCVNMPPVEPSKKSPMTSAADLNTKL NPQQATQLTTEDSTSLAATSENHLHTETTP TSDATISQQATDEHTTLLRPINRQTTQTTTE KKPTGATTKKDKEKETTTRTTSTAATQTL NTTNQTSNGREATTTSARSRNGATTQNSD QTIQAADPSSKPYHTQTNTTTAHNTDTSSL SS	21	69
AF017149	8913-10727	gb:AF017149|Organism:Hendra virus|Strain Name:UNKNOWN-AF017149|Protein Name:glycoprotein|Gene Symbol:G	MMADSKLVSLNNNLSGKIKDQGKVIKNY YGTMDIKKINDGLLDSKILGAFNTVIALLG SIIIIVMNIMIIQNYTRTTDNQALIKESLQSV QQQIKALTDKIGTEIGPKVSLIDTSSTITIPA NIGLLGSKISQSTSSINENVNDKCKFTLPPL KIHECNISCPNPLPFREYRPISQGVSDLVGL PNQICLQKTTSTILKPRLISYTLPINTREGVC ITDPLLAVDNGFFAYSHLEKIGSCTRGIAK QRIIGVGEVLDRGDKVPSMFMTNVWTPPN PSTIHHCSSTYHEDFYYTLCAVSHVGDPIL NSTSWTESLSLIRLAVRPKSDSGDYNQKYI AITKVERGKYDKVMPYGPSGIKQGDTLYF PAVGFLPRTEFQYNDSNCPIIHCKYSKAEN CRLSMGVNSKSHYILRSGLLKYNLSLGGDI ILQFIEIADNRLTIGSPSKIYNSLGQPVFYQA SYSWDTMIKLGDVDTVDPLRVQWRNNSV ISRPGQSQCPRFNVCPEVCWEGTYNDAFLI DRLNWVSAGVYLNSNQTAENPVFAVFKD NEILYQVPLAEDDTNAQKTITDCFLLENVI WCISLVEIYDTGDSVIRPKLFAVKIPAQCSE S	14	70
AF212302	8943-10751	gb:AF212302|Organism:Nipah virus|Strain Name:UNKNOWN-AF212302|Protein	MPAENKKVRFENTTSDKGKIPSKVIKSYY GTMDIKKINEGLLDSKILSAFNTVIALLGSI VIIVMNIMIIQNYTRSTDNQAVIKDALQGIQ QQIKGLADKIGTEIGPKVSLIDTSSTITIPAN IGLLGSKISQSTASINENVNEKCKFTLPPLKI HECNISCPNPLPFREYRPQTEGVSNLVGLP NNICLQKTSNQILKPKLISYTLPVVGQSGT	14	71

		Name:attachment glycoprotein|Gene Symbol:G	CITDPLLAMDEGYFAYSHLERIGSCSRGVS KQRIIGVGEVLDRGDEVPSLFMTNVWTPP NPNTVYHCSAVYNNEFYYVLCAVSTVGD PILNSTYWSGSLMMTRLAVKPKSNGGGY NQHQLALRSIEKGRYDKVMPYGPSGIKQG DTLYFPAVGFLVRTEFKYNDSNCPITKCQY SKPENCRLSMGIRPNSHYILRSGLLKYNLS DGENPKVVFIEISDQRLSIGSPSKIYDSLGQ PVFYQASFSWDTMIKFGDVLTVNPLVVN WRNNTVISRPGQSQCPRFNTCPEICWEGV YNDAFLIDRINWISAGVFLDSNQTAENPVF TVFKDNEILYRAQLASEDTNAQKTITNCFL LKNKIWCISLVEIYDTGDNVIRPKLFAVKIP EQCT		
EU439428	6751-8638	gb:EU439428:6751-8638|Organism:Swine parainfluenza virus 3|Strain Name:92-7783_ISU-92|Protein Name:hemagglutinin-neuraminidase HN|Gene Symbol:HN	MEYWKHTNSTKDTNNELGTTRDRHSSKA TNIIMYIFWTTTSTILSVIFIMILINLIQENNH NKLMLQEIKKEFAVIDTKIQKTSDDISTSIQ SGINTRLLTIQSHVQNYIPLSLTQQMSDLR KFINDLTTKREHQEVPIQRMTHDSGIEPLN PDKFWRCTSGNPSLTSSPKIRLIPGPGLLAT STTVNGCIRIPSLAINNLIYAYTSNLITQGC QDIGKSYQVLQIGIITINSDLVPDLNPRVTH TFNIDDNRKSCSLALLNTDVYQLCSTPKV DERSDYASTGIEDIVLDIVTSNGLIITTRFTN NNITFDKPYAALYPSVGPGIYYKDKVIFLG YGGLEHEENGDVICNTTGCPGKTQRDCNQ ASYSPWFSNRRMVNSIIVVDKSIDTTFSLR VWTIPMRQNYWGSEGRLLLLGDRIYIYTR STSWHSKLQLGVIDISDYNNIRINWTWHN VLSRPGNDECPWGHSCPDGCITGVYTDAY PLNPSGSVVSSVILDSQKSRENPIITYSTAT NRVNELAIYNRTLPAAYTTTNCITHYDKG YCFHIVEINHRSLNTFQPMLFKTEVPKNCS	14	72
KF530164	6157-6906	gb:KF530164:6157-6906|Organism:Human metapneumovirus|Strain Name:HMPV/AUS/172832788/2004/B|Protein Name:attachment glycoprotein G|Gene Symbol:G	MEVRVENIRAIDMFKAKIKNRIRSSRCYRN ATLILIGLTALSMALNIFLIIDHATLRNMIK TENCANMPSAEPSKKTPMTSTAGPSTKPN PQQATQWTTENSTSPAATLEGHPYTGTTQ TPDTTAPQQTTDKHTALPKSTNEQITQTTT EKKTTRATTQKREKRKENTNQTTSTAATQ TTNTTNQTRNASETITTSDGPRIDTTTQSSE QTARATEPGSSPYHARRGAGPR	14	73
AB910309	6960-8747	gb:AB910309:6960-8747|Organism:Feline morbillivirus|Strain Name:SS1|Protein Name:hemaggl	MKNINIKYYKDSNRYLGKILDEHKIVNSQL YSLSIKVITIIAIIVSLIATIMTIINATSGRTTL NSNTDILLNQRDEIHSIHEMIFDRVYPLITA MSTELGLHIPTLLDELTKAIDQKIKIMNPPV DTVTSDLSWCIKPPNGIIIDPKGYCESMELS KTYKLLLDQLDVSRKKSLTINRKNINQCQ LVDDSEIIFATVNIQSTPRFLNFGHTVSNQR ITFGQGTYSSTYILTIQEDGITDVQYRVFEI GYISDQFGVFPSLIVSRVLPIRMVLGMESC	12	74

		utinin protein|Gene Symbol:H	TLTSDRQGGYFLCMNTLTRSIYDYVNIRDL KSLYITLPHYGKVNYTYFNFGKIRSPHEID KLWLTSDRGQIISGYFAAFVTITIRNYNNY PYKCLNNPCFDNSENYCRGWYKNITGTDD VPILAYLLVEMYDEEGPLITLVAIPPYNYT APSHNSLYYDDKINKLIMTTSHIGYIQINEV HEVIVGDNLKAILLNRLSDEHPNLTACRLN QGIKEQYKSDGMIISNSALIDIQERMYITVK AIPPVGNYNFTVELHSRSNTSYILLPKQFN AKYDKLHLECFNWDKSWWCALIPQFSLS WNESLSVDTAIFNLINCK		
AB759118	7116-8957	gb:AB759118:7116-8957|Organism:Avian paramyxovirus 6|Strain Name:red-necked stint/Japan/8KS0813/2008|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	MASPSELNRSQATLYEGDPNSKRTWRTVY RASTLILDLAILCVSIVAIVRMSTLTPSDVT DSISSSITSLSDTYQSVWSDTHQKVNSIFKE VGISIPVTLDKMQVEMGTAVNIITDAVRQL QGVNGSAGFSITNSPEYSGGIDALIYPQKSL NGKSLAISDLLEHPSFIPAPTTSHGCTRIPTF HLGYRHWCYSHNTIESGCHDAGESIMYLS MGAVGVGHQGKPVFTTSAAVILDDGKNR KSCSVVANPNGCDVLCSLVKQTEDQDYA DPTPTPMIHGRLHFNGTYTESMLDQSLFTG HWVAQYPAVGSGSVSHGRLFFPLYGGISK SSSLFPKLRAHAYFTHNEELECKNLTSKQR EDLFNAYMPGKIAGSLWAQGIVICNLTTL ADCKIAVANTSTMMMAAEGRLQLVQDK VVLYQRSSSWWPVLIYYDILVSELVNARH LDIVNWVPYPQSKFPRPTWTKGLCEKPSIC PAVCVTGVYQDVWVVSVGDFSNETVVIG GYLEAASERKDPWIAAANQYNWLTRRQL FTAQTEAAYSSTTCFRNTHQDKVFCLTIME VTDNLLGDWRIAPLLYEVTVVDRQQSSRK AVAMSEAHRTRFKYYSPENKFTPQH	11	75
AY141760	6791-8485	gb:AY141760|Organism:Fer-de-Lance paramyxovirus|Strain Name:ATCC VR-895|Protein Name:hemagglutinin-neuraminidase protein HN|Gene Symbol:HN	MDPKSYYCNEDLRSDGGEKSPGGDLYKGI ILVSTVISLIIAIISLAFIIDNKINIQSLDPLRG LEDSYLVPIKDKSESISQDIQEGIFPRLNLIT AATTTTIPRSIAIQTKDLSDLIMNRCYPSVV NNDTSCDVLAGAIHSNLFSQLDPSTYWTC SSGTPTMNQTVKLLPDNSQIPGSTYSTGCV RIPTFSLGSMIYSYSHNVIYEGCNDHSKSSQ YWQLGYISTSKTGEPLQQVSRTLTLNNGL NRKSCSTVAQGRGAYLLCTNVVEDERTD YSTEGIQDLTLDYIDIFGAERSYRYTNNEV DLDRPYAALYPSVGSGTVYNDRILFLGYG GLMTPYGDQAMCQAPECTSATQEGCNSN QLIGYFSGRQIVNCIIEIITVGTEKPIIRVRTI PNSQVWLGAEGRIQTLGGVLYLYIRSSGW HALAQTGIILTLDPIRISWIVNTGYSRPGNG PCSASSRCPAQCITGVYTDIFPLSQNYGYL ATVTLLSGVDRVNPVISYGTSTGRVADSQ LTSSSQVAAYTTTTCFTFNQKGYCYHIIEL SPATLGIFQPVLVVTEIPKICS	8	76
EU877976	6248-8161	gb:EU877976:6248-8161|Organism:Avian paramyxovirus	MQGNMEGSRDNLTVDDELKTTWRLAYR VVSLLLMVSALIISIVILTRDNSQSIITAINQS SDADSKWQTGIEGKITSIMTDTLDTRNAAL LHIPLQLNTLEANLLSALGGNTGIGPGDLE HCRYPVHDTAYLHGVNRLLINQTADYTAE	8	77

		4|Strain Name:APMV-4/KR/YJ/06|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	GPLDHVNFIPAPVTTTGCTRIPSFSVSSSIW CYTHNVIETGCNDHSGSNQYISMGVIKRA GNGLPYFSTVVSKYLTDGLNRKSCSVAAG SGHCYLLCSLVSEPEPDDYVSPDPTPMRLG VLTWDGSYTEQAVPERIFKNIWSANYPGV GSGAIVGNKVLFPFYGGVRNGSTPEVMNR GRYYYIQDPNDYCPDPLQDQILRAEQSYY PTRFGRRMVMQGVLACPVSNNSTIASQCQ SYYFNNSLGFIGAESRIYYLNGNIYLYQRS SSWWPHPQIYLLDSRIASPGTQNIDSGVNL KMLNVTVITRPSSGFCNSQSRCPNDCLFGV YSDIWPLSLTSDSIFAFTMYLQGKTTRIDPA WALFSNHAIGHEARLFNKEVSAAYSTTTC FSDTIQNQVYCLSILEVRSELLGAFKIVPFL YRVL		
AB176531	6821-8536	gb:AB176531:6821-8536|Organism:Human parainfluenza virus 2|Strain Name:Nishio|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MEDYSNLSLKSIPKRTCRIIFRTATILGICTL IVLCSSILHEIIHLDVSSGLMDSDDSQQGIIQ PIIESLKSLIALANQILYNVAIIPLKIDSIETV IFSALKDMHTGSMSNTNCTPGNLLLHDAA YINGINKFLVLKSYNGTPKYGPLLNIPSFIP SATSPNGCTRIPSFSLIKTHWCYTHNVMLG DCLDFTTSNQYLAMGIIQQSAAAFPIFRTM KTIYLSDGINRKSCSVTAIPGGCVLYCYVA TRSEKEDYATTDLAELRLAFYYYNDTFIER VISLPNTTGQWATINPAVGSGIYHLGFILFP VYGGLISGTPSYNKQSSRYFIPKHPNITCAG NSSEQAAAARSSYVIRYHSNRLIQSAVLIC PLSDMHTARCNLVMFNNSQVMMGAEGR LYVIDNNLYYYQRSSSWWSASLFYRINTD FSKGIPPIIEAQWVPSYQVPRPGVMPCNAT SFCPANCITGVYADVWPLNDPEPTSQNAL NPNYRFAGAFLRNESNRTNPTFYTASASA LLNTTGFNNTNHKAAYTSSTCFKNTGTQK IYCLIIIEMGSSLLGEFQIIPFLRELIP	7	78
AF052755	6584-8281	gb:AF052755|Organism:Parainfluenza virus 5|Strain Name:W3A|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MVAEDAPVRATCRVLFRTTTLIFLCTLLAL SISILYESLITQKQIMSQAGSTGSNSGLGSIT DLLNNILSVANQIIYNSAVALPLQLDTLEST LLTAIKSLQTSDKLEQNCSWSAALINDNRY INGINQFYFSIAEGRNLTLGPLLNMPSFIPT ATTPEGCTRIPSFSLTKTHWCYTHNVILNG CQDHVSSNQFVSMGIIEPTSAGFPFFRTLKT LYLSDGVNRKSCSISTVPGGCMMYCFVST QPERDDYFSAAPPEQRIIIMYYNDTIVERII NPPGVLDVWATLNPGTGSGVYYLGWVLF PIYGGVIKGTSLWNNQANKYFIPQMVAAL CSQNQATQVQNAKSSYYSSWFGNRMIQS GILACPLRQDLTNECLVLPFSNDQVLMGA EGRLYMYGDSVYYYQRSNSWWPMTMLY KVTITFTNGQPSAISAQNVPTQQVPRPGTG DCSATNRCPGFCLTGVYADAWLLTNPSST STFGSEATFTGSYLNTATQRINPTMYIANN TQIISSQQFGSSGQEAAYGHTTCFRDTGSV MVYCIYIIELSSSLLGQFQIVPFIRQVTLS	7	79
BK005918	6560-8290	gb:BK005918|Organism:Porcine	MSQLGTDQIMHLAQPAIARRTWRLCFRIF ALFILIAIVITQIFMLTFDHTLLTTTQFLTSIG NLQSTITSWTPDVQAMLSISNQLIYTTSITL	7	80

		rubulavirus|Strain Name:UNKNOWN-BK005918|Protein Name:attachment protein|Gene Symbol:HN	PLKISTTEMSILTAIRDHCHCPDCSSACPTR QMLLNDPRYMSGVNQFIGAPTESINITFGP LFGIPSFIPTSTTTQGCTRIPSFALGPSHWCY THNFITAGCADGGHSNQYLAMGTIQSASD GSPLLITARSYYLSDGVNRKSCSIAVVPGG CAMYCYVATRSETDYYAGNSPPQQLLTL VFSNDTIIERTIHPTGLANGWVMLVPGVGS GTLYNEYLLFPAYGGMQQILANQSGEINQ FFTPYNATVRCAMAQPQFSQRAAASYYPR YFSNRWIRSAIVACPYRAIYQTQCTLIPLPN RMVMMGSEGRIFTLGDRLFYYQRSSSWW PYPLLYQVGLNFLTTPPSVSSMTQVPLEHL ARPGKGGCPGNSHCPATCVTGVYADVWP LTDPRSGVGGTSLVAAGGLDSTSERMAPV NYLAIGESLLSKTYLLSKTQPAAYTTTTCF RDTDTGKIYCITIAELGKVLLGEFQIVPFLR EIKIQSRY		
EU338414	6015-7913	gb:EU338414:6015-7913|Organism:Avian paramyxovirus 2|Strain Name:APMV-2/Chicken/California/Yucaipa/56|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	MDFPSRENLAAGDISGRKTWRLLFRILTLS IGVVCLAINIATIAKLDHLDNMASNTWTTT EADRVISSITTPLKVPVNQINDMFRIVALDL PLQMTSLQKEITSQVGFLAESINNVLSKNG SAGLVLVNDPEYAGGIAVSLYQGDASAGL NFQPISLIEHPSFVPGPTTAKGCIRIPTFHMG PSHWCYSHNIIASGCQDASHSSMYISLGVL KASQTGSPIFLTTASHLVDDNINRKSCSIVA SKYGCDILCSIVIETENEDYRSDPATSMIIG RLFFNGSYTESKINTGSIFSLFSANYPAVGS GIVVGDEAAFPIYGGVKQNTWLFNQLKDF GYFTHNDVYKCNRTDIQQTILDAYRPPKIS GRLWVQGILLCPVSLRPDPGCRLKVFNTS NVMMGAEARLIQVGSTVYLYQRSSSWWV VGLTYKLDVSEITSQTGNTLNHVDPIAHTK FPRPSFRRDACARPNICPAVCVSGVYQDIW PISTATNNSNIVWVGQYLEAFYSRKDPRIG IATQYEWKVTNQLFNSNTEGGYSTTTCFR NTKRDKAYCVVISEYADGVFGSYRIVPQLI EIRTTTGKSE	7	81
KC403973	6234-6964	gb:KC403973:6234-6964|Organism:Human metapneumovirus|Strain Name:HMPV/USA/TN-82-518/1982/A|Protein Name:attachment glycoprotein G|Gene Symbol:G|Segment: 8	MEVKVENIRTIDMLKARVKNRVARSKCFK NASLILIGITTLSIALNIYLIINYTMQENTSES EHHTSSSPMESSRETPTVPIDNSDTNPSSQY PTQQSTEGSTLYFAASASSPETEPTSTPDTT SRPPFVDTHTTPPSASRTKTSPAVHTKNNP RISSRTHSPPWAMTRTVRRTTTLRTSSIRK RSSTASVQPDSSATTHKHEEASPVSPQTSA STTRPQRKSMEASTSTTYNQTS	6	82
KF015281	4511-5844	gb:KF015281:4511-5844|Organism:Canine	MRPAEQLIQENYKLTSLSMGRNFEVSGST TNLNFERTQYPDTFRAVVKVNQMCKLIAG VLTSAAVAVCVGVIMYSVFTSNHKANSM QNATIRNSTSAPPQPTAGPPTTEQGTTPKFT	6	83

		pneumovirus|Strain Name:dog/Bari/100-12/ITA/2012|Protein Name:attachment protein|Gene Symbol:G	KPPTKTTTHHEITEPAKMVTPSEDPYQCSS NGYLDRPDLPEDFKLVLDVICKPPGPEHHS TNCYEKREINLGSVCPDLVTMKANMGLN NGGGEEAAPYIEVITLSTYSNKRAMCVHN GCDQGFCFFLSGLSTDQKRAVLELGGQQA IMELHYDSYWKHYWSNSNCVVPRTNCNL TDQTVILFPSFNNKNQSQCTTCADSAGLD NKFYLTCDGLSRNLPLVGLPSLSPQAHKA ALKQSTGTTTAPTPETRNPTPAPRRSKPLS RKKRALCGVDSSREPKPTMPYWCPMLQL FPRRSNS		
KF973339	4624-5310	gb:KF973339:4624-5310|Organism:Respiratory syncytial virus type A|Strain Name:RSV-A/US/BID-V7358/2002|Protein Name:truncated attachment glycoprotein|Gene Symbol:G	MSKTKDQRAAKTLEKTWDTLNHLLFISSC LYKSNLKSIAQITLSILAMTIPTSLIIVATTFI ASANNKVTPTTAIIQDATSQIKNTTPTHLT QNPQPGISFFNLSGTISQTTAILAPTTPSVEP ILQSTTVKTKNTTTTQIQPSKLTTKQRQNK PPNKPNDDFHFEVFNFVPCSICSNNPTCWA ICKRIPSKKPGKKTTTKPTKKQTIKTTKKD LKPQTTKPKEAPTT	6	84
FJ215864	6383-8116	gb:FJ215864:6383-8116|Organism:Avian paramyxovirus 8|Strain Name:pintail/Wakuya/20/78|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MSNIASSLENIVEQDSRKTTWRAIFRWSVL LITTGCLALSIVSIVQIGNLKIPSVGDLADE VVTPLKTTLSDTLRNPINQINDIFRIVALDIP LQVTSIQKDLASQFSMLIDSLNAIKLGNGT NLIIPTSDKEYAGGIGNPVFTVDAGGSIGFK QFSLIEHPSFIAGPTTTRGCTRIPTFHMSESH WCYSHNIIAAGCQDASASSMYISMGVLHV SSSGTPIFLTTASELIDDGVNRKSCSIVATQ FGCDILCSIVIEKEGDDYWSDTPTPMRHGR FSFNGSFVETELPVSSMFSSFSANYPAVGS GEIVKDRILFPIYGGIKQTSPEFTELVKYGL FVSTPTTVCQSSWTYDQVKAAYRPDYISG RFWAQVILSCALDAVDLSSCIVKIMNSSTV MMAAEGRIIKIGIDYFYYQRSSSWWPLAF VTKLDPQELADTNSIWLTNSIPIPQSKFPRP SYSENYCTKPAVCPATCVTGVYSDIWPLT SSSSLPSIIWIGQYLDAPVGRTYPRFGIANQ SHWYLQEDILPTSTASAYSTTTCFKNTARN RVFCVTIAEFADGLFGEYRITPQLYELVRN N	5	85
JX857409	6619-8542	gb:JX857409:6619-8542|Organism:Porcine parainfluenza virus 1|Strain Name:S206N|Protein Name:haemagglutinin	MEETKVKTSEYWARSPQIHATNHPNVQN REKIKEILTILISFISSLSLVLVIAVLIMQSLH NGTILRCKDVGLESINKSTYSISNAILDVIK QELITRIINTQSSVQVALPILINKKIQDLSLII EKSSKVHQNSPTCSGVAALTHVEGIKPLDP DDYWRCPSGEPYLEDELTLSLIPGPSMLAG TSTIDGCVRLPSLAIGKSLYAYSSNLITKGC QDIGKSYQVLQLGIITLNSDLHPDLNPIISH TYDINDNRKSCSVAVSETKGYQLCSMPRV NEKTDYTSDGIEDIVFDVLDLKGSSRSFKF SNNDINFDHPFSALYPSVGSGIIWKNELYF	5	86

		protein|Gene Symbol:H	LGYGALTTALQGNTKCNLMGCPGATQDN CNKFISSSWLYSKQMVNVLIQVKGYLSSK PSIIVRTIPITENYVGAEGKLVGTRERIYIYT RSTGWHTNLQIGVLNINHPITITWTDHRVL SRPGRSPCAWNNKCPRNCTTGVYTDAYPI SPDANYVATVTLLSNSTRNNPTIMYSSSDR VYNMLRLRNTELEAAYTTTSCIVHFDRGY CFHIIEINQKELNTLQPMLFKTAIPKACRIS NL		
KF9 082 38	7510-9249	gb:KF908238:7 510-9249|Organism :Human parainfluenza virus 4b|Strain Name:QLD-01|Protein Name:hemaggl utinin-neuraminidase protein|Gene Symbol:HN	MQDSRGNTQIFSQANSMVKRTWRLLFRIV TLILLISIFVLSLIIVLQSTPGNLQSDVDIIRK ELDELMENFETTSKSLLSVANQITYDVSVL TPIRQEATETNIIAKIKDHCKDRVVKGEST CTLGHKPLHDVSFLNGFNKFYFTYRDNVQ IRLNPLLDYPNFIPTATTPHGCIRIPSFSLSQ THWCYTHNTILRGCEDTASSKQYVSLGTL QTLENGDPYFKVEYSHYLNDRKNRKSCSV VAVLDGCLLYCVIMTKNETENFKDPQLAT QLLTYISYNGTIKERIINPPGSSRDWVHISP GVGSGILYSNYIIFPLYGGLMENSMIYNNQ SGKYFFPNSTKLPCSNKTSEKITGAKDSYTI TYFSKRLIQSAFLICDLRQFLSEDCEILIPSN DHMLVGAEGRLYNIENNIFYYQRGSSWW PYPSLYRIKLNSNKKYPRIIEIKFTKIEIAPRP GNKDCPGNKACPKECITGVYQDIWPLSYP NTAFPHKKRAYYTGFYLNNSLARRNPTFY TADNLDYHQQERLGKFNLTAGYSTTTCFK QTTTARLYCLYILEVGDSVIGDFQIFPFLRS IDQAIT	5	87
KT0 717 57	6066-7962	gb:KT071757:6 066-7962|Organism :Avian paramyxovirus 2|Strain Name:APMV-2/Emberiza spodocephala/China/Daxing'anling/974/2013|Protein Name:hemaggl utinin-neuraminidase protein|Gene Symbol:HN	MDALSRENLTEISQGGRRTWRMLFRILTL VLTLVCLAINIATIAKLDSIDTSKVQTWTTT ESDRVIGSLTDTLKIPINQVNDMFRIVALDL PLQMTTLQKEIASQVGFLAESINNFLSKNG SAGSVLVNDPEYAGGIGTSLFHGDSASGL DFEAPSLIEHPSFIPGPTTAKGCIRIPTFHMS ASHWCYSHNIIASGCQDAGHSSMYISMGV LKATQAGSPSFLTTASQLVDDKLNRKSCSI ISTTYGCDILCSLVVENEDADYRSDPPTDM ILGRLFFNGTYSESKLNTSAIFQLFSANYPA VGSGIVLGDEIAFPVYGGVKQNTWLFNQL KDYGYFAHNNVYKCNNSNIHQTVLNAYR PPKISGRLWSQVVLICPMRLFINTDCRIKVF NTSTVMMGAEARLIQVGSDIYLYQRSSSW WVVGLTYKLDFQELSSKTGNILNNVSPIA HAKFPRPSYSRDACARPNICPAVCVSGVY QDIWPISTAHNLSQVVWVGQYLEAFYARK DPWIGIATQYDWKKNVRLFNANTEGGYS TTTCFRNTKRDKAFCVIISEYADGVFGSYR IVPQLIEIRTTSKKGLPS	5	88
LC0 411 32	6605-8437	gb:LC041132:6 605-8437|Organism :Avian paramyxovirus goose/Shimane/67/2000|Strain	MQPGISEVSFVNDERSERGTWRLLFRILTI VLCLTSIGIGIPALIYSKEAATSGDIDKSLEA VKTGMSTLSSKIDESINTEQKIYRQVILEAP VSQLNMESNILSAITSLSYQIDGTSNSSGCG SPMHDQDFVGGINKEIWTTDNVNLGEITL TPFLEHLNFIPAPTTGNGCTRIPSFDLGLTH WCYTHNVILSGCQDYSSSFQYIALGVLKIS	4	89

		Name:goose/Shimane/67/2000|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	ATGHVFLSTMRSINLDDERNRKSCSISATSIGCDIICSLVTEREVDDYNSPAATPMIHGRLDFSGKYNEVDLNVGQLFGDWSANYPGVGGGSFLNGRVWFPIYGGVKEGTPTFKENDGRYAIYTRYNDTCPDSESEQVSRAKSSYRPSYFGGKLVQQAVLSIKIDDTLGLDPVLTISNNSITLMGAESRVLQIEEKLYFYQRGTSWFPSLIMYPLTVDDKMVRFEPPTIFDQFTRPGNHPCSADSRCPNACVTGVYTDGYPIVFHNNHSIAAVYGMQLNDVTNRLNPRSAVWYGVSMSNVIRVSSSTTKAAYTTSTCFKVKKTQRVYCLSIGEIGNTLFGEFRIVPLLLEVYSEKGKSLKSSFDGWEDISINNPLRPLDNHRVDPILISNYTSSWP
AF092942	4705-5478	gb:AF092942|Organism:Bovine respiratory syncytial virus|Strain Name:ATue51908|Protein Name:attachment glycoprotein|Gene Symbol:G	MSNHTHHLKFKTLKRAWKASKYFIVGLSCLYKFNLKSLVQTALTTLAMITLTSLVITAIIYISVGNAKAKPTSKPTIQQTQQPQNHTSPFFTEHNYKSTHTSIQSTTLSQLPNTDTTRETTYSHSINETQNRKIKSQSTLPATRKPPINPSGSNPPENHQDHNNSQTLPYVPCSTCEGNLACLSLCQIGPERAPSRAPTITLKKTPKPKTTKKPTKTTIHHRTSPEAKLQPKNNTAAPQQGILSSPEHHTNQSTTQI	3	90
AF326114	6691-847	gb:AF326114|Organism:Menangle virus|Strain Name:UNKNOWN-AF326114|Protein Name:attachment protein|Gene Symbol:HN	MWNSIPQLVSDHEEAKGKFTDIPLQDDTDSQHPSGSKSTCRTLFRTVSIILSLVILVLGVTSTMFSAKYSGGCATNSQLLGVSNLINQIQKSIDSLISEVNQVSITTAVTLPIKIMDFGKSVTDQVTQMIRQCNTVCKGPGQKPGSQNVRIMPSNNLSTFQNINMSARGIAYQDVPLTFVRPIKNPQSCSRFPSYSVSFGVHCFANAVTDQTCELNQNTFYRVVLSVSKGNISDPSSLETKAETRTPKGTPVRTCSIISSVYGCYLLCSKATVPESEEMKTIGFSQMFILYLSMDSKRIIYDNIVSSTSAIWSGLYPGEGAGIWHMGQLFFPLWGGIPFLTPLGQKILNSTLDIPEVGSKCKSDLTSNPAKTKDMLFSPYYGENVMVFGFLTCYLLSNVPTNCHADYLNSTVLGFGSKAQFYDYRGIVYMYIQSAGWYPFTQIFRITLQLKQNRLQAKSIKRIEVTSTTRPGNRECSVLRNCPYICATGLFQVPWIVNSDAITSKEVDNMVFVQAWAADFTEFRKGILSLCSQVSCPINDLLSKDNSYMRDTTTYCFPQTVPNILSCTSFVEWGGDSGNPINILEIHYEVIFVAS	3	91
GU206351	7500-9714	gb:GU206351:7500-9714|Organism:Avian paramyxovirus 5|Strain Name:budgerigar/Kunitachi/74|Protein Name:hemagglutinin	MDKSYYTEPEDQRGNSRTWRLLFRLIVLTLLCLIACTSVSQLFYPWLPQVLSTLISLNSSIITSSNGLKKEILNQNIKEDLIYREVAINIPLTLDRVTVEVGTAVNQITDALRQLQSVNGSAAFALSNSPDYSGGIEHLVFQRNTLINRSVSVSDLIEHPSFIPTPTTQHGCTRIPTFHLGTRHWCYSHNIIGQGCADSGASMMYISMGALGVSSLGTPTFTTSATSILSDSLNRKSCSIVATTEGCDVLCSIVTQTEDQDYADHTPTPMIHGRLWFNGTYTERSLSQSLFLGTWAAQYPAVGSGIMTPGRVIFPFYGGVIPNSPLFLDLERF	3	92

		neuraminidase protein|Gene Symbol:HN	ALFTHNGDLECRNLTQYQKEAIYSAYKPP KIRGSLWAQGFIVCSVGDMGNCSLKVINT STVMMGAEGRLQLVGDSVMYYQRSSSW WPVGILYRLSLVDIIARDIQVVINSEPLPLS KFPRPTWTPGVCQKPNVCPAVCVTGVYQ DLWAISAGETLSEMTFFGGYLEASTQRKD PWIGVANQYSWFMRRRLFKTSTEAAYSSS TCFRNTRLDRNFCLLIFELTDNLLGDWRIV PLLFELTIV		
JQ0017 76	8170-10275	gb:JQ001776:8170-10275|Organism:Cedar virus|Strain Name:CG1a|Protein Name:attachment glycoprotein|Gene Symbol:G	MLSQLQKNYLDNSNQQGDKMNNPDKKLS VNFNPLELDKGQKDLNKSYYVKNKNYNV SNLLNESLHDIKFCIYCIFSLLIIITIINIITISIV ITRLKVHEENNGMESPNLQSIQDSLSSLTN MINTEITPRIGILVTATSVTLSSSINYVGTKT NQLVNELKDYITKSCGFKVPELKLHECNIS CADPKISKSAMYSTNAYAELAGPPKIFCKS VSKDPDFRLKQIDYVIPVQQDRSICMNNPL LDISDGFFTYIHYEGINSCKKSDSFKVLLSH GEIVDRGDYRPSLYLLSSHYHPYSMQVINC VPVTCNQSSFVFCHISNNTKTLDNSDYSSD EYYITYFNGIDRPKTKKIPINNMTADNRYI HFTFSGGGGVCLGEEFIIPVTTVINTDVFTH DYCESFNCSVQTGKSLKEICSESLRSPTNSS RYNLNGIMIISQNNMTDFKIQLNGITYNKL SFGSPGRLSKTLGQVLYYQSSMSWDTYLK AGFVEKWKPFTPNWMNNTVISRPNQGNC PRYHKCPEICYGGTYNDIAPLDLGKDMYV SVILDSDQLAENPEITVFNSTTILYKERVSK DELNTRSTTTSCFLFLDEPWCISVLETNRF NGKSIRPEIYSYKIPKYC	3	93
KP271123	6644-8431	gb:KP271123:6644-8431|Organism:Teviot virus|Strain Name:Geelong|Protein Name:attachment protein|Gene Symbol:HN	MWSTQASKHPAMVNSATNLVDIPLDHPSS AQFPINRKRTGRLIYRLFSILCNLILISILISL VVIWSRSSRDCAKSDGLSSVDNQLSSLSRS INSLITEVNQISVTTAINLPIKLSEFGKSVVD QVTQMIRQCNAACKGPGEKPGIQNVRINIP NNFSTYSELNRTANSLNFQSRTALFARPNP YPKTCSRFPSYSVYFGIHCFSHAVTDSSCE LSDSTYYRLVIGVADKNLSDPADVKYIGE TTTPVRVQTRGCSVVSSIYGCYLLCSKSNQ DYQDDFREQGFHQMFILFLSRELKTTFFDD MVSSTTVTWNGLYPGEGSGIWHMGHLVF PLWGGIRFGTHASEGILNSTLELPPVGPSC KRSLADNGLINKDVLFSPYFGDSVMVFAY LSCYMLSNVPTHCQVETMNSSVLGFGSRA QFYDLKGIVYLYIQSAGWFSYTQLFRLSLQ SKGYKLSVKQIKRIPISSTSRPGTEPCDIIHN CPYTCATGLFQAPWIVNGDSIRDRDVRNM AFVQAWSGAINTFQRPFMSICSQYSCPLSE LLDSESSIMRSTTTYCFPSLTESILQCVSFIE WGGPVGNPISINEVYSSISFRPD	3	94
AY286409	7644-9542	gb:AY286409|Organism:Mossman virus|Strain Name:UNKNOWN-	MVDPPAVSYYTGTGRNDRVKVVTTQSTN PYWAHNPNQGLRRLIDMVVNVIMVTGVIF ALINIILGIVIISQSAGSRQDTSKSLDIIQHVD SSVAITKQIVMENLEPKIRSILDSVSFQIPKL LSSLLGPGKTDPPIALPTKASTPVIPTEYPSL NTTTCLRIEESVTQNAAALFNISFDLKTVM	2	95

		AY286409|Protein Name:attachment glycoprotein|Gene Symbol:G	YELVTRTGGCVTLPSYSELYTRVRTFSTAI RNPKTCQRAGQETDLNLIPAFIGTDTGILIN SCVRQPVIATGDGIYALTYLTMRGTCQDH RHAVRHFEIGLVRRDAWWDPVLTPIHHFT EPGTPVFDGCSLTVQNQTALALCTLTTDG PETDIHNGASLGLALVHFNIRGEFSKHKVD PRNIDTQNQGLHLVTTAGKSAVKKGILYS FGYMVTRSPEPGDSKCVTEECNQNNQEKC NAYSKTTLDPDKPRSMIIFQIDVGAEYFTV DKVVVVPRTQYYQLTSGDLFYTGEENDLL YQLHNKGWYNKPIRGRVTFDGQVTLHEH SRTYDSLSNQRACNPRLGCPSTCELTSMAS YFPLDKDFKAAVGVIALRNGMTPIITYSTD DWRNHWKYIKNADLEFSESSLSCYSPNPP LDDYVLCTAVITAKVMSNTNPQLLATSW YQYDKCHT		
AY900001	7809-9938	gb:AY900001|Organism:J-virus|Strain Name:UNKNOWN-AY900001|Protein Name:attachment glycoprotein|Gene Symbol:G	MNPVAMSNFYGINQADHLREKGDQPEKG PSVLTYVSLITGLLSLFTIIALNVTNIIYLTG SGGTMATIKDNQQSMSGSMRDISGMLVE DLKPKTDLINSMVSYTIPSQISAMSAMIKN EVLRQCTPSFMFNNTICPIAEHPVHTSYFEE VGIEAISMCTGTNRKLVVNQGINFVEYPSF IPGSTKPGGCVRLPSFSLGLEVFAYAHAIT QDDCTSSSTPDYYFSVGRIADHGTDVPVFE TLAEWFLDDKMNRRSCSVTAAGKGGWL GCSILVGSFTDELTSPEVNRISLSYMDTFGK KKDWLYTGSEVRADQSWSALFFSVGSGV VIGDTVYFLVWGGLNHPINVDAMCRAPG CQSPTQSLCNYAIKPQEWGGNQIVNGILHF KHDTNEKPTLHVRTLSPDNNWMGAEGRL FHFHNSGKTFIYTRSSTWHTLPQVGILTLG WPLSVQWVDITSISRPGQSPCEYDNRCPHQ CVTGVYTDLFPLGVSYEYSVTAYLDQVQS RMNPKIALVGAQEKIYEKTITTNTQHADY TTTSCFAYKLRVWCVSIVEMSPGVITTRQP VPFLYHLNLGCQDTSTGSLTPLDAHGGTY LNTDPVGNKVDCYFVLHEGQIYFGMSVGP INYTYSIVGRSREIGANMNVSLNQLCHSVY TEFLKEKEHPGTRNNIDVEGWLLKRIETLN GTKIFGLDDLEGSGPGHQSGPEDPSIAPIGH N	2	96
EF199772	6150-6944	gb:EF199772:6150-6944|Organism:Avian metapneumovirus|Strain Name:PL-2|Protein Name:attachment glycoprotein|Gene Symbol:G	MEVKVENVGKSQELKVKVKNFIKRSDCK KKLFALILGLVSFELTMNIMLSVMYVESNE ALSLCRIQGTPAPRDNKTNTENATKETTLH TTTTTRDPEVRETKTTKPQANEGATNPSR NLTTKGDKHQTTRATTEAELEKQSKQTTE PGTSTQKHTPARPSSKSPTTTQATAQPTTP TAPKASTAPKNRQATTKKTETDTTTASRA RNTNNPTETATTTPKATTETGKGKEGPTQ HTTKEQPETTARETTTPQPRRTAGASPRAS	2	97
JF424833	5981-7156	gb:JF424833:5981-7156|Organism	MGSKLYMVQGTSAYQTAVGFWLDIGRRY ILAIVLSAFGLTCTVTIALTVSVIVEQSVLE ECRNYNGGDRDWWSTTQEQPTTAPSATP	2	98

		:Avian metapneumovirus|Strain Name:IT/Ty/A/259-01/03|Protein Name:attachment protein|Gene Symbol:G	AGNYGGLQTARTRKSESCLHVQISYGDM YSRSDTVLGGFDCMGLLVLCKSGPICQRD NQVDPTALCHCRVDLSSVDCCKVNKISTN SSTTSEPQKTNPAWPSQDNTDSDPNPQGIT TSTATLLSTSLGLMLTSKTGTHKSGPPQAL PGSNTNGKTTTDRELGSTNQPNSTTNGQH NKHTQRMTLPPSYDNTRTILQHTTPWEKT FSTYKPTHSPTNESDQSLPTTQNSINCEHFD PQGKEKICYRVGSYNSNITKQCRIDVPLCS TYNTVCMKTYYTEPFNCWRRIWRCLCDD GVGLVEWCCTS		
JN689227	7918-12444	gb:JN689227:7918-12444|Organism:Tailam virus|Strain Name:TL8K|Protein Name:attachment glycoprotein|Gene Symbol:G	MSQLAAHNLAMSNFYGIHQGGQSTSQKE EEQPVQGVIRYASMIVGLLSLFTIIALNVTN IIYMTESGGTMQSIKNAQGSIDGSMKDLSG TIMEDIKPKTDLINSMVSYNIPAQLSMIHQI IKNDVLKQCTPSFMFNNTICPLAENPTHSR YFEEVNLDSISECSGNEMSLELGTEPEFIEY PSFAPGSTKPGSCVRLPSFSLSSTVFAYTHT IMGHGCSELDVGDHYLAIGRIADAGHEIPQ FETISSWFINDKINRRSCTVAAGVMETWM GCVIMTETFYDDLDSLDTGKITISYLDVFG RKKEWIYTRSEILYDYTYTSVYFSIGSGVV VGDTVYFLLWGSLSSPIEETAYCYAPGCSN YNQRMCNEAQRPAKFGHRQMANAILRFK TNSMGKPSISVRTLSPTVIPFGTEGRLIYSD FTKIIYLYLRSTSWYVLPLTGLLILGPPVSIS WVTQEAVSRPGEYPCGASNRCPKDCITGV YTDLFPLGARYEYAVTVYLNAETYRVNPT LALIDRSKIIARKKITTESQKAGYTTTTCFV FKLRIWCMSVVELAPATMTAFEPVPFLYQ LDLTCKRNNGTTAMQFSGQDGMYKSGRY KSPRNECFFEKVSNKYYFVVSTPEGIQPYE VRDLTPERVSHVIMYISDVCAPALSAFKKL IPAMRPITTLTIGNWQFRPVDISGGLRVNIY RNLTRYGDLSMSAPEDPGTDTFPGTHAPS KGHEEVGHYTLPNEKLSEVTTAAVKTKES LNLIPDTKDTRGEEENGSGLNEIITGHTTPG HIKTHPAETKVTKHTVIIPQIEEDGSGATTS TELQDETGYHTEDYNTTNTNGSLTAPNER NNYTSGDHTVSGEDITHTITVSDRTKTTQT LPTDNTFNQTPTKIQEGSPKSESTPKDYTAI ESEDSHFTDPTLIRSTPEGTIVQVIGDQFHS AVTQLGESNAIGNSEPIDQGNNLIPTTDRG TMDNTSSQSHSSTTSTQGSHSAGHGSQSN MNLTALADTDSVTDQSTSTQEIDHEHENV SSILNPLSRHTRVMRDTVQEALTGAWGFIR GMIP	2	99
KC562242	6178-6926	gb:KC562242:6178-6926|Organism:Human metapneumovirus|Strain Name:HMPV/USA/C1-334/2004/B|Pro	MEVRVENIRAIDMFKAKIKNRIRNSRCYR NATLILIGLTALSMALNIFLIIDHATLRNMI KTENCANMPSAEPSKKTPMTSIAGPSTKPN PQQATQWTTENSTSPAATLEGHPYTGTTQ TPDTTAPQQTTDKHTALPKSTNEQITQTTT EKKTTRATTQKRKKEKKTQTKPQVQLQP KQPTPPTKSEMQVRQSQHPTDPELTPLPKA VNRQPGQQNQAPHHIMHGEVQDPGERNT QVSHPSS	2	100

		tein Name:attachment glycoprotein G|Gene Symbol:G			
KC915036	6154-7911	gb:KC915036:6154-7911|Organism:Avian metapneumovirus type C|Strain Name:GDY|Protein Name:attachment glycoprotein|Gene Symbol:G	MEVKIENVGKSQELRVKVKNFIKRSDCKK KLFALILGLISFDITMNIMLSVMYVESNEA LSSCRVQGTPAPRDNRTNTENTAKETTLH TMTTTRNTEAGGTKTTKPQADERATSPSK NPTIGADKHKTTRATTEAEQEKQSKQTTE PGTSTPKHIPARPSSKSPATTKTTTQPTTPT VAKGGTAPKNRQTTTKKTEADTPTTSRAK QTNKPTGTETTPPRATTETDKDKEGPTQH TTKEQPETTAGGTTTPQPRRTTSRPAPTTN TKEGAETTGTRTTKSTQTSASPPRPTRSTPS KTATGTNKRATTTKGPNTASTDRRQQTRT TPKQDQQTQTKAKTTTNKAHAKAATTPE HNTDTTDSMKENSKEDKTTRDPSSKATTK QENTSKGTTATNLGNNTEAGARTPPTTTP TRHTTEPATSTAGGHTKARTTRWKSTAAR QPTRNNTTADTKTAQSKQTTPAQLGNNTT PENTTPPDNKSNSQTNVAPTEEIEIGSSLWR RRYVYGPCRENALEHPMNPCLKDNTTWI YLDNGRNLPAGYYDSKTDKIICYGIYRGN SYCYGRIECTCKNGTGLLSYCCNSYNWS	2	101
LC168749	7239-9196	gb:LC168749:7239-9196|Organism:Rinderpest morbillivirus|Strain Name:Lv|Protein Name:H protein|Gene Symbol:H	MSSPRDRVNAFYKDNLQFKNTRVVLNKE QLLIERPYMLLAVLFVMFLSLVGLLAIAGI RLHRAAVNTAEINSGLTTSIDITKSIEYQVK DVLTPLFKIIGDEVGLRTPQRFTDLTKFISD KIKFLNPDKEYDFRDINWCISPPERIKINYD QYCAHTAAEELITMLVNSSLAGTAVLRTS LVNLGRSCTGSTTTKGQFSNMSLALSGIYS GRGYNISSMITITEKGMYGSTYLVGKHNQ GARRPSTAWQRDYRVFEVGIIRELGVGTP VFHMTNYLELPRQPELEICMLALGEFKLA ALCLADNSVALHYGGLRDDHKIRFVKLG VWPSPADSDTLATLSAVDPTLDGLYITTHR GIIAAGKAVWAVPVTRTDDQRKMGQCRR EACREKPPPFCNSTDWEPLEAGRIPAYGIL TIRLGLADKPEIDIISEFGPLITHDSGMDLYT PLDGNEYWLTIPPLQNSALGTVNTLVLEPS LKISPNILTLPIRSGGGDCYTPTYLSDLADD DVKLSSNLVILPSRNLQYVSATYDTSRVEH AIVYYIYSTGRLSSYYYPVKLPIKGDPVSL QIGCFPWGLKLWCHHFCSVIDSGTGKQVT HTGAVGIEITCNSR	2	102
LC187310	8144-9871	gb:LC187310:8144-9871|Organism:Avian paramyxovirus 10|Strain Name:rAPMV-10-FI324/YmHA|Protein	MDSSQMNILDAMDRESSKRTWRGVFRVT TIIMVVTCVVLSAITLSKVAHPQGFDTNEL GNGIVDRVSDKITEALTVPNNQIGEIFKIVA LDLHVLVSSSQQAIAGQIGMLAESINSILSQ NGSASTILSSSPEYAGGIGVPLFSNKLTNGT VIKPITLIEHPSFIPGPTTIGGCTRIPTFHMAS SHWCYSHNIIEKGCKDSGISSMYISLGVLQ VLKKGTPVFLVTASAVLSDDRNRKSCSIIS SRFGCEILCSLVTEAESDDYKSDTPTGMVH GRLYFNGTYREGLVDTETIFRDFSANYPG	2	103

		Name:hemagglutinin-neuraminidase|Gene Symbol:HN	VGSGEIVEGHIHFPIYGGVKQNTGLYNSLT PYWLDAKNKYDYCKLPYTNQTIQNSYKPP FIHGRFWAQGILSCELDLFNLGNCNLKIIRS DKVMMGAESRLMLVGSKLLMYQRASSW WPLGITQEIDIAELHSSNTTILREVKPILSSK FPRPSYQPNYCTKPSVCPAVCVTGVYTDM WPISITGNISDYAWISHYLDAPTSRQQPRIG IANQYFWIHQTTIFPTNTQSSYSTTTCFRNQ VRSRMFCLSIAEFADGVFGEFRIVPLLYEL RV		
NC_004074	6590-8563	gb:NC_004074:6590-8563|Organism:Tioman virus|Strain Name:UNKNOWN-NC_004074|Protein Name:attachment protein|Gene Symbol:HN	MWATSESKAPIPANSTLNLVDVPLDEPQTI TKHRKQKRTGRLVFRLLSLVLSLMTVILV LVILASWSQKINACATKEGFNSLDLQISGL VKSINSLITEVNQISITTAINLPIKLSDFGKSI VDQVTQMIRQCNAVCKGPGEKPGIQNIRI NIPNNFSTYLELNNTVKSIELQRRPALLARP NPIPKSCSRFPSYSVNFGIHCFAHAITDQSC ELSDKTYYRLAIGISDKNLSDPSDVKYIGE AFTPMGLQARGCSVISSIYGCYLLCSKSNQ GYEADFQTQGFHQMYILFLSRDLKTTLFN DMISSTTVVWNGLYPGEGAGIWHMGYLIF PLWGGIKIGTPASTSILNSTLDLPLVGPSCK STLEENNLINKDVLFSPYFGESVMVFGFLS CYMLSNVPTHCQVEVLNSSVLGFGSRSQL MDLKGIVYLYIQSAGWYSYTQLFRLSLQS RGYKLTVKQIRRIPISSTTRPGTAPCDVVH NCPYTCATGLFQAPWIVNGDSILDRDVRN LVFVQAWSGNFNTFQKGLISICNQYTCPLT TLLDNDNSIMRSTTTYCYPSLSEYNLQCQS FIEWGGPVGNPIGILEVHYIIKFK	2	104
NC_005283	7091-8905	gb:NC_005283:7091-8905|Organism:Dolphin morbillivirus|Strain Name:UNKNOWN-NC_005283|Protein Name:haemagglutinin protein|Gene Symbol:H	MSSPRDKVDAFYKDIPRPRNNRVLLDNER VIIERPLILVGVLAVMFLSLVGLLAIAGVRL QKATTNSIEVNRKLSTNLETTVSIEHHVKD VLTPLFKIIGDEVGLRMPQKLTEIMQFISNK IKFLNPDREYDFNDLHWCVNPPDQVKIDY AQYCNHIAAEELIVTKFKELMNHSLDMSK GRIFPPKNCSGSVITRGQTIKPGLTLVNIYT TRNFEVSFMVTVISGGMYGKTYFLKPPEP DDPFEFQAFRIFEVGLVRDVGSREPVLQMT NFMVIDEDEGLNFCLLSVGELRLAAVCVR GRPVVTKDIGGYKDEPFKVVTLGIIGGGLS NQKTEIYPTIDSSIEKLYITSHRGIIRNSKAR WSVPAIRSDDKDKMEKCTQALCKSRPPPS CNSSDWEPLTSNRIPAYAYIALEIKEDSGLE LDITSNYGPLIIHGAGMDIYEGPSSNQDWL AIPPLSQSVLGVINKVDFTAGFDIKPHTLTT AVDYESGKCYVPVELSGAKDQDLKLESNL VVLPTKDFGYVTATYDTSRSEHAIVYYVY DTARSSSYFFPFRIKARGEPIYLRIECFPWS RQLWCHHYCMINSTVSNEIVVVDNLVSIN MSCSR	2	105
NC_007803	7978-12504	gb:NC_007803:7978-12504|Organism:Beilong virus|Strain	MSQLAAHNLAMSNFYGTHQGDLSGSQKG EEQQVQGVIRYVSMIVSLLSLFTIIALNVTN IIYMTESGGTMQSIKTAQGSIDGSMREISG VIMEDVKPKTDLINSMVSYNIPAQLSMIHQ IIKNDVPKQCTPSFMFNNTICPLAENPTHSR	2	106

		Name:Li|Protein Name:attachment glycoprotein|Gene Symbol:G	YFEEVNLDSISECSGPDMHLGLGVNPEFIE FPSFAPGSTKPGSCVRLPSFSLSTTVFAYTH TIMGHGCSELDVGDHYFSVGRIADAGHEIP QFETISSWFINDKINRRSCTVAAGAMEAW MGCVIMTETFYDDRNSLDTGKLTISYLDV FGRKKEWIYTRSEILYDYTYTSVYFSVGSG VVVGDTVYFLIWGSLSSPIEETAYCFAPDC SNYNQRMCNEAQRPSKFGHRQMVNGILK FKTTSTGKPLLSVGTLSPSVVPFGSEGRLM YSEITKIIYLYLRSTSWHALPLTGLFVLGPP TSISWIVQRAVSRPGEFPCGASNRCPKDCV TGVYTDLFPLGSRYEYAATVYLNSETYRV NPTLALINQTNIIASKKVTTESQRAGYTTTT CFVFKLRVWCISVVELAPSTMTAYEPIPFL YQLDLTCKGKNGSLAMRFAGKEGTYKSG RYKSPRNECFFEKVSNKYYFIVSTPEGIQP YEIRDLTPDRMPHIIMYISDVCAPALSAFK KLLPAMRPITTLTIGNWQFRPVEVSGGLRV NIGRNLTKEGDLTMSAPEDPGSNTFPGNHI PGNGILDAGYYTVEYPKE		
NC_009489	6559-8512	gb:NC_009489:6559-8512|Organism:Mapuera virus|Strain Name:BeAnn 370284|Protein Name:attachment protein|Gene Symbol:HN	MASLQSEPGSQKPHYQSDDQLVKRTWRSF FRFSVLVVTITSLALSIITLIGVNRISTAKQIS NAFAAIQANILSSIPDIRPINSLLNQLVYTSS VTLPLRISSLESNVLAAIQEACTYRDSQSSC SATMSVMNDQRYIEGIQVYSGSFLDLQKH TLSPPIAFPSFIPTSTTTVGCTRIPSFSLTKTH WCYTHNYIKTGCRDATQSNQYIALGTIYT DPDGTPGFSTSRSQYLNDGVNRKSCSISAV PMGCALYCFISVKEEVDYYKGTVPPAQTLI LFFFNGTVHEHRIVPSSMNSEWVMLSPGV GSGVFYNNYIIFPLYGGMTKDKAEKRGEL TRFFTPKNSRSLCKMNDSVFSNAAQSAYY PPYFSSRWIRSGLLACNWNQIITTNCEILTF SNQVMMMGAEGRLILINDDLFYYQRSTS WWPRPLVYKLDIELNYPDSHIQRVDQVEV TFPTRPGWGGCVGNNFCPMICVSGVYQD VWPVTNPVNTTDSRTLWVGGTLLSNTTRE NPASVVTSGGSISQTVSWFNQTVPGAYSTT TCFNDQVQGRIFCLIIFEVGGGLLGEYQIVP FLKELKYQGAVHA	2	107
NC_017937	6334-8544	gb:NC_017937:6334-8544|Organism:Nariva virus|Strain Name:UNKNOWN-NC_017937|Protein Name:attachment protein|Gene Symbol:H	MAPINYPASYYTNNAERPVVITTKSTESKG QRPLPLGNARFWEYFGHVCGTLTFCMSLI GIIVGIIALANYSSDKDWKGRIGGDIQVTR MATEKTVKLILEDTTPKLRNILDSVLFQLP KMLASIASKINTQTPPPPTTSGHSTALATQ CSSNCENRPEIGYDYLRQVEQSLQRITNISI QLLEASEIHSMAGAYPNALYKIRTQDSWS VTAKECPLQAFQPNLNLIPAMIGTATGALI RNCVRQPVIVVDDGVYMLTYLAMRGSCQ DHQKSVRHFEMGVITSDPFGDPVPTPLRH WTKRALPAYDGCALAVKGHAGFALCTET SVGPLRDRTAKRKPNIVLFKASLVGELSER VIPPQSWLSGFSFFSVYTVAGKGYAYHSKF HAFGNVVRVGQSEYQAKCRGTGCPTANQ DDCNTAQRVSQEDNTYLHQAILSVDIDSVI DPEDVVYVIERDQYYQASAGDLYRVPETG	2	108

			EILYNLHNGGWSNEVQVGRIQPSDRFYMR EIQLTSTRVPAPNGCNRVKGCPGGCVAVIS PAFTPMHPEFNVGVGIFPMNQPHNPSIMH VQQQTELFWKPIVGGNITLHESSIACYSTV PPNPSYDLCIGVMTLLLHQGQLPQFQALS WYQPTMCNGNAPQNRRALIPVIVEDSKA MSVSSDAPRTP		
NC_025256	9117-11015	gb:NC_025256:9117-11015|Organism:Bat Paramyxovirus Eid_hel/GH-M74a/GHA/2009|Strain Name:BatPV/Eid_hel/GH-M74a/GHA/2009|Protein Name:glycoprotein|Gene Symbol:G	MPQKTVEFINMNSPLERGVSTLSDKKTLN QSKITKQGYFGLGSHSERNWKKQKNQND HYMTVSTMILEILVVLGIMFNLIVLTMVYY QNDNINQRMAELTSNITVLNLNLNQLTNKI QREIIPRITLIDTATTITIPSAITYILATLTTRI SELLPSINQKCEFKTPTLVLNDCRINCTPPL NPSDGVKMSSLATNLVAHGPSPCRNFSSV PTIYYYRIPGLYNRTALDERCILNPRLTISST KFAYVHSEYDKNCTRGFKYYELMTFGEIL EGPEKEPRMFSRSFYSPTNAVNYHSCTPIV TVNEGYFLCLECTSSDPLYKANLSNSTFHL VILRHNKDEKIVSMPSFNLSTDQEYVQIIPA EGGGTAESGNLYFPCIGRLLHKRVTHPLC KKSNCSRTDDESCLKSYYNQGSPQHQVVN CLIRIRNAQRDNPTWDVITVDLTNTYPGSR SRIFGSFSKPMLYQSSVSWHTLLQVAEITD LDKYQLDWLDTPYISRPGGSECPFGNYCP TVCWEGTYNDVYSLTPNNDLFVTVYLKSE QVAENPYFAIFSRDQILKEFPLDAWISSART TTISCFMFNNEIWCIAALEITRLNDDIIRPIY YSFWLPTDCRTPYPHTGKMTRVPLRSTYN Y	2	109
NC_025347	6398-8418	gb:NC_025347:6398-8418|Organism:Avian paramyxovirus 7|Strain Name:APMV-7/dove/Tennessee/4/75|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	MESIGKGTWRTVYRVLTILLDVVIIILSVIA LISLGLKPGERIINEVNGSIHNQLVPLSGITS DIQAKVSSIYRSNLLSIPLQLDQINQAISSSA RQIADTINSFLALNGSGTFIYTNSPEFANGF NRAMFPTLNQSLNMLTPGNLIEFTNFIPTPT TKSGCIRIPSFSMSSSHWCYTHNIIASGCQD HSTSSEYISMGVVEVTDQAYPNFRTTLSIT LADNLNRKSCSIAATGFGCDILCSVVTETE NDDYQSPEPTQMIYGRLFFNGTYSEMSLN VNQMFADWVANYPAVGSGVELADFVIFP LYGGVKITSTLGASLSQYYYIPKVPTVNCS ETDAQQIEKAKASYSPPKVAPNIWAQAVV RCNKSVNLANSCEILTFNTSTMMMGAEGR LLMIGKNVYFYQRSSSYWPVGIIYKLDLQE LTTFSSNQLLSTIPIPFEKFPRPASTAGVCSK PNVCPAVCQTGVYQDLWVLYDLGKLENT TAVGLYLNSAVGRMNPFIGIANTLSWYNT TRLFAQGTPASYSTTTCFKNTKIDTAYCLSI LELSDSLLGSWRITPLLYNITLSIMS	2	110
NC_025348	6590-8548	gb:NC_025348:6590-8548|Organism:Tuhoko virus 2|Strain Name:UNKNOWN-NC_025348|Pr	MPPVPTVSQSIDEGSFTDIPLSPDDIKHPLS KKTCRKLFRIVTLIGVGLISILTIISLAQQTGI LRKVDSSDFQSYVQESFKQVLNLMKQFSS NLNSLIEITSVTLPFRIDQFGTDIKTQVAQL VRQCNAVCRGPIKGPTTQNIVYPALYETSL NKTLETKNVRIQEVRQEVDPVPGPGLSNG CTRNPSFSVYHGVWCYTHATSIGNCNGSL GTSQLFRIGNVLEGDGGAPYHKSLATHLL	2	111

		otein Name:hemagglutinin-neuraminidase| Gene Symbol:HN	TTRNVSRQCSATASYYGCYFICSEPVLTER DDYETPGIEPITIFRLDPDGNWVVFPNINRF TEYSLKALYPGIGSGVLFQGKLIFPMYGGI DKERLSALGLGNIGLIERRMADTCNHTEK ELGRSFPGAFSSPYYHDAVMLNFLLICEMI ENLPGDCDLQILNPTNMSMGSESQLSVLD NELFLYQRSASWWPYTLIYRLNMRYTGK YLKPKSIIPMVIKSNTRPGYEGCNHERVCP KVCVTGVFQAPWILSIGRDHKERVSNVTY MVAWSMDKSDRTYPAVSVCGSDTCKLTV PLGDSKVHSAYSVTRCYLSRDHMSAYCLV IFELDARPWAEMRIQSFLYKLILT		
NC_025350	6451-8341	gb:NC_025350:6451-8341|Organism:Tuhoko virus 3|Strain Name:UNKNOWN-NC_025350|Protein Name:hemagglutinin-neuraminidase| Gene Symbol:HN	MHNRTQSVSSIDTSSDVYLPRRKKAVTKF TFKKIFRVLILTLLLSIIIIIAVIFPKIDHIRETC DNSQILETITNQNSEIKNLINSAITNLNVLLT STTVDLPIKLNNFGKSIVDQVTMMVRQCN AVCRGPGDRPTQNIELFKGLYHTSPPSNTS TKLSMITEASNPDDIVPRPGKLLGCTRFPSF SVHYGLWCYGHMASTGNCSGSSPSVQIIRI GSIGTNKDGTPKYVIIASASLPETTRLYHCS VTMTSIGCYILCTTPSVSETDDYSTMGIEK MSISFLSLDGYLTQLGQPTGLDNQNLYAL YPGPGSGVIFRDFLIFPMMGGIRLMDAQK MLNRNITYRGFPPSETCTESELKLKQEVAN MLTSPYYGEVLVLNFLYVCSLLDNIPGDCS VQLIPPDNMTLGAESRLYVLNGSLIMYKR GSSWWPYTELYQINYRVNNRAFRVRESVR INTTSTSRPGVQGCNLEKVCPKVCVSGIYQ SPGIISAPVNPTRQEEGLLYFLVWTSSMSSR TGPLSSLCDHSTCRITYPIGDDTIFIGYTDSS CFMSSIKEGIYCIAFLELDNQPYSMMAIRSL SYIIN	2	112
NC_025352	8716-11257	gb:NC_025352:8716-11257|Organism:Mojiang virus|Strain Name:Tonggua n1|Protein Name:attachment glycoprotein|Gene Symbol:G	MATNRDNTITSAEVSQEDKVKKYYGVET AEKVADSISGNKVFILMNTLLILTGAIITITL NITNLTAAKSQQNMLKIIQDDVNAKLEMF VNLDQLVKGEIKPKVSLINTAVSVSIPGQIS NLQTKFLQKYVYLEESITKQCTCNPLSGIF PTSGPTYPPTDKPDDDTTDDDKVDTTIKPI EYPKPDGCNRTGDHFTMEPGANFYTVPNL GPASSNSDECYTNPSFSIGSSIYMFSQEIRK TDCTAGEILSIQIVLGRIVDKGQQGPQASPL LVWAVPNPKIINSCAVAAGDEMGWVLCS VTLTAASGEPIPHMFDGFWLYKLEPDTEV VSYRITGYAYLLDKQYDSVFIGKGGGIQK GNDLYFQMYGLSRNRQSFKALCEHGSCL GTGGGGYQVLCDRAVMSFGSEESLITNAY LKVNDLASGKPVIIGQTFPPSDSYKGSNGR MYTIGDKYGLYLAPSSWNRYLRFGITPDIS VRSTTWLKSQDPIMKILSTCTNTDRDMCP EICNTRGYQDIFPLSEDSEYYTYIGITPNNG GTKNFVAVRDSDGHIASIDILQNYYSITSA TISCFMYKDEIWCIAITEGKKQKDNPQRIY AHSYKIRQMCYNMKSATVTVGNAKNITIR RY	2	113

NC_025363	6503-8347	gb:NC_025363:6503-8347|Organism:Avian paramyxovirus 12|Strain Name:Wigeon/Italy/3920_1/2005|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	MESATSQVSFENDKTSDRRTWRAVFRVL MIILALSSLCVTVAALIYSAKAAIPGNIDAS EQRILSSVEAVQVPVSRLEDTSQKIYRQVIL EAPVTQLNMETNILNAITSLSYQIDASANS SGCGAPVHDSDFTGGVGRELLQEAEVNLT IIRPSKFLEHLNFIPAPTTGNGCTRIPSFDLG QTHWCYTHNVVLNGCRDRGHSFQYVALG ILRTSATGSVFLSTLRSVNLDDDRNRKSCS VSATPIGCEMLCSLVTETEEGDYDSIDPTP MVHGRLGFDGKYREVDLSEKEIFADWRA NYPAVGGGAFFGNRVWFPVYGGLKEGTQ SERDAEKGYAIYKRFNNTCPDDNTTQIAN AKASYRPSRFGGRFIQQGILSFKVEGNLGS DPILSLTDNSITLMGAEARVMNIENKLYLY QRGTSWFPSALVYPLDVANTAVKVRAPYI FDKFTRPGGHPCSASSRCPNVCVTGVYTD AYPLVFSRSHDIVAVYGMQLAAGTARLDP QAAIWYGNEMSTPTKVSSSTTKAAYTTST CFKVTKTKRIYCISIAEIGNTLFGEFRIVPLL IEVQKTPLTRRSELRQQMPQPPIDLVIDNPF CAPSGNLSRKNAIDEYANSWP	2	114
NC_025373	6619-8605	gb:NC_025373:6619-8605|Organism:Avian paramyxovirus 3|Strain Name:turkey/Wisconsin/68|Protein Name:hemagglutinin|Gene Symbol:HN	MEPTGSKVDIVPSQGTKRTCRTFYRLLILIL NLIIIILTIISIYVSISTDQHKLCNNEADSLLH SIVEPITVPLGTDSDVEDELREIRRDTGINIP IQIDNTENIILTTLASINSNIARLHNATDESP TCLSPVNDPRFIAGINKITKGSMIYRNFSNL IEHVNFIPSPTTLSGCTRIPSFSLSKTHWCYS HNVISTGCQDHAASSQYISIGIVDTGLNNE PYLRTMSSRLLNDGLNRKSCSVTAGAGVC WLLCSVVTESESADYRSRAPTAMILGRFN FYGDYTESPVPASLFSGRFTANYPGVGSGT QLNGTLYFPIYGGVVNDSDIELSNRGKSFR PRNPTNPCPDPEVTQSQRAQASYYPTRFGR LLIQQAILACRISDTTCTDYYLLYFDNNQV MMGAEARIYYLNNQMYLYQRSSSWWPH PLFYRFSLPHCEPMSVCMITDTHLILTYATS RPGTSICTGASRCPNNCVDGVYTDVWPLT EGTTQDPDSYYTVFLNSPNRRISPTISIYSY NQKISSRLAVGSEIGAAYTTSTCFSRTDTG ALYCITIIEAVNTIFGQYRIVPILVQLISD	2	115
NC_025386	7541-9403	gb:NC_025386:7541-9403|Organism:Salem virus|Strain Name:UNKNOWN-NC_025386|Protein Name:attachment glycoprotein|Gene Symbol:G	MKAMHYYKNDFADPGTNDNSSDLTTNPFI SNQIKSNLSPPVLAEGHLSPSPIPKFRKILLT ISFVSTIVVLTVILLVLTIRILTIIEASAGDEK DIHTILSSLLNTFMNEYIPVFKNLVSIISLQIP QMLIDLKTSSTQMMQSLKTFPRDLETLST VTQSVAVLLEKAKSTIPDINKFYKNVGKV TFNDPNIKVLTLEVPAWLPIVRQCLKQDFR QVISNSTGFALIGALPSQLFNEFEGYPSLAI VSEVYAITYLKGVMFENQENFLYQYFEIG TISPDGYNKPYFLRHTSVMLSTFKLSGKCT AAVDYRGGIFLCTPSPKIPKILQNPPDLPTL TVVSIPFDGRYTIRNISLMLTDEADIIYDLD TLQGRGVLQAMRFYALVRVISSSSPRHFPF CKNSWCPTADDKICDQSRRLGADGNYPV MYGLISIPAHSSYQGNVSLKLIDPKYYAYT RDASLFYNSMTDTYHYSFGTRGWVSRPII	2	116

			GELLLGDDIVLTRYTVRSVSRATAGDCTT VSMCPQACSGGMNSIFYPLNFDKPQVTGV AIRQYERQQEGIIVVTMNDHYYYSVPIIKN GTLLISSVTDCFWLMGDLWCMSLMEKNN LPLGVRSLAHLTWNIHWSCS		
NC_025390	6647-8386	gb:NC_025390:6647-8386|Organism:Avian paramyxovirus 9|Strain Name:duck/New York/22/1978|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	MESGISQASLVNDNIELRNTWRTAFRVVS LLLGFTSLVLTACALHFALNAATPADLSSI PVAVDQSHHEILQTLSLMSDIGNKIYKQVA LDSPVALLNTESTLMSAITSLSYQINNAAN NSGCGAPVHDKDFINGVAKELFVGSQYNA SNYRPSRFLEHLNFIPAPTTGKGCTRIPSFD LAATHWCYTHNVILNGCNDHAQSYQYISL GILKVSATGNVFLSTLRSINLDDDENRKSC SISATPLGCDLLCAKVTEREEADYNSDAAT RLVHGRLGFDGVYHEQALPVESLFSDWV ANYPSVGGGSYFDNRVWFGVYGGIRPGS QTDLLQSEKYAIYRRYNNTCPDNNPTQIER AKSSYRPQRFGQRLVQQAILSIRVEPSLGN DPKLSVLDNTVVLMGAEARIMTFGHVAL MYQRGSSYFPSALLYPLSLTNGSAAASKPF IFEQYTRPGSPPCQATARCPNSCVTGVYTD AYPLFWSEDHKVNGVYGMMLDDITSRLN PVAAIFDRYGRSRVTRVSSSSTKAAYTTNT CFKVVKTKRVYCLSIAEIENTLFGEFRITPL LSEIIFDPNLEPSDTSRN	2	117
NC_025403	6692-8645	gb:NC_025403:6692-8645|Organism:Achimota virus 1|Strain Name:UNKNOWN-NC_025403|Protein Name:attachment protein|Gene Symbol:HN	MATNLSTITNGKFSQNSDEGSLTELPFFEH NRKVATTKRTCRFVFRSVITLCNLTILIVTV VVLFQQAGFIKRTESNQVCETLQNDMHGV VTMSKGVITTLNNLIEITSVNLPFQMKQFG QGIVTQVTQMVRQCNAVCKGPTIGPDIQNI VYPASYESMIKHPVNNSNILLSEIRQPLNFV PNTGKLNGCTRTPSFSVYNGFWCYTHAES DWNCNGSSPYMQVFRVGVVTSDYDYNVI HKTLHTKTSRLANVTYQCSTISTGYECYFL CSTPNVDEITDYKTPGIESLQIYKIDNRGTF AKFPITDQLNKELLTALYPGPGNGVLYQG RLLFPMHGGMQSSELNKVNLNNTVLSQFN DNKGCNATEIKLESEFPGTFTSPYYSNQVM LNYILICEMIENLPGNCDLQIVAPKNMSMG SESQLYSINNKLYLYQRSSSRWPYPLIYEV GTRLTNRQFRLRAINRFLIKSTTRPGSEGC NIYRVCPKVCVTGVYQAPWILHVSKAGSQ SIAKVLYAVAWSKDHMSRKGPLFSICDND TCFLTKSLASEHVHSGYSITRCYLENSERHI ICVVIMELDASPWAEMRIQSVIYNITLPS	2	118
NC_025404	6655-8586	gb:NC_025404:6655-8586|Organism:Achimota virus 2|Strain Name:UNKNOWN-NC_025404|Protein Name:attachme	MDNSMSISTISLDAQPRIWSRHESRRTWRN IFRITSLVLLGVTVIICIWLCCEVARESELEL LASPLGALIMAINTIKSSVVKMTTELNQVT FTTSIILPNKVDQFGQNVVSQVAQLVKQC NAVCRGHQDTPELEQFINQKNPTWILQPN YTTKLTNLHEIDSIIPLVDYPGFSKSCTRFPS FSEGSKFWCFTYAVVKEPCSDISSSIQVVK YGAIKANHSDGNPYLVLGTKVLDDGKFR RGCSITSSLYGCYLLCSTANVSEVNDYAHT PAYPLTLELISKDGITTDLSPTYTVQLDKW SALYPGIGSGVIFKGYLMFPVYGGLPFKSP	2	119

		nt protein|Gene Symbol:HN	LISASWVGPGNKWPVDFSCSEDQYSTFNF SNPYSALYSPHFSNNIVVSALFVCPLNENL PYSCEVQVLPQGNLTIGAEGRLYVIDQDL YYYQRSTSWWPYLQLYKLNIRITNRVFRV RSLSLLPIKSTTRPGYGNCTYFKLCPHICVT GVYQSPWLISIRDKRPHEEKNILYFIGWSP DEQIRQNPLVSLCHETACFINRSLATNKTH AGYSESHCVQSFERNKLTCTVFYELTAKP WAEMRVQSLLFQVDFL		
NC_025410	6799-8869	gb:NC_025410:6799-8869|Organism:Tuhoko virus 1|Strain Name:UNKNOWN-NC_025410|Protein Name:hemagglutinin-neuraminidase| Gene Symbol:HN	MDSRSDSFTDIPLDNRIERTVTSKKTWRSIF RVTAIILLIICVVVSSISLNQHNDAPLNGAG NQATSGFMDAIKSLEKLMSQTINELNQVV MTTSVQLPNRITKFGQDILDQVTQMVRQC NAVCRGPGVGPSIQNYVIQGHAPTVSFDPI SAEYQKFVFGITEKTLITAYHNPWECLRFP SQHLFDTTWCVSYQILTQNCSDHGPRITVI QLGEIMIANNLSTVFRDPVIKYIRHHIWLRS CSVVAYYSQCTIFCTSTNKSEPSDYADTGY EQLFLATLQSDGTFTEHSMHGVNIVHQWN AIYGGVGNGVIIGRNMLIPLYGGINYYDHN TTIVQTVDLRPYPIPDSCSQTDNYQTNYLP SMFTNSYYGTNLVVSGYLSCRLMAGTPTS CSIRVIPIENMTMGSEGQFYLINNQLYYYK RSSNWIRDTQVYLLSYSDKGNIIEITSAERY IFKSVTSPDEGDCVTNHGCPSNCIGGLFQA PWILNDFKLCGSNITCPKIVTVWADQPDK RSNPMLSIAETDKLLLHKSYINYHTAVGYS TVLCFDSPKLNLKTCVVLQELMSDDKLLI RISYSIVSIMVE	2	120
NC_028249	7059-9010	gb:NC_028249:7059-9010|Organism:Phocine distemper virus|Strain Name:PDV/Wadden_Sea.NLD/1988|Protein Name:hemagglutinin protein|Gene Symbol:H	MFSHQDKVGAFYKNNARANSSKLSLVTD EVEERRSPWFLSILLILLVGILILLAITGIRFH QVVKSNLEFNKLLIEDMEKTKAVHHQVK DVLTPLFKIIGDEVGLRLPQKLNEIKQFIVQ KTNFFNPNREFDFRELHWCINPPSKVKVNF TQYCEITEFKEATRSVANSILLLTLYRGRD DIFPPYKCRGATTSMGNVFPLAVSLSMSLI SKPSEVINMLTAISEGIYGKTYLLVTDDTE ENFETPEIRVFEIGFINRWLGDMPLFQTTN YRIISNNSNTKICTIAVGELALASLCTKESTI LLNLGDEESQNSVLVVILGLFGATHMDQL EEVIPVAHPSIEKIHITNHRGFIKDSVATWM VPALALSEQGEQINCLRSACKRRTYPMCN QTSWEPFGDKRLPSYGRLTLSLDVSTDLSI NVSVAQGPIIFNGDGMDYYEGTLLNSGWL TIPPKNGTILGLINQASKGDQFIVTPHILTFA PRESSTDCHLPIQTYQIQDDDVLLESNLVV LPTQSFEYVVATYDVSRSDHAIVYYVYDP ARTVSYTYPFRLRTKGRPDILRIECFVWDG HLWCHQFYRFQLDATNSTSVVENLIRIRFS CDRLDP	2	121
NC_028362	6951-8675	gb:NC_028362:6951-8675|Organism:Caprine parainfluenza virus 3|Strain	MEYWGHTNNPDKINRKVGVDQVRDRSKT LKIITFIISMMTSIMSTVALILILIMFIQNNNN NRIILQELRDETDAIEARIQKASNDIGVSIQS GINTRLLTIQNHVQNYIPLALTQQVSSLRES INDVITKREETQSKMPIQRMTHDDGIEPLIP DNFWKCPSGIPTISASPKIRLIPGPGLLATST	2	122

		Name:JS2013|Protein Name:hemagglutinin-neuraminidase|Gene Symbol:HN	TINGCIRLPSLVINNLIYAYTSNLITQGCQDIGKSYQVLQIGIITINSDLVPDLNPRITHTFDIDDNRKSCSLALRNADVYQLCSTPKVDERSDYSSIGIEDIVLDIVTSEGTVSTTRFTNNNITFDKPYAALYPSVGPGIYYDNKIIFLGYGGLEHEENGDVICNITGCPGKTQHDCNQASYSPWFSNRRMVNAIILVNKGLNKVPSLQVWTIPMRQNYWGSEGRLLLLGNKIYIYTRSTSWHSKLQLGTLDISNYNDIRIRWTHHDVLSRPGSEECPWGNTCPRGCITGVYNDAYPLNPSGSVVSSVILDSRTSRENPIITYSTDTSRVNELAIRNNTLSAAYTTTNCVTHYGKGYCFHIIEINHKSLNTLQPMLFKTEIPKSCN		
AB548428	5999-7261	gb:AB548428:5999-7261|Organism:Avian metapneumovirus|Strain Name:VCO3/60616|Protein Name:attachment glycoprotein|Gene Symbol:G	MGSELYIIEGVSSSEIVLKQVLRRSKKILLGLVLSALGLTLTSTIVISICISVEQVKLRQCVDTYWAENGSLHPGQSTENTSTRGKTTTKDPRRLQATGAGKFESCGYVQVVDGDMHDRSYAVLGGVDCLGLLALCESGPICQGDTWSEDGNFCRCTFSSHGVSCCKKPKSKATTAQRNSKPANSKSTPPVHSDRASKEHNPSQGEQPRRGPTSSKTTIASTPSTEDTAKPTISKPKLTIRPSQRGPSGSTKAASSTPSHKTNTRGTSKTTDQRPRTGPTPERPRQTHSTATPPPTTPIHKGRAPTPKPTTDLKVNPREGSTSPTAIQKNPTTQSNLVDCTLSDPDEPQRICYQVGTYNPSQSGTCNIEVPKCSTYGHACMATLYDTPFNCWRRTRRCICDSGGELIEWCCTSQ	1	123
AF079780	8118-10115	gb:AF079780|Organism:Tupaia paramyxovirus|Strain Name:UNKNOWN-AF079780|Protein Name:hemagglutinin|Gene Symbol:H	MDYHSHTTQTGSNETLYQDPLQSQSGSRDTLDGPPSTLQHYSNPPPYSEEDQGIDGPQRSQPLSTPHQYDRYYGVNIQHTRVYNHLGTIYKGLKLAFQILGWVSVIITMIITVTTLKKMSDGNSQDSAMLKSLDENFDAIQEVANLLDNEVRPKLGVTMTQTTFQLPKELSEIKRYLLRLERNCPVCGTEATPQGSKGNASGDTAFCPPCLTRQCSEDSTHDQGPGVEGTSRNHKGKINFPHILQSDDCGRSDNLIVYSINLVPGLSFIQLPSGTKHCIIDVSYTFSDTLAGYLIVGGVDGCQLHNKAIIYLSLGYYKTKMIYPPDYIAIATYTYDLVPNLRDCSIAVNQTSLAAICTSKKTKENQDFSTSGVHPFYIFTLNTDGIFTVTVIEQSQLKLDYQYAALYPATGPGIFIGDHLVFLMWGGLMTKAEGDAYCQASGCNDAHRTSCNIAQMPSAYGHRQLVNGLLMLPIKELGSHLIQPSLETISPKINWAGGHGRLYYNWEINTTYIYIEGKTWRSRPNLGIISWSKPLSIRWIDHSVARRPGARPCDSANDCPEDCLVGGYYDMFPMSSDYKTAITIIPTHHQWPSSPALKLFNTNREVRVVMILRPPNNVKKTTISCIRIMQTNWCLGFIIFKEGNNAWGQIYSYIYQVESTCPNTK	1	124
AY590688	6138-7935	gb:AY590688:6138-7935|Organism:Avian metapneumovirus|Strain	MEVKVENVGKSQELKVKVKNFIKRSDCKKKLFALILGLVSFELTMNIMLSVMYVESNEALSLCRIQGTPAPRDNKTNTENATKETTLHTTTTTRDPEVRETKTTKPQANEGATNPSRNLTTKGDKHQTTRATTEAELEKQSKQTTEPGTSTQKHTPTRPSSKSPTTTQAIAQLTTPT	1	125

		Name:Colorado |Protein Name:attachment glycoprotein|Gene Symbol:G	TPKASTAPKNRQATTKKTETDTTTASRAR NTNNPTETATTTPKATTETGKSKEGPTQHT TKEQPETTAGETTTPQPRRTASRPAPTTKIE EEAETTKTRTTKSTQTSTGPPRPTGGAPSG AATEGSGRAAAAGGPSAASAGGRRRTEA AAERDRRTRAGAGPTAGGARARTAAASE RGADTAGSAGGGPGGDGATGGLSGGAPA EREDASGGTAAAGPGDGTEADGRAPPAA ALAGRTTESAAGAAGDSGRAGTAGWGSA ADGRSTGGNAAAEAGAAQSGRAAPRQPS GGTAPESTAPPNSGGSGRADAAPTEEVGV GSGLWRGRYVCGPCGESVPEHPMNPCFG DGTAWICSDDGGSLPAGCYDGGTDGVVC CGVCGGNSCCCGRVECTCGGGAGLLSCC CGSYSWS		
EU403085	6620-8593	gb:EU403085:6620-8593|Organism:Avian paramyxovirus 3|Strain Name:APMV3/PKT/Netherland/449/75|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MESPPSGKDAPAFREPKRTCRLCYRATTLS LNLTIVVLSIISIYVSTQTGANNSCVNPTIVT PDYLTGSTTGSVEDLADLESQLREIRRDTG INLPVQIDNTENLILTTLASINSNLRFLQNA TTESQTCLSPVNDPRFVAGINRIPAGSMAY NDFSNLIEHVNFIPSPTTLSGCTRIPSFSLSK THWCYTHNVISNGCLDHAASSQYISIGIVD TGLNNEPYFRTMSSKSLNDGLNRKSCSVT AAANACWLLCSVVTEYEAADYRSRTPTA MVLGRFDFNGEYTEIAVPSSLFDGRFASNY PGVGSGTQVNGTLYFPLYGGVLNGSDIET ANKGKSFRPQNPKNRCPDSEAIQSFRAQDS YYPTRFGKVLIQQAIIACRISNKSCTDFYLL YFDNNRVMMGAEARLYYLNNQLYLYQR SSSWWPHPLFYSISLPSCQALAVCQITEAH LTLTYATSRPGMSICTGASRCPNNCVDGV YTDVWPLTKNDAQDPNLFYTVYLNNSTR RISPTISLYTYDRRIKSKLAVGSDIGAAYTT STCFGRSDTGAVYCLTIMETVNTIFGQYRI VPILLRVTSR	1	126
FJ977568	6139-7936	gb:FJ977568:6139-7936|Organism:Avian metapneumovirus|Strain Name:aMPV/MN/turkey/2a/97|Protein Name:attachment glycoprotein|Gene Symbol:G	MEVKVENVGKSQELKVKVKNFIKRSDCK KKLFALILGLVSFELTMNIMLSVMYVESNE ALSLCRIQGTPAPRDNKTNTENATKETTLH TTTTTRDPEVRETKTTKPQANEGATNPSR NLTTKGDKHQTTRATTEAELEKQSKQTTE PGTSTQKHTPARPSSKSPTTTQATAQPTTP TAPKASTAPKNRQATTKKTETDTTTASRA RNTNNPTETATTTPKATTETGKGKEGPTQ HTTKEQPETTARETTTPQPRRTASRPAPTT KIEEEAETTKTRTTKNTQTSTGPPRPTRSTP SKTATENNKRTTTTKRPNTASTDSRQQTR TTAEQDQQTQTRAKPTTNGAHPQTTTTPE HNTDTTNSTKGSPKEDKTTRDPSSKTPTEQ EDASKGTAAANPGGSAEADRRAPPATTPT GRTTESAAGTTGDDSGAETTRRRSAADRR PTGGSTAAEAGTAQSGRATPKQPSGGTAA GNTAPPNNESSGRADAAPAEEAGVGPSIR RGRHACGPRRESAPEHPTNPCPGDGTAWT RSDGGGNLPAGRHDSGADGAARRGARGG NPRRRGRAERTRGGGAGPPSCRCGSHNRS	1	127

HG934339	5997-7166	gb:HG934339:5997-7166|Organism:Avian metapneumovirus type D|Strain Name:Turkey/1985/Fr85.1|Protein Name:attachment glycoprotein|Gene Symbol:G	MGAKLYAISGASDAQLMKKTCAKLLEKVVPIIILAVLGITGTTTIALSISISIERAVLSDCTTQLRNGTTSGSLSNPTRSTTSTAVTTRDIRGLQTTRTRELKSCSNVQIAYGYLHDSSNPVLDSIGCLGLLALCESGPFCQRNYNPRDRPKCRCTLRGKDISCCKEPPTAVTTSKTTPWGTEVHPTYPTQVTPQSQPATMAHQTATANQRSSTTEPVGSQGNTTSSNPEQQTEPPPSPQHPPTTTSQDQSTETADGQEHTPTRKTPTATSNRRSPTPKRQETGRATPRNTATTQSGSSPPHSSPPGVDANMEGQCKELQAPKPNSVCKGLDIYREALPRGCDKVLPLCKTSTIMCVDAYYSKPPICFGYNQRCFCMETFGPIEFCCKS	1	128
JN032116	4659-5252	gb:JN032116:4659-5252|Organism:Respiratory syncytial virus|Strain Name:B/WI/629-12/06-07|Protein Name:attachment glycoprotein|Gene Symbol:G	MSKNKNQRTARTLEKTWDTLNHLIVISSCLYKLNLKSIAQIALSVLAMIISTSLIIAAIIFIISANHKVTLTTVTVQTIKNHTEKNITTYLTQVSPERVSPSKQPTTTPPIHTNSATISPNTKSEIHHTTAQTKGRTSTPTQNNKPNTKPRPKNPPKKDDYHFEVFNFVPCSICGNNQLCKSICKTIPSNKPRKNQP	1	129
KX258200	6254-7996	gb:KX258200:6254-7996|Organism:Avian paramyxovirus 14|Strain Name:APMV14/duck/Japan/11OG0352/2011|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MEGSRTVIYQGDPNEKNTWRLVFRTLTLILNLAILSVTIASIIITSKITLSEVTTLKTEGVEEVITPLMATLSDSVQQEKMIYKEVAISIPLVLDKIQTDVGTSVAQITDALRQIQGVNGTQAFALSNAPEYSGGIEVPLFQIDSFVNKSMSISGLLEHASFIPSPTTLHGCTRIPSFHLGPRHWCYTHNIIGSRCRDEGFSSMYISIGAITVNRDGNPLFITTASTILADDNNRKSCSIIASSYGCDLLCSIVTESENDDYANPNPTKMVHGRFLYNGSYVEQALPNSLFQDKWVAQYPGVGSGITTHGKVLFPIYGGIKKNTQLFYELSKYGFFAHNKELECKNMTEEQIRDIKAAYLPSKTSGNLFAQGIIYCNISKLGDCNVAVLNTSTTMMGAEGRLQMMGEYVYYYQRSSSWWPVGIVYKKSLAELMNGINMEVLSFEPIPLSKFPRPTWTAGLCQKPSICPDVCVTGVYTDLFSVTIGSTTDKDTYFGVYLDSATERKDPWVAAADQYEWRNRVRLFESTTEAAYTTSTCFKNTVNNRVFCVSIVELRENLLGDWKIVPLLFQIGVSQGPPPK	1	130
KX940961	7978-12504	gb:KX940961:7978-12504|Organism:Beilong virus|Strain Name:ERN081008_1S|Protein Name:attachment	MSQLAAHNLAMSNFYGTHQGDLSGSQKGEEQQVQGVIRYVSMIVGLLSLFTIIALNVTNIIYMTESGGTMQSIKTAQGSIDGSMREISGVIMEDVKPKTDLINSMVSYNIPAQLSMIHQIIKNDVLKQCTPSFMFNNTICPLAENPTHSRYFEEVNLDSISECSGPDMHLGLGVNPEFIEFPSFAPGSTKPGSCVRLPSFSLSTTVFAYTHTIMGHGCSELDVGDHYFSVGRIADAGHEIPQFETISSWFINDKINRRSCTVAAGAMEA	1	131

		glycoprotein|Gene Symbol:G	WMGCVIMTETFYDDLNSLDTGKLTISYLD VFGRKKEWIYTRSEILYDYTYTSVYFSVGS GVVVGDTVYFLIWGSLSSPIEETAYCFAPD CSNYNQRMCNEAQRPSKFGHRQMVNGIL KFKTTSTGKPLLSVGTLSPSVVPFGSEGRL MYSEITKIIYLYLRSTSWHALPLTGLFVLGP PTSISWIVQRAVSRPGEFPCGASNRCPKDC VTGVYTDLFPLGSRYEYAATVYLNSETYR VNPTLALINQTNIIASKKVTTESQRAGYTT TTCFVFKLRVWCISVVELAPSTMTAYEPIP FLYQLDLTCKGKNGSLAMRFTGKEGTYKS GRYKSPRNECFFEKVSNKYYFIVSTPEGIQ PYEIRDLTPDRMPHIIMYISDVCAPALSAFK KLLPAMRPITTLTIGNWQFRPVEVSGGLRV SIGRNLTKEGDLTMSAPEDPGSNTFPGGHI PGNGLFDAGYYTVEYPKEWKQTTPKPSEG GNIIDKNKTPVIPSRDNPTSDSSIPHRESIEP VRPTREVLKSSDYVTIVSTDSGSGSGDFAT GVPWTGVSPKAPQNGINLPGTELPHPTVL DRINTPAPSDPKVSADSDHTRDTIDPTALS KPLNHDTTGDTDTRINTGTATYGFTPGRE ATSSGKLANDLTNSTSVPSEAHPSASTSEA SKPEKNTDNRVTQDPTSGTAERPTTNAPV DGKHSTQLTDARPNTADPERTSQHSSSTTR DEVKPSLPSTTEASTHQRTEAATPPELVNN TLNPPSTQVRSVRSLMQDAIAQAWNFVRG VTP
KY511044	6454-8310	gb:KY511044:6454-8310|Organism:Avian paramyxovirus UPO216|Strain Name:APMV-15/WB/Kr/UPO216/2014|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	MERGISEVALANDRTEEKNTWRLIFRITVL VVSVITLGLTAASLVYSMNAAQPADFDGII PAVQQVGTSLTNSIGGMQDVLDRTYKQV ALESPLTLLNMESTIMNAITSLSYKINNGG NSSGCGAPIHDPEYIGGIGKELLIDDNVDV TSFYPSAFKEHLNFIPAPTTGAGCTRIPSFD LSATHYCYTHNVILSGCQDHSHSHQYIAL GVLKLSDTGNVFFSTLRSINLDDTANRKSC SISATPLGCDILCSKVTETELEDYKSEEPTP MVHGRLSFDGTYSEKDLDVNNLFSDWTA NYPSVGGGSYIGNRVWYAVYGGLKPGSN TDQSQRDKYVIYKRYNNTCPDPEDYQINK AKSSYTPSYFGSKRVQQAILSIAVSPTLGSD PVLTPLSNDVVLMGAEGRVMHIGGYTYL YQRGTSYYSPALLYPLNIQDKSATASSPYK FDAFTRPGSVPCQADARCPQSCVTGVYTD PYPLIFAKDHSIRGVYGMMLNDVTARLNPI AAVFSNISRSQITRVSSSSTKAAYTTSTCFK VIKTNRIYCMSIAEISNTLFGEFRIVPLLVEI LSNGGNTARSAGGTPVKESPKGWSDAIAE PLFCTPTNVTRYNADIRRYAYSWP	1	132
NC_025360	8127-10158	gb:NC_025360:8127-10158|Organism:Atlantic salmon paramyxovirus|Strain Name:ASPV/Y	MPPAPSPVHDPSSFYGSSLFNEDTASRKGT SEEIHLLGIRWNTVLIVLGLILAIIGIGIGASS FSASGITGNTTKEIRLIVEEMSYGLVRISDS VRQEISPKVTLLQNAVLSSIPALVTTETNTII NAVKNHCNSPPTPPPPTEAPLKKHETGMA PLDPTTYWTCTSGTPRFYSSPNATFIPGPSP LPHTATPGGCVRIPSMHIGSEIYAYTSNLIA SGCQDIGKSYQNVQIGVLDRTPEGNPEMS	1	133

		rkje371/95|Protein Name:hemagglutinin-neuraminidase protein|Gene Symbol:HN	PMLSHTFPINDNRKSCSIVTLKRAAYIYCS QPKVTEFVDYQTPGIEPMSLDHINANGTTK TWIYSPTEVVTDVPYASMYPSVGSGVVID GKLVFLVYGGLLNGIQVPAMCLSPECPGID QAACNASQYNQYLSGRQVVNGIATVDLM NGQKPHISVETISPSKNWFGAEGRLVYMG GRLYIYIRSTGWHSPIQIGVIYTMNPLAITW VTNTVLSRPGSAGCDWNNRCPKACLSGV YTDAYPISPDYNHLATMILHSTSTRSNPVM VYSSPTNMVNYAQLTTTAQIAGYTTTSCF TDNEVGYCATALELTPGTLSSVQPILVMT KIPKECV		

ii) Lipid Fusogens

In some embodiments, the fusosome may be treated with fusogenic lipids, such as saturated fatty acids. In some embodiments, the saturated fatty acids have between 10-14 carbons. In some embodiments, the saturated fatty acids have longer-chain carboxylic acids. In some embodiments, the saturated fatty acids are mono-esters.

In some embodiments, the fusosome may be treated with unsaturated fatty acids. In some embodiments, the unsaturated fatty acids have between C16 and C18 unsaturated fatty acids. In some embodiments, the unsaturated fatty acids include oleic acid, glycerol mono-oleate, glycerides, diacylglycerol, modified unsaturated fatty acids, and any combination thereof.

Without wishing to be bound by theory, in some embodiments negative curvature lipids promote membrane fusion. In some embodiments, the fusosome comprises one or more negative curvature lipids, e.g., negative curvature lipids that are exogenous relative to the source cell, in the membrane. In embodiments, the negative curvature lipid or a precursor thereof is added to media comprising source cells or fusosomes. In embodiments, the source cell is engineered to express or overexpress one or more lipid synthesis genes. The negative curvature lipid can be, e.g., diacylglycerol (DAG), cholesterol, phosphatidic acid (PA), phosphatidylethanolamine (PE), or fatty acid (FA).

Without wishing to be bound by theory, in some embodiments positive curvature lipids inhibit membrane fusion. In some embodiments, the fusosome comprises reduced levels of one or more positive curvature lipids, e.g., exogenous positive curvature lipids, in the membrane. In embodiments, the levels are reduced by inhibiting synthesis of the lipid, e.g., by knockout or knockdown of a lipid synthesis gene, in the source cell. The positive curvature lipid can be, e.g.,

lysophosphatidylcholine (LPC), phosphatidylinositol (PtdIns), lysophosphatidic acid (LPA), lysophosphatidylethanolamine (LPE), or monoacylglycerol (MAG).

iii) Chemical fusogens

In some embodiments, the fusosome may be treated with fusogenic chemicals. In some embodiments, the fusogenic chemical is polyethylene glycol (PEG) or derivatives thereof.

In some embodiments, the chemical fusogen induces a local dehydration between the two membranes that leads to unfavorable molecular packing of the bilayer. In some embodiments, the chemical fusogen induces dehydration of an area near the lipid bilayer, causing displacement of aqueous molecules between cells and allowing interaction between the two membranes together.

In some embodiments, the chemical fusogen is a positive cation. Some nonlimiting examples of positive cations include Ca2+, Mg2+, Mn2+, Zn2+, La3+, Sr3+, and H+.

In some embodiments, the chemical fusogen binds to the target membrane by modifying surface polarity, which alters the hydration-dependent intermembrane repulsion.

In some embodiments, the chemical fusogen is a soluble lipid soluble. Some nonlimiting examples include oleoylglycerol, dioleoylglycerol, trioleoylglycerol, and variants and derivatives thereof.

In some embodiments, the chemical fusogen is a water-soluble chemical. Some nonlimiting examples include polyethylene glycol, dimethyl sulphoxide, and variants and derivatives thereof.

In some embodiments, the chemical fusogen is a small organic molecule. A nonlimiting example includes n-hexyl bromide.

In some embodiments, the chemical fusogen does not alter the constitution, cell viability, or the ion transport properties of the fusogen or target membrane.

In some embodiments, the chemical fusogen is a hormone or a vitamin. Some nonlimiting examples include abscisic acid, retinol (vitamin A1), a tocopherol (vitamin E), and variants and derivatives thereof.

In some embodiments, the fusosome comprises actin and an agent that stabilizes polymerized actin. Without wishing to be bound by theory, stabilized actin in a fusosome can

promote fusion with a target cell. In embodiments, the agent that stabilizes polymerized actin is chosen from actin, myosin, biotin-streptavidin, ATP, neuronal Wiskott–Aldrich syndrome protein (N-WASP), or formin. See, e.g., Langmuir. 2011 Aug 16;27(16):10061-71 and Wen et al., Nat Commun. 2016 Aug 31;7. In embodiments, the fusosome comprises actin that is exogenous or overexpressed relative to the source cell, e.g., wild-type actin or actin comprising a mutation that promotes polymerization. In embodiments, the fusosome comprises ATP or phosphocreatine, e.g., exogenous ATP or phosphocreatine.

iv) Small molecule fusogens

In some embodiments, the fusosome may be treated with fusogenic small molecules. Some nonlimiting examples include halothane, nonsteroidal anti-inflammatory drugs (NSAIDs) such as meloxicam, piroxicam, tenoxicam, and chlorpromazine.

In some embodiments, the small molecule fusogen may be present in micelle-like aggregates or free of aggregates.

v) Fusogen modifications

In some embodiments, the fusogen is linked to a cleavable protein. In some cases, a cleavable protein may be cleaved by exposure to a protease. An engineered fusion protein may bind any domain of a transmembrane protein. The engineered fusion protein may be linked by a cleavage peptide to a protein domain located within the intermembrane space. The cleavage peptide may be cleaved by one or a combination of intermembrane proteases (e.g. HTRA2/OMI which requires a non-polar aliphatic amino acid - valine, isoleucine or methionine are preferred - at position P1, and hydrophilic residues - arginine is preferred - at the P2 and P3 positions).

In some embodiments the fusogen is linked to an affinity tag. In some embodiments the affinity tag aids in fusosome separation and isolation. In some embodiments the affinity tag is cleavable. In some embodiments the affinity tag is non-covalently linked to the fusogen. In some embodiments the affinity tag is present on the fusosome and separate from the fusogen.

In some embodiments, fusogen proteins are engineered by any methods known in the art or any method described herein to comprise a proteolytic degradation sequence, e.g., a mitochondrial or cytosolic degradation sequence. Fusogen proteins may be engineered to include,

but is not limited to a proteolytic degradation sequence, e.g., a Caspase 2 protein sequence (e.g., Val-Asp-Val-Ala-Asp-|- (SEQ ID NO: 155)) or other proteolytic sequences (see, for example, Gasteiger et al., The Proteomics Protocols Handbook; 2005: 571-607), a modified proteolytic degradation sequence that has at least 75%, 80%, 85%, 90%, 95% or greater identity to the wildtype proteolytic degradation sequence, a cytosolic proteolytic degradation sequence, e.g., ubiquitin, or a modified cytosolic proteolytic degradation sequence that has at least 75%, 80%, 85%, 90%, 95% or greater identity to the wildtype proteolytic degradation sequence. In some embodiments, a composition comprises mitochondria in a source cell or chondrisome comprising a protein modified with a proteolytic degradation sequence, e.g., at least 75%, 80%, 85%, 90%, 95% or greater identity to the wildtype proteolytic degradation sequence, a cytosolic proteolytic degradation sequence, e.g., ubiquitin, or a modified cytosolic proteolytic degradation sequence that has at least 75%, 80%, 85%, 90%, 95% or greater identity to the wildtype proteolytic degradation sequence.

In some embodiments, the fusogen may be modified with a protease domain that recognizes specific proteins, e.g., over-expression of a protease, e.g., an engineered fusion protein with protease activity. For example, a protease or protease domain from a protease, such as MMP, mitochondrial processing peptidase, mitochondrial intermediate peptidase, inner membrane peptidase.

See, Alfonzo, J.D. & Soll, D. Mitochondrial tRNA import – the challenge to understand has just begun. Biological Chemistry 390: 717-722. 2009; Langer, T. et al. Characterization of Peptides Released from Mitochondria. THE JOURNAL OF BIOLOGICAL CHEMISTRY. Vol. 280, No. 4. 2691–2699, 2005; Vliegh, P. et al. Synthetic therapeutic peptides: science and market. Drug Discovery Today. 15(1/2). 2010; Quiros P.M.m et al., New roles for mitochondrial proteases in health, ageing and disease. Nature Reviews Molecular Cell Biology. V16, 2015; Weber-Lotfi, F. et al. DNA import competence and mitochondrial genetics. Biopolymers and Cell. Vol. 30. N 1. 71–73, 2014.

III. Positive target cell-specific regulatory element

In some embodiments, a fusosome described herein, e.g. a virus, e.g., a retrovirus, contains a nucleic acid (e.g., the gene encoding the exogenous agent), e.g. a retroviral nucleic

acid, that comprises a positive target cell-specific regulatory element such as a tissue-specific promoter, a tissue-specific enhancer, a tissue-specific splice site, a tissue-specific site extending half-life of an RNA or protein, a tissue-specific mRNA nuclear export promoting site, a tissue-specific translational enhancing site, or a tissue-specific post-translational modification site.

In some embodiments, a fusosome, e.g. virus, e.g. retrovirus, described herein contains a nucleic acid, e.g. a retroviral nucleic acid, that can comprise regions, e.g., non-translated regions such as origins of replication, selection cassettes, promoters, enhancers, translation initiation signals (Shine Dalgarno sequence or Kozak sequence), introns, a polyadenylation sequence, 5′ and 3′ untranslated regions—which interact with host cellular proteins to carry out transcription and translation, and which are capable of directing, increasing, regulating, or controlling the transcription or expression of an operatively linked polynucleotide. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including ubiquitous promoters and inducible promoters may be used.

In particular embodiments, control elements are capable of directing, increasing, regulating, or controlling the transcription or expression of an operatively linked polynucleotide in a cell-specific manner. In particular embodiments, a nucleic acid, e.g. retroviral nucleic acids, comprise one or more expression control sequences that are specific to particular cells, cell types, or cell lineages e.g., target cells; that is, expression of polynucleotides operatively linked to an expression control sequence specific to particular cells, cell types, or cell lineages is expressed in target cells and not (or at a lower level) in non-target cells.

In particular embodiments, a nucleic acid, e.g. a retroviral nucleic acid, can include exogenous, endogenous, or heterologous control sequences such as promoters and/or enhancers.

In embodiments, the promoter comprises a recognition site to which an RNA polymerase binds. An RNA polymerase initiates and transcribes polynucleotides operably linked to the promoter. In particular embodiments, promoters operative in mammalian cells comprise an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated and/or another sequence found 70 to 80 bases upstream from the start of transcription, a CNCAAT region where N may be any nucleotide.

In embodiments, an enhancer comprises a segment of DNA which contains sequences capable of providing enhanced transcription and in some instances can function independent of

orientation relative to another control sequence. An enhancer can function cooperatively or additively with promoters and/or other enhancer elements. In some embodiments, a promoter/enhancer segment of DNA contains sequences capable of providing both promoter and enhancer functions.

Illustrative ubiquitous expression control sequences include, but are not limited to, a cytomegalovirus (CMV) immediate early promoter, a viral simian virus 40 (SV40) (e.g., early or late), a Moloney murine leukemia virus (MoMLV) LTR promoter, a Rous sarcoma virus (RSV) LTR, a herpes simplex virus (HSV) (thymidine kinase) promoter, H5, P7.5, and P11 promoters from vaccinia virus, an elongation factor 1-alpha (EF1a) promoter, early growth response 1 (EGR1), ferritin H (FerH), ferritin L (FerL), Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), eukaryotic translation initiation factor 4A1 (EIF4A1), heat shock 70 kDa protein 5 (HSPA5), heat shock protein 90 kDa beta, member 1 (HSP90B1), heat shock protein 70 kDa (HSP70), β-kinesin (β-KIN), the human ROSA 26 locus Orions et al., *Nature Biotechnology* 25, 1477-1482 (2007)), a Ubiquitin C promoter (UBC), a phosphoglycerate kinase-1 (PGK) promoter, a cytomegalovirus enhancer/chicken β-actin (CAG) promoter, a β-actin promoter and a myeloproliferative sarcoma virus enhancer, negative control region deleted, d1587rev primer-binding site substituted (MND) promoter (Challita et al., *J Virol.* 69(2):748-55 (1995)).

In some embodiments, a promoter may be paired with a heterologous gene to impart the regulatory functions of that promoter on the heterologous gene. In some embodiments, the cis-regulatory elements from a first gene's promoter may be linked to segments of a different gene's promoter to create chimeric promoters that have properties of both promoters.

In some embodiments, the promoter is a tissue-specific promoter, e.g., a promoter that drives expression in CNS cells, e.g., pan-neuronal cells, GABAergic neurons, Glutamatergic neurons, Cholinergic neurons, Dopaminergic neurons, Serotonergic neurons, astrocytes, microglia, oligodendrocytes, or choroid plexus cells. Various suitable CNS cell-specific promoters are described in Table 3 below. In some embodiments, a fusosome (e.g., viral vector) described herein comprises, in its nucleic acid, a promoter having a sequence of a promoter in Table 3, or transcriptionally active fragment thereof, or a variant having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto. In some embodiments, a fusosome (e.g., viral vector) described herein comprises, in its nucleic acid, a promoter having transcription factor binding sites from the region within 3 kb of the transcriptional start site for the genes listed

in Table 3. In some embodiments, a fusosome (e.g., viral vector) described herein comprises, in its nucleic acid, a region within 2.5 kb, 2 kb, 1.5 kb, 1 kb, or 0.5 kb immediately upstream of the transcriptional start site of a gene listed in Table 3, or a transcriptionally active fragment thereof, or a variant having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity thereto.

Table 3. *Exemplary promoters, e.g., CNS cell-specific promoters*

Target cell type	**Exemplary promoters**
Pan-neuronal	SYN, NSE, CaMKII, aTubulin, PDGF
GABAergic neurons	fSST, fNPY, GAD67, DLX5/6
Glutamatergic neurons	VGLUT1, Dock10
Cholinergic neurons	ChAT, VAChT
Dopaminergic neurons	Drd1a
Serotonergic neurons	TPH-2
Astrocytes	GFAP, EAAT1, GS
Microglia	CX3CR1, TMEM119
Oligodendrocytes	MBP, CNP
Choroid plexus	CRFR2β

In some embodiments, the CNS cell-specific promoter is a promoter described in Hioki et al., Ther. 2007 Jun;14(11):872-82, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a SYN, NSE, CaMKII, aTubulin, or PDGF promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Nathanson et al., Front. Neural Circuits, 2009, 3:19. doi: 10.3389/neuro.04.019.2009, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a fSST or fNPY promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Delzor et al., Hum Gene Ther Methods. 2012 Aug;23(4):242-54, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a GAD67 or DLX5/6 promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Egashira et al., Sci Rep. 2018 Oct 11;8(1):15156, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a VGLUT1 or Dock10 promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Naciff et al., J. Neurochem., 1999 Jan;72(1):17-28, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a ChAT

promoter. In some embodiments, the CNS cell-specific promoter is a VAChT promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Delzor et al., Hum Gene Ther Methods. 2012 Aug; 23(4): 242–254, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a Drd1a promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Benzekhroufa et al., Gene Ther. 2009 May;16(5):681-8, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a TPH-2 promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Merienne et al., Gene Ther. 2015 Oct;22(10):830-9, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a GFAP, EAAT1, or GS promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Immgen consortium, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a CX3CR1 promoter. In some embodiments, the CNS cell-specific promoter is a TMEM119 promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in McIver et al., J Neurosci Res. 2005 Nov 1;82(3):397-403, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a MBP promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Kagiava et al., J Gene Med. 2014 Nov-Dec;16(11-12):364-73, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a MBP or CNP promoter. In some embodiments, the CNS cell-specific promoter is a promoter described in Regev et al., Proc Natl Acad Sci U S A. 2010 Mar 2;107(9):4424-9, herein incorporated by reference in its entirety, e.g., the CNS cell-specific promoter is a CRFR2β promoter. In some embodiments, the CNS cell-specific promoter is a transcriptionally active fragment of any of the foregoing. In some embodiments, the CNS-cell specific promoter is a variant having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to any of the foregoing.

An internal ribosome entry site (IRES) typically promotes direct internal ribosome entry to the initiation codon, such as ATG, of a cistron (a protein encoding region), thereby leading to the cap-independent translation of the gene. See, e.g., Jackson et al, (1990) Trends Biochem Sci 15(12):477-83) and Jackson and Kaminski. (1995) RNA 1 (10):985-1000. In particular embodiments, a vector includes one or more exogenous genes encoding one or more exogenous agents. In particular embodiments, to achieve efficient translation of each of the plurality of

exogenous protein agents, the polynucleotide sequences can be separated by one or more IRES sequences or polynucleotide sequences encoding self-cleaving polypeptides.

The nucleic acid, e.g. retroviral nucleic acids herein, can also comprise one or more Kozak sequences, e.g., a short nucleotide sequence that facilitates the initial binding of mRNA to the small subunit of the ribosome and increases translation. The consensus Kozak sequence is (GCC)RCCATGG, where R is a purine (A or G) (Kozak, (1986) Cell. 44(2):283-92, and Kozak, (1987) Nucleic Acids Res. 15(20): 8125-48).

Promoters responsive to a heterologous transcription factor and inducer

In some embodiments, a nucleic acid, retroviral nucleic acid, comprises an element allowing for conditional expression of the exogenous agent, e.g., any type of conditional expression including, but not limited to, inducible expression; repressible expression; cell type-specific expression, or tissue-specific expression. In some embodiments, to achieve conditional expression of the exogenous agent, expression is controlled by subjecting a cell, tissue, or organism to a treatment or condition that causes the exogenous agent to be expressed or that causes an increase or decrease in expression of the exogenous agent.

Illustrative examples of inducible promoters/systems include, but are not limited to, steroid-inducible promoters such as promoters for genes encoding glucocorticoid or estrogen receptors (inducible by treatment with the corresponding hormone), metallothionine promoter (inducible by treatment with various heavy metals), MX-1 promoter (inducible by interferon), the "GeneSwitch" mifepristone-regulatable system (Sirin et al., 2003, *Gene,* 323:67), the cumate inducible gene switch (WO 2002/088346), tetracycline-dependent regulatory systems, etc.

Transgene expression may be activated or repressed by the presence or absence of an inducer molecule. In some cases the inducer molecule activates or represses gene expression in a graded manner, and in some cases the inducer molecules activates or represses gene expression in an all-or-nothing manner.

A commonly used inducible promoter/system is tetracycline (Tet)-regulated system. The Tet system is based on the coexpression of two elements in the respective target cell: (i) the tetracycline response element containing repeats of the Tet-operator sequences (TetO) fused to a minimal promoter and connected to a gene of interest (e.g., a gene encoding the exogenous agent) and (ii) the transcriptional transactivator (tTA), a fusion protein of the Tet-repressor

(TetR) and the transactivation domain of the herpes simplex virus derived VP16 protein. Whereas in the originally described version, transgene expression was active in the absence of tetracycline or its potent analogue doxycycline (Do), referred to as Tet-OFF system, modification of four amino acids within the transactivator protein resulted in a reverse tTA (rtTA), which only binds to TetO in the presence of Dox (Tet-ON system). In some embodiments, in the transactivator, the VP16 domain has been replaced by minimal activation domains, potential splice-donor and splice acceptor sites have been removed, and the protein has been codon optimization, resulting in the improved Transactivator variant rtTA2S-M2 with higher sensitivity to Dox and lower baseline activity. Furthermore, different Tet-responsive promoter elements have been generated, including modification in the TetO with 36-nucleotide spacing from neighboring operators to enhance regulation. Additional modifications may be useful to further reduce basal activity and increase the expression dynamic range. As an example, the pTet-T11 (short: TII) variant displays a high dynamic range and low background activity.

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

Riboswitches to regulate exogenous agent expression

Some of the compositions and methods provided herein include one or more riboswitches or polynucleotides that include one or more riboswitch. Riboswitches are a common feature in bacteria to regulate gene expression and are a means to achieve RNA control of biological functions. Riboswitches can be present in the 5'-untranslated region of mRNAs and can allow for regulatory control over gene expression through binding of a small molecule ligand that induces

or suppresses a riboswitch activity. In some embodiments, the riboswitch controls a gene product involved in the generation of the small molecule ligand. Riboswitches typically act in a cis-fashion, although riboswitches have been identified that act in a trans-fashion. Natural riboswitches consist of two domains: an aptamer domain that binds the ligand through a three-dimensional folded RNA structure and a function switching domain that induces or suppresses an activity in the riboswitch based on the absence or presence of the ligand. Thus, there are two ligand sensitive conformations achieved by the riboswitch, representing on and off states (Garst et al., 2011). The function switching domain can affect the expression of a polynucleotide by regulating: an internal ribosome entry site, pre-mRNA splice donor accessibility in the retroviral gene construct, translation, termination of transcription, transcript degradation, miRNA expression, or shRNA expression (Dambach and Winkler 2009). The aptamer and function switching domains can be used as modular components allowing for synthetic RNA devices to control gene expression either as native aptamers, mutated/evolved native aptamers, or totally synthetic aptamers that are identified from screening random RNA libraries (McKeague et al 2016).

The purine riboswitch family represents one of the largest families with over 500 sequences found (Mandal et al 2003; US20080269258; and WO2006055351). The purine riboswitches share a similar structure consisting of three conserved helical elements/stem structures (PI, P2, P3) with intervening loop/junction elements (Jl-2, L2, J2-3, L3, J3-1). The aptamer domains of the purine family of riboswitches naturally vary in their affinity/regulation by various purine compounds such as adenine, guanine, adenosine, guanosine, deoxyadenosine, deoxyguanosine, etc. due to sequence variation (Kim et al. 2007)

In some embodiments, a nucleic acid, e.g. retroviral nucleic acid, described herein comprises a polynucleotide encoding the exogenous agent operably linked to a promoter and a riboswitch. The riboswitch include one or more of, e.g., all of: a.) an aptamer domain, e.g., an aptamer domain capable of binding a nucleoside analogue antiviral drug and having reduced binding to guanine or 2'-deoxyguanosine relative to the nucleoside analogue antiviral drug; and b.) a function switching domain, e.g., a function switching domain capable of regulating expression of the exogenous agent, wherein binding of the nucleoside analogue by the aptamer domain induces or suppresses the expression regulating activity of the function switching domain, thereby regulating expression of the exogenous agent. In some embodiments, the

exogenous agent can be a polypeptide, an miRNA, or an shRNA. For example, in an embodiment, the riboswitch is operably linked to a nucleic acid encoding a chimeric antigen receptor (CAR). In non-limiting illustrative examples provided herein, the exogenous gene encodes one or more engineered signaling polypeptides. For instance, the riboswitch and the target polynucleotide encoding one or more engineered signaling polypeptides can be found in the genome of a source cell, in a replication incompetent recombinant retroviral particle, in a T cell and/or in an NK cell.

The aptamer domains can be used, e.g., as modular components and combined with any of the function switching domains to affect the RNA transcript. In any of the embodiments disclosed herein, the riboswitch can affect the RNA transcript by regulating any of the following activities: internal ribosomal entry site (IRES), pre-mRNA splice donor accessibility, translation, termination of transcription, transcript degradation, miRNA expression, or shRNA expression. In some embodiments, the function switching domain can control binding of an anti-IRES to an IRES (see, e.g. Ogawa, RNA (2011), 17:478- 488, the disclosure of which is incorporated by reference herein in its entirety). In any of the embodiments disclosed herein, the presence or absence of the small molecule ligand can cause the riboswitch to affect the RNA transcript. In some embodiments, the riboswitch can include a ribozyme. Riboswitches with ribozymes can inhibit or enhance transcript degradation of target polynucleotides in the presence of the small molecule ligand. In some embodiments, the ribozyme can be a pistol class of ribozyme, a hammerhead class of ribozyme, a twisted class of ribozyme, a hatchet class of ribozyme, or the HDV (hepatitis delta virus).

IV. Non-target cell-specific regulatory element

In some embodiments, the non-target cell specific regulatory element or negative TCSRE comprises a tissue-specific miRNA recognition sequence, tissue-specific protease recognition site, tissue-specific ubiquitin ligase site, tissue-specific transcriptional repression site, or tissue-specific epigenetic repression site.

In some embodiments, a non-target cell comprises an endogenous miRNA. In some embodiments, a fusosome described herein, e.g. a virus, e.g., a retrovirus, contains a nucleic acid, e.g. retroviral nucleic acid (e.g., the gene encoding the exogenous agent) that may comprise a recognition sequence for that miRNA. Thus, if the nucleic acid, retroviral nucleic acid, enters

the non-target cell, the miRNA can downregulate expression of the exogenous agent. This helps achieve additional specificity for the target cell versus non-target cells.

In some embodiments, the miRNA is a small non-coding RNAs of 20-22 nucleotides, typically excised from ~70 nucleotide foldback RNA precursor structures known as pre-miRNAs. In general, miRNAs negatively regulate their targets in one of two ways depending on the degree of complementarity between the miRNA and the target. First, miRNAs that bind with perfect or nearly perfect complementarity to protein-coding mRNA sequences typically induce the RNA-mediated interference (RNAi) pathway. miRNAs that exert their regulatory effects by binding to imperfect complementary sites within the 3′ untranslated regions (UTRs) of their mRNA targets, typically repress target-gene expression post-transcriptionally, apparently at the level of translation, through a RISC complex that is similar to, or possibly identical with, the one that is used for the RNAi pathway. Consistent with translational control, miRNAs that use this mechanism reduce the protein levels of their target genes, but the mRNA levels of these genes are only minimally affected. miRNAs (e.g., naturally occurring miRNAs or artificially designed miRNAs) can specifically target any mRNA sequence. For example, in one embodiment, the skilled artisan can design short hairpin RNA constructs expressed as human miRNA (e.g., miR-30 or miR-21) primary transcripts. This design adds a Drosha processing site to the hairpin construct and has been shown to greatly increase knockdown efficiency (Pusch et al., 2004). The hairpin stem consists of 22-nt of dsRNA (e.g., antisense has perfect complementarity to desired target) and a 15-19-nt loop from a human miR. Adding the miR loop and miR30 flanking sequences on either or both sides of the hairpin results in greater than 10-fold increase in Drosha and Dicer processing of the expressed hairpins when compared with conventional shRNA designs without microRNA. Increased Drosha and Dicer processing translates into greater siRNA/miRNA production and greater potency for expressed hairpins.

Hundreds of distinct miRNA genes are differentially expressed during development and across tissue types. Several studies have suggested important regulatory roles for miRNAs in a broad range of biological processes including developmental timing, cellular differentiation, proliferation, apoptosis, oncogenesis, insulin secretion, and cholesterol biosynthesis. (See Bartel 2004 Cell 116:281-97; Ambros 2004 Nature 431 :350-55; Du et al. 2005 Development 132:4645-52; Chen 2005 N. Engl. J. Med. 353:1768-71; Krutzfeldt et al. 2005 Nature 438:685-89.) Molecular analysis has shown that miRNAs have distinct expression profiles in different

tissues. Computational methods have been used to analyze the expression of approximately 7,000 predicted human miRNA targets. The data suggest that miRNA expression broadly contributes to tissue specificity of mRNA expression in many human tissues. (See Sood et al. 2006 PNAS USA 103(8):2746-51.)

Thus, an miRNA-based approach may be used for restricting expression of the exogenous agent to a target cell population by silencing exogenous agent expression in non-target cell types by using endogenous microRNA species. MicroRNA induces sequence-specific post-transcriptional gene silencing in many organisms, either by inhibiting translation of messenger RNA (mRNA) or by causing degradation of the mRNA. See, e.g., Brown et al. 2006 Nature Med. 12(5):585-91., and WO2007/000668, each of which is herein incorporated by reference in its entirety. In some embodiments, the nucleic acid, e.g. retroviral nucleic acid, comprises one or more of (e.g., a plurality of) tissue-specific miRNA recognition sequences. In some embodiments, the tissue-specific miRNA recognition sequence is about 20-25, 21-24, or 23 nucleotides in length. In embodiments, the tissue-specific miRNA recognition sequence has perfect complementarity to an miRNA present in a non-target cell. In some embodiments, the exogenous agent does not comprise GFP, e.g., does not comprise a fluorescent protein, e.g., does not comprise a reporter protein. In some embodiments, the off-target cells are not hematopoietic cell and/or the miRNA is not present in hematopoietic cells.

In some embodiments, a method herein comprises tissue-specific expression of an exogenous agent in a target cell comprising contacting a plurality of fusosomes, e.g. a virus, e.g. retroviral vectors, comprising a nucleotide encoding the exogenous agent and at least one tissue-specific microRNA (miRNA) target sequence with a plurality of cells comprising target cells and non-target cells, wherein the exogenous agent is preferentially expressed in, e.g., restricted, to the target cell.

For example, the nucleic acid, e.g. retroviral nucleic acid, can comprise at least one miRNA recognition sequence operably linked to a nucleotide sequence having a corresponding miRNA in a non-target cell, e.g., a hematopoietic progenitor cell (HSPC), hematopoietic stem cell (HSC), which prevents or reduces expression of the nucleotide sequence in the non-target cell but not in a target cell, e.g., differentiated cell. In some embodiments, the nucleic acid, e.g. retroviral nucleic acid, comprises at least one miRNA sequence target for a miRNA which is present in an effective amount (e.g., concentration of the endogenous miRNA is sufficient to

reduce or prevent expression of a transgene) in the non-target cell, and comprises a transgene. In embodiments, the miRNA used in this system is strongly expressed in non-target cells, such as HSPC and HSC, but not in differentiated progeny of e.g. the myeloid and lymphoid lineage, preventing or reducing expression of a transgene in sensitive stem cell populations, while maintaining expression and therapeutic efficacy in the target cells.

In some embodiments, the negative TSCRE or NTSCRE comprises an miRNA recognition site. Exemplary miRNAs are provided in Table 4 below. In some embodiments, the nucleic acid (e.g., fusosome nucleic acid or retroviral nucleic acid) comprises a sequence that is complementary to a miRNA of Table 4, or has at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% complementarity thereto. In some embodiments, the nucleic acid (e.g., fusosome nucleic acid or retroviral nucleic acid) comprises a sequence that is perfectly complementary to a seed sequence within an endogenous miRNA, e.g., miRNA of Table 4. In some embodiments, the miRNA comprises the sequence set forth in any one of SEQ ID NOS: 156-162. In embodiments, the seed sequence is at least 6, 7, 8, 9, or 10 nucleotides in length.

Table 4. *Exemplary miRNAs.*

Target cell type	**miRNA name and silenced cell type (non-target cell type)**	**SEQUENCE**	**SEQ ID NO**
Pan-neuronal	miR-338-3p (for de-targeting from oligodendrocytes);miR-9 and miR-125b-5p (for de-targeting from astrocytes); miR-342-3p (for de-targeting from microglia)	miR-338-3p: uccagcaucagugauuuuguug	156
GABAergic neurons		miR-9-5p: ucuuugguuaucuagcuguauga	157
Glutamatergic neurons		miR-9-3p: auaaagcuagauaaccgaaagu	158
Cholinergic neurons		miR-125b-5p: ucccugagacccuaacuuguga	159
Dopaminergic neurons		miR-342-3p: ucucacacagaaaucgcacccgu	160
Serotonergic neurons			
Astrocytes	miR-124 (for de-targeting from neuronal lineage)	miR-124-3p: uaaggcacgcggugaaugccaa	161
Microglia			
Oligodendrocytes		miR-124-5p: cguguucacagcggaccuugau	162
Choroid plexus			

In some embodiments, the negative TSCRE or NTSCRE comprises an miRNA recognition site for an miRNA described herein. Exemplary miRNAs include those found in Butovsky et al., Nat Neurosci. 2014 Jan;17(1):131-43, herein incorporated by reference in its entirety, e.g., miR-338-3p, miR-9, miR-125b-5p, or miR-342-3p. Additional exemplary miRNAs can be found in Delzor et al., Curr. Drug Targets, 2013 Oct;14(11):1336-46, herein incorporated by reference in its entirety, e.g., miR-124.

In some embodiments, a fusosome described herein comprises a nucleic acid comprising a payload gene and a positive target cell-specific regulatory element, e.g., wherein the target cell is a neuron, e.g., a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, or a Serotonergic neuron. In some embodiments, the nucleic acid further comprises a non-target cell-specific regulatory element (NTCSRE), e.g.,

wherein the NTSCRE comprises an miRNA recognition site for an miRNA expressed in a glial cell, e.g., an astrocyte, a microglial cell, or an oligodendrocyte.

In some embodiments, a fusosome described herein comprises a nucleic acid comprising a payload gene and a positive target cell-specific regulatory element, e.g., wherein the target cell is a glial cell, e.g., an astrocyte, a microglial cell, or an oligodendrocyte. In some embodiments, the nucleic acid further comprises a non-target cell-specific regulatory element (NTCSRE), e.g., wherein the NTSCRE comprises an miRNA recognition site for an miRNA expressed in a neuron, e.g., a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, or a Serotonergic neuron.

In some embodiments, the negative TSCRE or NTSCRE comprises an miRNA recognition site for an miRNA described herein. Exemplary miRNAs include those found in Griffiths-Jones et al. Nucleic Acids Res. 2006 Jan 1, 34; Chen and Lodish, Semin Immunol. 2005 Apr;17(2):155-65; Chen et al. Science. 2004 Jan 2;303(5654):83-6; Barad et al. Genome Res. 2004 Dec; 14(12): 2486–2494; Krichevsky et al., RNA. 2003 Oct;9(10):1274-81; Kasashima et al. Biochem Biophys Res Commun. 2004 Sep 17;322(2):403-10; Houbaviy et al., Dev Cell. 2003 Aug;5(2):351-8; Lagos-Quintana et al., Curr Biol. 2002 Apr 30;12(9):735-9; Calin et al., Proc Natl Acad Sci U S A. 2004 Mar 2;101(9):2999-3004; Sempere et al. Genome Biol. 2004; 5(3): R13; Metzler et al., Genes Chromosomes Cancer. 2004 Feb;39(2):167-9; Calin et al., Proc Natl Acad Sci U S A. 2002 Nov 26;99(24):15524-9; Mansfield et al. Nat Genet. 2004 Oct;36(10):1079-83; Michael et al. Mol Cancer Res. 2003 Oct;1(12):882-91; and at www.miRNA.org.

In some embodiments, the negative TSCRE or NTSCRE comprises an miRNA recognition site for an miRNA selected from miR-1b, miR-189b , miR-93, miR-125b, miR-130 , miR-32, miR-128, miR-22, miR124a, miR-296, miR-143, miR-15, miR-141, miR-143, miR-16, miR-127, miR99a, miR-183, miR-19b, miR-92, miR-9, miR-130b , miR-21 , miR-30b, miR-16, miR-99a , miR-212, miR-30c, miR-213, miR-20, miR-155, miR-152, miR-139, miR-30b, miR-7, miR-30c , miR-18, miR-137, miR-219, miR-1d, miR-178, miR-24, miR-122a, miR-215, miR-124a, miR-190, miR-149, miR-193, let-7a, miR-132, miR-27a, miR-9*, miR-200b, miR-266, miR-153, miR-135, miR-206, miR-24, miR-19a, miR-199, miR-26a, miR-194, miR-125a, miR-15a, miR-145, miR-133, miR-96, miR-131, miR-124b, miR-151, miR-7b, miR-103, and miR-208.

In some embodiments, the nucleic acid (e.g., retroviral nucleic acid) comprises two or more miRNA recognition sites. In some embodiments, the first miRNA recognition site and second miRNA recognition site are recognized by the same miRNA, and in some embodiments, the first miRNA recognition site and second miRNA recognition site are recognized by different miRNAs. In some embodiments, the first miRNA recognition site and second miRNA recognition site are recognized by miRNAs present in the same non-target cell, and in some embodiments, the first miRNA recognition site and second miRNA recognition site are recognized by miRNAs present in different non-target cells. In some embodiments, one or both of the first miRNA recognition site and second miRNA recognition site are recognized by miRNAs of Table 4. In some embodiments, one or more of the miRNA recognition sites on the fusosome nucleic acid (e.g. retroviral nucleic acid) are transcribed in *cis* with the exogenous agent. In some embodiments, one or more of the miRNA recognition sites on the fusosome nucleic acid (e.g., retroviral nucleic acid) are situated downstream of the poly A tail sequence, e.g., between the poly A tail sequence and the WPRE. In some embodiments, one or more of the miRNA recognition sites on the fusosome nucleic acid (e.g., retroviral nucleic acid) are situated downstream of the WPRE.

V. Immune modulation

In some embodiments, a fusosome, e.g. retroviral vector, or VLP, described herein comprises elevated CD47. See, e.g., US Pat. 9,050,269, which is herein incorporated by reference in its entirety. In some embodiments, a fusosome, e.g. a retroviral vector or VLP, described herein comprises elevated Complement Regulatory protein. See, e.g., ES2627445T3 and US6790641, each of which is incorporated herein by reference in its entirety. In some embodiments, a fusosome, e.g. a retroviral vector, or VLP, described herein lacks or comprises reduced levels of an MHC protein, e.g., an MHC-1 class 1 or class II. See, e.g., US20170165348, which is herein incorporated by reference in its entirety.

Sometimes fusosomes, e.g. retroviral vectors, or VLPs, are recognized by the subject's immune system. In the case of enveloped viral vector particles (e.g., retroviral vector particles), membrane-bound proteins that are displayed on the surface of the viral envelope may be recognized and the viral particle itself may be neutralised. Furthermore, on infecting a target cell, the viral envelope becomes integrated with the cell membrane and as a result viral envelope proteins may become displayed on or remain in close association with the surface of the cell. The

immune system may therefore also target the cells which the viral vector particles have infected. Both effects may lead to a reduction in the efficacy of exogenous agent delivery by viral vectors.

A viral particle envelope typically originates in a membrane of the source cell. Therefore, membrane proteins that are expressed on the cell membrane from which the viral particle buds may be incorporated into the viral envelope.

The immune modulating protein CD47

The internalization of extracellular material into cells is commonly performed by a process called endocytosis (Rabinovitch, 1995, Trends Cell Biol. 5(3):85-7; Silverstein, 1995, Trends Cell Biol. 5(3):141-2; Swanson et al., 1995, Trends Cell Biol. 5(3):89-93; Allen et al., 1996, J. Exp. Med. 184(2):627-37). Endocytosis may fall into two general categories: phagocytosis, which involves the uptake of particles, and pinocytosis, which involves the uptake of fluid and solutes.

Professional phagocytes have been shown to differentiate from non-self and self, based on studies with knockout mice lacking the membrane receptor CD47 (Oldenborg et al., 2000, Science 288(5473):2051-4). CD47 is a ubiquitous member of the Ig superfamily that interacts with the immune inhibitory receptor SIRPα (signal regulatory protein) found on macrophages (Fujioka et al., 1996, Mol. Cell. Biol. 16(12):6887-99; Veillette et al., 1998, J. Biol. Chem. 273(35):22719-28; Jiang et al., 1999, J. Biol. Chem. 274(2):559-62). Although CD47-SIRPα interactions appear to deactivate autologous macrophages in mouse, severe reductions of CD47 (perhaps 90%) are found on human blood cells from some Rh genotypes that show little to no evidence of anemia (Mouro-Chanteloup et al., 2003, Blood 101(1):338-344) and also little to no evidence of enhanced cell interactions with phagocytic monocytes (Arndt et al., 2004, Br. J. Haematol. 125(3):412-4).

In some embodiments, a fusosome, e.g. a retroviral vector, or VLP (e.g., a viral particle having a radius of less than about 1 μm, 400 nm, or 150 nm), comprises at least a biologically active portion of CD47, e.g., on an exposed surface of the fusosome, e.g. retroviral vector, or VLP. In some embodiments, the fusosome, e.g. retroviral vector (e.g., lentivirus), or VLP, includes a lipid coat. In embodiments, the amount of the biologically active CD47 in the fusosome, e.g. retroviral vector, or VLP, is between about 20-250, 20-50, 50-100, 100-150, 150-200, or 200-250 molecules/μm^2. In some embodiments, the CD47 is human CD47.

A method described herein can comprise evading phagocytosis of a particle by a phagocytic cell. The method may include expressing at least one peptide including at least a biologically active portion of CD47 in a fusosome, e.g. a retroviral vector, or VLP, so that, when the fusosome, e.g. retroviral vector, or VLP, comprising the CD47 is exposed to a phagocytic cell, the fusosome, e.g. viral particle, evades phacocytosis by the phagocytic cell, or shows decreased phagocytosis compared to an otherwise similar unmodified fusosome, e.g. retroviral vector, or VLP. In some embodiments, the half-life of the fusosome, e.g. retroviral vector, or VLP, in a subject is extended compared to an otherwise similar unmodified fusosome, e.g. retroviral vector, or VLP.

MHC deletion

The major histocompatibility complex class I (MHC-I) is a host cell membrane protein that can be incorporated into viral envelopes and, because it is highly polymorphic in nature, it is a major target of the body's immune response (McDevitt H. O. (2000) *Annu. Rev. Immunol.* 18: 1-17). MHC-I molecules exposed on the plasma membrane of source cells can be incorporated in the viral particle envelope during the process of vector budding. These MHC-I molecules derived from the source cells and incorporated in the viral particles can in turn be transferred to the plasma membrane of target cells. Alternatively, the MHC-I molecules may remain in close association with the target cell membrane as a result of the tendency of viral particles to absorb and remain bound to the target cell membrane.

The presence of exogenous MHC-I molecules on or close to the plasma membrane of transduced cells may elicit an alloreactive immune response in subjects. This may lead to immune-mediated killing or phagocytosis of transduced cells either upon ex vivo gene transfer followed by administration of the transduced cells to the subject, or upon direct in vivo administration of the viral particles. Furthermore, in the case of in vivo administration of MHC-I bearing viral particles into the bloodstream, the viral particles may be neutralised by pre-existing MHC-I specific antibodies before reaching their target cells.

Accordingly, in some embodiments, a source cell is modified (e.g., genetically engineered) to decrease expression of MHC-I on the surface of the cell. In embodiments, the source comprises a genetically engineered disruption of a gene encoding β2-microglobulin (β2M). In embodiments, the source cell comprises a genetically engineered disruption of one or

more genes encoding an MHC-I α chain. The cell may comprise genetically engineered disruptions in all copies of the gene encoding β2-microglobulin. The cell may comprise genetically engineered disruptions in all copies of the genes encoding an MHC-I α chain. The cell may comprise both genetically engineered disruptions of genes encoding β2-microglobulin and genetically engineered disruptions of genes encoding an MHC-I α chain. In some embodiments, the retroviral vector or VLP comprises a decreased number of surface-exposed MHC-I molecules. The number of surface-exposed MHC-I molecules may be decreased such that the immune response to the MHC-I is decreased to a therapeutically relevant degree. In some embodiments, the enveloped viral vector particle is substantially devoid of surface-exposed MHC-I molecules.

HLA-G/E overexpression

In some embodiments, a retroviral vector or VLP displays on its envelope a tolerogenic protein, e.g., an ILT-2 or ILT-4 agonist, e.g., HLA-E or HLA-G or any other ILT-2 or ILT-4 agonist. In some embodiments, a retroviral vector or VLP has increased expression of HLA-E, HLA-G, ILT-2 or ILT-4 compared to a reference retrovirus, e.g., an unmodified retrovirus otherwise similar to the retrovirus.

In some embodiments, a retrovirus composition has decreased MHC Class I compared to an unmodified retrovirus and increased HLA-G compared to an unmodified retrovirus.

In some embodiments, the retroviral vector or VLP has an increase in expression of HLA-G or HLA-E, e.g., an increase in expression of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of HLA-G or HLA-E, compared to a reference retrovirus, e.g., an unmodified retrovirus otherwise similar to the retrovirus, wherein expression of HLA-G or HLA-E is assayed i*n vitro* using flow cytometry, e.g., FACS.

In some embodiments, the retrovirus with increased HLA-G expression demonstrates reduced immunogenicity, e.g., as measured by reduced immune cell infiltration, in a teratoma formation assay.

Complement regulatory proteins

Complement activity is normally controlled by a number of complement regulatory proteins (CRPs). These proteins prevent spurious inflammation and host tissue damage. One

group of proteins, including CD55/ decay accelerating factor (DAF) and CD46/membrane cofactor protein (MCP), inhibits the classical and alternative pathway C3/C5 convertase enzymes. Another set of proteins including CD59 regulates MAC assembly. CRPs have been used to prevent rejection of xenotransplanted tissues and have also been shown to protect viruses and viral vectors from complement inactivation.

Membrane resident complement control factors include, e.g., decay-accelerating factor (DAF) or CD55, factor H (FH)–like protein-1 (FHL-1), C4b-binding protein (C4BP), Complement receptor 1 (CD35), membrane cofactor protein (MCP) or CD46, and CD59 (protectin) (e.g., to prevent the formation of membrane attack complex (MAC) and protect cells from lysis).

Albumin binding protein

In some embodiments the lentivirus binds albumin. In some embodiments the lentivirus comprises on its surface a protein that binds albumin. In some embodiments the lentivirus comprises on its surface an albumin binding protein. In some embodiments the albumin binding protein is streptococcal Albumin Binding protein. In some embodiments the albumin binding protein is streptococcal Albumin Binding Domain.

Expression of non-fusogen proteins on the lentiviral envelope

In some embodiments the lentivirus is engineered to comprise one or more proteins on its surface. In some embodiments the proteins affect immune interactions with a subject. In some embodiments the proteins affect the pharmacology of the lentivirus in the subject. In some embodiments the protein is a receptor. In some embodiments the protein is an agonist. In some embodiments the protein is a signaling molecule. In some embodiments, the protein on the lentiviral surface comprises OKT3 or IL7.

In some embodiments, comprises a mitogenic transmembrane protein and/or a cytokine-based transmembrane protein is present in the source cell, which can be incorporated into the retrovirus when it buds from the source cell membrane. The mitogenic transmembrane protein and/or a cytokine-based transmembrane protein can be expressed as a separate cell surface molecule on the source cell rather than being part of the viral envelope glycoprotein.

Exemplary Features

In some embodiments of any of the aspects described herein, the fusosome, e.g. retroviral vector, VLP, or pharmaceutical composition is substantially non-immunogenic. Immunogenicity can be quantified, e.g., as described herein.

In some embodiments, a retroviral vector or VLP fuses with a target cell to produce a recipient cell. In some embodiments, a recipient cell that has fused to one or more retroviral vectors or VLPs is assessed for immunogenicity. In embodiments, a recipient cell is analyzed for the presence of antibodies on the cell surface, e.g., by staining with an anti-IgM antibody. In other embodiments, immunogenicity is assessed by a PBMC cell lysis assay. In embodiments, a recipient cell is incubated with peripheral blood mononuclear cells (PBMCs) and then assessed for lysis of the cells by the PBMCs. In other embodiments, immunogenicity is assessed by a natural killer (NK) cell lysis assay. In embodiments, a recipient cell is incubated with NK cells and then assessed for lysis of the cells by the NK cells. In other embodiments, immunogenicity is assessed by a CD8+ T-cell lysis assay. In embodiments, a recipient cell is incubated with CD8+ T-cells and then assessed for lysis of the cells by the CD8+ T-cells.

In some embodiments, the retroviral vector or VLP comprises elevated levels of an immunosuppressive agent (e.g., immunosuppressive protein) as compared to a reference retroviral vector or VLP, e.g., one produced from an unmodified source cell otherwise similar to the source cell, or a HEK293 cell. In some embodiments, the elevated level is at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold. In some embodiments, the retroviral vector or VLP comprises an immunosuppressive agent that is absent from the reference cell. In some embodiments, the retroviral vector or VLP comprises reduced levels of an immunostimulatory agent (e.g., immunostimulatory protein) as compared to a reference retroviral vector or VLP, e.g., one produced from an unmodified source cell otherwise similar to the source cell, or a HEK293 cell. In some embodiments, the reduced level is at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, or 99% compared to the reference retroviral vector or VLP. In some embodiments, the immunostimulatory agent is substantially absent from the retroviral vector or VLP.

In some embodiments, the retroviral vector or VLP, or the source cell from which the retroviral vector or VLP is derived, has one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, or more of the following characteristics:

a. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of MHC class I or MHC class II, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a source cell otherwise similar to the source cell, or a HeLa cell, or a HEK293 cell;

b. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of one or more co-stimulatory proteins including but not limited to: LAG3, ICOS-L, ICOS, Ox40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3, or B7-H4, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, or a HEK cell, or a reference cell described herein;

c. expression of surface proteins which suppress macrophage engulfment e.g., CD47, e.g., detectable expression by a method described herein, e.g., more than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, or more expression of the surface protein which suppresses macrophage engulfment, e.g., CD47, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a Jurkat cell, or a HEK293 cell;

d. expression of soluble immunosuppressive cytokines, e.g., IL-10, e.g., detectable expression by a method described herein, e.g., more than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, or more expression of soluble immunosuppressive cytokines, e.g., IL-10, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, or a HEK293 cell;

e. expression of soluble immunosuppressive proteins, e.g., PD-L1, e.g., detectable expression by a method described herein, e.g., more than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, or more expression of soluble immunosuppressive proteins, e.g., PD-L1, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, or a HEK293 cell;

f. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of soluble immune stimulating cytokines, e.g., IFN-gamma or TNF-a, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, or a HEK293 cell or a U-266 cell;

g. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of endogenous immune-stimulatory antigen, e.g., Zg16 or Hormad1, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, or a HEK293 cell or an A549 cell, or a SK-BR-3 cell;

h. expression of, e.g., detectable expression by a method described herein, HLA-E or HLA-G, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a or a Jurkat cell;

i. surface glycosylation profile, e.g., containing sialic acid, which acts to, e.g., suppress NK cell activation;

j. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of TCRα/β, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell;

k. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of ABO blood groups, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a HeLa cell;

l. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of Minor Histocompatibility Antigen (MHA), compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell; or

m. has less than 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less, of mitochondrial MHAs, compared to a reference retroviral vector or VLP e.g., an

unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell, or has no detectable mitochondrial MHAs.

In embodiments, the co-stimulatory protein is 4-1BB, B7, SLAM, LAG3, HVEM, or LIGHT, and the reference cell is HDLM-2. In some embodiments, the co-stimulatory protein is BY-H3 and the reference cell is HeLa. In some embodiments, the co-stimulatory protein is ICOSL or B7-H4, and the reference cell is SK-BR-3. In some embodiments, the co-stimulatory protein is ICOS or OX40, and the reference cell is MOLT-4. In some embodiments, the co-stimulatory protein is CD28, and the reference cell is U-266. In some embodiments, the co-stimulatory protein is CD30L or CD27, and the reference cell is Daudi.

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition does not substantially elicit an immunogenic response by the immune system, e.g., innate immune system. In embodiments, an immunogenic response can be quantified, e.g., as described herein. In some embodiments, an immunogenic response by the innate immune system comprises a response by innate immune cells including, but not limited to NK cells, macrophages, neutrophils, basophils, eosinophils, dendritic cells, mast cells, or gamma/delta T cells. In some embodiments, an immunogenic response by the innate immune system comprises a response by the complement system which includes soluble blood components and membrane bound components.

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition does not substantially elicit an immunogenic response by the immune system, e.g., adaptive immune system. In some embodiments, an immunogenic response by the adaptive immune system comprises an immunogenic response by an adaptive immune cell including, but not limited to a change, e.g., increase, in number or activity of T lymphocytes (e.g., CD4 T cells, CD8 T cells, and or gamma-delta T cells), or B lymphocytes. In some embodiments, an immunogenic response by the adaptive immune system includes increased levels of soluble blood components including, but not limited to a change, e.g., increase, in number or activity of cytokines or antibodies (e.g., IgG, IgM, IgE, IgA, or IgD).

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition is modified to have reduced immunogenicity. In some embodiments, the retroviral vector, VLP, or pharmaceutical composition has an immunogenicity less than 5%, 10%, 20%, 30%, 40%, or 50% lesser than the immunogenicity of a reference retroviral vector or VLP, e.g., an unmodified

retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell.

In some embodiments of any of the aspects described herein, the retroviral vector, VLP, or pharmaceutical composition is derived from a source cell, e.g., a mammalian cell, having a modified genome, e.g., modified using a method described herein, to reduce, e.g., lessen, immunogenicity. Immunogenicity can be quantified, e.g., as described herein.

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition is derived from a mammalian cell depleted of, e.g., with a knock out of, one, two, three, four, five, six, seven or more of the following:

a. MHC class I, MHC class II or MHA;
b. one or more co-stimulatory proteins including but not limited to: LAG3, ICOS-L, ICOS, Ox40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3, or B7-H4;
c. soluble immune-stimulating cytokines e.g., IFN-gamma or TNF-a;
d. endogenous immune-stimulatory antigen, e.g., Zg16 or Hormad1;
e. T-cell receptors (TCR);
f. The genes encoding ABO blood groups, e.g., *ABO* gene;
g. transcription factors which drive immune activation, e.g., NFkB;
h. transcription factors that control MHC expression e.g., class II trans-activator (CIITA), regulatory factor of the Xbox 5 (RFX5), RFX-associated protein (RFXAP), or RFX ankyrin repeats (RFXANK; also known as RFXB); or
i. TAP proteins, e.g., TAP2, TAP1, or TAPBP, which reduce MHC class I expression.

In some embodiments, the retroviral vector or VLP is derived from a source cell with a genetic modification which results in increased expression of an immunosuppressive agent, e.g., one, two, three or more of the following (e.g., wherein before the genetic modification the cell did not express the factor):

a. surface proteins which suppress macrophage engulfment, e.g., CD47; e.g., increased expression of CD47 compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell;

b. soluble immunosuppressive cytokines, e.g., IL-10, e.g., increased expression of IL-10 compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell;

c. soluble immunosuppressive proteins, e.g., PD-1, PD-L1, CTLA4, or BTLA; e.g., increased expression of immunosuppressive proteins compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the cell source, a HEK293 cell, or a Jurkat cell;

d. a tolerogenic protein, e.g., an ILT-2 or ILT-4 agonist, e.g., HLA-E or HLA-G or any other endogenous ILT-2 or ILT-4 agonist, e.g., increased expression of HLA-E, HLA-G, ILT-2 or ILT-4 compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell; or

e. surface proteins which suppress complement activity, e.g., complement regulatory proteins, e.g. proteins that bind decay-accelerating factor (DAF, CD55), e.g. factor H (FH)-like protein-1 (FHL-1), e.g. C4b-binding protein (C4BP), e.g. complement receptor 1 (CD35), e.g. Membrane cofactor protein (MCP, CD46), eg. Profectin (CD59), e.g. proteins that inhibit the classical and alternative compelement pathway CD/C5 convertase enzymes, e.g. proteins that regulate MAC assembly; e.g. increased expression of a complement regulatory protein compared to a reference retroviral vector or VLP, e.g. an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell.

In some embodiments, the increased expression level is at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold higher as compared to a reference retroviral vector or VLP.

In some embodiments, the retroviral vector or VLP is derived from a source cell modified to have decreased expression of an immunostimulatory agent, e.g., one, two, three, four, five, six, seven, eight or more of the following:

a. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of MHC class I or MHC class II, compared to a reference retroviral vector or VLP, e.g.,

an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a HeLa cell;

b. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of one or more co-stimulatory proteins including but not limited to: LAG3, ICOS-L, ICOS, Ox40L, OX40, CD28, B7, CD30, CD30L 4-1BB, 4-1BBL, SLAM, CD27, CD70, HVEM, LIGHT, B7-H3, or B7-H4, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a reference cell described herein;

c. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of soluble immune stimulating cytokines, e.g., IFN-gamma or TNF-a, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a U-266 cell;

d. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of endogenous immune-stimulatory antigen, e.g., Zg16 or Hormad1, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or an A549 cell or a SK-BR-3 cell;

e. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of T-cell receptors (TCR) compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell;

f. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of ABO blood groups, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a HeLa cell;

g. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of transcription factors which drive immune activation, e.g., NFkB; compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell

h. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of transcription factors that control MHC expression, e.g., class II trans-activator

(CIITA), regulatory factor of the Xbox 5 (RFX5), RFX-associated protein (RFXAP), or RFX ankyrin repeats (RFXANK; also known as RFXB) compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a Jurkat cell; or

i. less than 50%, 40%, 30%, 20%, 15%, 10%, or 5% or lesser expression of TAP proteins, e.g., TAP2, TAP1, or TAPBP, which reduce MHC class I expression compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, a HEK293 cell, or a HeLa cell.

In some embodiments, a retroviral vector, VLP, or pharmaceutical composition derived from a mammalian cell, e.g., a HEK293, modified using shRNA expressing lentivirus to decrease MHC Class I expression, has lesser expression of MHC Class I compared to an unmodified retroviral vector or VLP, e.g., a retroviral vector or VLP from a cell (e.g., mesenchymal stem cell) that has not been modified. In some embodiments, a retroviral vector or VLP derived from a mammalian cell, e.g., a HEK293, modified using lentivirus expressing HLA-G to increase expression of HLA-G, has increased expression of HLA-G compared to an unmodified retroviral vector or VLP, e.g., from a cell (e.g., a HEK293) that has not been modified.

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition is derived from a source cell, e.g., a mammalian cell, which is not substantially immunogenic, wherein the source cells stimulate, e.g., induce, T-cell IFN-gamma secretion, at a level of 0 pg/mL to >0 pg/mL, e.g., as assayed i*n vitro,* by IFN-gamma ELISPOT assay.

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition is derived from a source cell, e.g., a mammalian cell, wherein the mammalian cell is from a cell culture treated with an immunosuppressive agent, e.g., a glucocorticoid (e.g., dexamethasone), cytostatic (e.g., methotrexate), antibody (e.g., Muromonab-CD3), or immunophilin modulator (e.g., Ciclosporin or rapamycin).

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition is derived from a source cell, e.g., a mammalian cell, wherein the mammalian cell comprises an exogenous agent, e.g., a therapeutic agent.

In some embodiments, the retroviral vector, VLP, or pharmaceutical composition is derived from a source cell, e.g., a mammalian cell, wherein the mammalian cell is a recombinant cell.

In some embodiments, the retroviral vector, VLP, or pharmaceutical is derived from a mammalian cell genetically modified to express viral immunoevasins, e.g., hCMV US2, or US11.

In some embodiments, the surface of the retroviral vector or VLP, or the surface of the source cell, is covalently or non-covalently modified with a polymer, e.g., a biocompatible polymer that reduces immunogenicity and immune-mediated clearance, e.g., PEG.

In some embodiments, the surface of the retroviral vector or VLP, or the surface of the source cell is covalently or non-covalently modified with a sialic acid, e.g., a sialic acid comprising glycopolymers, which contain NK-suppressive glycan epitopes.

In some embodiments, the surface of the retroviral vector or VLP, or the surface of the source cell is enzymatically treated, e.g., with glycosidase enzymes, e.g., α-N-acetylgalactosaminidases, to remove ABO blood groups

In some embodiments, the surface of the retroviral vector or VLP, or the surface of the source cell is enzymatically treated, to give rise to, e.g., induce expression of, ABO blood groups which match the recipient's blood type.

Parameters for assessing immunogenicity

In some embodiments, the retroviral vector or VLP is derived from a source cell, e.g., a mammalian cell which is not substantially immunogenic, or modified, e.g., modified using a method described herein, to have a reduction in immunogenicity. Immunogenicity of the source cell and the retroviral vector or VLP can be determined by any of the assays described herein.

In some embodiments, the retroviral vector or VLP has an increase, e.g., an increase of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more, in in vivo graft survival compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell.

In some embodiments, the retroviral vector or VLP has a reduction in immunogenicity as measured by a reduction in humoral response following one or more implantation of the retroviral vector or VLP into an appropriate animal model, e.g., an animal model described

herein, compared to a humoral response following one or more implantation of a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, into an appropriate animal model, e.g., an animal model described herein. In some embodiments, the reduction in humoral response is measured in a serum sample by an anti-cell antibody titre, e.g., anti-retroviral or anti-VLP antibody titre, e.g., by ELISA. In some embodiments, the serum sample from animals administered the retroviral vector or VLP has a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of an anti-retroviral or anti-VLP antibody titer compared to the serum sample from animals administered an unmodified retroviral vector or VLP. In some embodiments, the serum sample from animals administered the retroviral vector or VLP has an increased anti- retroviral or anti-VLP antibody titre, e.g., increased by 1%, 2%, 5%, 10%, 20%, 30%, or 40% from baseline, e.g., wherein baseline refers to serum sample from the same animals before administration of the retroviral vector or VLP.

In some embodiments, the retroviral vector or VLP has a reduction in macrophage phagocytosis, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in macrophage phagocytosis compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein the reduction in macrophage phagocytosis is determined by assaying the phagocytosis index *in vitro*, e.g., as described in Example 8. In some embodiments, the retroviral vector or VLP has a phagocytosis index of 0, 1, 10, 100, or more, e.g., as measured by an assay of Example 8, when incubated with macrophages in an *in vitro* assay of macrophage phagocytosis.

In some embodiments, the source cell or recipient cell has a reduction in cytotoxicity mediated cell lysis by PBMCs, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in cell lysis compared to a reference cell, e.g., an unmodified cell otherwise similar to the source cell, or a recipient cell that received an unmodified retroviral vector or VLP, or a mesenchymal stem cells, e.g., using an assay of Example 17. In embodiments, the source cell expresses exogenous HLA-G.

In some embodiments, the source cell or recipient cell has a reduction in NK-mediated cell lysis, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in NK-mediated cell lysis compared to a reference cell, e.g., an unmodified cell otherwise similar to the source cell, or a recipient cell that received an unmodified retroviral vector or VLP,

wherein NK-mediated cell lysis is assayed i*n vitro,* by a chromium release assay or europium release assay, e.g., using an assay of Example 18.

In some embodiments, the source cell or recipient cell has a reduction in CD8+ T-cell mediated cell lysis, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in CD8 T cell mediated cell lysis compared to a reference cell, e.g., an unmodified cell otherwise similar to the source cell, or a recipient cell that received an unmodified retroviral vector or VLP, wherein CD8 T cell mediated cell lysis is assayed *in vitro,* by an assay of Example 19.

In some embodiments, the source cell or recipient cell has a reduction in CD4+ T-cell proliferation and/or activation, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more compared to a reference cell, e.g., an unmodified cell otherwise similar to the source cell, or a recipient cell that received an unmodified retroviral vector or VLP, wherein CD4 T cell proliferation is assayed i*n vitro* (*e.g.* co-culture assay of modified or unmodified mammalian source cell, and CD4+T-cells with CD3/CD28 Dynabeads).

In some embodiments, the retroviral vector or VLP causes a reduction in T-cell IFN-gamma secretion, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in T-cell IFN-gamma secretion compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein T-cell IFN-gamma secretion is assayed *in vitro, e.g.,* by IFN-gamma ELISPOT.

In some embodiments, the retroviral vector or VLP causes a reduction in secretion of immunogenic cytokines, e.g., a reduction of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in secretion of immunogenic cytokines compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein secretion of immunogenic cytokines is assayed *in vitro* using ELISA or ELISPOT.

In some embodiments, the retroviral vector or VLP results in increased secretion of an immunosuppressive cytokine, e.g., an increase of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in secretion of an immunosuppressive cytokine compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein secretion of the immunosuppressive cytokine is assayed *in vitro* using ELISA or ELISPOT.

In some embodiments, the retroviral vector or VLP has an increase in expression of HLA-G or HLA-E, e.g., an increase in expression of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of HLA-G or HLA-E, compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein expression of HLA-G or HLA-E is assayed *in vitro* using flow cytometry, e.g., FACS. In some embodiments, the retroviral vector or VLP is derived from a source cell which is modified to have an increased expression of HLA-G or HLA-E, e.g., compared to an unmodified cell, e.g., an increased expression of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of HLA-G or HLA-E, wherein expression of HLA-G or HLA-E is assayed *in vitro* using flow cytometry, e.g., FACS. In some embodiments, the retroviral vector or VLP derived from a modified cell with increased HLA-G expression demonstrates reduced immunogenicity.

In some embodiments, the retroviral vector or VLP has or causes an increase in expression of T cell inhibitor ligands (e.g. CTLA4, PD1, PD-L1), e.g., an increase in expression of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of T cell inhibitor ligands as compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein expression of T cell inhibitor ligands is assayed i*n vitro* using flow cytometry, e.g., FACS.

In some embodiments, the retroviral vector or VLP has a decrease in expression of co-stimulatory ligands, e.g., a decrease of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more in expression of co-stimulatory ligands compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell, wherein expression of co-stimulatory ligands is assayed *in vitro* using flow cytometry, e.g., FACS.

In some embodiments, the retroviral vector or VLP has a decrease in expression of MHC class I or MHC class II, e.g., a decrease in expression of 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more of MHC Class I or MHC Class II compared to a reference retroviral vector or VLP, e.g., an unmodified retroviral vector or VLP from a cell otherwise similar to the source cell or a HeLa cell, wherein expression of MHC Class I or II is assayed i*n vitro* using flow cytometry, e.g., FACS.

In some embodiments, the retroviral vector or VLP is derived from a cell source, e.g., a mammalian cell source, which is substantially non-immunogenic. In some embodiments, immunogenicity can be quantified, e.g., as described herein. In some embodiments, the mammalian cell source comprises any one, all or a combination of the following features:

a. wherein the source cell is obtained from an autologous cell source; e.g., a cell obtained from a recipient who will be receiving, e.g., administered, the retroviral vector or VLP;

b. wherein the source cell is obtained from an allogeneic cell source which is of matched, e.g., similar, gender to a recipient, e.g., a recipient described herein who will be receiving, e.g., administered; the retroviral vector or VLP;

c. wherein the source cell is obtained is from an allogeneic cell source is which is HLA matched with a recipient's HLA, e.g., at one or more alleles;

d. wherein the source cell is obtained is from an allogeneic cell source which is an HLA homozygote;

e. wherein the source cell is obtained is from an allogeneic cell source which lacks (or has reduced levels compared to a reference cell) MHC class I and II; or

f. wherein the source cell is obtained is from a cell source which is known to be substantially non-immunogenic including but not limited to a stem cell, a mesenchymal stem cell, an induced pluripotent stem cell, an embryonic stem cell, a sertoli cell, or a retinal pigment epithelial cell.

In some embodiments, the subject to be administered the retroviral vector or VLP has, or is known to have, or is tested for, a pre-existing antibody (e.g., IgG or IgM) reactive with a retroviral vector or VLP. In some embodiments, the subject to be administered the retroviral vector or VLP does not have detectable levels of a pre-existing antibody reactive with the retroviral vector or VLP. Tests for the antibody are described, e.g., in Example 13.

In some embodiments, a subject that has received the retroviral vector or VLP has, or is known to have, or is tested for, an antibody (e.g., IgG or IgM) reactive with a retroviral vector or VLP. In some embodiments, the subject that received the retroviral vector or VLP (e.g., at least once, twice, three times, four times, five times, or more) does not have detectable levels of antibody reactive with the retroviral vector or VLP. In embodiments, levels of antibody do not rise more than 1%, 2%, 5%, 10%, 20%, or 50% between two timepoints, the first timepoint being before the first administration of the retroviral vector or VLP, and the second timepoint being

after one or more administrations of the retroviral vector or VLP. Tests for the antibody are described, e.g., in Example 14.

VI. Exogenous agents

In some embodiments, fusosome, e.g. a retroviral vector, VLP, or pharmaceutical composition described herein contains an exogenous agent. In some embodimetns, the fusosome, e.g. a retroviral vector, VLP, or pharmaceutical composition described herein contains a nucleic acid that encodes an exogenous agent.

A. Exogenous protein agents

In some embodiments, the exogenous agent comprises a cytosolic protein, e.g., a protein that is produced in the recipient cell and localizes to the recipient cell cytoplasm. In some embodiments, the exogenous agent comprises a secreted protein, e.g., a protein that is produced and secreted by the recipient cell. In some embodiments, the exogenous agent comprises a nuclear protein, e.g., a protein that is produced in the recipient cell and is imported to the nucleus of the recipient cell. In some embodiments, the exogenous agent comprises an organellar protein (e.g., a mitochondrial protein), e.g., a protein that is produced in the recipient cell and is imported into an organelle (e.g., a mitochondrial) of the recipient cell. In some embodiments, the protein is a wild-type protein or a mutant protein. In some embodiments the protein is a fusion or chimeric protein.

In some embodiments, the exogenous agent is encoded by a gene from among SYNE1, SETX, FMR1, SLC6A8, UBE3A, SOD1, TDP43, C9orf72, FXN, MECP2, ASPA, ALDH7A1, TPP1, FUCA1, GALC, HEXA, HEXB, MANBA, ARSA, GNPTAB, or MCOLN1. In some embodiments, the exogenous agent is encoded by a gene from among In some embodiments, the exogenous agent is encoded by a gene from among SYNE1, SETX, FMR1, SLC6A8, UBE3A, SOD1, TDP43, C9orf72, FXN, MECP2, ASPA, or ALDH7A1. In some embodiments, the exogenous agent is encoded by a gene from among TPP1, FUCA1, GALC, HEXA, HEXB, MANBA, ARSA, GNPTAB, or MCOLN1.

In some embodiments, the exogenous agent comprises a protein of Table 5 below. In some embodiments, the exogenous agent comprises the wild-type human sequence of any of the proteins of Table 5, a functional fragment thereof (e.g., an enzymatically active fragment

thereof), or a functional variant thereof. In some embodiments, the exogenous agent comprises an amino acid seqence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence of Table 5, e.g., a Uniprot Protein Accession Number sequence of column 2 of Table 5 or an amino acid sequence of column 3 of Table 5. In some embodiments, the payload gene encodes an an amino acid seqence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequnce of Table 5.

Table 5. *CNS diseases or disorders. The first column lists exogenous agents that can be delivered to treat the indications in the fourth column, according to the methods and uses herein. Each Uniprot accession number of Table 5 is herein incorporated by reference in its entirety.*

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
SYNE1	Q8NF91	MATSRGASRCPRDIANVMQRLQDEQEIVQKRT FTKWINSHLAKRKPPMVVDDLFEDMKDGVKL LALLEVLSGQKLPCEQGRRMKRIHAVANIGTA LKFLEGRKIKLVNINSTDIADGRPSIVLGLMWTI ILYFQIEELTSNLPQLQSLSSSASSVDSIVSSETPS PPSKRKVTTKIQGNAKKALLKWVQYTAGKQT GIEVKDFGKSWRSGVAFHSVIHAIRPELVDLET VKGRSNRENLEDAFTIAETELGIPRLLDPEDVD VDKPDEKSIMTYVAQFLKHYPDIHNASTDGQE DDEILPGFPSFANSVQNFKREDRVIFKEMKVWI EQFERDLTRAQMVESNLQDKYQSFKHFRVQY EMKRKQIEHLIQPLHRDGKLSLDQALVKQSWD RVTSRLFDWHIQLDKSLPAPLGTIGAWLYRAE VALREEITVQQVHEETANTIQRKLEQHKDLLQ NTDAHKRAFHEIYRTRSVNGIPVPPDQLEDMA ERFHFVSSTSELHLMKMEFLELKYRLLSLLVLA ESKLKSWIIKYGRRESVEQLLQNYVSFIENSKFF EQYEVTYQILKQTAEMYVKADGSVEEAENVM KFMNETTAQWRNLSVEVRSVRSMLEEVISNW DRYGNTVASLQAWLEDAEKMLNQSENAKKD FFRNLPHWIQQHTAMNDAGNFLIETCDEMVSR DLKQQLLLLNGRWRELFMEVKQYAQADEMD RMKKEYTDCVVTLSAFATEAHKKLSEPLEVSF MNVKLLIQDLEDIEQRVPVMDAQYKIITKTAH LITKESPQEEGKEMFATMSKLKEQLTKVKECY SPLLYESQQLLIPLEELEKQMTSFYDSLGKINEII TVLEREAQSSALFKQKHQELLACQENCKKTLT LIEKGSQSVQKFVTLSNVLKHFDQTRLQRQIAD IHVAFQSMVKKTGDWKKHVETNSRLMKKFEE SRAELEKVLRIAQEGLEEKGDPEELLRRHTEFF SQLDQRVLNAFLKACDELTDILPEQEQQGLQE AVRKLHKQWKDLQGEAPYHLLHLKIDVEKNR FLASVEECRTELDRETKLMPQEGSEKIIKEHRV FFSDKGPHHLCEKRLQLIEELCVKLPVRDPVRD	Spinocerebellar Ataxia, Autosomal Recessive, Type 1	134

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		TPGTCHVTLKELRAAIDSTYRKLMEDPDKWK DYTSRFSEFSSWISTNETQLKGIKGEAIDTANH GEVKRAVEEIRNGVTKRGETLSWLKSRLKVLT EVSSENEAQKQGDELAKLSSSFKALVTLLSEVE KMLSNFGDCVQYKEIVKNSLEELISGSKEVQE QAEKILDTENLFEAQQLLLHHQQKTKRISAKK RDVQQQIAQAQQGEGGLPDRGHEELRKLESTL DGLERSRERQERRIQVTLRKWERFETNKETVV RYLFQTGSSHERFLSFSSLESLSSELEQTKEFSK RTESIAVQAENLVKEASEIPLGPQNKQLLQQQA KSIKEQVKKLEDTLEEDIKTMEMVKTKWDHF GSNFETLSVWITEKEKELNALETSSSAMDMQIS QIKVTIQEIESKLSSIVGLEEEAQSFAQFVTTGES ARIKAKLTQIRRYGEELREHAQCLEGTILGHLS QQQKFEENLRKIQQSVSEFEDKLAVPIKICSSAT ETYKVLQEHMDLCQALESLSSAITAFSASARK VVNRDSCVQEAAALQQQYEDILRRAKERQTA LENLLAHWQRLEKELSSFLTWLERGEAKASSP EMDISADRVKVEGELQLIQALQNEVVSQASFY SKLLQLKESLFSVASKDDVKMMKLHLEQLDE RWRDLPQIINKRINFLQSVVAEHQQFDELLLSF SVWIKLFLSELQTTSEISIMDHQVALTRHKDHA AEVESKKGELQSLQGHLAKLGSLGRAEDLHLL QGKAEDCFQLFEEASQVVERRQLALSHLAEFL QSHASLSGILRQLRQTVEATNSMNKNESDLIEK DLNDALQNAKALESAAVSLDGILSKAQYHLKI GSSEQRTSCRATADQLCGEVERIQNLLGTKQSE ADALAVLKKAFQDQKEELLKSIEDIEERTDKER LKEPTRQALQQRLRVFNQLEDELNSHEHELCW LKDKAKQIAQKDVAFAPEVDREINRLEVTWD DTKRLIHENQGQCCGLIDLMREYQNLKSAVSK VLENASSVIVTRTTIKDQEDLKWAFSKHETAK NKMNYKQKDLDNFTSKGKHLLSELKKIHSSDF SLVKTDMESTVDKWLDVSEKLEENMDRLRVS LSIWDDVLSTRDEIEGWSNNCVPQMAENISNL DNHLRAEELLKEFESEVKNKALRLEELHSKVN DLKELTKNLETPPDLQFIEADLMQKLEHAKEIT EVAKGTLKDFTAQSTQVEKFINDITTWFTKVEE SLMNCAQNETCEALKKVKDIQKELQSQQSNIS STQENLNSLCRKYHSAELESLGRAMTGLIKKH EAVSQLCSKTQASLQESLEKHFSESMQEFQEW FLGAKAAAKESSDRTGDSKVLEAKLHDLQNIL DSVSDGQSKLDAVTQEGQTLYAHLSKQIVSSIQ EQITKANEEFQAFLKQCLKDKQALQDCASELG SFEDQHRKLNLWIHEMEERFNTENLGESKQHIP EKKNEVHKVEMFLGELLAARESLDKLSQRGQ LLSEEGHGAGQEGRLCSQLLTSHQNLLRMTKE KLRSCQVALQEHEALEEALQSMWFWVKAIQD RLACAESTLGSKDTLEKRLSQIQDILLMKGEGE VKLNMAIGKGEQALRSSNKEGQRVIQTQLETL KEVWADIMSSSVHAQSTLESVISQWNDYVERK NQLEQWMESVDQKIEHPLQPQPGLKEKFVLLD		

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		HLQSILSEAEDHTRALHRLIAKSRELYEKTEDE SFKDTAQEELKTQFNDIMTVAKEKMRKVEEIV KDHLMYLDAVHEFTDWLHSAKEELHRWSDM SGDSSATQKKLSKIKELIDSREIGASRLSRVESL APEVKQNTTASGCELMHTEMQALRADWKQW EDSVFQTQSCLENLVSQMALSEQEFSGQVAQL EQALEQFSALLKTWAQQLTLLEGKNTDEEIVE CWHKGQEILDALQKAEPRTEDLKSQLNELCRF SRDLSTYSGKVSGLIKEYNCLCLQASKGCQNK EQILQQRFRKAFRDFQQWLVNAKITTAKCFDIP QNISEVSTSLQKIQEFLSESENGQHKLNMMLSK GELLSTLLTKEKAKGIQAKVTAAKEDWKNFHS NLHQKESALENLKIQMKDFEVSAEPIQDWLSK TEKMVHESSNRLYDLPAKRREQQKLQSVLEEI HCYEPQLNRLKEKAQQLWEGQAASKSFRHRV SQLSSQYLALSNLTKEKVSRLDRIVAEHNQFSL GIKELQDWMTDAIHMLDSYCHPTSDKSVLDSR TLKLEALLSVKQEKEIQMKMIVTRGESVLQNT SPEGIPTIQQQLQSVKDMWASLLSAGIRCKSQL EGALSKWTSYQDGVRQFSGWMDSMEANLNE SERQHAELRDKTTMLGKAKLLNEEVLSYSSLL ETIEVKGAGMTEHYVTQLELQDLQERYRAIQE RAKEAVTKSEKLVRLHQEYQRDLKAFEVWLG QEQEKLDQYSVLEGDAHTHETTLRDLQELQV HCAEGQALLNSVLHTREDVIPSGIPQAEDRALE SLRQDWQAYQHRLSETRTQFNNVVNKLRLME QKFQQVDEWLKTAEEKVSPRTRRQSNRATKEI QLHQMKKWHEEVTAYRDEVEEVGARAQEILD ESHVNSRMGCQATQLTSRYQALLLQVLEQIKF LEEEIQSLEESESSLSSYSDWYGSTHKNFKNVA TKIDKVDTVMMGKKLKTLEVLLKDMEKGHSL LKSAREKGERAVKYLEEGEAERLRKEIHDHME QLKELTSTVRKEHMTLEKGLHLAKEFSDKCKA LTQWIAEYQEILHVPEEPKMELYEKKAQLSKY KSLQQTVLSHEPSVKSVREKGEALLELVQDVT LKDKIDQLQSDYQDLCSIGKEHVFSLEAKVKD HEDYNSELQEVEKWLLQMSGRLVAPDLLETSS LETITQQLAHHKAMMEEIAGFEDRLNNLQMK GDTLIGQCADHLQAKLKQNVHAHLQGTKDSY SAICSTAQRMYQSLEHELQKHVSRQDTLQQCQ AWLSAVQPDLEPSPQPPLSRAEAIKQVKHFRAL QEQARTYLDLLCSMCDLSNASVKTTAKDIQQT EQTIEQKLVQAQNLTQGWEEIKHLKSELWIYL QDADQQLQNMKRRHSELELNIAQNMVSQVKD FVKKLQSKQASVNTIIEKVNKLTKKEESPEHKE INHLNDQWLDLCRQSNNLCLQREEDLQRTRD YHDCMNVVEVFLEKFTTEWDNLARSDAESTA VHLEALKKLALALQERKYAIEDLKDQKQKMIE HLNLDDKELVKEQTSHLEQRWFQLEDLIKRKI QVSVTNLEELNVVQSRFQELMEWAEEQQPNIA EALKQSPPPDMAQNLLMDHLAICSELEAKQML LKSLIKDADRVMADLGLNERQVIQKALSDAQS		

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		HVNCLSDLVGQRRKYLNKALSEKTQFLMAVF QATSQIQQHERKIMFREHICLLPDDVSKQVKTC KSAQASLKTYQNEVTGLWAQGRELMKEVTEQ EKSEVLGKLQELQSVYDSVLQKCSHRLQELEK NLVSRKHFKEDFDKACHWLKQADIVTFPEINL MNESSELHTQLAKYQNILEQSPEYENLLLTLQR TGQTILPSLNEVDHSYLSEKLNALPRQFNVIVA LAKDKFYKVQEAILARKEYASLIELTTQSLSEL EAQFLRMSKVPTDLAVEEALSLQDGCRAILDE VAGLGEAVDELNQKKEGFRSTGQPWQPDKML HLVTLYHRLKRQTEQRVSLLEDTTSAYQEHEK MCQQLERQLKSVKEEQSKVNEETLPAEEKLK MYHSLAGSLQDSGIVLKRVTIHLEDLAPHLDPL AYEKARHQIQSWQGELKLLTSAIGETVTECESR MVQSIDFQTEMSRSLDWLRRVKAELSGPVYLD LNLQDIQEEIRKIQIHQEEVQSSLRIMNALSHKE KEKFTKAKELISADLEHSLAELSELDGDIQEAL RTRQATLTEIYSQCQRYYQVFQAANDWLEDA QELLQLAGNGLDVESAEENLKSHMEFFSTEDQ FHSNLEELHSLVATLDPLIKPTGKEDLEQKVAS LELRSQRMSRDSGAQVDLLQRCTAQWHDYQK AREEVIELMNDTEKKLSEFSLLKTSSSHEAEEK LSEHKALVSVVNSFHEKIVALEEKASQLEKTG NDASKATLSRSMTTVWQRWTRLRAVAQDQE KILEDAVDEWTGFNNKVKKATEMIDQLQDKL PGSSAEKASKAELLTLLEYHDTFVLELEQQQSA LGMLRQQTLSMLQDGAAPTPGEEPPLMQEITA MQDRCLNMQEKVKTNGKLVKQELKDREMVE TQINSVKCWVQETKEYLGNPTIEIDAQLEELQI LLTEATNHRQNIEKMAEEQKEKYLGLYTILPSE LSLQLAEVALDLKIRDQIQDKIKEVEQSKATSQ ELSRQIQKLAKDLTTILTKLKAKTDNVVQAKT DQKVLGEELDGCNSKLMELDAAVQKFLEQNG QLGKPLAKKIGKLTELHQQTIRQAENRLSKLN QAASHLEEYNEMLELILKWIEKAKVLAHGTIA WNSASQLREQYILHQTLLEESKEIDSELEAMTE KLQYLTSVYCTEKMSQQVAELGRETEELRQMI KIRLQNLQDAAKDMKKFEAELKKLQAALEQA QATLTSPEVGRLSLKEQLSHRQHLLSEMESLKP KVQAVQLCQSALRIPEDVVASLPLCHAALRLQ EEASRLQHTAIQQCNIMQEAVVQYEQYEQEM KHLQQLIEGAHREIEDKPVATSNIQELQAQISR HEELAQKIKGYQEQIASLNSKCKMLTMKAKH ATMLLTVTEVEGLAEGTEDLDGELLPTPSAHPS VVMMTAGRCHTLLSPVTEESGEEGTNSEISSPP ACRSPSPVANTDASVNQDIAYYQALSAERLQT DAAKIHPSTSASQEFYEPGLEPSATAKLGDLQR SWETLKNVISEKQRTLYEALERQQKYQDSLQSI STKMEAIELKLSESPEPGRSPESQMAEHQALM DEILMLQDEINELQSSLAEELVSESCEADPAEQ LALQSTLTVLAERMSTIRMKASGKRQLLEEKL NDQLEEQRQEQALQRYRCEADELDSWLLSTK		

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		ATLDTALSPPKEPMDMEAQLMDCQNMLVEIE QKVVALSELSVHNENLLLEGKAHTKDEAEQL AGKLRRLKGSLLELQRALHDKQLNMQGTAQE KEESDVDLTATQSPGVQEWLAQARTTWTQQR QSSLQQQKELEQELAEQKSLLRSVASRGEEILI QHSAAETSGDAGEKPDVLSQELGMEGEKSSAE DQMRMKWESLHQEFSTKQKLLQNVLEQEQEQ VLYSRPNRLLSGVPLYKGDVPTQDKSAVTSLL DGLNQAFEEVSSQSGGAKRQSIHLEQKLYDGV SATSTWLDDVEERLFVATALLPEETETCLFNQE ILAKDIKEMSEEMDKNKNLFSQAFPENGDNRD VIEDTLGCLLGRLSLLDSVVNQRCHQMKERLQ QILNFQNDLKVLFTSLADNKYIILQKLANVFEQ PVAEQIEAIQQAEDGLKEFDAGIIELKRRGDKL QVEQPSMQELSKLQDMYDELMMIIGSRRSGLN QNLTLKSQYERALQDLADLLETGQEKMAGDQ KIIVSSKEEIQQLLDKHKEYFQGLESHMILTETL FRKIISFAVQKETQFHTELMAQASAVLKRAHK RGVELEYILETWSHLDEDQQELSRQLEVVESSI PSVGLVEENEDRLIDRITLYQHLKSSLNEYQPK LYQVLDDGKRLLISISCSDLESQLNQLGECWLS NTNKMSKELHRLETILKHWTRYQSESADLIHW LQSAKDRLEFWTQQSVTVPQELEMVRDHLNA FLEFSKEVDAQSSLKSSVLSTGNQLLRLKKVDT ATLRSELSRIDSQWTDLLTNIPAVQEKLHQLQ MDKLPSRHAISEVMSWISLMENVIQKDEDNIK NSIGYKAIHEYLQKYKGFKIDINCKQLTVDFVN QSVLQISSQDVESKRSDKTDFAEQLGAMNKSW QILQGLVTEKIQLLEGLLESWSEYENNVQCLKT WFETQEKRLKQQHRIGDQASVQNALKDCQDL EDLIKAKEKEVEKIEQNGLALIQNKKEDVSSIV MSTLRELGQTWANLDHMVGQLKILLKSVLDQ WSSHKVAFDKINSYLMEARYSLSRFRLLTGSL EAVQVQVDNLQNLQDDLEKQERSLQKFGSITN QLLKECHPPVTETLTNTLKEVNMRWNNLLEEI AEQLQSSKALLQLWQRYKDYSKQCASTVQQQ EDRTNELLKAATNKDIADDEVATWIQDCNDLL KGLGTVKDSLFFLHELGEQLKQQVDASAASAI QSDQLSLSQHLCALEQALCKQQTSLQAGVLDY ETFAKSLEALEAWIVEAEEILQGQDPSHSSDLS TIQERMEELKGQMLKFSSMAPDLDRLNELGYR LPLNDKEIKRMQNLNRHWSLISSQTTERFSKLQ SFLLQHQTFLEKCETWMEFLVQTEQKLAVEIS GNYQHLLEQQRAHELFQAEMFSRQQILHSIIID GQRLLEQGQVDDRDEFNLKLTLLSNQWQGVI RRAQQRRGIIDSQIRQWQRYREMAEKLRKWL VEVSYLPMSGLGSVPIPLQQARTLFDEVQFKEK VFLRQQGSYILTVEAGKQLLLSADSGAEAALQ AELAEIQEKWKSASMRLEEQKKKLAFLLKDW EKCEKGIADSLEKLRTFKKKLSQSLPDHHEELH AEQMRCKELENAVGSWTDDLTQLSLLKDTLS AYISADDISILNERVELLQRQWEELCHQLSLRR		

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		QQIGERLNEWAVFSEKNKELCEWLTQMESKV SQNGDILIEEMIEKLKKDYQEEIAIAQENKIQLQ QMGERLAKASHESKASEIEYKLGKVNDRWQH LLDLIAARVKKLKETLVAVQQLDKNMSSLRT WLAHIESELAKPIVYDSCNSEEIQRKLNEQQEL QRDIEKHSTGVASVLNLCEVLLHDCDACATDA ECDSIQQATRNLDRRWRNICAMSMERRLKIEE TWRLWQKFLDDYSRFEDWLKSSERTAAFPSSS GVIYTVAKEELKKFEAFQRQVHECLTQLELINK QYRRLARENRTDSACSLKQMVHEGNQRWDN LQKRVTSILRRLKHFIGQREEFETARDSILVWL TEMDLQLTNIEHFSECDVQAKIKQLKAFQQEIS LNHNKIEQIIAQGEQLIEKSEPLDAAIIEEELDEL RRYCQEVFGRVERYHKKLIRLPLPDDEHDLSD RELELEDSAALSDLHWHDRSADSLLSPQPSSNL SLSLAQPLRSERSGRDTPASVDSIPLEWDHDYD LSRDLESAMSRALPSEDEEGQDDKDFYLRGAV GLSGDHSALESQIRQLGKALDDSRFQIQQTENII RSKTPTGPELDTSYKGYMKLLGECSSSIDSVKR LEHKLKEEEESLPGFVNLHSTETQTAGVIDRWE LLQAQALSKELRMKQNLQKWQQFNSDLNSIW AWLGDTEEELEQLQRLELSTDIQTIELQIKKLK ELQKAVDHRKAIILSINLCSPEFTQADSKESRDL QDRLSQMNGRWDRVCSLLEEWRGLLQDALM QCQGFHEMSHGLLLMLENIDRRKNEIVPIDSNL DAEILQDHHKQLMQIKHELLESQLRVASLQDM SCQLLVNAEGTDCLEAKEKVHVIGNRLKLLLK EVSRHIKELEKLLDVSSSQQDLSSWSSADELDT SGSVSPTSGRSTPNRQKTPRGKCSLSQPGPSVSS PHSRSTKGGSDSSLSEPGPGRSGRGFLFRVLRA ALPLQLLLLLLIGLACLVPMSEEDYSCALSNNF ARSFHPMLRYTNGPPPL		
SETX	Q7Z333	MSTCCWCTPGGASTIDFLKRYASNTPSGEFQT ADEDLCYCLECVAEYHKARDELPFLHEVLWE LETLRLINHFEKSMKAEIGDDDELYIVDNNGE MPLFDITGQDFENKLRVPLLEILKYPYLLLHER VNELCVEALCRMEQANCSFQVFDKHPGIYLFL VHPNEMVRRWAILTARNLGKVDRDDYYDLQE VLLCLFKVIELGLLESPDIYTSSVLEKGKLILLPS HMYDTTNYKSYWLGICMLLTILEEQAMDSLLL GSDKQNDFMQSILHTMEREADDDSVDPFWPA LHCFMVILDRLGSKVWGQLMDPIVAFQTIINN ASYNREIRHIRNSSVRTKLEPESYLDDMVTCSQ IVYNYNPEKTKKDSGWRTAICPDYCPNMYEE METLASVLQSDIGQDMRVHNSTFLWFIPFVQS LMDLKDLGVAYIAQVVNHLYSEVKEVLNQTD AVCDKVTEFFLLILVSVIELHRNKKCLHLLWVS SQQWVEAVVKCAKLPTTAFTRSSEKSSGNCSK GTAMISSLSLHSMPSNSVQLAYVQLIRSLLKEG YQLGQQSLCKRFWDKLNLFLRGNLSLGWQLT SQETHELQSCLKQIIRNIKFKAPPCNTFVDLTSA CKISPASYNKEESEQMGKTSRKDMHCLEASSP	Ataxia with Oculomotor Apraxia, Type 2	135

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		TFSKEPMKVQDSVLIKADNTIEGDNNEQNYIK DVKLEDHLLAGSCLKQSSKNIFTERAEDQIKIS TRKQKSVKEISSYTPKDCTSRNGPERGCDRGII VSTRLLTDSSTDALEKVSTSNEDFSLKDDALAK TSKRKTKVQKDEICAKLSHVIKKQHRKSTLVD NTINLDENLTVSNIESFYSRKDTGVQKGDGFIH NLSLDPSGVLDDKNGEQKSQNNVLPKEKQLK NEELVIFSFHENNCKIQEFHVDGKELIPFTEMTN ASEKKSSPFKDLMTVPESRDEEMSNSTSVIYSN LTREQAPDISPKSDTLTDSQIDRDLHKLSLLAQ ASVITFPSDSPQNSSQLQRKVKEDKRCFTANQN NVGDTSRGQVIIISDSDDDDDERILSLEKLTKQ DKICLEREHPEQHVSTVNSKEEKNPVKEEKTET LFQFEESDSQCFEFESSSEVFSVWQDHPDDNNS VQDGEKKCLAPIANTTNGQGCTDYVSEVVKK GAEGIEEHTRPRSISVEEFCEIEVKKPKRKRSEK PMAEDPVRPSSSVRNEGQSDTNKRDLVGNDFK SIDRRTSTPNSRIQRATTVSQKKSSKLCTCTEPI RKVPVSKTPKKTHSDAKKGQNRSSNYLSCRTT PAIVPPKKFRQCPEPTSTAEKLGLKKGPRKAYE LSQRSLDYVAQLRDHGKTVGVVDTRKKTKLIS PQNLSVRNNKKLLTSQELQMQRQIRPKSQKNR RRLSDCESTDVKRAGSHTAQNSDIFVPESDRSD YNCTGGTEVLANSNRKQLIKCMPSEPETIKAK HGSPATDDACPLNQCDSVVLNGTVPTNEVIVS TSEDPLGGGDPTARHIEMAALKEGEPDSSSDAE EDNLFLTQNDPEDMDLCSQMENDNYKLIELIH GKDTVEVEEDSVSRPQLESLSGTKCKYKDCLE TTKNQGEYCPKHSEVKAADEDVFRKPGLPPPA SKPLRPTTKIFSSKSTSRIAGLSKSLETSSALSPS LKNKSKGIQSILKVPQPVPLIAQKPVGEMKNSC NVLHPQSPNNSNRQGCKVPFGESKYFPSSSPVN ILLSSQSVSDTFVKEVLKWKYEMFLNFGQCGP PASLCQSISRPVPVRFHNYGDYFNVFFPLMVLN TFETVAQEWLNSPNRENFYQLQVRKFPADYIK YWEFAVYLEECELAKQLYPKENDLVFLAPERI NEEKKDTERNDIQDLHEYHSGYVHKFRRTSV MRNGKTECYLSIQTQENFPANLNELVNCIVISS LVTTQRKLKAMSLLGSRNQLARAVLNPNPMD FCTKDLLTTTSERIIAYLRDFNEDQKKAIETAY AMVKHSPSVAKICLIHGPPGTGKSKTIVGLLYR LLTENQRKGHSDENSNAKIKQNRVLVCAPSNA AVDELMKKIILEFKEKCKDKKNPLGNCGDINL VRLGPEKSINSEVLKFSLDSQVNHRMKKELPSH VQAMHKRKEFLDYQLDELSRQRALCRGGREI QRQELDENISKVSKERQELASKIKEVQGRPQKT QSIIILESHIICCTLSTSGGLLLESAFRGQGGVPFS CVIVDEAGQSCEIETLTPLIHRCNKLILVGDPKQ LPPTVISMKAQEYGYDQSMMARFCRLLEENVE HNMISRLPILQLTVQYRMHPDICLFPSNYVYNR NLKTNRQTEAIRCSSDWPFQPYLVFDVGDGSE RRDNDSYINVQEIKLVMEIIKLIKDKRKDVSFR		

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		NIGIITHYKAQKTMIQKDLDKEFDRKGPAEVDT VDAFQGRQKDCVIVTCVRANSIQGSIGFLASLQ RLNVTITRAKYSLFILGHLRTLMENQHWNQLI QDAQKRGAIIKTCDKNYRHDAVKILKLKPVLQ RSLTHPPTIAPEGSRPQGGLPSSKLDSGFAKTSV AASLYHTPSDSKEITLTVTSKDPERPPVHDQLQ DPRLLKRMGIEVKGGIFLWDPQPSSPQHPGATP PTGEPGFPVVHQDLSHIQQPAAVVAALSSHKPP VRGEPPAASPEASTCQSKCDDPEEELCHRREAR AFSEGEQEKCGSETHHTRRNSRWDKRTLEQED SSSKKRKLL		
FMR 1	Q06787	MEELVVEVRGSNGAFYKAFVKDVHEDSITVAF ENNWQPDRQIPFHDVRFPPPVGYNKDINESDE VEVYSRANEKEPCCWWLAKVRMIKGEFYVIE YAACDATYNEIVTIERLRSVNPNKPATKDTFH KIKLDVPEDLRQMCAKEAAHKDFKKAVGAFS VTYDPENYQLVILSINEVTSKRAHMLIDMHFRS LRTKLSLIMRNEEASKQLESSRQLASRFHEQFI VREDLMGLAIGTHGANIQQARKVPGVTAIDLD EDTCTFHIYGEDQDAVKKARSFLEFAEDVIQVP RNLVGKVIGKNGKLIQEIVDKSGVVRVRIEAEN EKNVPQEEEIMPPNSLPSNNSRVGPNAPEEKKH LDIKENSTHFSQPNSTKVQRVLVASSVVAGESQ KPELKAWQGMVPFVFVGTKDSIANATVLLDY HLNYLKEVDQLRLERLQIDEQLRQIGASSRPPP NRTDKEKSYVTDDGQGMGRGSRPYRNRGHGR RGPGYTSGTNSEASNASETESDHRDELSDWSL APTEEERESFLRRGDGRRRGGGGRGQGGRGR GGGFKGNDDHSRTDNRPRNPREAKGRTTDGS LQIRVDCNNERSVHTKTLQNTSSEGSRLRTGK DRNQKKEKPDSVDGQQPLVNGVP	Fragile X Syndrome	136
SLC6 A8	P48029	MAKKSAENGIYSVSGDEKKGPLIAPGPDGAPA KGDGPVGLGTPGGRLAVPPRETWTRQMDFIM SCVGFAVGLGNVWRFPYLCYKNGGGVFLIPY VLIALVGGIPIFFLEISLGQFMKAGSINVWNICP LFKGLGYASMVIVFYCNTYYIMVLAWGFYYL VKSFTTTLPWATCGHTWNTPDCVEIFRHEDCA NASLANLTCDQLADRRSPVIEFWENKVLRLSG GLEVPGALNWEVTLCLLACWVLVYFCVWKG VKSTGKIVYFTATFPYVVLVVLLVRGVLLPGA LDGIIYYLKPDWSKLGSPQVWIDAGTQIFFSYA IGLGALTALGSYNRFNNNCYKDAIILALINSGT SFFAGFVVFSILGFMAAEQGVHISKVAESGPGL AFIAYPRAVTLMPVAPLWAALFFFMLLLLGLD SQFVGVEGFITGLLDLLPASYYFRFQREISVALC CALCFVIDLSMVTDGGMYVFQLFDYYSASGTT LLWQAFWECVVVAWVYGADRFMDDIACMIG YRPCPWMKWCWSFFTPLVCMGIFIFNVVYYEP LVYNNTYVYPWWGEAMGWAFALSSMLCVPL HLLGCLLRAKGTMAERWQHLTQPIWGLHHLE YRAQDADVRGLTTLTPVSESSKVVVVESVM	Cerebral Creatine Deficiency Syndrome 1	137

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
UBE3A	Q05086	MEKLHQCYWKSGEPQSDDIEASRMKRAAAKH LIERYYHQLTEGCGNEACTNEFCASCPTFLRM DNNAAAIKALELYKINAKLCDPHPSKKGASSA YLENSKGAPNNSCSEIKMNKKGARIDFKDVTY LTEEKVYEILELCREREDYSPLIRVIGRVFSSAE ALVQSFRKVKQHTKEELKSLQAKDEDKDEDE KEKAACSAAAMEEDSEASSSRIGDSSQGDNNL QKLGPDDVSVDIDAIRRVYTRLLSNEKIETAFL NALVYLSPNVECDLTYHNVYSRDPNYLNLFIIV MENRNLHSPEYLEMALPLFCKAMSKLPLAAQ GKLIRLWSKYNADQIRRMMETFQQLITYKVIS NEFNSRNLVNDDDAIVAASKCLKMVYYANVV GGEVDTNHNEEDDEEPIPESSELTLQELLGEER RNKKGPRVDPLETELGVKTLDCRKPLIPFEEFI NEPLNEVLEMDKDYTFFKVETENKFSFMTCPFI LNAVTKNLGLYYDNRIRMYSERRITVLYSLVQ GQQLNPYLRLKVRRDHIIDDALVRLEMIAMEN PADLKKQLYVEFEGEQGVDEGGVSKEFFQLVV EEIFNPDIGMFTYDESTKLFWFNPSSFETEGQFT LIGIVLGLAIYNNCILDVHFPMVVYRKLMGKK GTFRDLGDSHPVLYQSLKDLLEYEGNVEDDM MITFQISQTDLFGNPMMYDLKENGDKIPITNEN RKEFVNLYSDYILNKSVEKQFKAFRRGFHMVT NESPLKYLFRPEEIELLICGSRNLDFQALEETTE YDGGYTRDSVLIREFWEIVHSFTDEQKRLFLQF TTGTDRAPVGGLGKLKMIIAKNGPDTERLPTS HTCFNVLLLPEYSSKEKLKERLLKAITYAKGFG ML	Angelman Syndrome	138
SOD1	P00441	MATKAVCVLKGDGPVQGIINFEQKESNGPVKV WGSIKGLTEGLHGFHVHEFGDNTAGCTSAGPH FNPLSRKHGGPKDEERHVGDLGNVTADKDGV ADVSIEDSVISLSGDHCIIGRTLVVHEKADDLG KGGNEESTKTGNAGSRLACGVIGIAQ	Amyotrophic Lateral Sclerosis	139
TDP43	Q13148	MSEYIRVTEDENDEPIEIPSEDDGTVLLSTVTAQ FPGACGLRYRNPVSQCMRGVRLVEGILHAPDA GWGNLVYVVNYPKDNKRKMDETDASSAVKV KRAVQKTSDLIVLGLPWKTTEQDLKEYFSTFG EVLMVQVKKDLKTGHSKGFGFVRFTEYETQV KVMSQRHMIDGRWCDCKLPNSKQSQDEPLRS RKVFVGRCTEDMTEDELREFFSQYGDVMDVFI PKPFRAFAFVTFADDQIAQSLCGEDLIIKGISVH ISNAEPKHNSNRQLERSGRFGGNPGGFGNQGG FGNSRGGGAGLGNNQGSNMGGGMNFGAFSIN PAMMAAAQAALQSSWGMMGMLASQQNQSG PSGNNQNQGNMQREPNQAFGSGNNSYSGSNS GAAIGWGSASNAGSGSGFNGGFGSSMDSKSSG WGM	Amyotrophic Lateral Sclerosis	140
C9orf72	Q96LT7	MSTLCPPPSPAVAKTEIALSGKSPLLAATFAYW DNILGPRVRHIWAPKTEQVLLSDGEITFLANHT LNGEILRNAESGAIDVKFFVLSEKGVIIVSLIFD GNWNGDRSTYGLSIILPQTELSFYLPLHRVCVD RLTHIIRKGRIWMHKERQENVQKIILEGTERME	Amyotrophic Lateral Sclerosis	141

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		DQGQSIIPMLTGEVIPVMELLSSMKSHSVPEEID IADTVLNDDDIGDSCHEGFLLNAISSHLQTCGC SVVVGSSAEKVNKIVRTLCLFLTPAERKCSRLC EAESSFKYESGLFVQGLLKDSTGSFVLPFRQVM YAPYPTTHIDVDVNTVKQMPPCHEHIYNQRRY MRSELTAFWRATSEEDMAQDTIIYTDESFTPDL NIFQDVLHRDTLVKAFLDQVFQLKPGLSLRSTF LAQFLLVLHRKALTLIKYIEDDTQKGKKPFKSL RNLKIDLDLTAEGDLNIIMALAEKIKPGLHSFIF GRPFYTSVQERDVLMTF		
FXN	Q16595	MWTLGRRAVAGLLASPSPAQAQTLTRVPRPAE LAPLCGRRGLRTDIDATCTPRRASSNQRGLNQI WNVKKQSVYLMNLRKSGTLGHPGSLDETTYE RLAEETLDSLAEFFEDLADKPYTFEDYDVSFGS GVLTVKLGGDLGTYVINKQTPNKQIWLSSPSS GPKRYDWTGKNWVYSHDGVSLHELLAAELTK ALKTKLDLSSLAYSGKDA	Friedreich's Ataxia	142
MECP2	P51608	MVAGMLGLREEKSEDQDLQGLKDKPLKFKKV KKDKKEEKEGKHEPVQPSAHHSAEPAEAGKA ETSEGSGSAPAVPEASASPKQRRSIIRDRGPMY DDPTLPEGWTRKLKQRKSGRSAGKYDVYLINP QGKAFRSKVELIAYFEKVGDTSLDPNDFDFTV TGRGSPSRREQKPPKKPKSPKAPGTGRGRGRP KGSGTTRPKAATSEGVQVKRVLEKSPGKLLVK MPFQTSPGGKAEGGGATTSTQVMVIKRPGRKR KAEADPQAIPKKRGRKPGSVVAAAAAEAKKK AVKESSIRSVQETVLPIKKRKTRETVSIEVKEVV KPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKG RSSSASSPPKKEHHHHHHHSESPKAPVPLLPPL PPPPPEPESSEDPTSPPEPQDLSSSVCKEEKMPR GGSLESDGCPKEPAKTQPAVATAATAAEKYK HRGEGERKDIVSSSMPRPNREEPVDSRTPVTER VS	Rett Syndrome	143
ASPA	P45381	MTSCHIAEEHIQKVAIFGGTHGNELTGVFLVKH WLENGAEIQRTGLEVKPFITNPRAVKKCTRYID CDLNRIFDLENLGKKMSEDLPYEVRRAQEINH LFGPKDSEDSYDIIFDLHNTTSNMGCTLILEDSR NNFLIQMFHYIKTSLAPLPCYVYLIEHPSLKYA TTRSIAKYPVGIEVGPQPQGVLRADILDQMRK MIKHALDFIHHFNEGKEFPPCAIEVYKIIEKVDY PRDENGEIAAIIHPNLQDQDWKPLHPGDPMFLT LDGKTIPLGGDCTVYPVFVNEAAYYEKKEAFA KTTKLTLNAKSIRCCLH	Canavan Disease	144
ALDH7A1	P49419	MWRLPRALCVHAAKTSKLSGPWSRPAAFMST LLINQPQYAWLKELGLREENEGVYNGSWGGR GEVITTYCPANNEPIARVRQASVADYEETVKK AREAWKIWADIPAPKRGEIVRQIGDALREKIQV LGSLVSLEMGKILVEGVGEVQEYVDICDYAVG LSRMIGGPILPSERSGHALIEQWNPVGLVGIITA FNFPVAVYGWNNAIAMICGNVCLWKGAPTTS LISVAVTKIIAKVLEDNKLPGAICSLTCGGADIG TAMAKDERVNLLSFTGSTQVGKQVGLMVQER	Pyridoxine-Dependent Epilepsy	145

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ ID NO
		FGRSLLELGGNNAIIAFEDADLSLVVPSALFAA VGTAGQRCTTARRLFIHESIHDEVVNRLKKAY AQIRVGNPWDPNVLYGPLHTKQAVSMFLGAV EEAKKEGGTVVYGGKVMDRPGNYVEPTIVTG LGHDASIAHTETFAPILYVFKFKNEEEVFAWNN EVKQGLSSSIFTKDLGRIFRWLGPKGSDCGIVN VNIPTSGAEIGGAFGGEKHTGGGRESGSDAWK QYMRRSTCTINYSKDLPLAQGIKFQ		

In some embodiments, the target cell is a cell in the cerebellum, e.g., for the treatment of Spinocerebellar Ataxia, Autosomal Recessive, Type 1, e.g., when the exogenous agent is SYNE1. In some embodiments, the target cell is a cell in the brain, spinal cord, and/or muscles, e.g., for the treatment of Ataxia with Oculomotor Apraxia, Type 2, e.g., when the exogenous agent is SETX. In some embodiments, the target cell is a neuron or an astrocyte, e.g., for the treatment of Fragile X Syndrome, e.g., when the exogenous agent is FMR1. In some embodiments, the target cell is a motor neuron, e.g., for the treatment of Amyotrophic Lateral Sclerosis, e.g., when the exogenous agent is SOD1, TDP43, or C9orf72. In some embodiments, the target cell is a cell in the nervous system, heart, liver, and/or skeletal muscle, e.g., for the treatment of Friedreich's Ataxia, e.g., when the exogenous agent is FXN. In some embodiments, the target cell is a neuron in the brain, e.g., for the treatment of Rett Syndrome, e.g., when the exogenous agent is MECP2. In some embodiments, the target cell is an oligodendrocyte or a neuron, e.g., for the treatment of Canavan Disease, e.g., when the exogenous agent is ASPA.

In some embodiments, the exogenous agent comprises a protein of Table 6 below. In some embodiments, the exogenous agent comprises the wild-type human sequence of any of the proteins of Table 6, a functional fragment thereof (e.g., an enzymatically active fragment thereof), or a functional variant thereof. In some embodiments, the exogenous agent comprises an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence of Table 6, e.g., a Uniprot Protein Accession Number sequence of column 2 of Table 6 or an amino acid sequence of column 3 of Table 6. In some embodiments, the payload gene encodes an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence of Table 6.

Table 6. Lysosomal storage diseases or disorders and/or CNS diseases or disorders. The first column lists exogenous agents that can be delivered to treat the indications in the fourth column, according to the methods and uses herein. Each Uniprot accession number of Table 6 is herein incorporated by reference in its entirety.

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ IS NO
TPP1	O14773	MGLQACLLGLFALILSGKCSYSPEPDQRRTLP PGWVSLGRADPEEELSLTFALRQQNVERLSEL VQAVSDPSSPQYGKYLTLENVADLVRPSPLTL HTVQKWLLAAGAQKCHSVITQDFLTCWLSIR QAELLLPGAEFHHYVGGPTETHVVRSPHPYQ LPQALAPHVDFVGGLHRFPPTSSLRQRPEPQV TGTVGLHLGVTPSVIRKRYNLTSQDVGSGTS NNSQACAQFLEQYFHDSDLAQFMRLFGGNFA HQASVARVVGQQGRGRAGIEASLDVQYLMS AGANISTWVYSSPGRHEGQEPFLQWLMLLSN ESALPHVHTVSYGDDEDSLSSAYIQRVNTEL MKAAARGLTLLFASGDSGAGCWSVSGRHQF RPTFPASSPYVTTVGGTSFQEPFLITNEIVDYIS GGGFSNVFPRPSYQEEAVTKFLSSSPHLPPSSY FNASGRAYPDVAALSDGYWVVSNRVPIPWV SGTSASTPVFGGILSLINEHRILSGRPPLGFLNP RLYQQHGAGLFDVTRGCHESCLDEEVEGQGF CSGPGWDPVTGWGTPNFPALLKTLLNP	Batten Disease (CLN2)	146
FUCA1	P04066	MRAPGMRSRPAGPALLLLLLFLGAAESVRRA QPPRRYTPDWPSLDSRPLPAWFDEAKFGVFIH WGVFSVPAWGSEWFWWHWQGEGRPQYQRF MRDNYPPGFSYADFGPQFTARFFHPEEWADL FQAAGAKYVVLTTKHHEGFTNWPSPVSWNW NSKDVGPHRDLVGELGTALRKRNIRYGLYHS LLEWFHPLYLLDKKNGFKTQHFVSAKTMPEL YDLVNSYKPDLIWSDGEWECPDTYWNSTNFL SWLYNDSPVKDEVVVNDRWGQNCSCHHGG YYNCEDKFKPQSLPDHKWEMCTSIDKFSWG YRRDMALSDVTEESEIISELVQTVSLGGNYLL NIGPTKDGLIVPIFQERLLAVGKWLSINGEAIY ASKPWRVQWEKNTTSVWYTSKGSAVYAIFL HWPENGVLNLESPITTSTTKITMLGIQGDLKW STDPDKGLFISLPQLPPSAVPAEFAWTIKLTGV K	Fucosidosis	147
GALC	P54803	MAEWLLSASWQRRAKAMTAAAGSAGRAAV PLLLCALLAPGGAYVLDDSDGLGREFDGIGA VSGGGATSRLLVNYPEPYRSQILDYLFKPNFG ASLHILKVEIGGDGQTTDGTEPSHMHYALDE NYFRGYEWWLMKEAKKRNPNITLIGLPWSFP GWLGKGFDWPYVNLQLTAYYVVTWIVGAK RYHDLDIDYIGIWNERSYNANYIKILRKMLNY QGLQRVKIIASDNLWESISASMLLDAELFKVV DVIGAHYPGTHSAKDAKLTGKKLWSSEDFST LNSDMGAGCWGRILNQNYINGYMTSTIAWN	Krabbe Disease	148

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ IS NO
		LVASYYEQLPYGRCGLMTAQEPWSGHYVVE SPVWVSAHTTQFTQPGWYYLKTVGHLEKGG SYVALTDGLGNLTIIIETMSHKHSKCIRPFLPY FNVSQQFATFVLKGSFSEIPELQVWYTKLGKT SERFLFKQLDSLWLLDSDGSFTLSLHEDELFT LTTLTTGRKGSYPLPPKSQPFPSTYKDDFNVD YPFFSEAPNFADQTGVFEYFTNIEDPGEHHFT LRQVLNQRPITWAADASNTISIIGDYNWTNLT IKCDVYIETPDTGGVFIAGRVNKGGILIRSARG IFFWIFANGSYRVTGDLAGWIIYALGRVEVTA KKWYTLTLTIKGHFTSGMLNDKSLWTDIPVN FPKNGWAAIGTHSFEFAQFDNFLVEATR		
HEX A	P06865	MTSSRLWFSLLLAAAFAGRATALWPWPQNF QTSDQRYVLYPNNFQFQYDVSSAAQPGCSVL DEAFQRYRDLLFGSGSWPRPYLTGKRHTLEK NVLVVSVVTPGCNQLPTLESVENYTLTINDD QCLLLSETVWGALRGLETFSQLVWKSAEGTF FINKTEIEDFPRFPHRGLLLDTSRHYLPLSSILD TLDVMAYNKLNVFHWHLVDDPSFPYESFTFP ELMRKGSYNPVTHIYTAQDVKEVIEYARLRGI RVLAEFDTPGHTLSWGPGIPGLLTPCYSGSEP SGTFGPVNPSLNNTYEFMSTFFLEVSSVFPDF YLHLGGDEVDFTCWKSNPEIQDFMRKKGFGE DFKQLESFYIQTLLDIVSSYGKGYVVWQEVF DNKVKIQPDTIIQVWREDIPVNYMKELELVTK AGFRALLSAPWYLNRISYGPDWKDFYIVEPL AFEGTPEQKALVIGGEACMWGEYVDNTNLV PRLWPRAGAVAERLWSNKLTSDLTFAYERLS HFRCELLRRGVQAQPLNVGFCEQEFEQT	Tay Sachs Disease	149
HEX B	P07686	MELCGLGLPRPPMLLALLLATLLAAMLALLT QVALVVQVAEAARAPSVSAKPGPALWPLPLS VKMTPNLLHLAPENFYISHSPNSTAGPSCTLL EEAFRRYHGYIFGFYKWHHEPAEFQAKTQVQ QLLVSITLQSECDAFPNISSDESYTLLVKEPVA VLKANRVWGALRGLETFSQLVYQDSYGTFTI NESTIIDSPRFSHRGILIDTSRHYLPVKIILKTLD AMAFNKFNVLHWHIVDDQSFPYQSITFPELSN KGSYSLSHVYTPNDVRMVIEYARLRGIRVLPE FDTPGHTLSWGKGQKDLLTPCYSRQNKLDSF GPINPTLNTTYSFLTTFFKEISEVFPDQFIHLGG DEVEFKCWESNPKIQDFMRQKGFGTDFKKLE SFYIQKVLDIIATINKGSIVWQEVFDDKAKLAP GTIVEVWKDSAYPEELSRVTASGFPVILSAPW YLDLISYGQDWRKYYKVEPLDFGGTQKQKQ LFIGGEACLWGEYVDATNLTPRLWPRASAVG ERLWSSKDVRDMDDAYDRLTRHRCRMVER GIAAQPLYAGYCNHENM	Sandhoff Disease	150
MAN BA	O00462	MRLHLLLLLALCGAGTTAAELSYSLRGNWSI CNGNGSLELPGAVPGCVHSALFQQGLIQDSY YRFNDLNYRWVSLDNWTYSKEFKIPFEISKW QKVNLILEGVDTVSKILFNEVTIGETDNMFNR YSFDITNVVRDVNSIELRFQSAVLYAAQQSKA	Beta-mannosidosis	151

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ IS NO
		HTRYQVPPDCPPLVQKGECHVNFVRKEQCSF SWDWGPSFPTQGIWKDVRIEAYNICHLNYFT FSPIYDKSAQEWNLEIESTFDVVSSKPVGGQVI VAIPKLQTQQTYSIELQPGKRIVELFVNISKNIT VETWWPHGHGNQTGYNMTVLFELDGGLNIE KSAKVYFRTVELIEEPIKGSPGLSFYFKINGFPI FLKGSNWIPADSFQDRVTSELLRLLLQSVVDA NMNTLRVWGGGIYEQDEFYELCDELGIMVW QDFMFACALYPTDQGFLDSVTAEVAYQIKRL KSHPSIIIWSGNNENEEALMMNWYHISFTDRP IYIKDYVTLYVKNIRELVLAGDKSRPFITSSPT NGAETVAEAWVSQNPNSNYFGDVHFYDYIS DCWNWKVFPKARFASEYGYQSWPSFSTLEK VSSTEDWSFNSKFSLHRQHHEGGNKQMLYQ AGLHFKLPQSTDPLRTFKDTIYLTQVMQAQC VKTETEFYRRSRSEIVDQQGHTMGALYWQLN DIWQAPSWASLEYGGKWKMLHYFAQNFFAP LLPVGFENENTFYIYGVSDLHSDYSMTLSVRV HTWSSLEPVCSRVTERFVMKGGEAVCLYEEP VSELLRRCGNCTRESCVVSFYLSADHELLSPT NYHFLSSPKEAVGLCKAQITAIISQQGDIFVFD LETSAVAPFVWLDVGSIPGRFSDNGFLMTEKT RTILFYPWEPTSKNELEQSFHVTSLTDIY		
ARSA	P15289	MGAPRSLLLALAAGLAVARPPNIVLIFADDLG YGDLGCYGHPSSTTPNLDQLAAGGLRFTDFY VPVSLCTPSRAALLTGRLPVRMGMYPGVLVP SSRGGLPLEEVTVAEVLAARGYLTGMAGKW HLGVGPEGAFLPPHQGFHRFLGIPYSHDQGPC QNLTCFPPATPCDGGCDQGLVPIPLLANLSVE AQPPWLPGLEARYMAFAHDLMADAQRQDRP FFLYYASHHTHYPQFSGQSFAERSGRGPFGDS LMELDAAVGTLMTAIGDLGLLEETLVIFTAD NGPETMRMSRGGCSGLLRCGKGTTYEGGVR EPALAFWPGHIAPGVTHELASSLDLLPTLAAL AGAPLPNVTLDGFDLSPLLLGTGKSPRQSLFF YPSYPDEVRGVFAVRTGKYKAHFFTQGSAHS DTTADPACHASSSLTAHEPPLLYDLSKDPGEN YNLLGGVAGATPEVLQALKQLQLLKAQLDA AVTFGPSQVARGEDPALQICCHPGCTPRPACC HCPDPHA	Metachromatic Leukodystrophy	152
GNPTAB	Q3T906	MLFKLLQRQTYTCLSHRYGLYVCFLGVVVTI VSAFQFGEVVLEWSRDQYHVLFDSYRDNIAG KSFQNRLCLPMPIDVVYTWVNGTDLELLKEL QQVREQMEEEQKAMREILGKNTTEPTKKSEK QLECLLTHCIKVPMLVLDPALPANITLKDLPS LYPSFHSASDIFNVAKPKNPSTNVSVVVFDST KDVEDAHSGLLKGNSRQTVWRGYLTTDKEV PGLVLMQDLAFLSGFPPTFKETNQLKTKLPEN LSSKVKLLQLYSEASVALLKLNNPKDFQELN KQTKKNMTIDGKELTISPAYLLWDLSAISQSK QDEDISASRFEDNEELRYSLRSIERHAPWVRNI FIVTNGQIPSWLNLDNPRVTIVTHQDVFRNLS	Mucolipidosis Type IIIa; Mucolipidosis Type IIIb	153

Gene	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number)	Disease / Disorder	SEQ IS NO
		HLPTFSSPAIESHIHRIEGLSQKFIYLNDDVMF GKDVWPDDFYSHSKGQKVYLTWPVPNCAEG CPGSWIKDGYCDKACNNSACDWDGGDCSGN SGGSRYIAGGGGTGSIGVGQPWQFGGGINSVS YCNQGCANSWLADKFCDQACNVLSCGFDAG DCGQDHFHELYKVILLPNQTHYIIPKGECLPY FSFAEVAKRGVEGAYSDNPIIRHASIANKWKT IHLIMHSGMNATTIHFNLTFQNTNDEEFKMQI TVEVDTREGPKLNSTAQKGYENLVSPITLLPE AEILFEDIPKEKRFPKFKRHDVNSTRRAQEEV KIPLVNISLLPKDAQLSLNTLDLQLEHGDITLK GYNLSKSALLRSFLMNSQHAKIKNQAIITDET NDSLVAPQEKQVHKSILPNSLGVSERLQRLTF PAVSVKVNGHDQGQNPPLDLETTARFRVETH TQKTIGGNVTKEKPPSLIVPLESQMTKEKKIT GKEKENSRMEENAENHIGVTEVLLGRKLQHY TDSYLGFLPWEKKKYFQDLLDEEESLKTQLA YFTDSKNTGRQLKDTFADSLRYVNKILNSKF GFTSRKVPAHMPHMIDRIVMQELQDMFPEEF DKTSFHKVRHSEDMQFAFSYFYYLMSAVQPL NISQVFDEVDTDQSGVLSDREIRTLATRIHELP LSLQDLTGLEHMLINCSKMLPADITQLNNIPP TQESYYDPNLPPVTKSLVTNCKPVTDKIHKA YKDKNKYRFEIMGEEEIAFKMIRTNVSHVVG QLDDIRKNPRKFVCLNDNIDHNHKDAQTVKA VLRDFYESMFPIPSQFELPREYRNRFLHMHEL QEWRAYRDKLKFWTHCVLATLIMFTIFSFFA EQLIALKRKIFPRRRIHKEASPNRIRV		
MCOLN1	Q9GZU1	MTAPAGPRGSETERLLTPNPGYGTQAGPSPAP PTPPEEEDLRRRLKYFFMSPCDKFRAKGRKPC KLMLQVVKILVVTVQLILFGLSNQLAVTFREE NTIAFRHLFLLGYSDGADDTFAAYTREQLYQ AIFHAVDQYLALPDVSLGRYAYVRGGGDPW TNGSGLALCQRYYHRGHVDPANDTFDIDPM VVTDCIQVDPPERPPPPPSDDLTLLESSSSYKN LTLKFHKLVNVTIHFRLKTINLQSLINNEIPDC YTFSVLITFDNKAHSGRIPISLETQAHIQECKH PSVFQHGDNSFRLLFDVVVILTCSLSFLLCARS LLRGFLLQNEFVGFMWRQRGRVISLWERLEF VNGWYILLVTSDVLTISGTIMKIGIEAKNLAS YDVCSILLGTSTLLVWVGVIRYLTFFHNYNILI ATLRVALPSVMRFCCCVAVIYLGYCFCGWIV LGPYHVKFRSLSMVSECLFSLINGDDMFVTFA AMQAQQGRSSLVWLFSQLYLYSFISLFIYMVL SLFIALITGAYDTIKHPGGAGAEESELQAYIAQ CQDSPTSGKFRRGSGSACSLLCCCGRDPSEEH SLLVN	Mucolipidosis Type IV	154

In some embodiments, the target cell is an oligodendrocyte, e.g., for the treatment of Krabbe Disease, e.g., when the exogenous agent is GALC. In some embodiments, the target cell

is a neuron, e.g., for the treatment of Tay Sachs Disease, e.g., when the exogenous agent is HEXA. In some embodiments, the target cell is a neuron, e.g., for the treatment of Sandhoff Disease, e.g., when the exogenous agent is HEXB.

In some embodiments, the protein agent is other than a clotting factor, e.g., other than Factor VII or Factor IX. In some embodiments, the protein agent is other than a reporter protein, e.g., fluorescent protein, e.g., GFP or luciferase. In some the protein agent is other than a cell surface receptor, an NGF receptor, galactocerebrosidase, gp91 phox, IFN-alpha, TK, GCV, and autoimmune antigen, cytokine, angiogenesis inhibitor, or anti-cancer agent, or a fragment or variant thereof.

VII. Insulator elements

In some embodiments, a fusosome, retroviral or lentiviral vector, or VLP further comprises one or more insulator elements, e.g., an insulator element described herein. Insulators elements may contribute to protecting lentivirus- expressed sequences, e.g., therapeutic polypeptides, from integration site effects, which may be mediated by cis-acting elements present in genomic DNA and lead to deregulated expression of transferred sequences (e.g., position effect; see, e.g., Burgess- Beusse et al, 2002, Proc. Natl. Acad. Sci., USA, 99: 16433; and Zhan et al, 2001, Hum. Genet., 109:471) or deregulated expression of endogenous sequences adjacent to the transferred sequences. In some embodiments, transfer vectors comprise one or more insulator element the 3' LTR and upon integration of the provirus into the host genome, the provirus comprises the one or more insulators at the 5' LTR and/or 3' LTR, by virtue of duplicating the 3' LTR. Suitable insulators include, but are not limited to, the chicken β-globin insulator (see Chung et al, 1993. Cell 74:505; Chung et al, 1997. N4S 94:575; and Bell et al., 1999. Cell 98:387, incorporated by reference herein) or an insulator from a human β-globin locus, such as chicken HS4. In some embodiments the insulator binds CCCTC binding factor (CTCF). In some embodiments the insulator is a barrier insulator. In some embodiments the insulator is an enhancer-blocking insulator. See, e.g., Emery et al., *Human Gene Therapy*, 2011, and in Browning and Trobridge, Biomedicines, 2016, both of which are included in their entirety by reference.

In some embodiments, insulators in the retroviral nucleic acid reduce genotoxicity in recipient cells. Genotoxicity can be measured, e.g., as described in Cesana et al, "Uncovering

and dissecting the genotoxicity of self-inactivating lentiviral vectors in vivo" Mol Ther. 2014 Apr;22(4):774-85. doi: 10.1038/mt.2014.3. Epub 2014 Jan 20.

VIII. Assessing fusosome content of target cell

The present disclosure also provides, in some aspects, a method of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell) in a subject, comprising providing a biological sample from a subject that has received a fusosome composition (e.g., a fusosome composition described herein), and performing an assay to determine one or more properties of the biological sample resulting from fusion of a target cell in the biological sample with a fusosome as described herein. In some aspects, the disclosure provides a method of measuring fusion with a target cell, e.g., as described in Example 71. In some embodiments, determining one or more properties of the biological sample comprises determining: the presence of a fusogen, the level of a cargo or payload, or an activity relating to a cargo or payload.

In some aspects, the present disclosure provides a method of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell) in a subject, comprising providing a biological sample from a subject that has received a fusosome composition, e.g., as described herein, and testing the biological sample for the presence of a fusogen, e.g., a fusogen described herein. In some instances, the level of the fusogen detected is greater (e.g., at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, 2000%, 3000%, 4000%, 5000%, 10,000%, 50,000%, or 100,000% greater) than that observed in a corresponding biological sample from a subject that has not received a fusosome composition. In some embodiments, the subject is the same subject prior to administration of the fusosome composition, and in some embodiments, the subject is a different subject.

In some aspects, the present disclosure provides a method of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell) in a subject, comprising providing a biological sample from a subject that has received a fusosome composition, e.g., as described herein, and testing the biological sample for the presence of a cargo or payload, e.g., delivered by a fusosome as described herein. In some instances, the level of the cargo or payload detected is greater (e.g., at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, 2000%, 3000%, 4000%,

5000%, 10,000%, 50,000%, or 100,000% greater) than that observed in a corresponding biological sample from a subject that has not received a fusosome composition. In some embodiments, the subject is the same subject prior to administration of the fusosome composition, and in some embodiments, the subject is a different subject.

In some aspects, the present disclosure provides a method of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell in a subject), comprising providing a biological sample from a subject that has received a fusosome composition, e.g., as described herein, and testing the biological sample for alteration of an activity relating to the fusosome composition, e.g., an activity relating to a cargo or payload delivered by the fusosome composition. In some instances, the level of the activity detected is increased, e.g., by at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, 2000%, 3000%, 4000%, 5000%, 10,000%, 50,000%, or 100,000%, relative to that of a corresponding biological sample from a subject that has not received a fusosome composition (e.g., the same subject prior to administration of the fusosome composition). In some instances, the level of the activity detected is decreased, e.g., by at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000%, 2000%, 3000%, 4000%, 5000%, 10,000%, 50,000%, or 100,000%, relative to that of a corresponding biological sample from a subject that has not received a fusosome composition. In some embodiments, the subject is the same subject prior to administration of the fusosome composition, and in some embodiments, the subject is a different subject.

In one aspect, the present disclosure provides a method of assessing fusosome fusion to a target cell in a subject, comprising providing a biological sample from a subject that has received a fusosome composition, e.g., as described herein, and assessing a level of unfused fusosomes in the biological sample.

In some embodiments of the methods of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell), resulting in formation of a recipient cell, in the subject, the method futher comprises collecting the biological sample from the subject. In embodiments, the biological sample includes one or more recipient cells.

In some embodiments of the methods of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell) in the subject, the method futher comprises separating recipient

cells in the biological sample from unfused fusosomes in the biological sample, e.g., by centrifugation. In some embodiments, the method futher comprises enriching recipient cells relative to unfused fusosomes in the biological sample, e.g., by centrifugation. In some embodiments, the method further comprises enriching target cells relative to non-target cells in the biological sample, e.g., by FACS.

In some embodiments of the methods of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell) in a subject, the activity relating to the fusosome composition is chosen from the presence or level of a metabolite, the presence or level of a biomarker (e.g., a protein level or post-translational modification, e.g., phosphorylation or cleavage).

In some embodiments of the methods of assessing fusosome content of a target cell (e.g., fusosome fusion to a target cell) in a subject, the activity relating to the fusosome composition is immunogenicity. In embodiments, the target cell is a CD3+ cell and the biological sample is a blood sample collected from the subject. In embodiments, blood cells are enriched from the blood sample, e.g., using a buffered ammonium chloride solution. In embodiments, enriched blood cells are incubated with an anti-CD3 antibody (e.g., a murine anti-CD3-FITC antibody) and CD3+ cells are selected, e.g., by fluorescence activated cell sorting. In embodiments, cells, e.g., sorted cells, e.g., CD3+ cells are analyzed for the presence of antibodies on the cell surface, e.g., by staining with an anti-IgM antibody. In some embodiments, if antibodies are present at a level above a reference level, the subject is identified as having an immune response against recipient cells.

In embodiments, immunogenicity is assayed by a cell lysis assay. In embodiments, recipient cells from the biological sample are co-incubated with immune effector cells capable of lysing other cells. In embodiments, the immune effector cells are from the subject or from a subject not administered the fusosome composition. For instance, in embodiments, immunogenicity is assessed by a PBMC cell lysis assay. In embodiments, recipient cells from the biological sample are co-incubated with peripheral blood mononuclear cells (PBMCs) from the subject or control PBMCs from a subject not administered the fusosome composition and then assessed for lysis of the recipient cells by PBMCs. In embodiments, immunogenicity is assessed by a natural killer (NK) cell lysis assay. In embodiments, recipient cells are co-incubated with NK cells from the subject or control NK cells from a subject not administered the fusosome composition and then assessed for lysis of the recipient cells by the NK cells. In

embodiments, immunogenicity is assessed by a CD8+ T-cell lysis assay. In embodiments, recipient cells are co-incubated with CD8+ T-cells from the subject or control CD8+ T-cells from a subject not administered the fusosome composition and then assessed for lysis of the target cells by the CD8+ T-cells. In some embodiments, if cell lysis occurs at a level above a reference level, the subject is identified as having an immune response against recipient cells.

In some embodiments, immunogenicity is assayed by phagocytosis of recipient cells, e.g., by macrophages. In embodiments, recipient cellsare not targeted by macrophages for phagocytosis. In embodiments, the biological sample is a blood sample collected from the subject. In embodiments, blood cells are enriched from the blood sample, e.g., using a buffered ammonium chloride solution. In embodiments, enriched blood cells are incubated with an anti-CD3 antibody (e.g., a murine anti-CD3-FITC antibody) and CD3+ cells are selected, e.g., by fluorescence activated cell sorting. In embodiments, fluorescently-labeled CD3+ cells are incubated with macrophages and then tested for intracellular fluorescence within the macrophages, e.g., by flow cytometry. In some embodiments, if macrophage phagocytosis occurs at a level above a reference level, the subject is identified as having an immune response against recipient cells.

IX. Physical and functional characteristics of fusosomes

In some embodiments, the fusosome is capable of delivering (e.g., delivers) an agent, e.g., a protein, nucleic acid (e.g., mRNA), organelle, or metabolite to the cytosol of a target cell. Similarly, in some embodiments, a method herein comprises delivering an agent to the cytosol of a target cell. In some embodiments, the agent is a protein (or a nucleic acid encoding the protein, e.g., an mRNA encoding the protein) which is absent, mutant, or at a lower level than wild-type in the target cell. In some embodiments, the target cell is from a subject having a genetic disease, e.g., a monogenic disease, e.g., a monogenic intracellular protein disease. In some embodiments, the agent comprises a transcription factor, e.g., an exogenous transcription factor or an endogenous transcription factor. In some embodiments, the fusosome further comprises, or the method further comprises delivering, one or more (e.g., at least 2, 3, 4, 5, 10, 20, or 50) additional transcription factors, e.g., exogenous transcription factors, endogenous transcription factors, or a combination thereof.

In some embodiments, the fusosome comprises (e.g., is capable of delivering to the target cell) a plurality of agents (e.g., at least 2, 3, 4, 5, 10, 20, or 50 agents), wherein each agent of the plurality acts on a step of a pathway in the target cell, e.g., wherein the pathway is a biosynthetic pathway, a catabolic pathway, or a signal transduction cascade. In embodiments, each agent in the plurality upregulates the pathway or downregulates the pathway. In some embodiments, the fusosome further comprises, or the method further comprises delivering, one more additional agents (e.g., comprises a second plurality of agents) that do not act on a step of the pathway, e.g., that act on a step of a second pathway. In some embodiments, the fusosome comprises (e.g., is capable of delivering to the target cell), or the method further comprises delivering, a plurality of agents (e.g., at least 2, 3, 4, 5, 10, 20, or 50 agents), wherein each agent of the plurality is part of a single pathway, e.g., wherein the pathway is a biosynthetic pathway, a catabolic pathway, or a signal transduction cascade. In some embodiments, the fusosome further comprises, or the method further comprises delivering, one more additional agents (e.g., comprises a second plurality of agents) that are not part of the single pathway, e.g., are part of a second pathway.

In some embodiments, the target cell comprises an aggregated or misfolded protein. In some embodiments, the fusosome is capable of reducing levels (e.g., reduces levels) of the aggregated or misfolded protein in the target cell, or a method herein comprises reducing levels of the aggregated or misfolded protein in the target cell.

In some embodiments, the agent is selected from a transcription factor, enzyme (e.g., nuclear enzyme or cytosolic enzyme), reagent that mediates a sequence specific modification to DNA (e.g., Cas9, ZFN, or TALEN), mRNA (e.g., mRNA encoding an intracellular protein), organelle, or metabolite.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) an agent, e.g., a protein, to the cell membrane of a target cell. Similarly, in some embodiments, a method herein comprises delivering an agent to the cell membrane of a target cell. In some embodiments, delivering the protein comprises delivering a nucleic acid (e.g., mRNA) encoding the protein to the target cell such that the target cell produces the protein and localizes it to the membrane. In some embodiments, the fusosome comprises, or the method further comprises delivering, the protein, and fusion of the fusosome with the target cell transfers the protein to the cell membrane of the target cell. In some embodiments, the agent comprises a cell surface ligand or an antibody that binds a cell surface receptor. In some embodiments, the fusosome further comprises, or the

method further comprises delivering, a second agent that comprises or encodes a second cell surface ligand or antibody that binds a cell surface receptor, and optionally further comprising or encoding one or more additional cell surface ligands or antibodies that bind a cell surface receptor (e.g., 1, 2, 3, 4, 5, 10, 20, 50, or more). In some embodiments, the first agent and the second agent form a complex, wherein optionally the complex further comprises one or more additional cell surface ligands. In some embodiments, the agent comprises or encodes a cell surface receptor, e.g., an exogenous cell surface receptor. In some embodiments, the fusosome further comprises, or the method further comprises delivering, a second agent that comprises or encodes a second cell surface receptor, and optionally further comprises or encodes one or more additional cell surface receptors (e.g., 1, 2, 3, 4, 5, 10, 20, 50, or more cell surface receptors).

In some embodiments, the first agent and the second agent form a complex, wherein optionally the complex further comprises one or more additional cell surface receptors. In some embodiments, the agent comprises or encodes an antigen or an antigen presenting protein.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) a secreted agent, e.g., a secreted protein to a target site (e.g., an extracellular region), e.g., by delivering a nucleic acid (e.g., mRNA) encoding the protein to the target cell under conditions that allow the target cell to produce and secrete the protein. Similarly, in some embodiments, a method herein comprises delivering a secreted agent as described herein. In embodiments, the secreted protein is endogenous or exogenous. In embodiments, the secreted protein comprises a protein therapeutic, e.g., an antibody molecule, a cytokine, or an enzyme. In embodiments, the secreted protein comprises an autocrine signalling molecule or a paracrine signalling molecule. In embodiments, the secreted agent comprises a secretory granule.

In some embodiments, the fusosome is capable of reprogramming (e.g., reprograms) a target cell (e.g., an immune cell), e.g., by delivering an agent selected from a transcription factor or mRNA, or a plurality of said agents. Similarly, in some embodiments, a method herein comprises reprogramming a target cell. In embodiments, reprogramming comprises inducing a pancreatic endocrine cell to take on one or more characteristics of a pancreatic beta cell, by inducing a non-dopaminergic neuron to take on one or more characteristics of a dopaminergic neuron, or by inducing an exhausted T cell to take on one or more characteristics of a non-exhausted T cell, e.g., a killer T cell. In some embodiments, the agent comprises an antigen. In

some embodiments, the fusosome comprises a first agent comprising an antigen and a second agent comprising an antigen presenting protein.

In some embodiments, the fusosome is capable of donating (e.g., donates) one or more cell surface receptors to a target cell (e.g., an immune cell). Similarly, in some embodiments, a method herein comprises donating one or more cell surface receptors.

In some embodiments, a fusosome is capable of modifying, e.g., modifies, a target tumor cell. Similarly, in some embodiments, a method herein comprises modifying a target tumor cell. In embodiments, the fusosome comprises an mRNA encoding an immunostimulatory ligand, an antigen presenting protein, a tumor suppressor protein, or a pro-apoptotic protein. In some embodiments, the fusosome comprises an miRNA capable of reducing levels in a target cell of an immunosuppressive ligand, a mitogenic signal, or a growth factor.

In some embodiments, a fusosome comprises an agent that is immunomodulatory, e.g., immunostimulatory.

In some embodiments, a fusosome is capable of causing (e.g., causes) the target cell to present an antigen. Similarly, in some embodiments, a method herein comprises presenting an antigen on a target cell.

In some embodiments, the fusosome promotes regeneration in a target tissue. Similarly, in some embodiments, a method herein comprises promoting regeneration in a target tissue. In embodiments, the target cell is a cardiac cell, e.g., a cardiomyocyte (e.g., a quiescent cardiomyocyte), a hepatoblast (e.g., a bile duct hepatoblast), an epithelial cell, a naïve T cell, a macrophage (e.g., a tumor infiltrating macrophage), or a fibroblast (e.g., a cardiac fibroblast). In embodiments, the source cell is a T cell (e.g., a Treg), a macrophage, or a cardiac myocyte.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) a nucleic acid to a target cell, e.g., to stably modify the genome of the target cell, e.g., for gene therapy. Similarly, in some embodiments, a method herein comprises delivering a nucleic acid to a target cell. In some embodiments, the target cell has an enzyme deficiency, e.g., comprises a mutation in an enzyme leading to reduced activity (e.g., no activity) of the enzyme.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) a reagent that mediates a sequence specific modification to DNA (e.g., Cas9, ZFN, or TALEN) in the target cell. Similarly, in some embodiments, a method herein comprises delivering the reagent to the target cell. In embodiments, the target cell is a CNS cell.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) a nucleic acid to a target cell, e.g., to transiently modify gene expression in the target cell.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) a protein to a target cell, e.g., to transiently rescue a protein deficiency. Similarly, in some embodiments, a method herein comprises delivering a protein to a target cell. In embodiments, the protein is a membrane protein (e.g., a membrane transporter protein), a cytoplasmic protein (e.g., an enzyme), or a secreted protein (e.g., an immunosuppressive protein).

In some embodiments, the fusosome is capable of delivering (e.g., delivers) an organelle to a target cell, e.g., wherein the target cell has a defective organelle network. Similarly, in some embodiments, a method herein comprises delivering an organelle to a target cell. In embodiments, the source cell is a hepatocyte, skeletal muscle cell, or neuron.

In some embodiments, the fusosome is capable of delivering (e.g., delivers) a nucleus to a target cell, e.g., wherein the target cell has a genetic mutation. Similarly, in some embodiments, a method herein comprises delivering a nucleus to a target cell. In some embodiments, the nucleus is autologous and comprises one or more genetic changes relative to the target cell, e.g., it comprises a sequence specific modification to DNA (e.g., Cas9, ZFN, or TALEN), or an artificial chromosome, an additional genetic sequence integrated into the genome, a deletion, or any combination thereof. In embodiments, the source of the autologous nucleus is a stem cell, e.g., a hematopoietic stem cell. In embodiments, the target cell is a muscle cell (e.g., a skeletal muscle cell or cardiomyocyte), a hepatocyte, or a neuron.

In some embodiments, the fusosome is capable of intracellular molecular delivery, e.g., delivers a protein agent to a target cell. Similarly, in some embodiments, a method herein comprises delivering a molecule to an intracellular region of a target cell. In embodiments, the protein agent is an inhibitor. In embodiments, the protein agent comprises a nanobody, scFv, camelid antibody, peptide, macrocycle, or small molecule.

In some embodiments, the fusosome is capable of causing (e.g., causes) a target cell to secrete a protein, e.g., a therapeutic protein. Similarly, in some embodiments, a method herein comprises causing a target cell to secrete a protein.

In some embodiments, the fusosome is capable of secreting (e.g., secretes) an agent, e.g., a protein. In some embodiments, the agent, e.g., secreted agent, is delivered to a target site in a subject. In some embodiments, the agent is a protein that can not be made recombinantly or is

difficult to make recombinantly. In some embodiments, the fusosome that secretes a protein is from a source cell selected from an MSC or a chondrocyte.

In some embodiments, the fusosome comprises on its membrane one or more cell surface ligands (e.g., 1, 2, 3, 4, 5, 10, 20, 50, or more cell surface ligands). Similarly, in some embodiments, a method herein comprises presenting one or more cell surface ligands to a target cell. In some embodiments, the fusosome having a cell surface ligand is from a source cell chosen from a neutrophil (e.g., and the target cell is a tumor-infiltrating lymphocyte), dendritic cell (e.g., and the target cell is a naïve T cell), or neutrophil (e.g., and the target is a tumor cell or virus-infected cell). In some embodiments the fusosome comprises a membrane complex, e.g., a complex comprising at least 2, 3, 4, or 5 proteins, e.g., a homodimer, heterodimer, homotrimer, heterotrimer, homotetramer, or heterotetramer. In some embodiments, the fusosome comprises an antibody, e.g., a toxic antibody, e.g., the fusosome is capable of delivering the antibody to the target site, e.g., by homing to a target site. In some embodiments, the source cell is an NK cell or neutrophil.

In some embodiments, a method herein comprises causing secretion of a protein from a target cell or ligand presentation on the surface of a target cell. In some embodiments, the fusosome is capable of causing cell death of the target cell. In some embodiments, the fusosome is from a NK source cell.

In some embodiments, a fusosome or target cell is capable of phagocytosis (e.g., of a pathogen). Similarly, in some embodiments, a method herein comprises causing phagocytosis.

In some embodiments, a fusosome senses and responds to its local environment. In some embodiments, the fusosome is capable of sensing level of a metabolite, interleukin, or antigen.

In embodiments, a fusosome is capable of chemotaxis, extravasation, or one or more metabolic activities. In embodiments, the metabolic activity is selected from kyneurinine, gluconeogenesis, prostaglandin fatty acid oxidation, adenosine metabolism, urea cycle, and thermogenic respiration. In some embodiments, the source cell is a neutrophil and the fusosome is capable of homing to a site of injury. In some embodiments, the source cell is a macrophage and the fusosome is capable of phagocytosis. In some embodiments, the source cell is a brown adipose tissue cell and the fusosome is capable of lipolysis.

In some embodiments, the fusosome comprises (e.g., is capable of delivering to the target cell) a plurality of agents (e.g., at least 2, 3, 4, 5, 10, 20, or 50 agents). In embodiments, the fusosome comprises an inhibitory nucleic acid (e.g., siRNA or miRNA) and an mRNA.

In some embodiments, the fusosome comprises (e.g., is capable of delivering to the target cell) a membrane protein or a nucleic acid encoding the membrane protein. In embodiments, the fusosome is capable of reprogramming or transdifferentiating a target cell, e.g., the fusosome comprises one or more agents that induce reprogramming or transdifferentiation of a target cell.

In some embodiments, the subject is in need of regeneration. In some embodiments, the subject suffers from cancer, an autoimmune disease, an infectious disease, a metabolic disease, a neurodegenerative disease, or a genetic disease (e.g., enzyme deficiency).

In some embodiments (e.g., embodiments for assaying non-endocytic delivery of cargo) cargo delivery is assayed using one or more of (e.g., all of) the following steps: (a) placing 30,000 HEK-293T target cells into a first well of a 96-well plate comprising 100 nM bafilomycin A1, and placing a similar number of similar cells into a second well of a 96-well plate lacking bafilomycin A1, (b) culturing the target cells for four hours in DMEM media at 37°C and 5% CO2, (c) contacting the target cells with 10 ug of fusosomes that comprise cargo, (d) incubating the target cells and fusosomes for 24 hrs at 37°C and 5% CO2, and (e) determining the percentage of cells in the first well and in the second well that comprise the cargo. Step (e) may comprise detecting the cargo using microscopy, e.g., using immunofluorescence. Step (e) may comprise detecting the cargo indirectly, e.g., detecting a downstream effect of the cargo, e.g., presence of a reporter protein. In some embodiments, one or more of steps (a)-(e) above is performed as described in Example 80.

In some embodiments, an inhibitor of endocytosis (e.g., chloroquine or bafilomycin A1) inhibits inhibits endosomal acidification. In some embodiments, cargo delivery is independent of lysosomal acidification. In some embodiments, an inhibitor of endocytosis (e.g., Dynasore) inhibits dynamin. In some embodiments, cargo delivery is independent of dynamin activity.

In some embodiments (e.g., embodiments for specific delivery of cargo to a target cell versus a non-target cell), cargo delivery is assayed using one or more of (e.g., all of) the following steps: (a) placing 30,000 HEK-293T target cells that over-express CD8a and CD8b into a first well of a 96-well plate and placing 30,000 HEK-293T non-target cells that do not over-express CD8a and CD8b into a second well of a 96-well plate, (b) culturing the cells for

four hours in DMEM media at 37ºC and 5% CO2, (c) contacting the target cells with 10 ug of fusosomes that comprise cargo, (d) incubating the target cells and fusosomes for 24 hrs at 37ºC and 5% CO2, and (e) determining the percentage of cells in the first well and in the second well that comprise the cargo. Step (e) may comprise detecting the cargo using microscopy, e.g., using immunofluorescence. Step (e) may comprise detecting the cargo indirectly, e.g., detecting a downstream effect of the cargo, e.g., presence of a reporter protein. In some embodiments, one or more of steps (a)-(e) above is performed as described in Example 71.

In some embodiments, the fusosome fuses at a higher rate with a target cell than with a non-target cell, e.g., by at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold, e.g., in an assay of Example 42 In some embodiments, the fusosome fuses at a higher rate with a target cell than with other fusosomes, e.g., by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, e.g., in an assay of Example 42. In some embodiments, the fusosome fuses with target cells at a rate such that an agent in the fusosome is delivered to at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, of target cells after 24, 48, or 72 hours, e.g., in an assay of Example 42. In embodiments, the amount of targeted fusion is about 30%-70%, 35%-65%, 40%-60%, 45%-55%, or 45%-50%, e.g., about 48.8% e.g., in an assay of Example 42. In embodiments, the amount of targeted fusion is about 20%-40%, 25%-35%, or 30%-35%, e.g., about 32.2% e.g., in an assay of Example 43.

In some embodiments, the fusogen is present at a copy number of at least, or no more than, 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies, e.g., as measured by an assay of Example 26. In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the fusogen comprised by the fusosome is disposed in the cell membrane. In embodiments, the fusosome also comprises fusogen internally, e.g., in the cytoplasm or an organelle. In some embodiments, the fusogen comprises (or is identified as comprising) about 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1%, 5%, 10%, 11%, 12%, 13%, 14%, 15%, 20%, or more, or about 1-30%, 5-20%, 10-15%, 12-15%, 13-14%, or 13.6% of the total protein in a fusosome, e.g., as determined according to the method described in Example 94 and/or by a mass spectrometry assay. In embodiments, the fusogen comprises(or is identified as comprising) about 13.6% of the total

protein in the fusosome. In some embodiments, the fusogen is (or is identified as being) more or less abundant than one or more additional proteins of interest, e.g., as determined according to the method described in Example 94. In an embodiment, the fusogen has (or is identified as having) a ratio to EGFP of about 140, 145, 150, 151, 152, 153, 154, 155, 156, 157 (e.g., 156.9), 158, 159, 160, 165, or 170. In another embodiment, the fusogen has (or is identified as having) a ratio to CD63 of about 2700, 2800, 2900, 2910 (e.g., 2912), 2920, 2930, 2940, 2950, 2960, 2970, 2980, 2990, or 3000, or about 1000-5000, 2000-4000, 2500-3500, 2900-2930, 2910-2915, or 2912.0, e.g., by a mass spectrometry assay. In an embodiment, the fusogen has (or is identified as having) a ratio to ARRDC1 of about 600, 610, 620, 630, 640, 650, 660 (e.g., 664.9), 670, 680, 690, or 700. In another embodiment, the fusogen has (or is identified as having) a ratio to GAPDH of about 50, 55, 60, 65, 70 (e.g., 69), 75, 80, or 85, or about 1-30%, 5-20%, 10-15%, 12-15%, 13-14%, or 13.6%. In another embodiment, the fusogen has (or is identified as having) a ratio to CNX of about 500, 510, 520, 530, 540, 550, 560 (e.g., 558.4), 570, 580, 590, or 600, or about 300-800, 400-700, 500-600, 520-590, 530-580, 540-570, 550-560, or 558.4, e.g., by a mass spectrometry assay.

In some embodiments, the fusosome comprises a therapeutic agent at a copy number of at least, or no more than, 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies, e.g., as measured by an assay of Example 88. In some embodiments, the fusosome comprises a protein therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies, e.g., as measured by an assay of Example 88. In some embodiments, the fusosome comprises a nucleic acid therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the fusosome comprises a DNA therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the fusosome comprises an RNA therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000,

10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the fusosome comprises an exogenous therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the fusosome comprises an exogenous protein therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the fusosome comprises an exogenous nucleic acid (e.g., DNA or RNA) therapeutic agent at a copy number of at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies. In some embodiments, the ratio of the copy number of the fusogen to the copy number of the therapeutic agent is between 1,000,000:1 and 100,000:1, 100,000:1 and 10,000:1, 10,000:1 and 1,000:1, 1,000:1 and 100:1, 100:1 and 50:1, 50:1 and 20:1, 20:1 and 10:1, 10:1 and 5:1, 5:1 and 2:1, 2:1 and 1:1, 1:1 and 1:2, 1:2 and 1:5, 1:5 and 1:10, 1:10 and 1:20, 1:20 and 1:50, 1:50 and 1:100, 1:100 and 1:1,000, 1:1,000 and 1:10,000, 1:10,000 and 1:100,000, or 1:100,000 and 1:1,000,000.

In some embodiments, the fusosome delivers to a target cell at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies of a therapeutic agent. In some embodiments, the fusosome delivers to a target cell at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies of a protein therapeutic agent. In some embodiments, the fusosome delivers to a target cell at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies of a nucleic acid therapeutic agent. In some embodiments, the fusosome delivers to a target cell at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies of an RNA therapeutic agent. In some embodiments, the fusosome delivers to a target cell at least 10, 50, 100, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, 100,000, 200,000,

500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000, 100,000,000, 500,000,000, or 1,000,000,000 copies of a DNA therapeutic agent.

In some embodiments, the fusosome delivers to a target cell at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo (e.g., a therapeutic agent, e.g., an endogenous therapeutic agent or an exogenous therapeutic agent) comprised by the fusosome. In some embodiments, the fusosomes that fuse with the target cell(s) deliver to the target cell an average of at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo (e.g., a therapeutic agent, e.g., an endogenous therapeutic agent or an exogenous therapeutic agent) comprised by the fusosomes that fuse with the target cell(s). In some embodiments, the fusosome composition delivers to a target tissue at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% of the cargo (e.g., a therapeutic agent, e.g., an endogenous therapeutic agent or an exogenous therapeutic agent) comprised by the fusosome composition.

In some embodiments, the fusosome comprises 0.00000001 mg fusogen to 1 mg fusogen per mg of total protein in fusosome, e.g., 0.00000001 - 0.0000001, 0.0000001 - 0.000001, 0.000001 - 0.00001, 0.00001 - 0.0001, 0.0001 - 0.001, 0.001 - 0.01, 0.01 - 0.1, or 0.1 – 1 mg fusogen per mg of total protein in fusosome. In some embodiments, the fusosome comprises 0.00000001 mg fusogen to 5 mg fusogen per mg of lipid in fusosome, e.g., 0.00000001 - 0.0000001, 0.0000001 - 0.000001, 0.000001 - 0.00001, 0.00001 - 0.0001, 0.0001 - 0.001, 0.001 - 0.01, 0.01 - 0.1, 0.1 – 1, or 1-5 mg fusogen per mg of lipid in fusosome.

In some embodiments, the cargo is a protein cargo. In embodiments, the cargo is an endogenous or synthetic protein cargo. In some embodiments, the fusosomes have (or are identified as having) at least 1, 2, 3, 4, 5, 10, 20, 50, 100, or more protein cargo molecules per fusosome. In an embodiment, the fusosomes have (or are identified as having) about 100, 110, 120, 130, 140, 150, 160, 166, 170, 180, 190, or 200 protein agent molecules per fusosome, e.g., as quantified according to the method described in Example 88. In some embodiments, the endogenous or synthetic protein cargo comprises (or is identified as comprising) about 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 1%, 5%, 10%, 15%, 20%, 25% or more of the total protein in a fusosome. In an embodiment, the synthetic protein cargo comprises (or is identified as comprising) about 13.6% of the total protein in a fusosome. In some embodiments, the synthetic protein cargo has (or is identified as having) a ratio to VSV-G of about 4×10^{-3}, 5×10^{-3}, 6×10^{-3}

(e.g., 6.37 x 10^{-3}), 7 x 10^{-3}, or 8 x 10^{-3}. In embodiments, the synthetic protein cargo has (or is identified as having) a ratio to CD63 of about 10, 15, 16, 17, 18 (e.g., 18.6), 19, 20, 25, or 30, or about 10-30, 15-25, 16-19, 18-19, or 18.6. In embodiments, the synthetic protein cargo has (or is identified as having) a ratio to ARRDC1 of about 2, 3, 4 (e.g., 4.24), 5, 6, or 7. In embodiments, the synthetic protein cargo has (or is identified as having) a ratio to GAPDH of about 0.1, 0.2, 0.3, 0.4 (e.g., 0.44), 0.5, 0.6, or 0.7. In embodiments, the synthetic protein cargo has (or is identified as having) a ratio to CNX of about 1, 2, 3 (e.g., 3.56), 4, 5, or 6. In embodiments, the synthetic protein cargo has (or is identified as having) a ratio to TSG101 of about 10, 15, 16, 17, 18, 19 (e.g., 19.52), 20, 21, 22, 23, 24, 25, or 30.

In some embodiments, the fusogen comprises (or is identified as comprising) at least 0.5%, 1%, 5%, 10%, or more of the total protein in a fusosome, e.g., by a mass spectrometry assay. In an embodiment, the fusogen comprises (or is identified as comprising) about1-30%, 5-20%, 10-15%, 12-15%, 13-14%, or 13.6% of the total protein in a fusosome, e.g., by a mass spectrometry assay. In some embodiments, the fusogen is more abundant than other proteins of interest. In embodiments, the fusogen has (or is identified as having) a ratio to a payload protein, e.g., EGFP, of about 145-170, 150-165, 155-160, 156.9, e.g., by a mass spectrometry assay. In embodiments, the fusogen has(or is identified as having) a ratio to CD63 of about 1000-5000, 2000-4000, 2500-3500, 2900-2930, 2910-2915, or 2912.0, e.g., by a mass spectrometry assay. In embodiments, the fusogen has a ratio to ARRDC1 of about 300-1000, 400-900, 500-800, 600-700, 640-690, 650-680, 660-670, or 664.9, e.g., by a mass spectrometry assay. In embodiments, the fusogen has(or is identified as having) a ratio to GAPDH of about 20-120, 40-100, 50-90, 60-80, 65-75, 68-70, or 69.0, e.g., by a mass spectrometry assay. In embodiments, the fusogen has a ratio to CNX of about 200-900, 300-800, 400-700, 500-600, 520-590, 530-580, 540-570, 550-560, or 558.4, e.g., by a mass spectrometry assay. In embodiments, the mass spectrometry essay is an assay of Example 94.

In some embodiments, the number of lipid species present in both of (e.g., shared between) the fusosomes and source cells is (or is identified as being) at least 300, 400, 500, 550, 560, or 569, or is between 500-700, 550-600, or 560-580, e.g., using a mass spectrometry assay. In embodiments, the number of lipid species present in fusosomes at a level at least 25% of the corresponding lipid level in the source cells (both normalized to total lipid levels within a sample) is (or is identified as being) at least 300, 400, 500, 530, 540, or 548, or is between 400-

700, 500-600, 520-570, 530-560, or 540-550, e.g., using a mass spectrometry assay. In some embodiments, the fraction of lipid species present in both of (e.g., shared between) the fusosomes and source cells to total lipid species in the source cell is (or is identified as being) about 0.4-1.0, 0.5-0.9, 0.6-0.8, or 0.7, or at least 0.4, 0.5, 0.6, or 0.7, e.g., using a mass spectrometry assay. In some embodiments, the mass spectrometry assay is an assay of Example 86.

In some embodiments, the number of protein species present in both of (e.g., shared between) the fusosomes are source cells is (or is identified as being) at least 500, 1000, 1100, 1200, 1300, 1400, 1487, 1500, or 1600, or is (or is identified as being) between 1200-1700, 1300-1600, 1400-1500, 1450-1500, or 1480-1490, e.g., using a mass spectrometry assay. In embodiments, the number of protein species present in fusosomes at a level at least 25% of the corresponding protein level in the source cells (both normalized to total protein levels within a sample) is (or is identified as being) at least 500, 600, 700, 800, 900, 950, 957, 1000, or 1200, e.g., using a mass spectrometry assay. In some embodiments, the fraction of protein species present in both of (e.g., shared between) the fusosomes and source cells to total protein species in the source cell is (or is identified as being) about 0.1-0.6, 0.2-0.5, 0.3-0.4, or 0.333, or at least about 0.1, 0.2, 0.3, 0.333, or 0.4, e.g., using a mass spectrometry assay. In embodiments, the mass spectrometry assay is an assay of Example 87.

In some embodiments, CD63 is (or is identified as being) present at less than 0.048%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, or 10% the amount of total protein in fusosomes, e.g., by a mass spectrometry assay, e.g., an assay of Example 89.

In some embodiments, the fusosomes are produced by extrusion through a filter, e.g., a filter of about 1-10, 2-8, 3-7, 4-6, or 5 um. In some embodiments, the fusosomes have (or is identified as having) an average diameter of about 1-5, 2-5, 3-5, 4-5, or 5 um. In some embodiments, the fusosomes have (or is identified as having) an average diameter of at least 1, 2, 3, 4, or 5 um.

In some embodiments, the fusosomes are enriched for (or are identified as being enriched for) one or more of (e.g., at least 2, 3, 4, 5, or all of) the following lipids compared to the source cells: cholesteryl ester, free cholesterol, ether-linked lyso-phosphatidylethanolamine, lyso-phosphatidylserine, phosphatidate, ether-linked phosphatidylethanolamine, phosphatidylserine, and sphingomyelin. In some embodiments, the fusosomes are depleted for (or are identified as

being depleted for) one or more of (e.g., at least 2, 3, 4, 5, or all of) the following lipids compared to the source cells: ceramide, cardiolipin, lyso-phosphatidylcholine, lyso-phosphatidylethanolamine, lyso-phosphatidylglycerol, lyso-phosphatidylinositol, ether-linked phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol, phosphatidylinositol, and triacylglycerol. In some embodiments, the fusosomes are enriched for (or are identified as being enriched for) one or more of the aforementioned enriched lipids and depleted for one or more of the aforementioned depleted lipids. In some embodiments, the fusosomes comprise (or are identified as comprising) the enriched lipid as a percentage of total lipid that is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 5-fold, or 10-fold greater than the corresponding level in source cells. In some embodiments, the fusosome comprise (or are identified as comprising) the depleted lipid as a percentage of total lipid at a level that is less than 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, or 10% of the corresponding level in the source cells. In embodiments, lipid enrichment is measured by a mass spectrometry assay, e.g., an assay of Example 96.

In some embodiments, CE lipid levels are (or are identified as being) about 2-fold greater in fusosomes than in exosomes and/or about 5, 6, 7, 8, 9, or 10-fold higher in fusosomes than in parental cells (relative to total lipid in a sample). In some embodiments, ceramide lipid levels are (or are identified as being) about 2, 3, 4, or 5-fold greater in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, cholesterol levels are (or are identified as being) about 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold greater in exosomes than in fusosomes and/or about 2-fold higher in fusosomes than in parental cells (relative to total lipid in a sample). In some embodiments, CL lipid levels are (or are identified as being) at least about 5, 10, 20, 30, or 40-fold greater in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, DAG lipid levels are (or are identified as being) about 2 or 3-fold greater in exosomes than in fusosomes and/or about 1.5 or 2-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PC lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 1.3, 1.4, 1.5, 1.6, 1.7, or 1.8-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PC O- lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 2-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PE lipid levels are (or are identified

as being) about 1.3, 1.4, 1.5, 1.6, 1.7, or 1.8-fold higher in fusosomes than in exosomes and/or about 1.3, 1.4, 1.5, 1.6, 1.7, or 1.8-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PE O- lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PG lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 2, 3, 4, 5, 6, 7, 8, 9, or 10-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PI lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 3, 4, 5, 6, or 7-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, PS lipid levels are (or are identified as being) (or are identified as being) about equal between exosomes and fusosomes and/or about 2-fold higher in fusosomes than in parental cells (relative to total lipid in a sample). In some embodiments, SM lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 2, 2.5, or 3-fold higher in fusosomes than in parental cells (relative to total lipid in a sample). In some embodiments, TAG lipid levels are (or are identified as being) about equal between exosomes and fusosomes and/or about 10, 20, 30, 40, 50, 60, 70 80, 90, 100-fold, or more higher in parental cells than in fusosomes (relative to total lipid in a sample).

In some embodiments, the fusosomes are (or are identified as being) enriched for one or more of (e.g., at least 2, 3, 4, 5, or all of) the following lipids compared to exosomes: cholesteryl ester, ceramide, diacylglycerol, lyso-phosphatidate, and phosphatidylethanolamine, and triacylglycerol. In some embodiments, the fusosomes are (or are identified as being) depleted for one or more of (e.g., at least 2, 3, 4, 5, or all of) the following lipids compared to exosomes (relative to total lipid in a sample): free cholesterol, hexosyl ceramide, lyso-phosphatidylcholine, ether-linked lyso-phosphatidylcholine, lyso-phosphatidylethanolamine, ether-linked lyso-phosphatidylethanolamine, and lyso-phosphatidylserine. In some embodiments, the fusosomes are (or are identified as being) enriched for one or more of the aforementioned enriched lipids and depleted for one or more of the aforementioned depleted lipids. In some embodiments, the fusosomes comprise (or are identified as comprising) the enriched lipid as a percentage of total lipid that is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 5-fold, or 10-fold greater than the corresponding level in exosomes. In some embodiments, the fusosome comprise

(or are identified as comprising) the depleted lipid as a percentage of total lipid at a level that is less than 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, or 10% of the corresponding level in exosomes. In embodiments, lipid enrichment is measured by a mass spectrometry assay, e.g., an assay of Example 96.

In some embodiments, ceramide lipid levels are (or are identified as being) about 2-fold higher in fusosomes than in exosomes and/or about 2-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, HexCer lipid levels are (or are identified as being) about 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold higher in exosomes than in fusosomes and/or about equal in parental cells and fusosomes (relative to total lipid in a sample). In some embodiments, LPA lipid levels are (or are identified as being) about 3 or 4-fold higher in fusosomes than in exosomes and/or about 1.3, 1.4, 1.5, 1.6, 1.7, or 1.8-fold higher in fusosomes than in parental cells (relative to total lipid in a sample). In some embodiments, LPC lipid levels are (or are identified as being) about 2-fold higher in exosomes than in fusosomes and/or about 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, LPC O- lipid levels are (or are identified as being) about 3 or 4-fold higher in exosomes than in fusosomes and/or about equal between parental cells and fusosomes (relative to total lipid in a sample). In some embodiments, LPE lipid levels are (or are identified as being) about 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold higher in exosomes than in fusosomes and/or about 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold higher in parental cells than in fusosomes (relative to total lipid in a sample). In some embodiments, LPE O- lipid levels are (or are identified as being) about 2 or 3-fold higher in exosomes than in fusosomes and/or about equal between parental cells and fusosomes (relative to total lipid in a sample). In some embodiments, LPS lipid levels are (or are identified as being) about 3-fold higher in exosomes than in fusosomes (relative to total lipid in a sample). In some embodiments, PA lipid levels are (or are identified as being) about 1.5, 1.6, 1.7, 1.8, 1.9, or 2-fold higher in fusosomes than in exosomes and/or about 2-fold higher in fusosomes than in parental cells (relative to total lipid in a sample). In some embodiments, PG lipid levels are (or are identified as being) about equal between fusosomes and exosomes and/or about 10, 11, 12, 13, 14, or 15-fold higher in parental cells than in fusosomes (relative to total lipid in a sample).

In some embodiments, the fusosome comprises a lipid composition substantially similar to that of the source cell or wherein one or more of CL, Cer, DAG, HexCer, LPA, LPC, LPE,

LPG, LPI, LPS, PA, PC, PE, PG, PI, PS, CE, SM and TAG is within 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, or 50% of the corresponding lipid level in the source cell. In embodiments, the lipid composition of fusosomes is similar to the cells from which they are derived. In embodiments, fusosomes and parental cells have (or are identified as having) a similar lipid composition if greater than or equal to about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, or 90% of the lipid species identified in any replicate sample of the parental cells are present (or are identified as being present) in any replicate sample of the fusosomes, e.g., as determined according to Example 86. In embodiments, of identified lipids, the average level in the fusosome is greater than about 10%, 15%, 20%, 25%, 30%, 35%, or 40% of the corresponding average lipid species level in the parental cell (relative to total lipid in a sample). In an embodiment, the lipid composition of the fusosome is enriched and/or depleted for specific lipids relative to the parental cell (relative to total lipid in a sample).

In some embodiments, the lipid composition of the fusosome is (or is identified as bring) enriched and/or depleted for specific lipids relative to the parental cell, e.g., as determined according to the method described in Example 96.

In some embodiments, the fusosome has (or is identified as having) a ratio of phosphatidylserine to total lipids that is greater than that of the parental cell. In embodiments, the fusosome has (or is identified as having) a ratio of phosphatidylserine to total lipids of about 110%, 115%, 120%, 121%, 122%, 123%, 124%, 125%, 130%, 135%, 140%, or more relative to that of the parental cell. In some embodiments, the fusosome is (or is identified as being) enriched for cholesteryl ester, free cholesterol, ether-linked lyso-phosphatidylethanolamine, lyso-phosphatidylserine, phosphatidate, ether-linked phosphatidylethanolamine, phosphatidylserine, and/or sphingomyelin relative to the parental cell. In some embodiments, the fusosomes is (or is identified as being) depleted for ceramide, cardiolipin, lyso-phosphatidylcholine, lyso-phosphatidylethanolamine, lyso-phosphatidylglycerol, lyso-phosphatidylinositol, ether-linked phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol, phosphatidylinositol, and/or triacylglycerol relative to the parental cell. In some embodiments, the fusosome is (or is identified as being) enriched for cholesteryl ester, ceramide, diacylglycerol, lyso-phosphatidate, phosphatidylethanolamine, and/or triacylglycerol relative to an exosome. In some embodiments, the fusosome is (or is identified as being) depleted for free cholesterol, hexosyl ceramide, lyso-phosphatidylcholine, ether-linked lyso-phosphatidylcholine, lyso-phosphatidylethanolamine,

ether-linked lyso-phosphatidylethanolamine, and/or lyso-phosphatidylserine relative to an exosome.

In some embodiments, the fusosome has a ratio of cardiolipin: ceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin: ceramide in the source cell; or has a ratio of cardiolipin: diacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin: diacylglycerol in the source cell; or has a ratio of cardiolipin: hexosylceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin: hexosylceramide in the source cell; or has a ratio of cardiolipin:lysophosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin: lysophosphatidate in the source cell; or has a ratio of cardiolipin: lyso-phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin: lyso-phosphatidylcholine in the source cell; or has a ratio of cardiolipin: lyso-phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin: lyso-phosphatidylethanolamine in the source cell; or has a ratio of cardiolipin: lyso-phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : lyso-phosphatidylglycerol in the source cell; or has a ratio of cardiolipin: lyso-phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : lyso-phosphatidylinositol in the source cell; or has a ratio of cardiolipin: lyso-phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : lyso-phosphatidylserine in the source cell; or has a ratio of cardiolipin: phosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidate in the source cell; or has a ratio of cardiolipin: phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidylcholine in the source cell; or has a ratio of cardiolipin: phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidylethanolamine in the source cell; or has a ratio of cardiolipin: phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidylglycerol in the source cell; or has a ratio of cardiolipin: phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidylinositol in the source cell; or has a ratio of cardiolipin: phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidylserine in the source cell; or has a ratio of cardiolipin: cholesterol ester that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : cholesterol ester in the source cell; or has a ratio of cardiolipin: sphingomyelin that is within 10%, 20%, 30%, 40%, or

50% of the ratio of cardiolipin : sphingomyelin in the source cell; or has a ratio of cardiolipin: triacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : triacylglycerol in the source cell; or has a ratio of phosphatidylcholine: ceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine: ceramide in the source cell; or has a ratio of phosphatidylcholine: diacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine: diacylglycerol in the source cell; or has a ratio of phosphatidylcholine: hexosylceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine: hexosylceramide in the source cell; or has a ratio of phosphatidylcholine:lysophosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine: lysophosphatidate in the source cell; or has a ratio of phosphatidylcholine: lyso-phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine: lyso-phosphatidylcholine in the source cell; or has a ratio of phosphatidylcholine: lyso-phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine: lyso-phosphatidylethanolamine in the source cell; or has a ratio of phosphatidylcholine: lyso-phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : lyso-phosphatidylglycerol in the source cell; or has a ratio of phosphatidylcholine: lyso-phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : lyso-phosphatidylinositol in the source cell; or has a ratio of phosphatidylcholine: lyso-phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : lyso-phosphatidylserine in the source cell; or has a ratio of phosphatidylcholine: phosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cardiolipin : phosphatidate in the source cell; or has a ratio of phosphatidylcholine: phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : phosphatidylethanolamine in the source cell; or has a ratio of cardiolipin: phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : phosphatidylglycerol in the source cell; or has a ratio of phosphatidylcholine: phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : phosphatidylinositol in the source cell; or has a ratio of phosphatidylcholine: phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : phosphatidylserine in the source cell; or has a ratio of phosphatidylcholine: cholesterol ester that is within 10%, 20%, 30%, 40%, or 50% of the ratio of

phosphatidylcholine : cholesterol ester in the source cell; or has a ratio of phosphatidylcholine: sphingomyelin that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : sphingomyelin in the source cell; or has a ratio of phosphatidylcholine: triacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylcholine : triacylglycerol in the source cell; or has a ratio of phosphatidylethanolamine: ceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine: ceramide in the source cell; or has a ratio of phosphatidylethanolamine: diacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine: diacylglycerol in the source cell; or has a ratio of phosphatidylethanolamine: hexosylceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine: hexosylceramide in the source cell; or has a ratio of phosphatidylethanolamine:lysophosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine: lysophosphatidate in the source cell; or has a ratio of phosphatidylethanolamine: lyso-phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine: lyso-phosphatidylcholine in the source cell; or has a ratio of phosphatidylethanolamine: lyso-phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine: lyso-phosphatidylethanolamine in the source cell; or has a ratio of phosphatidylethanolamine: lyso-phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : lyso-phosphatidylglycerol in the source cell; or has a ratio of phosphatidylethanolamine: lyso-phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : lyso-phosphatidylinositol in the source cell; or has a ratio of phosphatidylethanolamine: lyso-phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : lyso-phosphatidylserine in the source cell; or has a ratio of phosphatidylethanolamine: phosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : phosphatidate in the source cell; or has a ratio of phosphatidylethanolamine: phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : phosphatidylglycerol in the source cell; or has a ratio of phosphatidylethanolamine: phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : phosphatidylinositol in the source cell; or has a ratio of phosphatidylethanolamine: phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : phosphatidylserine in the source cell; or has a ratio of

phosphatidylethanolamine: cholesterol ester that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : cholesterol ester in the source cell; or has a ratio of phosphatidylethanolamine: sphingomyelin that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : sphingomyelin in the source cell; or has a ratio of phosphatidylethanolamine: triacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylethanolamine : triacylglycerol in the source cell; or has a ratio of phosphatidylserine: ceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine: ceramide in the source cell; or has a ratio of phosphatidylserine: diacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine: diacylglycerol in the source cell; or has a ratio of phosphatidylserine: hexosylceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine: hexosylceramide in the source cell; or has a ratio of phosphatidylserine:lysophosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine: lysophosphatidate in the source cell; or has a ratio of phosphatidylserine: lyso-phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine: lyso-phosphatidylcholine in the source cell; or has a ratio of phosphatidylserine: lyso-phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine: lyso-phosphatidylethanolamine in the source cell; or has a ratio of phosphatidylserine: lyso-phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : lyso-phosphatidylglycerol in the source cell; or has a ratio of phosphatidylserine: lyso-phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : lyso-phosphatidylinositol in the source cell; or has a ratio of phosphatidylserine: lyso-phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : lyso-phosphatidylserine in the source cell; or has a ratio of phosphatidylserine: phosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : phosphatidate in the source cell; or has a ratio of phosphatidylserine: phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : phosphatidylglycerol in the source cell; or has a ratio of phosphatidylserine: phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : phosphatidylinositol in the source cell; or has a ratio of phosphatidylserine: cholesterol ester that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : cholesterol ester in the source cell; or has a ratio of phosphatidylserine: sphingomyelin that is

within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : sphingomyelin in the source cell; or has a ratio of phosphatidylserine: triacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of phosphatidylserine : triacylglycerol in the source cell; or has a ratio of sphingomyelin: ceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin: ceramide in the source cell; or has a ratio of sphingomyelin: diacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin: diacylglycerol in the source cell; or has a ratio of sphingomyelin: hexosylceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin: hexosylceramide in the source cell; or has a ratio of sphingomyelin:lysophosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin: lysophosphatidate in the source cell; or has a ratio of sphingomyelin: lyso-phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin: lyso-phosphatidylcholine in the source cell; or has a ratio of sphingomyelin: lyso-phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin: lyso-phosphatidylethanolamine in the source cell; or has a ratio of sphingomyelin: lyso-phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : lyso-phosphatidylglycerol in the source cell; or has a ratio of sphingomyelin: lyso-phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : lyso-phosphatidylinositol in the source cell; or has a ratio of sphingomyelin: lyso-phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : lyso-phosphatidylserine in the source cell; or has a ratio of sphingomyelin: phosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : phosphatidate in the source cell; or has a ratio of sphingomyelin: phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : phosphatidylglycerol in the source cell; or has a ratio of sphingomyelin: phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : phosphatidylinositol in the source cell; or has a ratio of sphingomyelin: cholesterol ester that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : cholesterol ester in the source cell; or has a ratio of sphingomyelin: triacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of sphingomyelin : triacylglycerol in the source cell; or has a ratio of cholesterol ester: ceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester: ceramide in the source cell; or has a ratio of cholesterol ester: diacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio

of cholesterol ester: diacylglycerol in the source cell; or has a ratio of cholesterol ester: hexosylceramide that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester: hexosylceramide in the source cell; or has a ratio of cholesterol ester:lysophosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester: lysophosphatidate in the source cell; or has a ratio of cholesterol ester: lyso-phosphatidylcholine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester: lyso-phosphatidylcholine in the source cell; or has a ratio of cholesterol ester: lyso-phosphatidylethanolamine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester: lyso-phosphatidylethanolamine in the source cell; or has a ratio of cholesterol ester: lyso-phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : lyso-phosphatidylglycerol in the source cell; or has a ratio of cholesterol ester: lyso-phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : lyso-phosphatidylinositol in the source cell; or has a ratio of cholesterol ester: lyso-phosphatidylserine that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : lyso-phosphatidylserine in the source cell; or has a ratio of cholesterol ester: phosphatidate that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : phosphatidate in the source cell; or has a ratio of cholesterol ester: phosphatidylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : phosphatidylglycerol in the source cell; or has a ratio of cholesterol ester: phosphatidylinositol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : phosphatidylinositol in the source cell; or has a ratio of cholesterol ester: triacylglycerol that is within 10%, 20%, 30%, 40%, or 50% of the ratio of cholesterol ester : triacylglycerol in the source cell.

In some embodiments, the fusosome comprises a proteomic composition similar to that of the source cell, e.g., using an assay of Example 87. In some embodiments, the protein composition of fusosomes are similar to the parental cells from which they are derived. In some embodiments, the fractional content of each of a plurality of categories of proteins is determined as the sum of intensity signals from each category divided by the sum of the intensity signals of all identified proteins in the sample, e.g., as described in Example 87. In some embodiments, the fusosome comprises (or is identified as comprising) varying amounts of compartment-specific proteins relative to parental cells and/or exosomes, e.g., as determined according to the method described in Example 97. In some embodiments, fusosomes are (or are identified as being) depleted with endoplasmic reticulum protein compared to parental cells and exosomes. In

some embodiments, fusosomes are (or are identified as being) depleted for exosomal protein compared to exosomes. In some embodiments, fusosomes have (or are identified as having) less than 15%, 20%, or 25% of the protein in the fusosome as being exosomal protein. In some embodiments, fusosomes are (or are identified as being) depleted for mitochondrial protein compared to parental cells. In some embodiments, fusosomes are (or are identified as being)enriched for nuclear protein compared to parental cells. In some embodiments, fusosomes are (or are identified as being) enriched for ribosomal proteins compared to parental cells and exosomes. In some embodiments, at least 0.025%, 0.03%, 0.04%, 0.05%, 0.06%, 0.07%, 0.08%, 0.09%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7% 8%, 9% or 10% of the protein in the fusosome is ribosomal protein, or about 0.025-0.2%, 0.05-0.15%, 0.06-1.4%, 0.07%-1.3%, 0.08%-1.2%, 0.09%-1.1%, 1%-20%, 3%-15%, 5%-12.5%, 7.5%-11%, or 8.5%-10.5%, or 9%-10% of the protein in the fusosome is ribosomal protein.

In some embodiments, the fusosome comprises a ratio of lipids to proteins that is within 10%, 20%, 30%, 40%, or 50% of the corresponding ratio in the source cell, e.g., as measured using an assay of Example 40. In embodiments, the fusosome comprises (or is identified as comprising) a ratio of lipid mass to proteins approximately equal to the lipid mass to protein ratio for nucleated cells. In embodiments, the fusosome comprises (or is identified as comprising) a greater lipid:protein ratio than the parental cell. In embodiments, the fusosome comprises (or is identified as comprising) a lipid:protein ratio of about 110%, 115%, 120%, 125%, 130%, 131%, 132%, 132.5%, 133%, 134%, 135%, 140%, 145%, or 150% of the lipid:protein ratio of the parental cell. In some embodiments, the fusosome or fusosome composition has (or is identified as having) a phospholipid:protein ratio of about 100-180, 110-170, 120-160, 130-150, 135-145, 140-142, or 141 μmol/g, e.g., in an assay of Example 83. In some embodiments, the fusosome or fusosome composition has (or is identified as having) a phospholipid:protein ratio that is about 60-90%, 70-80%, or 75% of the corresponding ratio in the source cells, e.g., in an assay of Example 83.

In some embodiments, the fusosome comprises a ratio of proteins to nucleic acids (e.g., DNA or RNA) that is within 10%, 20%, 30%, 40%, or 50% of the corresponding ratio in the source cell, e.g., as measured using an assay of Example 41. In embodiments, the fusosome comprises (or is identified as comprising) a ratio of protein mass to DNA mass similar to that of a parental cell. In embodiments, the fusosome comprises (or is identified as comprising) a ratio

of protein:DNA that is about about 85%, 90%, 95%, 96%, 97%, 98%, 98.2%, 99%, 100%, 101%, 102%, 103%, 104%, 105%, or 110% of the parental cell. In some embodiments, the fusosome comprises a ratio of proteins to DNA that is greater than the corresponding ratio in the source cell, e.g., at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% greater, e.g., as measured using an assay of Example 41. In some embodiments, the fusosome or fusosome composition comprises (or is identified as comprising) a ratio of protein:DNA that is about 20-35, 25-30, 26-29, 27-28, or 27.8 g/g, e.g., by an assay of Example 84. In some embodiments, the fusosome or fusosome composition comprises (or is identified as comprising) a ratio of protein:DNA that is within about 1%, 2%, 5%, 10%, or 20% of the corresponding ratio in the source cells, e.g., by an assay of Exmple 84.

In some embodiments, the fusosome comprises a ratio of lipids to nucleic acids (e.g., DNA) that is within 10%, 20%, 30%, 40%, or 50% of the corresponding ratio in the source cell, e.g., as measured using an assay of Example 91. In some embodiments, the fusosome or fusosome composition comprises (or is identified as comprising) a ratio of lipids:DNA that is about 2.0-6.0, 3.0-5.0, 3.5-4.5, 3.8-4.0, or 3.92 μmol/mg, e.g., by an assay of Example 85. In some embodiments, the fusosome comprises a ratio of lipids to nucleic acids (e.g., DNA) that is greater than the corresponding ratio in the source cell, e.g., at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% greater, e.g., as measured using an assay of Example 91. In embodiments, the fusosome comprises (or is identified as comprising) a greater lipid:DNA ratio than the parental cell. In embodiments, the fusosome comprises about a 105%, 110%, 115%, 120%, 125%, 130%, 135%, 140%, 145%, 150%, or greater lipid:DNA ratio compared to the parental cell.

In some embodiments, the fusosome composition has a half-life in a subject, e.g., in a mouse, that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% of the half life of a reference cell composition, e.g., the source cell, e.g., by an assay of Example 60. In some embodiments, the fusosome composition has a half-life in a subject, e.g., in a mouse, that is at least 1 hour, 2 hours, 3 hours, 4 hours, 5 hours, 6 hours, 12 hours, or 24 hours, e.g., in a human subject or in a mouse, e.g., by an assay of Example 60. In embodiments, the fusosome composition has a half-life of at least 1, 2, 4, 6, 12, or 24 hours in a subject, e.g., in an assay of Example 79. In some embodiments, the therapeutic agent has a half-life in a subject that is longer than the half-life of the fusosome composition, e.g., by at least 10%, 20%, 50%, 2-

fold, 5-fold, or 10-fold. For instance, the fusosome may deliver the therapeutic agent to the target cell, and the therapeutic agent may be present after the fusosome is no longer present or detectable.

In some embodiments, the fusosome transports glucose (e.g., labeled glucose, e.g., 2-NBDG) across a membrane, e.g., by at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% more than a negative control, e.g., an otherwise similar fusosome in the absence of glucose, e.g., as measured using an assay of Example 50. In some embodiments, the fusosome transports (or is identified as transporting) glucose (e.g., labeled glucose, e.g., 2-NBDG) across a membrane at a greater level than otherwise similar fusosomes treated with phloretin, e.g., in an assay of Example 72. In embodiments, a fusosome not treated with phloretin transports (or is identified as not transporting) glucose at a level at least 1%, 2%, 3%, 5%, or 10% higher (and optionally up to 15% higher) than an otherwise similar fusosome treated with phloretin, e.g., in an assay of Example 72. In some embodiments, the fusosome comprises esterase activity in the lumen that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of that of the esterase activity in a reference cell, e.g., the source cell or a mouse embryonic fibroblast, e.g., using an assay of Example 51. In some embodiments, the fusosome comprises (or is identified as comprising) esterase activity in the lumen that is at least 10-fold, 20-fold, 50-fold, 100-fold, 200-fold, 500-fold, 1000-fold, 2000-fold, or 5000-fold higher than an unstained control, e.g., by an assay of Example 73. In some embodiments, the fusosome comprises (or is identified as comprising) esterase activity in the lumen that is about 10-100-fold lower than that of the source cells, e.g., by an assay of Example 73. In some embodiments, the fusosome comprises (or is identified as comprising) an acetylcholinesterase activity of about 1E5-1E6, 6E5-8E5, 6.5E5-7E5, or 6.83E5 exosome equivalents, e.g., by an assay of Example 74. In some embodiments, the fusosome comprises a metabolic activity level (e.g., citrate synthase activity) that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the metabolic activity level in a reference cell, e.g., the source cell, e.g., as described in Example 53. In some embodiments, the fusosome comprises a metabolic activity level (e.g., citrate synthase activity) that is at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the metabolic activity level in a reference cell, e.g., the source cell, e.g., as described in Example 53. In some embodiments, the fusosome comprises (or is identified as comprising) a citrate synthase activity

that is about 1E-2 - 2 E-2, 1.3E-2 - 1.8E-2, 1.4E-2 - 1.7E-2, 1.5E-2 - 1.6E-2, or 1.57E-2 umol/ug fusosome/min, e.g., by an assay of Example 75. In some embodiments, the fusosome comprises a respiration level (e.g., oxygen consumption rate), e.g., basal, uncoupled, or maximal respiration level, that is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the respiration level in a reference cell, e.g., the source cell, e.g., as described in Example 54. In some embodiments, the fusosome comprises a respiration level (e.g., oxygen consumption rate), e.g., basal, uncoupled, or maximal respiration level, that is at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% of the respiration level in a reference cell, e.g., the source cell, e.g., as described in Example 54. In embodiments, the fusosome comprises (or is identified as comprising) a basal respiration rate of about 8-15, 9-14, 10-13, 11-12, or 11.3 pmol/min/20µg fusosome, e.g., by an assay of Example 76. In embodiments, the fusosome comprises (or is identified as comprising) an uncoupled respiration rate of about 8-13, 9-12, 10-11, 10-10.2, or 10.1 pmol/min/20µg fusosome, e.g., by an assay of Example 76. In embodiments, the fusosome comprises (or is identified as comprising) a maximal respiration rate of about 15-25, 16-24, 17-23, 18-22, 19-21, or 20 pmol/min/20µg fusosome, e.g., by an assay of Example 76. In embodiments, the fusosome has (or is identified as having) a higher basal respiration rate than uncoupled respiration rate, e.g., by about 1%, 2%, 5%, or 10%, e.g., up to about 15%, e.g., by an assay of Example 76. In embodiments, the fusosome has (or is identified as having) a higher maxaimal respiration rate than basal respiration rate, e.g., by about 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, e.g., by an assay of Example 76. In some embodiments, the fusosome comprises an Annexin-V staining level of at most 18,000, 17,000, 16,000, 15,000, 14,000, 13,000, 12,000, 11,000, or 10,000 MFI, e.g., using an assay of Example 55, or wherein the fusosome comprises an Annexin-V staining level at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% lower than the Annexin-V staining level of an otherwise similar fusosome treated with menadione in the assay of Example 55, or wherein the fusosome comprises an Annexin-V staining level at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% lower than the Annexin-V staining level of a macrophage treated with menadione in the assay of Example 55. In embodiments, the fusosome comprises (or is identified as comprising) an Annexin V-staining level that is at least about 1%, 2%, 5%, or 10% lower than the Annexin V-staining level of an otherwise similar fusosome treated with antimycin A, e.g., in an assay of Example77. In embodiments, the fusosome

comprises (or is identified as comprising) an Annexin V-staining level that is within about 1%, 2%, 5%, or 10% of the Annexin V-staining level of an otherwise similar fusosome treated with antimycin A, e.g., in an assay of Example 77.

In some embodiments, the fusosome has a miRNA content level of at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater than that of the source cell, e.g., by an assay of Example 33. In some embodiments, the fusosome has a miRNA content level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater of the miRNA content level of the source cell (e.g., up to 100% of the miRNA content level of the source cell), e.g., by an assay of Example 33. In some embodiments, the fusosome has a total RNA content level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater of the total RNA content level of the source cell (e.g., up to 100% of the total RNA content level of the source cell), e.g., as measured by an assay of Example 66.

In some embodiments, the fusosome has a soluble : non-soluble protein ratio is within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater than that of the source cell, e.g., within 1%-2%, 2%-3%, 3%-4%, 4%-5%, 5%-10%, 10%-20%, 20%-30%, 30%-40%, 40%-50%, 50%-60%, 60%-70%, 70%-80%, or 80%-90% of that of the source cell, e.g., by an assay of Example 38. In embodiments, the fusosome has a soluble: non-soluble protein ratio of about 0.3-0.8, 0.4-0.7, or 0.5-0.6, e.g., about 0.563, e.g., by an assay of Example 38. In some embodiments, the population of fusosomes has (or is identified as having) a soluble:insoluble protein mass ratio of about 0.3-0.8, 0.4-0.7, 0.5-0.6, or 0.563, or greater than about 0.1, 0.2, 0.3, 0.4, or 0.5. In some embodiments, the population of fusosomes has (or is identified as having) a soluble:insoluble protein mass ratio that is greater than that of the source cells, e.g., at least 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, or 20-fold higher. In embodiments, the soluble:insoluble protein mass ratio is determined by an assay of Example 70. In embodiments, the soluble: insoluble protein mass ratio is (or is identified as being) lower in the fusosome population than in the parental cells. In embodiments, when the ratio of fusosomes to parental cells is (or is identified as being) about 3%, 4%, 5%, 6%, 7%, or 8%, the soluble: insoluble ratio of the population of fusosomes is (or is identified as being) about equal to the soluble: insoluble ratio of the parental cells.

In some embodiments, the fusosome has an LPS level less than 5%, 1%, 0.5%, 0.01%, 0.005%, 0.0001%, 0.00001% or less of the LPS content of the source cell, e.g., as measured by mass spectrometry, e.g., in an assay of Example 39. In some embodiments, the fusosome is capable of signal transduction, e.g., transmitting an extracellular signal, e.g., AKT phosphorylation in response to insulin, or glucose (e.g., labeled glucose, e.g., 2-NBDG) uptake in response to insulin, e.g., by at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% more than a negative control, e.g., an otherwise similar fusosome in the absence of insulin, e.g., using an assay of Example 49. In some embodiments, the fusosome targets a tissue, e.g., liver, lungs, heart, spleen, pancreas, gastrointestinal tract, kidney, testes, ovaries, brain, reproductive organs, central nervous system, peripheral nervous system, skeletal muscle, endothelium, inner ear, or eye, when administered to a subject, e.g., a mouse, e.g., wherein at least 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50%, 60%, 70%, 80%, or 90% of the fusosomes in a population of administered fusosomes are present in the target tissue after 24, 48, or 72 hours, e.g., by an assay of Example 64. In some embodiments, the fusosome has a juxtacrine-signaling level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% greater than the level of juxtacrine signaling induced by a reference cell, e.g., the source cell or a bone marrow stromal cell (BMSC), e.g., by an assay of Example 56. In some embodiments, the fusosome has a juxtacrine-signaling level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% (e.g., up to 100%) of the level of juxtacrine signaling induced by a reference cell, e.g., the source cell or a bone marrow stromal cell (BMSC), e.g., by an assay of Example56. In some embodiments, the fusosome has a paracrine-signaling level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% greater than the level of paracrine signaling induced by a reference cell, e.g., the source cell or a macrophage, e.g., by an assay of Example 57. In some embodiments, the fusosome has a paracrine-signaling level of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% (e.g., up to 100%) of the level of paracrine signaling induced by a reference cell, e.g., the source cell or a macrophage, e.g., by an assay of Example 57. In some embodiments, the fusosome polymerizes actin at a level within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% compared to the level of polymerized actin in a reference cell, e.g., the source cell or a C2C12 cell, e.g., by the assay of Example 58. In some embodiments, the fusosome polymerizes actin (or

is identified as polymerizing actin) at a level that is constant over time, e.g, over at least 3, 5, or 24 hours, e.g., by an assay of Example 81. In embodiments, the level of actin polymerization changes by less than 1%, 2%, 5%, 10%, or 20% over a 5-hour period, e.g. by the assay of Example 81. In some embodiments, the fusosome has a membrane potential within about 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% of the membrane potential of a reference cell, e.g., the source cell or a C2C12 cell, e.g., by an assay of Example 59, or wherein the fusosome has a membrane potential of about -20 to -150mV, -20 to -50mV, -50 to -100mV, or -100 to -150mV, or wherein the fusosome has a membrane potential of less than -1mv, -5mv, -10mv, -20mv, -30mv, -40mv, -50mv, -60mv, -70mv, -80mv, -90mv, -100mv. In some embodiments, the fusosome has (or is identified as having) a membrane potential of about -25 to -35, -27 to -32, -28 to -31, -29 to -30, or -29.6 millivolts, e.g., in an assay of Example 78. In some embodiments, the fusosome is capable of extravasation from blood vessels, e.g., at a rate at least 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% the rate of extravasation of the source cell, e.g., using an assay of Example 44, e.g., wherein the source cell is a neutrophil, lymphocyte, B cell, macrophage, or NK cell. In some embodiments, the fusosome is capable of chemotaxis, e.g., of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% (e.g., up to 100%) compared to a reference cell, e.g., a macrophage, e.g., using an assay of Example 45. In some embodiments, the fusosome is capable of phagocytosis, e.g., at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% (e.g., up to 100%) compared to a reference cell, e.g., a macrophage, e.g., using an assay of Example 47. In some embodiments, the fusosome is capable of crossing a cell membrane, e.g., an endothelial cell membrane or the blood brain barrier. In some embodiments, the fusosome is capable of secreting a protein, e.g., at a rate at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100% greater than a reference cell, e.g., a mouse embryonic fibroblast, e.g., using an assay of Example 48. In some embodiments, the fusosome is capable of secreting a protein, e.g., at a rate at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% (e.g., up to 100%) compared to a reference cell, e.g., a mouse embryonic fibroblast, e.g., using an assay of Example 48.

In some embodiments, the fusosome is not capable of transcription or has transcriptional activity of less than 1%, 2.5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of that of the transcriptional activity of a reference cell, e.g., the source cell, e.g., using an assay of

Example 24. In some embodiments, the fusosome is not capable of nuclear DNA replication or has nuclear DNA replication of less than 1%, 2.5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the nuclear DNA replication of a reference cell, e.g., the source cell, e.g., using an assay of Example 25. In some embodiments, the fusosome lacks chromatin or has a chromatin content of less than 1%, 2.5% 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the of the chromatin content of a reference cell, e.g., the source cell, e.g., using an assay of Example 32.

In some embodiments, a characteristic of a fusosome is described by comparison to a reference cell. In embodiments, the reference cell is the source cell. In embodiments, the reference cell is a HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cell. In some embodiments, a characteristic of a population of fusosomes is described by comparison to a population of reference cells, e.g., a population of source cells, or a population of HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cells.

In some embodiments, the fusosome meets a pharmaceutical or good manufacturing practices (GMP) standard. In some embodiments, the fusosome was made according to good manufacturing practices (GMP). In some embodiments, the fusosome has a pathogen level below a predetermined reference value, e.g., is substantially free of pathogens. In some embodiments, the fusosome has a contaminant level below a predetermined reference value, e.g., is substantially free of contaminants. In some embodiments, the fusosome has low immunogenicity, e.g., as described herein.

In some embodiments, immunogenicity of a fusosome composition is assayed by a serum inactivation assay (e.g., an assay that detects antibody-mediated neutralization or complement mediated degradation). In some embodiments, fusosomes are not inactivated by serum, or are inactivated at a level below a predetermined value. In some embodiments, serum of a fusosome-naïve subject (e.g., human or mouse) is contacted with a test fusosome composition. In some embodiments, the serum of a subject that has received one or more administrations of fusosomes, e.g., has received at least two administrations of fusosomes, is contacted with the test fusosome composition. In embodiments, serum-exposed fusosomes are then tested for ability to deliver a cargo to target cells. In some embodiments, the percent of cells that detectably comprise the cargo after treatment with serum-incubated fusosomes is at least 50%, 60%, 70%, 80%, 90%, or 95% the percent of cells that detectably comprise the cargo after treatment with positive control

fusosomes not contacted with serum. In some embodiments, serum inactivation is measured using an assay of Example 99.

In some embodiments, immunogenicity of a fusosome composition is assayed by detecting complement activation in response to the fusosomes. In some embodiments, the fusosomes do not activate complement, or activate complement at a level below a predetermined value. In some embodiments, serum of a fusosome-naïve subject (e.g., human or mouse) is contacted with a test fusosome composition. In some embodiments, the serum of a subject that has received one or more administrations of fusosomes, e.g., has received at least two administrations of fusosomes, is contacted with the test fusosome composition. In embodiments, the composition comprising serum and fusosom

es is then tested for an activated complement factor (e.g., C3a), e.g., by ELISA. In some embodiments, a fusosome comprising a modification described herein (e.g., elevated levels of a complement regulatory protein compared to a reference cell) undergoes reduced complement activation compared to an otherwise similar fusosome that lacks the modification, e.g., reduced by at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 98%, or 99%. In some embodiments, complement activation is measured using an assay of Example 100.

In some embodiments, a fusosome or population of fusosomes will not be substantially inactivated by serum. In some embodiments, a fusosome or population of fusosomes is resistant to serum inactivation, e.g., as quantified according to the method described in Example 99. In embodiments, the fusosome or population of fusosomes is not substantially inactivated by serum or is resistant to serum inactivation following multiple administrations of the fusosome or population of fusosomes to a subject, e.g., according to the methods described herein. In some embodiments, a fusosome is modified to have a reduced serum inactivation, e.g., compared to a corresponding unmodified fusosome, e.g., following multiple administrations of the modified fusosome, e.g., as quantified according to the method described in Example 99.

In some embodiments, a fusosome does not substantially induce complement activity, e.g., as measured according to the method described in Example 100. In some embodiments, a fusosome is modified to induce reduced complement activity compared to a corresponding unmodified fusosome. In embodiments, complement activity is measured by determining expression or activity of a complement protein (e.g., DAF, proteins that bind decay-accelerating factor (DAF, CD55), e.g., factor H (FH)-like protein-1 (FHL-1), C4b-binding protein (C4BP),

complement receptor 1 (CD35), Membrane cofactor protein (MCP, CD46), Profectin (CD59), proteins that inhibit the classical and alternative complement pathway CD/C5 convertase enzymes, or proteins that regulate MAC assembly) in a cell

In some embodiments, the source cell is an endothelial cell, a fibroblast, a blood cell (e.g., a macrophage, a neutrophil, a granulocyte, a leukocyte), a stem cell (e.g., a mesenchymal stem cell, an umbilical cord stem cell, bone marrow stem cell, a hematopoietic stem cell, an induced pluripotent stem cell e.g., an induced pluripotent stem cell derived from a subject's cells), an embryonic stem cell (e.g., a stem cell from embryonic yolk sac, placenta, umbilical cord, fetal skin, adolescent skin, blood, bone marrow, adipose tissue, erythropoietic tissue, hematopoietic tissue), a myoblast, a parenchymal cell (e.g., hepatocyte), an alveolar cell, a neuron (e.g., a retinal neuronal cell) a precursor cell (e.g., a retinal precursor cell, a myeloblast, myeloid precursor cells, a thymocyte, a meiocyte, a megakaryoblast, a promegakaryoblast, a melanoblast, a lymphoblast, a bone marrow precursor cell, a normoblast, or an angioblast), a progenitor cell (e.g., a cardiac progenitor cell, a satellite cell, a radial gial cell, a bone marrow stromal cell, a pancreatic progenitor cell, an endothelial progenitor cell, a blast cell), or an immortalized cell (e.g., HeLa, HEK293, HFF-1, MRC-5, WI-38, IMR 90, IMR 91, PER.C6, HT-1080, or BJ cell). In some embodiments, the source cell is other than a 293 cell, HEK cell, human endothelial cell, or a human epithelial cell, monocyte, macrophage, dendritic cell, or stem cell.

In some embodiments, the source cell expresses (e.g., overexpresses) ARRDC1 or an active fragment or variant thereof. In some embodiments, the fusosome or fusosome composition has a ratio of fusogen to ARRDC1 of about 1-3, 1-10, 1-100, 3-10, 4-9, 5-8, 6-7, 15-100, 60-200, 80-180, 100-160, 120-140, 3-100, 4-100, 5-100, 6-100, 15-100, 80-100, 3-200, 4-200, 5-200, 6-200, 15-200, 80-200, 100-200, 120-200, 300-1000, 400-900, 500-800, 600-700, 640-690, 650-680, 660-670, 100-10,000, or about 664.9, e.g., by a mass spectrometry assay. In some embodiments, the level of ARRDC1 as a percentage of total protein content is at least about 0.01%, 0.02%, 0.03%, 0.04%, 0.05%; 0.1%, 0.15%, 0.2%, 0.25%; 0.5%, 1%, 2%, 3%, 4%, 5%; or the level of ARRDC1 as a percentage of total protein content is about 0.05-1.5%, 0.1%-0.3%, 0.05-0.2%, 0.1-0.2%, 0.25-7.5%, 0.5%-1.5%, 0.25-1%, 0.5-1%, 0.05-1.5%, 10%-30%, 5-20%, or 10-20%, e.g., by mass spectrometry, e.g., as measured according to the method described in Example 98. In some embodiments, the fusosome or fusosome composition has a

ratio of fusogen to TSG101 of about 100-1,000, 100-400, 100-500, 200-400, 200-500, 200-1,000, 300-400, 1,000-10,000, 2,000-5,000, 3,000-4,000, 3,050-3,100, 3,060-3,070, or about 3,064, 10,000-100,000, 10,000-200,000, 10,000-500,000, 20,000-500,000, 30,000-400,000, e.g., using a mass spectrometry assay, e.g., an assay of Example 94. In some embodiments, the fusosome or fusosome composition has a ratio of cargo to tsg101 of about 1-3, 1-30, 1-20, 1-25, 1.5-30, 10-30, 15-25, 18-21, 19-20, 10-300, 10-200, 15-300, 15-200, 100-300, 100-200, 150-300, or about 19.5 , e.g., using a mass spectrometry assay, e.g., an assay of Example 95. In some embodiments, the level of TSG101 as a percentage of total protein content is at least about 0.0001%, 0.0002%, 0.0003%, 0.0004%, 0.0005%, 0.0006%, 0.0007%, 0.001%, 0.002%, 0.003%, 0.004%, 0.005%, 0.006%, 0.007%; 0.01%, 0.02%, 0.03%, 0.04%, 0.05%, 0.06%, 0.07%; or the level of TSG101 as a percentage of total protein content is about 0.0001-0.001, 0.0001-0.002, 0.0001-0.01, 0.0001-0.1, 0.001-0.01, 0.002-0.006, 0.003-0.005, 0.001-0.1, 0.01-0.1, 0.02-0.06, 0.03-0.05, or 0.004, e.g., by mass spectrometry, e.g., as measured according to the method described in Example 98.

In some embodiments, the fusosome comprises a cargo, e.g., a therapeutic agent, e.g., an endogenous therapeutic agent or an exogenous therapeutic agent. In some embodiments, the therapeutic agent is chosen from one or more of a protein, e.g., an enzyme, a transmembrane protein, a receptor, an antibody; a nucleic acid, e.g., DNA, a chromosome (e.g. a human artificial chromosome), RNA, mRNA, siRNA, miRNA, or a small molecule. In some embodiments, the therapeutic agent is an organelle other than a mitochondrion, e.g., an organelle selected from: nucleus, Golgi apparatus, lysosome, endoplasmic reticulum, vacuole, endosome, acrosome, autophagosome, centriole, glycosome, glyoxysome, hydrogenosome, melanosome, mitosome, cnidocyst, peroxisome, proteasome, vesicle, and stress granule. In some embodiments, the organelle is a mitochondrion.

In some embodiments, the fusosome enters the target cell by endocytosis, e.g., wherein the level of therapeutic agent delivered via an endocytic pathway is 0.01-0.6, 0.01-0.1, 0.1-0.3, or 0.3-0.6, or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than a chloroquine treated reference cell contacted with similar fusosomes, e.g., using an assay of Example 62. In some embodiments, at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% of fusosomes in a fusosome composition that enter a target cell enter via a non-endocytic pathway, e.g., the fusosomes enter the target cell via fusion

with the cell surface. In some embodiments, the level of a therapeutic agent delivered via a non-endocytic pathway

for a given fusosome is 0.1-0.95, 0.1-0.2, 0.2-0.3, 0.3-0.4, 0.4-0.5, 0.5-0.6, 0.6-0.7, 0.7-0.8, 0.8-0.9, 0.9-

0.95, or at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or

greater than a chloroquine treated reference cell, e.g., using an assay of Example 61. In some embodiments, at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% of fusosomes in a fusosome composition that enter a target cell enter the cytoplasm (e.g., do not enter an endosome or lysosome). In some embodiments, less than 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, or 1% of fusosomes in a fusosome composition that enter a target cell enter an endosome or lysosome. In some embodiments, the fusosome enters the target cell by a non-endocytic pathway, e.g., wherein the level of therapeutic agent delivered is at least 90%, 95%, 98%, or 99% that of a chloroquine treated reference cell, e.g., using an assay of Example 62. In an embodiment, a fusosome delivers an agent to a target cell via a dynamin mediated pathway. In an embodiment, the level of agent delivered via a dynamin mediated pathway is in the range of 0.01-0.6, or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than Dynasore treated target cells contacted with similar fusosomes, e.g., as measured in an assay of Example 63. In an embodiment, a fusosome delivers an agent to a target cell via macropinocytosis. In an embodiment, the level of agent delivered via macropinocytosis is in the range of 0.01-0.6, or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than EIPA treated target cells contacted with similar fusosomes, e.g., as measured in an assay of Example 63. In an embodiment, a fusosome delivers an agent to a target cell via an actin-mediated pathway. In an embodiment, the level of agent delivered via an actin-mediated pathway will be in the range of 0.01-0.6, or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than Latrunculin B treated target cells contacted with similar fusosomes, e.g., as measured in an assay of Example 63.

In some embodiments, the fusosome has a density of <1, 1-1.1, 1.05-1.15, 1.1-1.2, 1.15-1.25, 1.2-1.3, 1.25-1.35, or >1.35 g/ml, e.g., by an assay of Example 30.

In some embodiments, the fusosome composition comprises less than 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 4%, 5%, or 10% source cells by protein mass or less than 0.01%, 0.05%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 4%, 5%, or 10% of cells have a functional nucleus. In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 99% of fusosomes in the fusosome composition comprise an organelle, e.g., a mitochondrion.

In some embodiments, the fusosome further comprises an exogenous therapeutic agent. In some embodiments, the exogenous therapeutic agent is chosen from one or more of a protein, e.g., an enzyme, a transmembrane protein, a receptor, an antibody; a nucleic acid, e.g., DNA, a chromosome (e.g. a human artificial chromosome), RNA, mRNA, siRNA, miRNA, or a small molecule.

In embodiments, the fusosome enters the cell by endocytosis or a non-endocytic pathway.

In embodiments, the fusosome composition is stable at a temperature of less than 4 C for at least 1, 2, 3, 6, or 12 hours; 1, 2, 3, 4, 5, or 6 days; 1, 2, 3, or 4 weeks; 1, 2, 3, or 6 months; or 1, 2, 3, 4, or 5 years. In embodiments, the fusosome composition is stable at a temperature of less than -20 C for at least 1, 2, 3, 6, or 12 hours; 1, 2, 3, 4, 5, or 6 days; 1, 2, 3, or 4 weeks; 1, 2, 3, or 6 months; or 1, 2, 3, 4, or 5 years. In embodiments, the fusosome composition is stable at a temperature of less than -80 C for at least 1, 2, 3, 6, or 12 hours; 1, 2, 3, 4, 5, or 6 days; 1, 2, 3, or 4 weeks; 1, 2, 3, or 6 months; or 1, 2, 3, 4, or 5 years.

In embodiments, the fusosome has a size, or the population of fusosomes has an average size, within about 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, of that of the source cell, e.g., as measured by an assay of Example 27. In embodiments, the fusosome has a size, or the population of fusosomes has an average size, that is less than about 0.01%, 0.05%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, of that of the source cell, e.g., as measured by an assay of Example 27. In embodiments, the fusosomes have (or are identified as having) a size less than parental cells. In embodiments, the fusosomes have (or are identified as having) a size within about 50%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 80%, or 90% of parental cells. In embodiments, the fusosomes have (or are identified as having) less than about 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 1%, or less of the parental cell's variability in size distribution, e.g., within about 90% of the sample. In embodiments, the fusosomes have (or are identified as

having) about 40%, 45%, 50%, 55%, 56%, 57%, 58%, 59%, 60%, 65%, or 70% less of the parental cell's variability in size distribution, e.g., within about 90% of the sample. In some embodiments, fusosomes have (or are identified as having) an average size of greater than 30, 35, 40, 45, 50, 55, 60, 65, or 70 nm in diameter. In embodiments, fusosomes have an average size of about 100, 110, 120, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 140, or 150 nm in diameter. In embodiments, the fusosome has a size, or the population of fusosomes has an average size, within about 0.01%-0.05%, 0.05%-0.1%, 0.1%-0.5%, 0.5%- 1%, 1%-2%, 2%-3%, 3%-4%, 4%-5%, 5%-10%, 10%-20%, 20%-30%, 30%-40%, 40%-50%, 50%-60%, 60%-70%, 70%-80%, or 80%-90% the size of the source cell, e.g., as measured by an assay of Example 27. In embodiments, the fusosome has a size, or the population of fusosomes has an average size, that is less than about 0.01%-0.05%, 0.05%-0.1%, 0.1%-0.5%, 0.5%- 1%, 1%-2%, 2%-3%, 3%-4%, 4%-5%, 5%-10%, 10%-20%, 20%-30%, 30%-40%, 40%-50%, 50%-60%, 60%-70%, 70%-80%, or 80%-90% of the size of the source cell, e.g., as measured by an assay of Example 27. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of less than about 500 nm (e.g., less than about 10, 50, 100, 150, 200, 250, 300, 350, 400, or 450 nm), e.g., as measured by an assay of Example 69. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of about 80-180, 90-170, 100-160, 110-150, 120-140, or 130 nm, e.g., as measured by an assay of Example 69. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of between about 11,000 nm and 21,000 nm, e.g., as measured by an assay of Example 69. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, between about 10-22,000, 12-20,000, 14-18,720 nm, 20-16,000 nm, e.g., as measured by an assay of Example 69. In embodiments, the fusosome has a volume, or the population of fusosomes has an average volume, of about 0.01-0.1 μm^3, 0.02-1 μm^3, 0.03-1 μm^3, 0.04-1 μm^3, 0.05-0.09 μm^3, 0.06-0.08 μm^3, 0.07 μm^3, e.g., as measured by an assay of Example 69. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of at least about 10 nm, 20 nm, 30 nm, 40 nm, 50 nm, 60 nm, 70 nm, 80 nm, 90 nm, 100 nm, 150 nm, 200 nm, or 250 nm e.g., as measured by an assay of Example 29. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of about 10 nm, 20 nm, 30 nm, 40 nm, 50 nm, 60 nm, 70 nm, 80 nm, 90 nm, 100 nm, 150 nm, 200 nm, or 250 nm (e.g., ±20%) e.g., as measured by an assay of Example 29. In

embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of at least about 500 nm, 750 nm, 1,000 nm, 1,500 nm, 2,000 nm, 2,500 nm, 3,000 nm, 5,000 nm, 10,000 nm, or 20,000 nm, e.g., as measured by an assay of Example 29. In embodiments, the fusosome has a diameter, or the population of fusosomes has an average diameter, of about 500 nm, 750 nm, 1,000 nm, 1,500 nm, 2,000 nm, 2,500 nm, 3,000 nm, 5,000 nm, 10,000 nm, or 20,000 nm (e.g., ±20%), e.g., as measured by an assay of Example 29. In embodiments, the population of fusosomes has (or is identified as having) one or more of: a 10% quantile diameter of about 40-90 nm, 45-60 nm, 50-55 nm or 53 nm; a 25% quantile diameter of about 70-100 nm, 80-95 nm, 85-90 nm, or 88 nm; a 75% quantile diameter of about 200-250 nm, 210-240 nm, 220-230 nm, or 226 nm; or a 90% quantile of about 4000-5000 nm, 4300-4600 nm, 4400-4500 nm, 4450 nm, e.g., by an assay of Example 68.

In embodiments, the fusosome composition comprises (or is identified as comprising) a GAPDH concentration of about 35-40, 36-39, 37-38, or 37.2 ng/mL, e.g., in an assay of Example 82. In embodiments, the GAPDH concentration of the fusosome composition is (or is identified as being) within about 1%, 2%, 5%, 10%, or 20% of the GAPDH concentration of the source cells, e.g., in an assay of Example 82. In embodiments, the GAPDH concentration of the fusosome composition is (or is identified as being) at least 1%, 2%, 5%, 10%, or 20% lower than the the GAPDH concentration of the source cells, e.g., in an assay of Example 82. In embodiments, the the fusosome composition comprises (or is identified as comprising) less than about 30, 35, 40, 45, 46, 47, 48, 49, 50, 55, 60, 65, or 70 µg GAPDH per gram total protein. In embodiments, the fusosome composition comprises (or is identified as comprising) less than about 500, 250, 100, or 50 µg GAPDH per gram total protein. In embodiments, the parental cell comprises (or is identified as comprising) at least 1%, 2.5%, 5%, 10%, 15%, 20%, 30%, 30%, 50%, or more GAPDH per total protein than the fusosome composition.

In embodiments, the average fractional content of calnexin in the fusosome is (or is identified as being) less than about $1x10^{-4}$, $1.5x10^{-4}$, $2x10^{-4}$, $2.1x10^{-4}$, $2.2x10^{-4}$, $2.3x10^{-4}$,; $2.4x10^{-4}$, $2.43x10^{-4}$, $2.5x10^{-4}$, $2.6x10^{-4}$, $2.7x10^{-4}$, $2.8x10^{-4}$, $2.9x10^{-4}$, $3x10^{-4}$, $3.5x10^{-4}$, or $4x10^{-4}$. In embodiments, the fusosome comprises an amount of calnexin per total protein that is lower than that of the parental cell by about 70%, 75%, 80%, 85%, 88%, 90%, 95%, 99%, or more.

In some embodiments, fusosomes comprise or are enriched for lipids that affect membrane curvature (see, e.g., Thiam et al., Nature Reviews Molecular Cell Biology, 14(12):

775-785, 2013). Some lipids have a small hydrophilic head group and large hydrophobic tails, which facilitate the formation of a fusion pore by concentrating in a local region. In some embodiments, fusosomes comprise or are enriched for negative-curvature lipids, such as cholesterol, phosphatidylethanolamine (PE), diglyceride (DAG), phosphatidic acid (PA), fatty acid (FA). In some embodiments, fusosomes do not comprise, are depleted of, or have few positive-curvature lipids, such as lysophosphatidylcholine (LPC), phosphatidylinositol (Ptdlns), lysophosphatidic acid (LPA), lysophosphatidylethanolamine (LPE), monoacylglycerol (MAG).

In some embodiments, the lipids are added to a fusosome. In some embodiments, the lipids are added to source cells in culture which incorporate the lipids into their membranes prior to or during the formation of a fusosome. In some embodiments, the lipids are added to the cells or fusosomes in the form of a liposome. In some embodiments methyl-betacyclodextrane (mβ-CD) is used to enrich or deplete lipids (see, e.g., Kainu et al, Journal of Lipid Research, 51(12): 3533-3541, 2010).

X. Pharmaceutical compositions and methods of making them

The present disclosure also provides, in some aspects, a pharmaceutical composition comprising the fusosome composition described herein and pharmaceutically acceptable carrier. The pharmaceutical compositions can include any of the described fusosomes, e.g. retroviral vectors, or VLPs.

In some embodiments, one or more transducing units of retroviral vector are administered to the subject. In some embodiments, at least 1, 10, 100, 1000, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, or 10^{14}, transducing units per kg are administered to the subject. In some embodiments at least 1, 10, 100, 1000, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, or 10^{14}, transducing units per target cell per ml of blood are administered to the subject.

Concentration and purification of retroviral virus, e.g. lentivirus

In some embodiments, a retroviral vector formulation described herein can be produced by a process comprising one or more of, e.g., all of, the following steps (i) to (vi), e.g., in chronological order:

(i) culturing cells that produce retroviral vector;

(ii) harvesting the retroviral vector containing supernatant;

(iii) optionally clarifying the supernatant;

(iv) purifying the retroviral vector to give a retroviral vector preparation;

(v) optionally filter-sterilization of the retroviral vector preparation; and

(vi) concentrating the retroviral vector preparation to produce the final bulk product.

In some embodiments the process does not comprise the clarifying step (iii). In other embodiments the process does include the clarifying step (iii). In some embodiments, step (vi) is performed using ultrafiltration, or tangential flow filtration, more preferably hollow fiber ultrafiltration. In some embodiments, the purification method in step (iv) is ion exchange chromatography, more preferably anion exchange chromatography. In some embodiments, the filter-sterilisation in step (v) is performed using a 0.22 μm or a 0.2 μm sterilising filter. In some embodiments, step (iii) is performed by filter clarification. In some embodiments, step (iv) is performed using a method or a combination of methods selected from chromatography, ultrafiltration/diafiltration, or centrifugation. In some embodiments, the chromatography method or a combination of methods is selected from ion exchange chromatography, hydrophobic interaction chromatography, size exclusion chromatography, affinity chromatography, reversed phase chromatography, and immobilized metal ion affinity chromatography. In some embodiments, the centrifugation method is selected from zonal centrifugation, isopycnic centrifugation and pelleting centrifugation. In some embodiments, the ultrafiltration/diafiltration method is selected from tangential flow diafiltration, stirred cell diafiltration and dialysis. In some embodiments, at least one step is included into the process to degrade nucleic acid to improve purification. In some embodiments, said step is nuclease treatment.

In some embodiments, concentration of the vectors is done before filtration. In some embodiments, concentration of the vectors is done after filtration. In some embodiments, concentration and filtrations steps are repeated.

In some embodiments, the final concentration step is performed after the filter-sterilisation step. In some embodiments, the process is a large scale-process for producing clinical grade formulations that are suitable for administration to humans as therapeutics. In some embodiments, the filter-sterilisation step occurs prior to a concentration step. In some embodiments, the concentration step is the final step in the process and the filter-sterilisation step is the penultimate step in the process. In some embodiments, the concentration step is performed

using ultrafiltration, preferably tangential flow filtration, more preferably hollow fiber ultrafiltration. In some embodiments, the filter-sterilisation step is performed using a sterilising filter with a maximum pore size of about 0.22 μm. In another preferred embodiment the maximum pore size is 0.2 μm

In some embodiments, the vector concentration is less than or equal to about 4.6×10^{11} RNA genome copies per ml of preparation prior to filter-sterilisation. The appropriate concentration level can be achieved through controlling the vector concentration using, e.g. a dilution step, if appropriate. Thus, in some embodiments, a retroviral vector preparation is diluted prior to filter sterilisation.

Clarification may be done by a filtration step, removing cell debris and other impurities. Suitable filters may utilize cellulose filters, regenerated cellulose fibers, cellulose fibers combined with inorganic filter aids (e.g. diatomaceous earth, perlite, fumed silica), cellulose filters combined with inorganic filter aids and organic resins, or any combination thereof, and polymeric filters (examples include but are not limited to nylon, polypropylene, polyethersulfone) to achieve effective removal and acceptable recoveries. A multiple stage process may be used. An exemplary two or three-stage process would consist of a coarse filter(s) to remove large precipitate and cell debris followed by polishing second stage filter(s) with nominal pore sizes greater than 0.2 micron but less than 1 micron. The optimal combination may be a function of the precipitate size distribution as well as other variables. In addition, single stage operations employing a relatively small pore size filter or centrifugation may also be used for clarification. More generally, any clarification approach including but not limited to dead-end filtration, microfiltration, centrifugation, or body feed of filter aids (e.g. diatomaceous earth) in combination with dead-end or depth filtration, which provides a filtrate of suitable clarity to not foul the membrane and/or resins in the subsequent steps, will be acceptable to use in the clarification step of the present invention.

In some embodiments, depth filtration and membrane filtration is used. Commercially available products useful in this regard are for instance mentioned in WO 03/097797, p. 20-21. Membranes that can be used may be composed of different materials, may differ in pore size, and may be used in combinations. They can be commercially obtained from several vendors. In some embodiments, the filter used for clarification is in the range of 1.2 to 0.22 μm. In some

embodiments, the filter used for clarification is either a 1.2/0.45 μm filter or an asymmetric filter with a minimum nominal pore size of 0.22 μm

In some embodiments, the method employs nuclease to degrade contaminating DNA/RNA, i.e. mostly host cell nucleic acids. Exemplary nucleases suitable for use in the present invention include Benzonase® Nuclease (EP 0229866) which attacks and degrades all forms of DNA and RNA (single stranded, double stranded linear or circular) or any other DNase and/or RNase commonly used within the art for the purpose of eliminating unwanted or contaminating DNA and/or RNA from a preparation. In preferred embodiments, the nuclease is Benzonase® Nuclease, which rapidly hydrolyzes nucleic acids by hydrolyzing internal phosphodiester bonds between specific nucleotides, thereby reducing the size of the polynucleotides in the vector containing supernatant. Benzonase® Nuclease can be commercially obtained from Merck KGaA (code W214950). The concentration in which the nuclease is employed is preferably within the range of 1-100 units/ml.

In some embodiments, the vector suspension is subjected to ultrafiltration (sometimes referred to as diafiltration when used for buffer exchange) at least once during the process, e.g. for concentrating the vector and/or buffer exchange. The process used to concentrate the vector can include any filtration process (e.g., ultrafiltration (UF)) where the concentration of vector is increased by forcing diluent to be passed through a filter in such a manner that the diluent is removed from the vector preparation whereas the vector is unable to pass through the filter and thereby remains, in concentrated form, in the vector preparation. UF is described in detail in, e.g., Microfiltration and Ultrafiltration: Principles and Applications, L. Zeman and A. Zydney (Marcel Dekker, Inc., New York, N.Y., 1996); and in: Ultrafiltration Handbook, Munir Cheryan (Technomic Publishing, 1986; ISBN No. 87762-456-9). A suitable filtration process is Tangential Flow Filtration ("TFF") as described in, e.g., MILLIPORE catalogue entitled "Pharmaceutical Process Filtration Catalogue" pp. 177-202 (Bedford, Mass., 1995/96). TFF is widely used in the bioprocessing industry for cell harvesting, clarification, purification and concentration of products including viruses. The system is composed of three distinct process streams: the feed solution, the permeate and the retentate. Depending on application, filters with different pore sizes may be used. In some embodiments, the retentate contains the product (lentiviral vector). The particular ultrafiltration membrane selected may have a pore size sufficiently small to retain vector but large enough to effectively clear impurities. Depending on

the manufacturer and membrane type, for retroviral vectors nominal molecular weight cutoffs (NMWC) between 100 and 1000 kDa may be appropriate, for instance membranes with 300 kDa or 500 kDa NMWC. The membrane composition may be, but is not limited to, regenerated cellulose, polyethersulfone, polysulfone, or derivatives thereof. The membranes can be flat sheets (also called flat screens) or hollow fibers. A suitable UF is hollow fibre UF, e.g., filtration using filters with a pore size of smaller than 0.1 μm. Products are generally retained, while volume can be reduced through permeation (or be kept constant during diafiltration by adding buffer with the same speed as the speed with which the permeate, containing buffer and impurities, is removed at the permeate side).

The two most widely used geometries for TFF in the biopharmaceutical industry are plate & frame (flat screens) and hollow fiber modules. Hollow fiber units for ultrafiltration and microfiltration were developed by Amicon and Ramicon in the early 1970s (Cheryan, M. Ultrafiltration Handbook), even though now there are multiple vendors including Spectrum and GE Healthcare. The hollow fiber modules consist of an array of self-supporting fibers with a dense skin layer. Fiber diameters range from 0.5 mm-3 mm. In certain embodiments, hollow fibers are used for TFF. In certain embodiments, hollow fibers of 500 kDa (0.05 μm) pore size are used. Ultrafiltration may comprise diafiltration (DF). Microsolutes can be removed by adding solvent to the solution being ultrafiltered at a rate equal to the UF rate. This washes microspecies from the solution at a constant volume, purifying the retained vector.

UF/DF can be used to concentrate and/or buffer exchange the vector suspensions in different stages of the purification process. The method can utilize a DF step to exchange the buffer of the supernatant after chromatography or other purification steps, but may also be used prior to chromatography.

In some embodiments, the eluate from the chromatography step is concentrated and further purified by ultrafiltration-diafiltration. During this process the vector is exchanged into formulation buffer. Concentration to the final desired concentration can take place after the filter-sterilisation step. After said sterile filtration, the filter sterilised substance is concentrated by aseptic UF to produce the bulk vector product.

In embodiments, the ultrafiltration/diafiltration may be tangential flow diafiltration, stirred cell diafiltration and dialysis.

Purification techniques tend to involve the separation of the vector particles from the cellular milieu and, if necessary, the further purification of the vector particles. One or more of a variety of chromatographic methods may be used for this purification. Ion exchange, and more particularly anion exchange, chromatography is a suitable method, and other methods could be used. A description of some chromatographic techniques is given below.

Ion-exchange chromatography utilises the fact that charged species, such as biomolecules and viral vectors, can bind reversibly to a stationary phase (such as a membrane, or else the packing in a column) that has, fixed on its surface, groups that have an opposite charge. There are two types of ion exchangers. Anion exchangers are stationary phases that bear groups having a positive charge and hence can bind species with a negative charge. Cation exchangers bear groups with a negative charge and hence can bind species with positive charge. The pH of the medium has an influence on this, as it can alter the charge on a species. Thus, for a species such as a protein, if the pH is above the pI, the net charge will be negative, whereas below the pI, the net charge will be positive.

Displacement (elution) of the bound species can be effected by the use of suitable buffers. Thus commonly the ionic concentration of the buffer is increased until the species is displaced through competition of buffer ions for the ionic sites on the stationary phase. An alternative method of elution entails changing the pH of the buffer until the net charge of the species no longer favours biding to the stationary phase. An example would be reducing the pH until the species assumes a net positive charge and will no longer bind to an anion exchanger.

Some purification can be achieved if impurities are uncharged, or else if they bear a charge of opposite sign to that of the desired species, but the same sign to that on the ion exchanger. This is because uncharged species and those having a charge of the same sign to that an ion exchanger, will not normally bind. For different bound species, the strength of the binding varies with factors such as the charge density and the distribution of charges on the various species. Thus by applying an ionic or pH gradient (as a continuous gradient, or as a series of steps), the desired species might be eluted separately from impurities.

Size exclusion chromatography is a technique that separates species according to their size. Typically it is performed by the use of a column packed with particles having pores of a well-defined size. For the chromatographic separation, particles are chosen that have pore sizes that are appropriate with regard to the sizes of the species in the mixture to be separated. When

the mixture is applied, as a solution (or suspension, in the case of a virus), to the column and then eluted with buffer, the largest particles will elute first as they have limited (or no) access to the pores. Smaller particles will elute later as they can enter the pores and hence take a longer path through the column. Thus in considering the use of size exclusion chromatography for the purification of viral vectors, it would be expected that the vector would be eluted before smaller impurities such as proteins.

Species, such as proteins, have on their surfaces, hydrophobic regions that can bind reversibly to weakly hydrophobic sites on a stationary phase. In media having a relatively high salt concentration, this binding is promoted. Typically in HIC the sample to be purified is bound to the stationary phase in a high salt environment. Elution is then achieved by the application of a gradient (continuous, or as a series of steps) of decreasing salt concentration. A salt that is commonly used is ammonium sulphate. Species having differing levels of hydrophobicity will tend to be eluted at different salt concentrations and so the target species can be purified from impurities. Other factors, such as pH, temperature and additives to the elution medium such as detergents, chaotropic salts and organics can also influence the strength of binding of species to HIC stationary phases. One, or more, of these factors can be adjusted or utilised to optimise the elution and purification of product.

Viral vectors have on their surface, hydrophobic moieties such as proteins, and thus HIC could potentially be employed as a means of purification.

Like HIC, RPC separates species according to differences in their hydrophobicities. A stationary phase of higher hydrophobicity than that employed in HIC is used. The stationary phase often consists of a material, typically silica, to which are bound hydrophobic moieties such as alkyl groups or phenyl groups. Alternatively the stationary phase might be an organic polymer, with no attached groups. The sample-containing the mixture of species to be resolved is applied to the stationary phase in an aqueous medium of relatively high polarity which promotes binding. Elution is then achieved by reducing the polarity of the aqueous medium by the addition of an organic solvent such as isopropanol or acetonitrile. Commonly a gradient (continuous, or as a series of steps) of increasing organic solvent concentration is used and the species are eluted in order of their respective hydrophobicities.

Other factors, such as the pH of the elution medium, and the use of additives, can also influence the strength of binding of species to RPC stationary phases. One, or more, of these

factors can be adjusted or utilised to optimise the elution and purification of product. A common additive is trifluororacetic acid (TFA). This suppresses the ionisation of acidic groups such as carboxyl moieties in the sample. It also reduces the pH in the eluting medium and this suppresses the ionisation of free silanol groups that may be present on the surface of stationary phases having a silica matrix. TFA is one of a class of additives known as ion pairing agents. These interact with ionic groups, present on species in the sample, that bear an opposite charge. The interaction tends to mask the charge, increasing the hydrophobicity of the species. Anionic ion pairing agents, such as TFA and pentafluoropropionic acid interact with positively charged groups on a species. Cationic ion pairing agents such, as triethylamine, interact with negatively charged groups.

Viral vectors have on their surface, hydrophobic moieties such as proteins, and thus RPC, potentially, could be employed as a means of purification.

Affinity chromatography utilises the fact that certain ligands that bind specifically with biomolecules such as proteins or nucleotides, can be immobilised on a stationary phase. The modified stationary phase can then be used to separate the relevant biomolecule from a mixture. Examples of highly specific ligands are antibodies, for the purification of target antigens and enzyme inhibitors for the purification of enzymes. More general interactions can also be utilised such as the use of the protein A ligand for the isolation of a wide range of antibodies.

Typically, affinity chromatography is performed by application of a mixture, containing the species of interest, to the stationary phase that has the relevant ligand attached. Under appropriate conditions this will lead to the binding of the species to the stationary phase. Unbound components are then washed away before an eluting medium is applied. The eluting medium is chosen to disrupt the binding of the ligand to the target species. This is commonly achieved by choice of an appropriate ionic strength, pH or by the use of substances that will compete with the target species for ligand sites. For some bound species, a chaotropic agent such as urea is used to effect displacement from the ligand. This, however, can result in irreversible denaturation of the species.

Viral vectors have on their surface, moieties such as proteins, that might be capable of binding specifically to appropriate ligands. This means that, potentially, affinity chromatography could be used in their isolation.

Biomolecules, such as proteins, can have on their surface, electron donating moieties that can form coordinate bonds with metal ions. This can facilitate their binding to stationary phases carrying immobilised metal ions such as Ni^{2+}, Cu^{2+}, Zn^{2+} or Fe^{3+}. The stationary phases used in IMAC have chelating agents, typically nitriloacetic acid or iminodiacetic acid covalently attached to their surface and it is the chelating agent that holds the metal ion. It is necessary for the chelated metal ion to have at least one coordination site left available to form a coordinate bond to a biomolecule. Potentially there are several moieties on the surface of biomolecules that might be capable of bonding to the immobilised metal ion. These include histidine, tryptophan and cysteine residues as well as phosphate groups. For proteins, however, the predominant donor appears to be the imidazole group of the histidine residue. Native proteins can be separated using IMAC if they exhibit suitable donor moieties on their surface. Otherwise IMAC can be used for the separation of recombinant proteins bearing a chain of several linked histidine residues.

Typically, IMAC is performed by application of a mixture, containing the species of interest, to the stationary phase. Under appropriate conditions this will lead to the coordinate bonding of the species to the stationary phase. Unbound components are then washed away before an eluting medium is applied. For elution, gradients (continuous, or as a series of steps) of increasing salt concentration or decreasing pH may be used. Also a commonly used procedure is the application of a gradient of increasing imidazole concentration. Biomolecules having different donor properties, for example having histidine residues in differing environments, can be separated by the use of gradient elution.

Viral vectors have on their surface, moieties such as proteins, that might be capable of binding to IMAC stationary phases. This means that, potentially, IMAC could be used in their isolation.

Suitable centrifugation techniques include zonal centrifugation, isopycnic ultra and pelleting centrifugation.

Filter-sterilisation is suitable for processes for pharmaceutical grade materials. Filter-sterilisation renders the resulting formulation substantially free of contaminants. The level of contaminants following filter-sterilisation is such that the formulation is suitable for clinical use. Further concentration (e.g. by ultrafiltration) following the filter-sterilisation step may be performed in aseptic conditions. In some embodiments, the sterilising filter has a maximum pore size of 0.22 μm.

The retroviral vectors herein can also be subjected to methods to concentrate and purify a lentiviral vector using flow-through ultracentrifugation and high-speed centrifugation, and tangential flow filtration. Flow through ultracentrifugation can be used for the purification of RNA tumor viruses (Toplin et al, Applied Microbiology 15:582-589, 1967; Burger et al., Journal of the National Cancer Institute 45: 499-503, 1970). Flow-through ultracentrifugation can be used for the purification of Lentiviral vectors. This method can comprise one or more of the following steps. For example, a lentiviral vector can be produced from cells using a cell factory or bioreactor system. A transient transfection system can be used or packaging or producer cell lines can also similarly be used. A pre-clarification step prior to loading the material into the ultracentrifuge could be used if desired. Flow-through ultracentrifugation can be performed using continuous flow or batch sedimentation. The materials used for sedimentation are, e.g.: Cesium chloride, potassium tartrate and potassium bromide, which create high densities with low viscosity although they are all corrosive. CsCl is frequently used for process development as a high degree of purity can be achieved due to the wide density gradient that can be created (1.0 to 1.9 g/cm^3). Potassium bromide can be used at high densities, e.g., at elevated temperatures, such as 25° C., which may be incompatible with stability of some proteins. Sucrose is widely used due to being inexpensive, non-toxic and can form a gradient suitable for separation of most proteins, sub-cellular fractions and whole cells. Typically the maximum density is about 1.3 g/cm^3. The osmotic potential of sucrose can be toxic to cells in which case a complex gradient material can be used, e.g. Nycodenz. A gradient can be used with 1 or more steps in the gradient. An embodiment is to use a step sucrose gradient. The volume of material can be from 0.5 liters to over 200 liters per run. The flow rate speed can be from 5 to over 25 liters per hour. A suitable operating speed is between 25,000 and 40,500 rpm producing a force of up to 122,000×g. The rotor can be unloaded statically in desired volume fractions. An embodiment is to unload the centrifuged material in 100 ml fractions. The isolated fraction containing the purified and concentrated Lentiviral vector can then be exchanged in a desired buffer using gel filtration or size exclusion chromatography. Anionic or cationic exchange chromatography could also be used as an alternate or additional method for buffer exchange or further purification. In addition, Tangential Flow Filtration can also be used for buffer exchange and final formulation if required. Tangential Flow Filtration (TFF) can also be used as an alternative step to ultra or high speed centrifugation, where a two step TFF procedure would be implemented. The first step would

reduce the volume of the vector supernatant, while the second step would be used for buffer exchange, final formulation and some further concentration of the material. The TFF membrane can have a membrane size of between 100 and 500 kilodaltons, where the first TFF step can have a membrane size of 500 kilodaltons, while the second TFF can have a membrane size of between 300 to 500 kilodaltons. The final buffer should contain materials that allow the vector to be stored for long term storage.

In embodiments, the method uses either cell factories that contains adherent cells, or a bioreactor that contains suspension cells that are either transfected or transduced with the vector and helper constructs to produce lentiviral vector. Non limiting examples or bioreactors, include the Wave bioreactor system and the Xcellerex bioreactors. Both are disposable systems. However non-disposable systems can also be used. The constructs can be those described herein, as well as other lentiviral transduction vectors. Alternatively the cell line can be engineered to produce Lentiviral vector without the need for transduction or transfection. After transfection, the lentiviral vector can be harvested and filtered to remove particulates and then is centrifuged using continuous flow high speed or ultra centrifugation. A preferred embodiment is to use a high speed continuous flow device like the JCF-A zonal and continuous flow rotor with a high speed centrifuge. Also preferably is the use of Contifuge Stratus centrifuge for medium scale Lentiviral vector production. Also suitable is any continuous flow centrifuge where the speed of centrifugation is greater than 5,000×g RCF and less than 26,000×g RCF. Preferably, the continuous flow centrifugal force is about 10,500×g to 23,500×g RCF with a spin time of between 20 hours and 4 hours, with longer centrifugal times being used with slower centrifugal force. The lentiviral vector can be centrifuged on a cushion of more dense material (a non limiting example is sucrose but other reagents can be used to form the cushion and these are well known in the art) so that the Lentiviral vector does not form aggregates that are not filterable, as sometimes occurrs with straight centrifugation of the vector that results in a viral vector pellet. Continuous flow centrifugation onto a cushion allows the vector to avoid large aggregate formation, yet allows the vector to be concentrated to high levels from large volumes of transfected material that produces the Lentiviral vector. In addition, a second less-dense layer of sucrose can be used to band the Lentiviral vector preparation. The flow rate for the continuous flow centrifuge can be between 1 and 100 ml per minute, but higher and lower flow rates can also be used. The flow rate is adjusted to provide ample time for the vector to enter the core of

the centrifuge without significant amounts of vector being lost due to the high flow rate. If a higher flow rate is desired, then the material flowing out of the continuous flow centrifuge can be re-circulated and passed through the centrifuge a second time. After the virus is concentrated using continuous flow centrifugation, the vector can be further concentrated using Tangential Flow Filtration (TFF), or the TFF system can be simply used for buffer exchange. A non-limiting example of a TFF system is the Xampler cartridge system that is produced by GB-Healthcare. Preferred cartridges are those with a MW cut-off of 500,000 MW or less. Preferably a cartridge is used with a MW cut-off of 300,000 MW. A cartridge of 100,000 MW cut-off can also be used. For larger volumes, larger cartridges can be used and it will be easy for those in the art to find the right TFF system for this final buffer exchange and/or concentration step prior to final fill of the vector preparation. The final fill preparation may contain factors that stabilize the vector—sugars are generally used and are known in the art.

Protein content

In some embodiments the retroviral particle includes various source cell genome-derived proteins, exogenous proteins, and viral-genome derived proteins. In some embodiments the retroviral particle contains various ratios of source cell genome-derived proteins to viral-genome-derived proteins, source cell genome-derived proteins to exogenous proteins, and exogenous proteins to viral-genome derived proteins.

In some embodiments, the viral-genome derived proteins are GAG polyprotein precursor, HIV-1 Integrase, POL polyprotein precursor, Capsid, Nucleocapsid, p17 matrix, p6, p2, VPR, Vif.

In some embodiments, the source cell-derived proteins are Cyclophilin A, Heat Shock 70kD, Human Elongation Factor-1 Alpha (EF-1R), Histones H1, H2A, H3, H4, beta-globin, Trypsin Precursor, Parvulin, Glyceraldehyde-3-phosphate dehydrogenase, Lck, Ubiquitin, SUMO-1, CD48, Syntenin-1, Nucleophosmin, Heterogeneous nuclear ribonucleoproteins C1/C2, Nucleolin, Probable ATP-dependent helicase DDX48, Matrin-3, Transitional ER ATPase, GTP-binding nuclear protein Ran, Heterogeneous nuclear ribonucleoprotein U, Interleukin enhancer binding factor 2, Non-POU domain containing octamer binding protein, RuvB like 2, HSP 90-b, HSP 90-a, Elongation factor 2, D-3-phosphoglycerate dehydrogenase, a-enolase, C-1-tetrahydrofolate synthase, cytoplasmic, Pyruvate kinase, isozymes M1/M2, Ubiquitin activating

enzyme E1, 26S protease regulatory subunit S10B, 60S acidic ribosomal protein P2, 60S acidic ribosomal protein P0, 40S ribosomal protein SA, 40S ribosomal protein S2, 40S ribosomal protein S3, 60S ribosomal protein L4, 60S ribosomal protein L3, 40S ribosomal protein S3a, 40S ribosomal protein S7, 60S ribosomal protein L7a, 60S acidic ribosomal protein L31, 60S ribosomal protein L10a, 60S ribosomal protein L6, 26S proteasome non-ATPase regulatory subunit 1, Tubulin b-2 chain, Actin, cytoplasmic 1, Actin, aortic smooth muscle, Tubulin a-ubiquitous chain, Clathrin heavy chain 1, Histone H2B.b, Histone H4, Histone H3.1, Histone H3.3, Histone H2A type 8, 26S protease regulatory subunit 6A, Ubiquitin-4, RuvB like 1, 26S protease regulatory subunit 7, Leucyl-tRNA synthetase, cytoplasmic, 60S ribosomal protein L19, 26S proteasome non-ATPase regulatory subunit 13, Histone H2B.F, U5 small nuclear ribonucleoprotein 200 kDa helicase, Poly[ADP-ribose]polymerase-1, ATP-dependent DNA helicase II, DNA replication licensing factor MCM5, Nuclease sensitive element binding protein 1, ATP-dependent RNA helicase A, Interleukin enhancer binding factor 3, Transcription elongation factor B polypeptide 1, Pre-mRNA processing splicing factor 8, Staphylococcal nuclease domain containing protein 1, Programmed cell death 6-interacting protein, Mediator of RNA polymerase II transcription subunit 8 homolog, Nucleolar RNA helicase II, Endoplasmin, DnaJ homolog subfamily A member 1, Heat shock 70 kDa protein 1L, T-complex protein 1 e subunit, GCN1-like protein 1, Serotransferrin, Fructose bisphosphate aldolase A, Inosine-5'monophosphate dehydrogenase 2, 26S protease regulatory subunit 6B, Fatty acid synthase, DNA-dependent protein kinase catalytic subunit, 40S ribosomal protein S17, 60S ribosomal protein L7, 60S ribosomal protein L12, 60S ribosomal protein L9, 40S ribosomal protein S8, 40S ribosomal protein S4 X isoform, 60S ribosomal protein L11, 26S proteasome non-ATPase regulatory subunit 2, Coatomer a subunit, Histone H2A.z, Histone H1.2, Dynein heavy chain cytosolic. See: Saphire et al., *Journal of Proteome Research,* 2005, and Wheeler et al., *Proteomics Clinical Applications*, 2007.

In some embodiments the retroviral vector is pegylated.

Particle size

In some embodiments the median retroviral vector diameter is between 10 and 1000 nM, 25 and 500 nm 40 and 300 nm, 50 and 250 nm, 60 and 225 nm, 70 and 200 nm, 80 and 175 nm, or 90 and 150 nm.

In some embodiments, 90% of the retroviral vectors fall within 50% of the median diameter of the retrovirus. In some embodiments, 90% of the retroviral vectors fall within 25% of the median diameter of the retrovirus. In some embodiments, 90% of the retroviral vectors fall within 20% of the median diameter of the retrovirus. In some embodiments, 90% of the retroviral vectors fall within 15% of the median diameter of the retrovirus. In some embodiments, 90% of the retroviral vectors fall within 10% of the median diameter of the retrovirus.

XI. Indications and uses

The fusosomes, retroviral vectors, VLPs, or pharmaceutical compositions described herein can be administered to a subject, e.g., a mammal, e.g., a human. In such embodiments, the subject may be at risk of, may have a symptom of, or may be diagnosed with or identified as having, a particular disease or condition (e.g., a disease or condition described herein). In some embodiments, the disease is a genetic deficiency, e.g., a genetic deficiency listed in Table 5 or Table 6.

This disclosure also provides, in certain aspects, a method of administering a fusosome composition to a subject (e.g., a human subject), a target tissue, or a cell, comprising administering to the subject, or contacting the target tissue or the cell with a fusosome composition comprising a plurality of fusosomes described herein, a fusosome composition described herein, or a pharmaceutical composition described herein, thereby administering the fusosome composition to the subject.

This disclosure also provides, in certain aspects, a method of delivering a therapeutic agent (e.g., a polypeptide, a nucleic acid, a metabolite, an organelle, or a subcellular structure) to a subject, a target tissue, or a cell, comprising administering to the subject, or contacting the target tissue or the cell with, a plurality of fusosomes described herein, a fusosome composition comprising a plurality of fusosomes described herein, a fusosome composition described herein, or a pharmaceutical composition described herein, wherein the fusosome composition is administered in an amount and/or time such that the therapeutic agent is delivered.

This disclosure also provides, in certain aspects, a method of delivering a function to a subject, a target tissue, or a cell, comprising administering to the subject, or contacting the target tissue or the cell with, a plurality of fusosomes described herein, a fusosome composition comprising a plurality of fusosomes described herein, a fusosome composition described herein, or a pharmaceutical composition described herein, wherein the fusosome composition is administered in an amount and/or time such that the function is delivered.

Target cells from mammalian (e.g., human) tissue include cells from epithelial, connective, muscular, or nervous tissue or cells, and combinations thereof. Target mammalian (e.g., human) cells and organ systems include the cardiovascular system (heart, vasculature); digestive system (esophagus, stomach, liver, gallbladder, pancreas, intestines, colon, rectum and anus); endocrine system (hypothalamus, pituitary gland, pineal body or pineal gland, thyroid, parathyroids, adrenal glands); excretory system (kidneys, ureters, bladder); lymphatic system (lymph, lymph nodes, lymph vessels, tonsils, adenoids, thymus, spleen); integumentary system (skin, hair, nails); muscular system (e.g., skeletal muscle); nervous system (brain, spinal cord, nerves)'; reproductive system (ovaries, uterus, mammary glands, testes, vas deferens, seminal vesicles, prostate); respiratory system (pharynx, larynx, trachea, bronchi, lungs, diaphragm); skeletal system (bone, cartilage), and combinations thereof. In some embodiments, a non-target cells or organ system is chosen from the cardiovascular system (heart, vasculature); digestive system (esophagus, stomach, liver, gallbladder, pancreas, intestines, colon, rectum and anus); endocrine system (hypothalamus, pituitary gland, pineal body or pineal gland, thyroid, parathyroids, adrenal glands); excretory system (kidneys, ureters, bladder); lymphatic system (lymph, lymph nodes, lymph vessels, tonsils, adenoids, thymus, spleen); integumentary system (skin, hair, nails); muscular system (e.g., skeletal muscle); nervous system (brain, spinal cord, nerves)'; reproductive system (ovaries, uterus, mammary glands, testes, vas deferens, seminal vesicles, prostate); respiratory system (pharynx, larynx, trachea, bronchi, lungs, diaphragm); skeletal system (bone, cartilage), and combinations thereof.

The administration of a pharmaceutical composition described herein may be by way of oral, inhaled, transdermal or parenteral (including intravenous, intratumoral, intraperitoneal, intramuscular, intracavity, and subcutaneous) administration. The fusosomes may be administered alone or formulated as a pharmaceutical composition.

In embodiments, the fusosome composition mediates an effect on a target cell, and the effect lasts for at least 1, 2, 3, 4, 5, 6, or 7 days, 2, 3, or 4 weeks, or 1, 2, 3, 6, or 12 months. In some embodiments (e.g., wherein the fusosome composition comprises an exogenous protein), the effect lasts for less than 1, 2, 3, 4, 5, 6, or 7 days, 2, 3, or 4 weeks, or 1, 2, 3, 6, or 12 months.

In embodiments, the fusosome composition described herein is delivered ex-vivo to a cell or tissue, e.g., a human cell or tissue.

The fusosome compositions described herein can be administered to a subject, e.g., a mammal, e.g., a human. In such embodiments, the subject may be at risk of, may have a symptom of, or may be diagnosed with or identified as having, a particular disease or condition (e.g., a disease or condition described herein).

In some embodiments, the source of fusosomes are from the same subject that is administered a fusosome composition. In other embodiments, they are different. For example, the source of fusosomes and recipient tissue may be autologous (from the same subject) or heterologous (from different subjects). In either case, the donor tissue for fusosome compositions described herein may be a different tissue type than the recipient tissue. For example, the donor tissue may be muscular tissue and the recipient tissue may be connective tissue (e.g., adipose tissue). In other embodiments, the donor tissue and recipient tissue may be of the same or different type, but from different organ systems.

In some embodiments, the fusosome is co-administered with an inhibitor of a protein that inhibits membrane fusion. For example, Suppressyn is a human protein that inhibits cell-cell fusion (Sugimoto et al., "A novel human endogenous retroviral protein inhibits cell-cell fusion" Scientific Reports 3:1462 DOI: 10.1038/srep01462). Thus, in some embodiments, the fusosome is co-administered with an inhibitor of sypressyn, e.g., a siRNA or inhibitory antibody.

Compositions described herein may also be used to similarly modulate the cell or tissue function or physiology of a variety of other organisms including but not limited to: farm or working animals (horses, cows, pigs, chickens etc.), pet or zoo animals (cats, dogs, lizards, birds, lions, tigers and bears etc.), aquaculture animals (fish, crabs, shrimp, oysters etc.), plants species (trees, crops, ornamentals flowers etc), fermentation species (saccharomyces etc.). Fusosome compositions described herein can be made from such non-human sources and administered to a non-human target cell or tissue or subject.

Fusosome compositions can be autologous, allogeneic or xenogeneic to the target.

XII. Additional therapeutic agents

In some embodiments, the fusosome composition is co-administered with an additional agent, e.g., a therapeutic agent, to a subject, e.g., a recipient, e.g., a recipient described herein. In some embodiments, the co-administered therapeutic agent is an immunosuppressive agent, e.g., a glucocorticoid (e.g., dexamethasone), cytostatic (e.g., methotrexate), antibody (e.g., Muromonab-CD3), or immunophilin modulator (e.g., Ciclosporin or rapamycin). In embodiments, the immunosuppressive agent decreases immune mediated clearance of fusosomes. In some embodiments the fusosome composition is co-administered with an immunostimulatory agent, e.g., an adjuvant, an interleukin, a cytokine, or a chemokine.

In some embodiments, the fusosome composition and the immunosuppressive agent are administered at the same time, e.g., contemporaneously administered. In some embodiments, the fusosome composition is administered before administration of the immunosuppressive agent. In some embodiments, the fusosome composition is administered after administration of the immunosuppressive agent.

In some embodiments, the immunosuppressive agent is a small molecule such as ibuprofen, acetaminophen, cyclosporine, tacrolimus, rapamycin, mycophenolate, cyclophosphamide, glucocorticoids, sirolimus, azathriopine, or methotrexate.

In some embodiments, the immunosuppressive agent is an antibody molecule, including but not limited to: muronomab (anti-CD3), Daclizumab (anti-IL12), Basiliximab, Infliximab (Anti-TNFa), or rituximab (Anti-CD20).

In some embodiments, co-administration of the fusosome composition with the immunosuppressive agent results in enhanced persistence of the fusosome composition in the subject compared to administration of the fusosome composition alone. In some embodiments, the enhanced persistence of the fusosome composition in the co-administration is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or longer, compared to persistence of the fusosome composition when administered alone. In some embodiments, the enhanced persistence of the fusosome composition in the co-administration is at least 1, 2, 3, 4, 5, 6, 7, 10, 15, 20, 25, or 30 days or longer, compared to survival of the fusosome composition when administered alone.

EXAMPLES

The Examples below are set forth to aid in the understanding of the inventions, but are not intended to, and should not be construed to, limit its scope in any way.

Example 1. Assaying off-target cells to detect specificity of retroviral nucleic acid delivery

This Example describes quantification of a nucleic acid in off-target recipient cells by measuring vector copy number in single cells.

In an embodiment, treated mice have a similar vector copy number in off-target cells as those from untreated mice, e.g., no vector or a vector number similar to negative control levels. In an embodiment, treated mice have a similar percent of off-target cells that contain the vector as those from untreated mice, e.g., no cells or a cell number similar to negative control levels.

In this example, the off-target recipient cell is a CD11c+ cell. However, this protocol may be adapted to any cell type for which suitable surface markers exist and which can be isolated from the subject. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol.

Mice are treated with retroviral vector produced as described herein or with PBS (negative control). 28 days following treatment, peripheral blood is collected from mice that received retroviral vector and mice that received PBS treatment. Blood is collected into 1ml PBS containing 5 μM EDTA and mixed immediately to prevent clotting. The tubes are kept on ice and red blood cells are removed using a buffered ammonium chloride (ACK) solution. Cells are stained with a murine CD11c:APC-Cy7 antibody (Biolegend Catalog #: 117323) or an isotype control APC-Cy7 antibody (Biolegend Catalog #: 400230) at 4°C for 30 minutes in the dark, after being Fc blocked (Biolegend Catalog #: 101319) in cell staining buffer (Biolegend Catalog #: 420201) for 10 minutes. After being washed two times with PBS, cells are analyzed on a FACS Aria (BD Biosciences, San Jose, CA.) with 640nm laser excitation and emission collected at 780 -/+ 60 nm running the FACSDiva™ software (BD Biosciences, San Jose, CA) to set negative gates using the isotype control APC-Cy7 antibody labeled cells. APC-Cy7 positive cells are sorted into single wells of plate for vector copy number analysis.

Vector copy number is assessed using single-cell nested PCR. PCR is performed with qPCR using primers and probes specific to the vector and an endogenous control gene. Vector

copy number is determined by dividing the amount of vector qPCR signal by the amount of the endogenous control gene qPCR signal. A cell that received the vector will have a vector copy number of at least 1.0. Vector copy number is assessed across the population by averaging the vector copy number of the plurality of cells

In some embodiments, mice treated with retroviral vectors have a similar average vector copy number in off-target cells as those from mice treated with vehicle. In some embodiments, mice treated with treated with retroviral vectors have a similar percent of off-target cells that received the vector as those from mice treated with vehicle.

Example 2. Assaying off-target cells to detect specificity of delivery of an exogenous protein agent

This Example describes quantification of the expression of an exogenous agent in off-target recipient cells by exogenous agent expression in single cells.

In an embodiment, treated mice have similar exogenous agent expression in off-target cells as those from untreated mice. In an embodiment, treated mice have a similar percent of off-target cells that express the exogenous agent as those from untreated mice.

In this example, the off-target recipient cell is a CD11c+ cell. However, this protocol may be adapted to any cell type for which suitable surface markers exist and which can be isolated from the subject. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol. In this example the exogenous agent is a fluorescent protein and expression is measured via flow cytometry. In other embodiments, the expression of an exogenous protein agent may be measured with immunostaining for the protein. In other embodiments expression of the exogenous protein agent may be measured via microscopy or western blot.

Mice are treated with retroviral vector with a tdtomato fluorescent protein agent produced via any of the methods described in this application or with PBS (negative control). 28 days following treatment, peripheral blood is collected from mice that received retroviral vector and mice that received PBS treatment. Blood is collected into 1ml PBS containing 5 μM EDTA and mixed immediately to prevent clotting. The tubes are kept on ice and red blood cells are removed using a buffered ammonium chloride (ACK) solution. Cells are stained with a murine CD11c:APC-Cy7 antibody (Biolgend Catalog #: 117323)or isotype controls APC-Cy7 antibody

(Biolegend Catalog #: 400230) at 4°C for 30 minutes in the dark, after being Fc blocked (Biolegend Catalog #: 101319) in cell staining buffer (Biolegend Catalog #: 420201) for 10 minutes. After being washed two times with PBS, cells are analyzed on a FACS Aria (BD Biosciences, San Jose, CA.) running the FACSDiva™ software (BD Biosciences, San Jose, CA). A negative gate for CD11c is set using the isotype control APC-Cy7 antibody labeled cells and with a 640nm laser excitation and emission collected at 780 -/+ 60. A negative gate for tdtomato expression is set with cells isolated from mice treated with vehicle and with a 552 nm laser excitation and an emission collected at 585 -/+ 42 nm.

The percent of CD11c+ cells that are tdtomato positive is measured. In some embodiments, the percent of CD11c+ cells that are tdtomato positive is similar in cells from treated and untreated mice. The median tdtomato fluorescence level is measured in CD11c+ cells. In some embodiments, the median tdtomato fluorescence level in CD11c+ cells is similar in cells from treated and untreated mice.

Example 3. Assaying target cells to detect specificity of retroviral nucleic acid delivery

This Example describes quantification of a nucleic acid in target recipient cells by measuring vector copy number in single cells.

In an embodiment, treated mice have a greater vector copy number in target cells than those from untreated mice. In an embodiment, treated mice have a greater percent of target cells that contain the vector than those from untreated mice.

In this example, the target recipient cell is a CD3+ cell. However, this protocol may be adapted to any cell type for which suitable surface markers exist and which can be isolated from the subject. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol.

Mice are treated with retroviral vector and a blood sample is collected as described above in Example 1. Cells are stained with a murine CD3:APC-Cy7 antibody (Biolegend Catalog #: 100330) or an isotype control using the protocol described above in Example 1. Vector copy number is assessed using single-cell nested PCR as described in Example 1.

In some embodiments, mice treated with retroviral vectors have a greater average vector copy number in target cells than those from mice treated with vehicle. In some

embodiments, mice treated with treated with retroviral vectors have a greater percent of target cells that received the vector than those from mice treated with vehicle.

Example 4. Assaying target cells to detect specificity of delivery of an exogenous protein agent

This Example describes quantification of the expression of an exogenous protein agent in target recipient cells by exogenous protein agent expression in single cells.

In an embodiment, treated mice have greater exogenous protein agent expression in target cells than those from untreated mice. In an embodiment, treated mice have a greater percent of target cells that express the exogenous protein agent than those from untreated mice.

In this example, the target recipient cell is a CD3+ cell. However, this protocol may be adapted to any cell type for which suitable surface markers exist and which can be isolated from the subject. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol. In this example the exogenous protein agent is a fluorescent protein and expression is measured via flow cytometry. In other embodiments, the expression of an exogenous protein agent may be measured with immunostaining for the protein. In other embodiments expression of the exogenous protein agent may be measured via microscopy or western blot.

Mice are treated with retroviral vector and a blood sample is collected as described above in Example 2. Cells are stained with a murine CD3:APC-Cy7 antibody (Biolegend Catalog #: 100330) or isotype controls and analyzed by flow cytometry using the protocol described in Example 2.

The percent of CD3+ cells that are tdtomato positive is measured. In some embodiments, the percent of CD3+ cells that are tdtomato positive is greater in cells from treated than untreated mice. The median tdtomato fluorescence level is measured in CD3+ cells. In some embodiments, the median tdtomato fluorescence level in CD3+ cells is greater in cells from treated than untreated mice.

Example 5. Modification of retroviral vector with HLA-G or HLA-E for decreased cytotoxicity mediated by PBMC cell lysis

This Example describes retroviral vectors derived from cells modified to have decreased cytotoxicity due to cell lysis by peripheral blood mononuclear cells (PBMCs).

In an embodiment, cytotoxicity mediated cell lysis of retroviral vectors by PBMCs is a measure of immunogenicity of retroviral vectors, as lysis will reduce, e.g., inhibit or stop, the activity of a retroviral vector.

Retroviral vectors are created from: unmodified cells (hereinafter NMCs, positive control), cells that are transfected with HLA-G or HLA-E cDNA (hereinafter NMC-HLA-G), and cells transfected with an empty vector control (hereinafter NMC-empty vector, negative control).

PBMC mediated lysis of a retroviral vector is determined by europium release assays as described in Bouma, et al. Hum. Immunol. 35(2):85-92; 1992 & van Besouw et al. Transplantation 70(1):136-143; 2000. PBMCs (hereinafter effector cells) are isolated from an appropriate donor, and stimulated with allogeneic gamma irradiated PMBCs and 200IU/mL IL-2 (proleukin, Chiron BV Amsterdam, The Netherlands) in a round bottom 96 well plate for 7 days at 37°C. The retroviral vectors are labeled with europium-diethylenetriaminepentaacetate (DTPA) (sigma, St. Louis, MO, USA).

At day 7 cytotoxicity-mediated lysis assays is performed by incubating ^{63}Eu-labelled retroviral vector with effector cells in a 96-well plate for 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 24, or 48 hours after plating at effector/target ratios ranging from 1000:1-1:1 and 1:1.25-1:1000. After incubation, the plates are centrifuged and a sample of the supernatant is transferred to 96-well plates with low background fluorescence (fluoroimmunoplates, Nunc, Roskilde, Denmark).

Subsequently, enhancement solution (PerkinElmer, Groningen, The Netherlands) is added to each well. The released europium is measured in a time-resolved fluorometer (Victor 1420 multilabel counter, LKB-Wallac, Finland). Fluorescence is expressed in counts per second (CPS). Maximum percent release of europium by a target retroviral vector is determined by incubating an appropriate number (1x 10^2 -1x 10^8) of retroviral vectors with 1% triton (sigma-aldrich) for an appropriate amount of time. Spontaneous release of europium by target retroviral vector is measured by incubation of labeled target retroviral vector without effector cells. Percentage leakage is then calculated as: (spontaneous release/maximum release) x100%. The

percentage of cytotoxicity mediated lysis is calculated as %lysis= [(measured lysis-spontaneous lysis- spontaneous release)/(maximum release-spontaneous release)]x100%. The data is analyzed by looking at the percentage of lysis as a function of different effector target ratios.

In an embodiment, retroviral vectors generated from NMC-HLA-G cells will have a decreased percentage of lysis by target cells at specific timepoints as compared to retroviral vectors generated from NMCs or NMC-empty vector.

Example 6. Modification of retroviral vector with HLA-G or HLA-E for decreased NK lysis activity

This Example describes the generation of a retroviral vector composition derived from a cell source which has been modified to decrease cytotoxicity mediated cell lysis by NK cells. In an embodiment cytotoxicity mediated cell lysis of retroviral vectors by NK cells is a measure of immunogenicity for retroviral vectors.

Retroviral vectors are created from: unmodified cells (hereinafter NMCs, positive control), cells that are transfected with HLA-G or HLA-E cDNA (hereinafter NMC-HLA-G), and cells transfected with an empty vector control (hereinafter NMC-empty vector, negative control).

NK cell mediated lysis of a retroviral vector is determined by europium release assays as described in Bouma, et al. Hum. Immunol. 35(2):85-92; 1992 & van Besouw et al. Transplantation 70(1):136-143; 2000. NK cells (hereinafter effector cells) are isolated from an appropriate donor according to the methods in Crop et al. Cell transplantation (20):1547-1559; 2011, and stimulated with allogeneic gamma irradiated PMBCs and 200IU/mL IL-2 (proleukin, Chiron BV Amsterdam, The Netherlands) in a round bottom 96 well plate for 7 days at 37°C. The retroviral vectors are labeled with europium-diethylenetriaminepentaacetate (DTPA) (sigma, St. Louis, MO, USA). Cytotoxicity-mediated lysis assays and data analsysis are performed as described above in Example 5.

In an embodiment, retroviral vectors generated from NMC-HLA-G cells will have a decreased percentage of lysis by target cells at specific timepoints as compared to retroviral vectors generated from NMCs or NMC-empty vector.

Example 7. Modification of retroviral vector with HLA-G or HLA-E for decreased CD8 Killer T cell lysis

This Example describes the generation of a retroviral vector composition derived from a cell source which has been modified to decrease cytotoxicity mediated cell lysis by CD8+ T-cells. In an embodiment, cytotoxicity mediated cell lysis of retroviral vector by CD8+ T-cells is a measure of immunogenicity for retroviral vectors.

Retroviral vectors are created from: unmodified cells (hereinafter NMCs, positive control), cells that are transfected with HLA-G or HLA-E cDNA (hereinafter NMC-HLA-G), and cells transfected with an empty vector control (hereinafter NMC-empty vector, negative control).

CD8+ T cell mediated lysis of a retroviral vector is determined by europium release assays as described in Bouma, et al. Hum. Immunol. 35(2):85-92; 1992 & van Besouw et al. Transplantation 70(1):136-143; 2000. CD8+ T-cells (hereinafter effector cells) are isolated from an appropriate donor according to the methods in Crop et al. Cell transplantation (20):1547-1559; 2011, and stimulated with allogeneic gamma irradiated PMBCs and 200IU/mL IL-2 (proleukin, Chiron BV Amsterdam, The Netherlands) in a round bottom 96 well plate for 7 days at 37°C. The retroviral vectors are labeled with europium-diethylenetriaminepentaacetate (DTPA) (sigma, St. Louis, MO, USA). Cytotoxicity-mediated lysis assays and data analysis are performed as described above in Example 5.

In an embodiment, retroviral vectors generated from NMC-HLA-G cells will have a decreased percentage of lysis by target cells at specific timepoints as compared to retroviral vectors generated from NMCs or NMC-empty vector.

Example 8: Modification of retroviral vector with CD47 to evade macrophage phagocytosis

This Example describes quantification of the evasion of phagocytosis by modified retroviral vector. In an embodiment, modified retroviral vector will evade phagocytosis by macrophages.

Cells engage in phagocytosis, engulfing particles, enabling the sequestration and destruction of foreign invaders, like bacteria or dead cells. In some embodiments, phagocytosis of lentiviral vectors by macrophages would reduce their activity. In some embodiments, phagocytosis of lentiviral vectors is a measure of immunogenicity of retroviral vectors.

Retroviral vectors are produced from cells which lack CD47 (hereinafter NMC, positive control), cells that are transfected with CD47 cDNA (hereinafter NMC-CD47), and cells transfected with an empty vector control (hereinafter NMC-empty vector, negative control). Prior to retroviral vector production, the cells are labeled with CSFE.

Reduction of macrophage mediated immune clearance is determined with a phagocytosis assay according to the following protocol. Macrophages are plated immediately after harvest in confocal glass bottom dishes. Macrophages are incubated in DMEM+10%FBS+1%P/S for 1h to attach. An appropriate number of retroviral vectors produced from NMC, NMC-CD47, NMC-empty vector are added to the macrophages as indicated in the protocol, and are incubated for 2h, tools.thermofisher.com/content/sfs/manuals/mp06694.pdf.

After 2h, the dish is gently washed and intracellular fluorescence is examined. Intracellular fluorescence emitted by engulfed retroviral particles is imaged by confocal microscopy at 488 excitation. The number of phagocytotic positive macrophage is quantified using imaging software. The data is expressed as the phagocytic index = (total number of engulfed cells/total number of counted macrophages) × (number of macrophages containing engulfed cells/total number of counted macrophages) × 100.

In an embodiment, the phagocytic index will be reduced when macrophages are incubated with retroviral vectors derived from NMC-CD47, versus those derived from NMC, or NMC-empty vector.

Example 9: Modification of retroviral vector with complement regulatory proteins to evade complement

This Example describes quantification of complement activity against a retroviral vector using an in vitro assay. In some embodiments a modified retroviral vector described herein will have reduced complement activity compared to an unmodified retroviral vector.

In this Example, serum from a mouse is assessed for complement activity against a retroviral vector. The example measures the level of complement C3a, which is a central node in all complement pathways. The methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol.

In this example, retroviral vectors are generated from HEK293 cells transfected with a cDNA coding for complement regulatory protein DAF (HEK293-DAF retroviral vector) or HEK

293 cells not expressing a complementary regulatory protein (HEK293 retroviral vector). In other embodiments, other complement regulatory proteins may be used, such as proteins that bind decay-accelerating factor (DAF, CD55), e.g. factor H (FH)-like protein-1 (FHL-1), e.g. C4b-binding protein (C4BP), e.g. complement receptor 1 (CD35), e.g. Membrane cofactor protein (MCP, CD46), eg. Profectin (CD59), e.g. proteins that inhibit the classical and alternative complement pathway CD/C5 convertase enzymes, e.g. proteins that regulate MAC assembly.

Serum is recovered from naïve mice, mice that are administered HEK293-DAF retroviral vector, or mice that are administered HEK293 retroviral vector. Sera are collected from mice by collecting fresh whole blood and allowing it to clot completely for several hours. Clots are pelleted by centrifugation and the serum supernatants are removed. A negative control is heat inactivated mouse serum. Negative control samples are heated at 56 degrees Celsius for 1 hour. Serum may be frozen in aliquots.

The different retroviral vectors are tested for the dose at which 50% of cells in a target cell population receive the exogenous agent in the retroviral vector. The retroviral vector may contain any of the exogenous agents described herein. Many methods for assaying retroviral delivery of an exogenous agent to recipient cells are also described herein. In this particular example, the exogenous agent is Cre protein (encoded by the retroviral nucleic acid) and the target cells are RPMI8226 cells which stably-express a "LoxP-GFP-stop-LoxP-RFP" cassette under a CMV promoter, which upon recombination by Cre switches from GFP to RFP expression, as a marker of delivery. The identified dose at which 50% of the recipient cells are RFP positive is used for further experiments. In some embodiments, the identified dose at which 50% of recipient cells receive the exogenous agent will be similar across retroviral vectors.

Two-fold dilutions in phosphate-buffered saline (PBS, pH 7.4) of the retroviral vectors, starting at the dose of retroviral vectors at which 50% of the target cells receive the exogenous agent, are mixed with a 1:10 dilution of the sera from mice treated with the same retroviral vectors or naïve mice (assay volume, 20 µl) and incubated for 1 h at 37°C. The samples are further diluted 1:500 and used in an enzyme-linked immunosorbent assay (ELISA) specific for C3a. The ELISA is mouse complement C3a ELISA Kit product LS-F4210 sold by LifeSpan BioSciences Inc, which measures the concentration of C3a in a sample. The dose of retroviral vector at which 200 pg/ml of C3a is present is compared across sera isolated from mice.

In some embodiments, the dose of retroviral vector at which 200 pg/ml of C3a is present will be greater for HEK293-DAF retroviral vector incubated with HEK-293 DAF mouse sera than for HEK293 retroviral vector incubated with HEK293 mouse sera, indicating that complement activity targeting retroviral vector is greater in mice treated with HEK293 retroviral vector than HEK293-DAF retroviral vector. In some embodiments, the dose of retroviral vector at which 200 pg/ml of C3a is present will be greater for HEK293-DAF retroviral vector incubated with naive mouse sera than for HEK293 retroviral vector incubated with naive mouse sera, indicating that complement activity targeting retroviral vector is greater in mice treated with HEK293 retroviral vector than HEK293-DAF retroviral vector.

Example 10: Modification of retroviral vector to knockdown immunogenic protein to reduce immunogenicity

This Example describes the generation of a retroviral vector composition derived from a cell source which has been modified to reduce expression of a molecule which is immunogenic, and quantification of the reduced expression. In an embodiment, a retroviral vector can be derived from a cell source, which has been modified to reduce expression of a molecule which is immunogenic.

Therapies that stimulate an immune response can reduce the therapeutic efficacy or cause toxicity to the recipient. Thus, immunogenicity is an important property for a safe and effective therapeutic retroviral vectors. Expression of certain immune activating agents can create an immune response. MHC class I represents one example of an immune activating agent.

Retroviral vectors are produced from unmodified cells which normally express MHC-1 (hereinafter NMC, positive control), cells that are transfected with a DNA coding for a shRNA targeting MHC class I (hereinafter NMC- shMHC class I), and cells transfected with a DNA coding for non-targeted scrambled shRNA vector control (hereinafter NMC-vector control, negative control). Prior to retroviral production, the cells are labeled with CSFE.

Retroviral vectors are assayed for expression of MHC class I using flow cytometry. An appropriate number of retroviral vectors are washed and resuspended in PBS, held on ice for 30 minutes with 1: 10-1: 4000 dilution of fluorescently conjugated monoclonal antibodies against MHC class I (Harlan Sera-Lab, Belton, UK). Retroviral vectors are washed three times in PBS and resuspended in PBS. Nonspecific fluorescence is determined, using equal aliquots of

retroviral vector preparation incubated with and appropriate fluorescently conjugated isotype control antibody at equivalent dilutions. Retroviral vectors are assayed in a flow cytometer (FACSort, Becton-Dickinson) and the data is analyzed with flow analysis software (Becton-Dickinson).

The mean fluorescence data of the retroviral vectors derived from NMCs, NMC- shMHC class I, and NMC-vector control, is compared. In an embodiment, retroviral vectors derived from NMC- shMHC class I will have lower expression of MHC class I compared to NMCs and NMC-vector control.

Example 11: Measuring pre-existing serum inactivation of retroviral vectors

This Example describes quantification of pre-existing serum inactivation of retroviral vectors using an *in vitro* delivery assay.

In some embodiments, a measure of immunogenicity for retroviral vectors is serum inactivation. Serum inactivation of retroviral vectors may be due to antibody-mediated neutralization or complement mediated degradation. In an embodiment, some recipients of a retroviral vectors described herein will have factors in their serum which bind to and inactivate retroviral vectors.

In this Example, a retroviral vector naïve mouse is assessed for the presence of factors that inactivate retroviral vectors in serum. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol.

The negative control is heat inactivated mouse serum and the positive control is serum derived from a mouse that has received multiple injections of retroviral vector generated from a xenogeneic source cell. Sera are collected from mice by collecting fresh whole blood and allowing it to clot completely for several hours. Clots are pelleted by centrifugation and the serum supernatants are removed. Negative control samples are heated at 56 degrees Celsius for 1 hour. Serum may be frozen in aliquots.

The retroviral vectors are tested for the dose at which 50% of cells in a target cell population receive the exogenous agent in the retroviral vector, as described above in Example 9.

To assess serum inactivation of retroviral vectors, retroviral vectors are diluted 1:5 into normal or heat-inactivated serum (or medium containing 10% heat-inactivated FBS as the no-serum control) and the mixture is incubated at 37° C for 1 h. Following the incubation, medium

is added to the reaction for an additional 1:5 dilution and then serially diluted twice at a 1:10 ratio. Following this step, the retroviral vectors should be present at the previously identified dose at which 50% of the recipient cells have received the exogenous agent (e.g. are RFP positive).

Retroviral vectors that have been exposed to serum are then incubated with target cells. The percent of cells which receive the exogenous agent, and thus are RFP positive, is calculated. In some embodiments, the percent of cells which receive the exogenous agent will not be different between retroviral vector samples that have been incubated with serum and heat-inactivated serum from retroviral vector naïve mice, indicating that there is not serum inactivation of retroviral vector. In some embodiments, the percent of cells which receive the exogenous agent will not be different between retroviral vector samples that have been incubated with serum from retroviral vector naïve mice and no-serum control incubations, indicating that there is not serum inactivation of retroviral vectors. In some embodiments, the percent of cells which receive the exogenous agent will be less in retroviral vector samples that have been incubated with positive control serum than in retroviral vector samples that have been incubated with serum from retroviral vector naïve mice, indicating that there is not serum inactivation of retroviral vectors.

Example 12: Measuring serum inactivation of retroviral vectors after multiple administrations

This Example describes quantification of serum inactivation of retroviral vectors using an *in vitro* delivery assay following multiple administrations of the retroviral vectors. In an embodiment, a modified retroviral vector, e.g., modified by a method described herein, will have a reduced (e.g., reduced compared to administration of an unmodified retroviral vector) serum inactivation following multiple (e.g., more than one, e.g., 2 or more) administrations of the modified retroviral vector. In an embodiment, a retroviral vector described herein will not be inactivated by serum following multiple administrations.

In some embodiments, a measure of immunogenicity for retroviral vector is serum inactivation. In an embodiment, repeated injections of a retroviral vector can lead to the development of anti-retroviral vector antibodies, e.g., antibodies that recognize retroviral vectors.

In an embodiment, antibodies that recognize retroviral vectors can bind in a manner that can limit retroviral vector activity or longevity and mediate complement degradation.

In this Example, serum inactivation is examined after one or more administrations of retroviral vectors. Retroviral vectors are produced by any one of the previous Examples. In this example, retroviral are created from: cells that are transfected with HLA-G or HLA-E cDNA (hereinafter NMC-HLA-G), and cells transfected with an empty vector control (hereinafter NMC-empty vector, negative control). In some embodiments, retroviral vectors are derived from cells that are expressing other immunoregulatory proteins.

Serum is drawn from different cohorts: mice injected systemically and/or locally with 1, 2, 3, 5, or 10 injections of vehicle (retroviral vector naïve group), HEK293-HLA-G retroviral vector, or HEK293 retroviral vector. Sera are collected from mice by collecting fresh whole blood and allowing it to clot completely for several hours. Clots are pelleted by centrifugation and the serum supernatants are removed. A negative control is heat inactivated mouse serum. Negative control samples are heated at 56 degrees Celsius for 1 hour. Serum may be frozen in aliquots.

The retroviral vectors are tested for the dose at which 50% of cells in a target cell population receive the exogenous agent in the retroviral vector, as described above in Example 9.

To assess serum inactivation of retroviral vectors, retroviral vectors are exposed to serum and incubated with target cells as described in Example 11 above.

The percent of cells which receive the exogenous agent, and thus are RFP positive, is calculated. In some embodiments, the percent of cells which receive the exogenous agent will not be different between retroviral vector samples that have been incubated with serum and heat-inactivated serum from mice treated with HEK293-HLA-G retroviral vectors, indicating that there is not serum inactivation of retroviral vectors or an adaptive immune response. In some embodiments, the percent of cells which receive the exogenous agent will not be different between retroviral vector samples that have been incubated from mice treated 1, 2, 3, 5 or 10 times with HEK293-HLA-G retroviral vectors, indicating that there is not serum inactivation of retroviral vectors or an adaptive immune response. In some embodiments, the percent of cells which receive the exogenous agent will not be different between retroviral vector samples that have been incubated with serum from mice treated with vehicle and from mice treated with HEK293-HLA-G retroviral vectors, indicating that there is not serum inactivation of retroviral

vectors or an adaptive immune response. In some embodiments, the percent of cells which receive the exogenous agent will be less for retroviral vectors derived from HEK293 than for HEK293-HLA-G retroviral vectors indicating that there is not serum inactivation of HEK293-HLA-G retroviral vectors or an adaptive immune response.

Example 13: Measuring pre-existing IgG and IgM antibodies reactive against retroviral vectors

This Example describes quantification of pre-existing anti- retroviral vector antibody titers measured using flow cytometry.

In some embodiments, a measure of immunogenicity for a retroviral vector is antibody responses. Antibodies that recognize retroviral vector can bind in a manner that can limit retroviral vector activity or longevity. In an embodiment, some recipients of a retroviral vector described herein will have pre-existing antibodies which bind to and recognize retroviral vector.

In this Example, anti- retroviral vector antibody titers are tested using retroviral vector produced using a xenogeneic source cell. In this Example, a retroviral vector naïve mouse is assessed for the presence of anti- retroviral vector antibodies. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol.

The negative control is mouse serum which has been depleted of IgM and IgG, and the positive control is serum derived from a mouse that has received multiple injections of retroviral vector generated from a xenogeneic source cell.

To assess the presence of pre-existing antibodies which bind to retroviral vector, sera from retroviral vector -naïve mice is first decomplemented by heating to 56°C for 30 min and subsequently diluted by 33% in PBS containing 3% FCS and 0.1% NaN3. Equal amounts of sera and retroviral vector ($1x10^2$ - $1x10^8$ retroviral vectors per mL) suspensions are incubated for 30 min at 4°C and washed with PBS through a calf-serum cushion.

IgM xenoreactive antibodies are stained by incubation of the retroviral vector with PE-conjugated goat antibodies specific for the Fc portion of mouse IgM (BD Bioscience) at 4°C for 45 min. Notably, anti-mouse IgG1 or IgG2 secondary antibodies may also be used. Retroviral vector from all groups are washed twice with PBS containing 2% FCS and then analyzed on a FACS system (BD Biosciences). Fluorescence data are collected by use of logarithmic amplification and expressed as mean fluorescent intensity.

In an embodiment, the negative control serum will show negligible fluorescence comparable to the no serum or secondary alone controls. In an embodiment, the positive control will show more fluorescence than the negative control, and more than the no serum or secondary alone controls. In an embodiment, in cases where immunogenicity occurs, serum from retroviral vector -naïve mice will show more fluorescence than the negative control. In an embodiment, in cases where immunogenicity does not occur, serum from retroviral vector -naïve mice will show similar fluorescence compared to the negative control.

Example 14: Measuring IgG and IgM antibody responses after multiple administrations of retroviral vectors

This Example describes quantification of the humoral response of a modified retroviral vector following multiple administrations of the modified retroviral vector. In an embodiment, a modified retroviral vector, e.g., modified by a method described herein, will have a reduced (e.g., reduced compared to administration of an unmodified retroviral vector) humoral response following multiple (e.g., more than one, e.g., 2 or more), administrations of the modified retroviral vector.

In some embodiments, a measure of immunogenicity for a retroviral vector is the antibody responses. In an embodiment, repeated injections of a retroviral vector can lead to the development of anti-retroviral vector antibodies, e.g., antibodies that recognize retroviral vector. In an embodiment, antibodies that recognize retroviral vector can bind in a manner that can limit retroviral vector activity or longevity.

In this Example, anti- retroviral vector antibody titers are examined after one or more administrations of retroviral vector. Retroviral vector is produced by any one of the previous Examples. In this example, retroviral are created from: cells that are not transfected with an immunomodulatory protein (NMCs), cells that are transfected with HLA-G or HLA-E cDNA (hereinafter NMC-HLA-G), and cells transfected with an empty vector control (hereinafter NMC-empty vector, negative control). In some embodiments, retroviral vectors are derived from cells that are expressing other immunoregulatory proteins.

Serum is drawn from different cohorts: mice injected systemically and/or locally with 1, 2, 3, 5, 10 injections of vehicle (retroviral vector naïve group), NMC retroviral vector, NMC-HLA-G retroviral vector, or NMC-empty vectors retroviral vector.

To assess the presence and abundance of anti- retroviral vector antibodies, sera from the mice is first decomplemented by heating to 56°C for 30 min and subsequently diluted by 33% in PBS with 3% FCS and 0.1% NaN3. Equal amounts of sera and retroviral vector ($1x10^2$ - $1x10^8$ retroviral vector per mL) are incubated for 30 min at 4°C and washed with PBS through a calf-serum cushion.

Retroviral vector reactive IgM antibodies are stained by incubation of the retroviral vector with PE-conjugated goat antibodies specific for the Fc portion of mouse IgM (BD Bioscience) at 4°C for 45 min. Notably, anti-mouse IgG1 or IgG2 secondary antibodies may also be used. Retroviral vector from all groups are washed twice with PBS containing 2% FCS and then analyzed on a FACS system (BD Biosciences). Fluorescence data are collected by use of logarithmic amplification and expressed as mean fluorescent intensity.

In an embodiment, NMC-HLA-G retroviral vectors will have decreased anti-viral IgM (or IgG1/2) antibody titers (as measured by fluorescence intensity on FACS) after injections, as compared to NMC retroviral vectors or NMC-empty retroviral vectors.

Example 15: Measuring IgG and IgM titers antibody responses to retroviral vector recipient cells

This Example describes quantification of antibody titers against recipient cells (cells that have fused with retroviral vectors) using flow cytometry. In some embodiments, a measure of the immunogenicity of recipient cells is the antibody response. Antibodies that recognize recipient cells can bind in a manner that can limit cell activity or longevity. In an embodiment, recipient cells will not be targeted by an antibody response, or an antibody response will be below a reference level.

In this Example, anti-recipient cell antibody titers in a subject (e.g., human, rat, or monkey) are tested. In addition, the protocol may be adapted to any cell type for which suitable surface markers exist. In this example, the target recipient cell is a CD3+ cell.

Mice are treated with retroviral vectors produced via any of the methods described in this application or with PBS (negative control) daily for 5 days. 28 days following the final treatment, peripheral blood is collected from mice that received retroviral vectors and mice that received PBS treatment. Blood is collected into 1ml PBS containing 5 μM EDTA and mixed immediately to prevent clotting. The tubes are kept on ice and red blood cells are removed using

a buffered ammonium chloride (ACK) solution. Cells are stained with a murine CD3-FITC antibody (Thermo Fisher Catalog #:11-0032-82), at 4°C for 30 minutes in the dark, after being blocked with bovine serum albumin for 10 minutes. After being washed two times with PBS, cells are analyzed on a LSR II (BD Biosciences, San Jose, CA) with 488nm laser excitation and emission collected at 530+/-30nmrunning the FACSDiva™ software (BD Biosciences, San Jose, CA). CD3+ cells are sorted.

The sorted CD3+ cells are then stained with IgM antibodies by incubation of the reaction mixture with PE-conjugated goat antibodies specific for the Fc portion of mouse IgM (BD Bioscience) at 4°C for 45 min. Notably, anti-mouse IgG1 or IgG2 secondary antibodies may also be used. Cells from all groups are washed twice with PBS containing 2% FCS and then analyzed on a FACS system (BD Biosciences). Fluorescence data are collected by use of logarithmic amplification and expressed as mean fluorescent intensity. The mean fluorescence intensity is calculated for the sorted CD3 cells from mice treated with retroviral vectors and the mice treated with PBS.

A low mean fluorescence intensity is indicative of a low humoral response against the recipient cells. Mice treated with PBS are expected to have low mean fluorescence intensity. In an embodiment, the mean fluorescence intensity will be similar for recipient cells from mice treated with retroviral vectors and mice treated with PBS.

Example 16: Measuring phagocytic response to retroviral vector recipient cells

This Example describes quantification of macrophage response against recipient cells with a phagocytosis assay.

In some embodiments, a measure of the immunogenicity of recipient cells is the macrophage response. Macrophages engage in phagocytosis, engulfing cells and enabling the sequestration and destruction of foreign invaders, like bacteria or dead cells. In some embodiments, phagocytosis of recipient cells by macrophages would reduce their activity.

In an embodiment, recipient cells are not targeted by macrophages. In this Example, the macrophage response against recipient cells in a subject is tested. In addition, the protocol may be adapted to any cell type for which suitable surface markers exist. In this example, the target recipient cell is a CD3+ cell.

Mice are treated with retroviral vectors produced via any of the methods described in this application or with PBS (negative control) daily for 5 days. 28 days following the final treatment, peripheral blood is collected from mice that received retroviral vectors and mice that received PBS treatment. Blood is collected into 1ml PBS containing 5 μM EDTA and mixed immediately to prevent clotting. The tubes are kept on ice and red blood cells are removed using a buffered ammonium chloride (ACK) solution.

Cells are stained with a murine CD3-FITC antibody (Thermo Fisher Catalog #:11-0032-82), at 4°C for 30 minutes in the dark, after being blocked with bovine serum albumin for 10 minutes. After being washed two times with PBS, cells are analyzed on a LSR II (BD Biosciences, San Jose, CA) with 488nm laser excitation and emission collected at 530+/-30nmrunning the FACSDiva™ software (BD Biosciences, San Jose, CA.). CD3+ cells are then sorted.

A phagocytosis assay is run to assess macrophage mediated immune clearance according to the following protocol. Macrophages are plated immediately after harvest in confocal glass bottom dishes. Macrophages are incubated in DMEM+10%FBS+1%P/S for 1h to attach. An appropriate number of sorted and FITC-stained CD3+ cells derived from mice that received retroviral vectors and PBS are added to the macrophages as indicated in the protocol, and are incubated for 2h, e.g., as described in the Vybrant™ Phagocytosis Assay Kit product information insert (Molecular Probes, revised 18-March-2001, found at tools.thermofisher.com/content/sfs/manuals/mp06694.pdf).

After 2h, the dish is gently washed and intracellular fluorescence is examined. To identify macrophages, cells are first incubated with Fc-receptor blocking antibody (eBioscence cat. no. 14-0161-86, clone 93) for 15 min on ice to block the binding of labeled mAbs to Fc receptors, which are abundantly expressed on macrophages. Following this step anti-F4/80-PE (ThermoFisher cat. No. 12-4801-82, clone BM8) and anti-CD11b-PerCP-Cy5.5 (BD Biosciences cat. No. 550993, clone M1/70) conjugated antibodies are added to stain macrophage surface antigens. Cells are incubated for 30 min in the dark at 4C followed by centrifugation and washing in PBS. The cells are then resuspended in PBS. Flow cytometry of samples is then performed and macrophages are identified via positive fluorescence signal for F4/80-PE and CD11b-PerCP-Cy5.5 using 533nm and 647nm laser excitation, respectively. After gating for macrophages, intracellular fluorescence emitted by engulfed recipient cells is assessed by 488nm

laser excitation. The number of phagocytotic positive macrophage is quantified using imaging software. The data is expressed as the phagocytic index = (total number of engulfed cells/total number of counted macrophages) × (number of macrophages containing engulfed cells/total number of counted macrophages) × 100.

A low phagocytic index is indicative of low phagocytosis and targeting by macrophages. Mice treated with PBS are expected to have a low phagocytic index. In an embodiment, the phagocytic index will be similar for recipient cells derived from mice treated with retroviral vectors and mice treated with PBS.

Example 17: Measuring PBMC response to retroviral vector recipient cells

This Example describes quantification of a PBMC response against recipient cells with a cell lysis assay.

In some embodiments, a measure of the immunogenicity of recipient cells is the PBMC response. In an embodiment, cytotoxicity mediated cell lysis of recipient cells by PBMCs is a measure of immunogenicity, as lysis will reduce, e.g., inhibit or stop, the activity of a retroviral vector.

In an embodiment, recipient cells do not elicit a PBMC response. In this Example, the PBMC response against recipient cells in a subject is tested.

In addition, the protocol may be adapted to any cell type for which suitable surface markers exist. In this example, the target recipient cell is a CD3+ cell.

Mice are treated with retroviral vector produced via any of the methods described in this application or with PBS (negative control) daily for 5 days. 28 days following the final treatment, peripheral blood is collected from mice that received retroviral vector and mice that received PBS treatment. Blood is collected into 1ml PBS containing 5 μM EDTA and mixed immediately to prevent clotting. The tubes are kept on ice and red blood cells are removed using a buffered ammonium chloride (ACK) solution. Cells are stained with a murine CD3:APC-Cy7 antibody (Biolgend Catalog #: 100330) or an isotype control APC-Cy7 (IC:APC-Cy7) antibody (Biolgend Catalog #: 400230) at 4°C for 30 minutes in the dark, after being Fc blocked (Biolgend Catalog #: 101319) in cell staining buffer (Biolgend Catalog #: 420201) for 10 minutes. After being washed two times with PBS, cells are analyzed on a FACS Aria (BD Biosciences, San Jose, CA.) with 640nm laser excitation and emission collected at 780 -/+ 60 nm

running the FACSDiva™ software (BD Biosciences, San Jose, CA) to set negative gates using the isotype control APC-Cy7 antibody labelled cells and then APC-Cy7 postive cells are sorted and collected. Sorted CD3+ cells are then labelled with either CellMask™ Green Plasma membrane Stain (CMG, ThermoFisher Catalog #: C37608) or DMSO as the negative control.

7 days prior to the isolation of CD3+ cells from the mice treated with retroviral vector or PBS, PBMCs are isolated from mice treated with retroviral vector or PBS according to the methods in Crop et al. Cell transplantation (20):1547-1559; 2011 and simulated in the presence of IL-2 recombinant mouse protein (R&D Systems Catalog #: 402-ML-020) and CD3/CD28 beads (ThermoFisher Catalog #: 11456D) in a round bottom 96 well plate for 7 days at 37C. At day 7, the stimulated PBMCs are co-incubated with CD3+/CMG+ or CD3+/DMSO control cells for 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 24, 48 hours at a plating ratio of PBMC:CD3+/CMG+ or PBMC: CD3+/DMSO control cells ranging from 1000:1-1:1 and 1:1.25-1:1000. As a negative control a set of wells would receive CD3+/CMG+ and CD3+/DMSO control cells only, no PBMCs. After incubation, the plates are centrifuged and processed so that they are labelled with either murine CD3:APC-Cy7 antibody or an IC:APC-Cy7 antibody as per above. After being washed two times with PBS, cells are re-suspended in PBS and analyzed on a FACS Aria (APC-Cy7: 640nm laser excitation/emission collected at 780 -/+ 60 nm and CMG 561 nm laser excitation/emission collected at 585 -/+ 16 nm) running the FACSDiva™ software (BD Biosciences, San Jose, CA). The FSC/SSC event data would then be used initially to set the gate for events labelled “cells”. This “cells” gate would be then used to display events to set the PMT voltage for the the 640nm and 561 nm laser analyzing samples labelled with IC:APC-Cy7/DMSO only. This sample would also be used to set the gates for negative cells for both APC-Cy7 and CMG. The CD3+/CMG+ cells that did not receive any PBMCs would then used to set the positive gates for CD3+ and CMG+ cells.

The data is analyzed by looking at the percentage of CD3+/CMG+ cells in the population of total cells. When comparing treatment groups, a relatively lower percentage of CD3+/CMG+ cells at any given assay ratio of PBMC:CD3+/CMG+ cells is indicative of recipient cell lysis. In an embodiment, the percent of CD3+/CMG+ will be similar for recipient cells derived from mice treated with retroviral vector and mice treated with PBS.

Example 18: Measuring NK cell response to retroviral vector recipient cells

This Example describes quantification of a natural killer cell response against recipient cells with a cell lysis assay.

In some embodiments, a measure of the immunogenicity of recipient cells is the natural killer cell response. In an embodiment, cytotoxicity mediated cell lysis of recipient cells by natural killer cells is a measure of immunogenicity, as lysis will reduce, e.g., inhibit or stop, the activity of a retroviral vector.

In an embodiment, recipient cells do not elicit a natural killer cell response. In this Example, the natural killer response against recipient cells in a subject is tested. In addition, the protocol may be adapted to any cell type for which suitable surface markers exist. In this example, the target recipient cell is a CD3+ cell.

Mice are treated with retroviral vector, a blood sample is drawn, and CD3+ cells are sorted as described above in Example 17. NK cells are isolated, cultured with the CD3+ cells, and analyzed by FACS according to the protocol described above in Example 17 except that NK cells are used in place of the PBMC cells used in Example 17.

The data is analyzed by looking at the percentage of CD3+/CMG+ cells in the population of total cells. When comparing treatment groups, a relatively lower percentage of CD3+/CMG+ cells at any given assay ratio of NK cells:CD3+/CMG+ cells is indicative of recipient cell lysis. In an embodiment, the percent of CD3+/CMG+ will be similar for recipient cells derived from mice treated with retroviral vector and mice treated with PBS.

Example 19: Measuring CD8 T cell response to retroviral vector recipient cells

This Example describes quantification of a CD8+ T cell response against recipient cells (cells that have fused with retroviral vectors) with a cell lysis assay.

In some embodiments, a measure of the immunogenicity of recipient cells is the CD8+ T cell response. In an embodiment, cytotoxicity mediated cell lysis of recipient cells by CD8+ T cells is a measure of immunogenicity, as lysis will reduce, e.g., inhibit or stop, the activity of a retroviral vector.

In an embodiment, recipient cells do not elicit a CD8+ T cell response. In this Example, the CD8+ T cell response against recipient cells in a subject is tested. In addition, the protocol

may be adapted to any cell type for which suitable surface markers exist. In this example, the target recipient cell is a CD3+ cell.

Mice are treated with retroviral vector, a blood sample is drawn, and CD3+ cells are sorted as described above in Example 17. CD8+ T cells are isolated, cultured with the CD3+ cells, and analyzed by FACS according to the protocol described above in Example 17 except that CD8+ T cells are used in place of the PBMC cells used in Example 17.

The data is analyzed by looking at the percentage of CD3+/CMG+ cells in the population of total cells. When comparing treatment groups, a relatively lower percentage of CD3+/CMG+ cells at any given assay ratio of CD8+ cells:CD3+/CMG+ cells is indicative of recipient cell lysis. In an embodiment, the percent of CD3+/CMG+ will be similar for recipient cells derived from mice treated with retroviral vectors and mice treated with PBS.

Example 20: Measuring CNS Cell-Specific Promoter Activity

This Example describes the measurement of the activity of a CNS cell-specific promoter (a positive TCSRE) in CNS cells compared to non-target cells.

The two cell types are cultured separately and treated with a retroviral vector produced as described herein. The retroviral vector is pseudotyped with a VSV-G and codes for tdtomato fluorescent protein reporter under the control of a CNS cell-specific promoter, e.g., a CNS cell-specific promoter of Table 3.

Two days after transduction, gene expression in the cells is measured via flow cytometry and the average vector copy number in the cells is measured with quantitative PCR. The median tdtomato gene expression per cell in the cell population is normalized to the population vector copy number.

In some embodiments, the population of CNS cells will have a greater ratio of tdtomato expression to vector copy number than the population of non-target cells. This will demonstrate that the CNS cell-specific promoter is more active in CNS cells.

Example 21: Measuring Change in Expression from Restrictive microRNA

This Example describes the measurement of the activity of a non-target-cell-restrictive microRNA (a NTCSRE) in CNS cells compared to non-target cells.

The two cell types are cultured separately and treated with a retroviral vector produced as described herein. The retroviral vector is pseudotyped with a VSV-G and codes for tdtomato fluorescent protein reporter under the control of a ubiquitously active promoter and a non-target-cell-restrictive microRNA, e.g., a microRNA of Table 4.

Two days after transduction, gene expression in the cells is measured via flow cytometry and the average vector copy number in the cells is measured with quantitative PCR. The median tdtomato gene expression per cell in the cell population is normalized to the population vector copy number.

In some embodiments, the population of CNS cells will have a greater ratio of tdtomato expression to vector copy number than the population of non-target cells. This will demonstrate that the non-target-cell restrictive microRNA decreases expression in non-target cells.

Example 22: Treatment for Friedreich's Ataxia with VSV-G pseudotype *in vitro*

This example describes delivery of a therapeutic transgene to cells *in vitro*. In this example, the therapeutic transgene is *frataxin (fxn)*.

Neurons are cultured from brains of YG8R mice and transduced with a retroviral vector produced as described herein or with PBS. The retroviral vector is pseudotyped with VSV-G and carries the *fxn* gene under the control of a neuronal-specific promoter (a positive TCSRE) and microglial cell restrictive microRNA sequence (a NTCSRE).

Following sufficient time for Fxn expression, the cells are prepared for imaging. The cells are fixed, permeabilized, blocked, and immunostained with an anti-Fxn antibody (for example, abcam catalog number ab175402). Following immunostaining, the cells are counterstained with a secondary antibody conjugated to Alexa Fluor 488 (for example, abcam catalog number ab150077). The cells are imaged on a Zeiss LSM 710 confocal microscope with a 63x oil immersion objective while maintained at 37C and 5% CO_2. Alexa Fluor is subjected to 488nm laser excitation and emission captured at 510±15nm. The Alexa Flour average intensity per cell is calculated to determine the level of Fxn expression per cell. For each group at least 30-40 cells are imaged and analyzed.

In some embodiments, the level of Fxn expression per cell will be higher in neurons treated with the retroviral vector encoding the Fxn gene than in neurons treated with PBS.

Example 23: Treatment for Friedreich's Ataxia with VSV-G pseudotyped retrovirus *in vivo*

This example describes delivery of a therapeutic transgene to cells *in vivo*. In this example, the therapeutic transgene is *frataxin* (*fxn*).

YG8R mice are injected intracerebellarly with retroviral vector pseudotyped with VSV-G and carrying the *fxn* gene agent under the control of a neuronal-specific promoter (a positive TCSRE) and microglial cell restrictive microRNA sequence (a NTCSRE). The retroviral vector is produced via any of the methods described in this application. Negative control mice are treated with PBS.

28 days following treatment, brain slices are obtained from mice treated with retrovirus or PBS and stained for Fxn expression as described in previous examples. In some embodiments, the level of Fxn expression per cell will be higher in brains from mice treated with the retroviral vector encoding the Fxn gene than in brains of mice treated with PBS.

In a separate group of mice, 28 days after treatment with retrovirus or PBS the mice are subjected to neurobehavioral tests as described in Rocca et al., 2017, Sci. Transl. Med. 9(413), doi: 10.1126/scitranslmed.aaj2347. In some embodiments, mice treated with retrovirus display significant improvement in neurobehavioral testing as compared to mice treated with PBS.

Example 24: Lack of transcriptional activity in fusosomes

This Example quantifies transcriptional activity in fusosomes compared to parent cells, e.g., source cells, used for fusosome generation. In an embodiment, transcriptional activity will be low or absent in fusosomes compared to the parent cells, e.g., source cells.

Fusosomes are a chassis for the delivery of therapeutic agent. Therapeutic agents, such as miRNA, mRNAs, proteins and/or organelles that can be delivered to cells or local tissue environments with high efficiency could be used to modulate pathways that are not normally active or active at pathological low or high levels in recipient tissue. In an embodiment, the observation that fusosomes are not capable of transcription, or that fusosomes have transcriptional activity of less than their parent cell, will demonstrate that removal of nuclear material has sufficiently occurred.

Fusosomes are prepared by any one of the methods described in previous Examples. A sufficient number of fusosomes and parent cells used to generate the fusosomes are then plated

into a 6 well low-attachment multiwell plate in DMEM containing 20% Fetal Bovine Serum, 1x Penicillin/Streptomycin and the fluorescent-taggable alkyne–nucleoside EU for 1hr at 37°C and 5% CO2. For negative controls, a sufficient number of fusosomes and parent cells are also plated in multiwell plate in DMEM containing 20% Fetal Bovine Serum, 1x Penicillin/Streptomycin but with no alkyne–nucleoside EU.

After the 1 hour incubation the samples are processed following the manufacturer's instructions for an imaging kit (ThermoFisher Scientific). The cell and fusosome samples including the negative controls are washed thrice with 1xPBS buffer and resuspended in 1xPBS buffer and analyzed by flow cytometry (Becton Dickinson, San Jose, CA, USA) using a 488nm argon laser for excitation, and the 530+/-30nm emission. BD FACSDiva software was used for acquisition and analysis. The light scatter channels are set on linear gains, and the fluorescence channels on a logarithmic scale, with a minimum of 10,000 cells analyzed in each condition.

In an embodiment, transcriptional activity as measured by 530+/-30nm emission in the negative controls will be null due to the omission of the alkyne–nucleoside EU. In some embodiments, the fusosomes will have less than about 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, 1% or less transcriptional activity than the parental cells.

See also, Proc Natl Acad Sci U S A, 2008, Oct 14;105(41):15779-84. doi: 10.1073/pnas.0808480105. Epub 2008 Oct 7.

Example 25: Lack of DNA replication or replication activity

This Example quantifies DNA replication in fusosomes. In an embodiment, fusosomes will replicate DNA at a low rate compared to cells.

Fusosomes are prepared by any one of the methods described in previous Examples. Fusosome and parental cell DNA replication activity is assessed by incorporation of a fluorescent-taggable nucleotide (ThermoFisher Scientific # C10632). Fusosomes and an equivalent number of cells are incubated with EdU at a final concentration of 10 μM for 2hr, after preparation of an EdU stock solution with in dimethylsulfoxide. The samples are then fixed for 15 min using 3.7% PFA, washed with 1xPBS buffer, pH 7.4 and permeabilized for 15 min in 0.5% detergent solution in 1xPBS buffer, pH 7.4.

After permeabilization, fusosomes and cells in suspension in PBS buffer containing 0.5% detergent are washed with 1xPBS buffer, pH 7.4 and incubated for 30 min at 21°C in

reaction cocktail, 1xPBS buffer, CuSO4 (Component F), azide-fluor 488, 1x reaction buffer additive.

A negative control for fusosome and cell DNA replication activity is made with samples treated the same as above but with no azide-fluor 488 in the 1x reaction cocktail.

The cell and fusosome samples are then washed and resuspended in 1xPBS buffer and analyzed by flow cytometry. Flow cytometry is done with a FACS cytometer (Becton Dickinson, San Jose, CA, USA) with 488nm argon laser excitation, and a 530+/-30nm emission spectrum is collected. FACS analysis software is used for acquisition and analysis. The light scatter channels are set on linear gains, and the fluorescence channels on a logarithmic scale, with a minimum of 10,000 cells analyzed in each condition. The relative DNA replication activity is calculated based on the median intensity of azide-fluor 488 in each sample. All events are captured in the forward and side scatter channels (alternatively, a gate can be applied to select only the fusosome population). The normalized fluorescence intensity value for the fusosomes is determined by subtracting from the median fluorescence intensity value of the fusosome the median fluorescence intensity value of the respective negative control sample. Then the normalized relative DNA replication activity for the fusosomes samples is normalized to the respective nucleated cell samples in order to generate quantitative measurements for DNA replication activity.

In an embodiment, fusosomes have less DNA replication activity than parental cells. See, also, Salic, 2415–2420, doi: 10.1073/pnas.0712168105.

Example 26: Quantification of fusogens

This example describes quantification of the absolute number of fusogens per fusosome.

A fusosome composition is produced by any one of the methods described in the previous Examples, except the fusosome is engineered as described in a previous Example to express a fusogen (VSV-G) tagged with GFP. In addition, a negative control fusosome is engineered with no fusogen (VSV-G) or GFP present.

The fusosomes with the GFP-tagged fusogen and the negative control(s) are then assayed for the absolute number of fusogens as follows. Commercially acquired recombinant GFP is serially diluted to generate a calibration curve of protein concentration. The GFP fluorescence of the calibration curve and a sample of fusosomes of known quantity is then measured in a

fluorimeter using a GFP light cube (469/35 excitation filter and a 525/39 emission filter) to calculate the average molar concentration of GFP molecules in the fusosome preparation. The molar concentration is then converted to the number of GFP molecules and divided by the number of fusosomes per sample to achieve an average number of GFP-tagged fusogen molecules per fusosome and thus provides a relative estimate of the number of fusogens per fusosome.

In an embodiment, GFP fluorescence will be higher in the fusosomes with GFP tag as compared to the negative controls, where no fusogen or GFP is present. In an embodiment, GFP fluorescence is relative to the number of fusogen molecules present.

Alternatively, individual fusosomes are isolated using a single cell prep system (Fluidigm) per manufacturer's instructions, and qRT-PCR is performed using a commercially available probeset (Taqman) and master mix designed to quantify fusogen or GFP cDNA levels based upon the C_t value. A RNA standard of the same sequence as the cloned fragment of the fusogen gene or the GFP gene is generated by synthesis (Amsbio) and then added to single cell prep system qRT-PCR experimental reaction in serial dilutions to establish a standard curve of C_t vs concentration of fusogen or GFP RNA.

The C_t value from fusosomes is compared to the standard curve to determine the amount of fusogen or GFP RNA per fusosome.

In an embodiment, fusogen and GFP RNA will be higher in the fusosomes with engineered to express the fusogens as compared to the negative controls, where no fusogen or GFP is present.

Fusogens may further be quantified in the lipid bilayer by analyzing the lipid bilayer structure as previously described and quantifying fusogens in the lipid bilayer by LC-MS as described in other Examples herein.

Example 27: Measuring the average size of fusosomes

This Example describes measurement of the average size of fusosomes.

Fusosomes are prepared by any one of the methods described in previous Examples. The fusosomes measured to determine the average size using commercially available systems (iZON Science). The system is used with software according to manufacturer's instructions and a nanopore designed to analyze particles within the 40 nm to 10 μm size range. Fusosomes and

parental cells are resuspended in phosphate-buffered saline (PBS) to a final concentration range of 0.01-0.1 µg protein/mL. Other instrument settings are adjusted as indicated in the following table:

Table 7: Fusosome measurement parameters and settings

Measurement Parameter	Setting
Pressure	6
Nanopore type	NP300
Calibration sample	CPC400_6P
Gold standard analysis	no
Capture assistant	none

All fusosomes are analyzed within 2 hours of isolation. In an embodiment, the fusosomes will have a size within about 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater than the parental cells.

Example 28: Measuring the average size distribution of fusosomes

This Example describes measurement of the size distribution of fusosomes.

Fusosomes are generated by any one of the methods described in previous Examples, and are tested to determine the average size of particles using a commercially available system, such as described in a previous Example. In an embodiment, size thresholds for 10%, 50%, and 90% of the fusosomes centered around the median are compared to parental cells to assess fusosome size distribution.

In an embodiment, the fusosomes will have less than about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, or less of the parental cell's variability in size distribution within 10%, 50%, or 90% of the sample.

Example 29: Average volume of fusosomes

This example describes measurement of the average volume of fusosomes. Without wishing to be bound by theory, varying the size (e.g., volume) of fusosomes can make them versatile for distinct cargo loading, therapeutic design or application.

Fusosomes are prepared as described in previous Examples. The positive control is HEK293 cells or polystyrene beads with a known size. The negative control is HEK293 cells that are passed through a 36 gauge needle approximately 50 times.

Analysis with a transmission electron microscope, as described in a previous Example, is used to determine the size of the fusosomes. The diameter of the fusosome is measured and volume is then calculated.

In an embodiment, fusosomes will have an average size of approximately 50nm or greater in diameter.

Example 30: Average density of fusosomes

Fusosome density is measured via a continuous sucrose gradient centrifugation assay as described in Théry et al., Curr Protoc Cell Biol. 2006 Apr; Chapter 3:Unit 3.22. Fusosomes are obtained as described in previous Examples.

First, a sucrose gradient is prepared. A 2 M and a 0.25 sucrose solution are generated by mixing 4 ml HEPES/sucrose stock solution and 1 ml HEPES stock solution or 0.5 ml HEPES/sucrose stock solution and 4.5 ml HEPES stock solution, respectively. These two fractions are loaded into the gradient maker with all shutters closed, the 2 M sucrose solution in the proximal compartment with a magnetic stir bar, and the 0.25 M sucrose solution in the distal compartment. The gradient maker is placed on a magnetic stir plate, the shutter between proximal and distal compartments is opened and the magnetic stir plate is turned on. HEPES stock solution is made as follows: 2.4 g N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES; 20mMfinal), 300 H2O, adjust pH to 7.4 with 10 N NaOH and finally adjust volume to 500 ml with H2O. HEPES/sucrose stock solution is made as follows: 2.4 g hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES; 20 mM final), 428 g protease-free sucrose (ICN; 2.5 M final), 150 ml H2O, adjust pH to 7.4 with 10 N NaOH and finally adjust volume to 500 ml with H2O.

The fusosomes are resuspended in 2 ml of HEPES/sucrose stock solution and are poured on the bottom of an SW 41 centrifuge tube. The outer tubing is placed in the SW 41 tube, just above the 2 ml of fusosomes. The outer shutter is opened, and a continuous 2 M (bottom) to 0.25 M (top) sucrose gradient is slowly poured on top of the fusosomes. The SW 41 tube is lowered as the gradient is poured, so that the tubing is always slightly above the top of the liquid.

All tubes with gradients are balanced with each other, or with other tubes having the same weight of sucrose solutions. The gradients are centrifuged overnight (≥14 hr) at 210,000 x g, 4°C, in the SW 41 swinging-bucket rotor with the brake set on low.

With a micropipettor, eleven 1-ml fractions, from top to bottom, are collected and placed in a 3-ml tube for the TLA-100.3 rotor. The samples are set aside and, in separate wells of a 96-well plate, 50 µl of each fraction is used to measure the refractive index. The plate is covered with adhesive foil to prevent evaporation and stored for no more than 1 hour at room temperature. A refractometer is used to measure the refractive index (hence the sucrose concentration, and the density) of 10 to 20 µl of each fraction from the material saved in the 96-well plate.

A table for converting the refractive index into g/ml is available in the ultracentrifugation catalog downloadable from the Beckman website.

Each fraction is then prepared for protein content analysis. Two milliliters of 20 mM HEPES, pH 7.4, is added to each 1-ml gradient fraction, and mixed by pipetting up and down two to three times. One side of each tube is marked with a permanent marker, and the tubes are placed marked side up in a TLA-100.3 rotor.

The 3 ml-tubes with diluted fractions are centrifuged for 1 hr at 110,000 x g, 4◦C. The TLA-100.3 rotor holds six tubes, so two centrifugations for each gradient is performed with the other tubes kept at 4°C until they can be centrifuged.

The supernatant is aspirated from each of the 3-ml tubes, leaving a drop on top of the pellet. The pellet most probably is not visible, but its location can be inferred from the mark on the tube. The invisible pellet is resuspended and transferred to microcentrifuge tubes. Half of each resuspended fraction is used for protein contentment analysis by bicinchoninic acid assay, described in another Example. This provides a distribution across the various gradient fractions of the fusosome preparation. This distribution is used to determine the average density of the fusosomes. The second half volume fraction is stored at -80°C and used for other purposes (e.g. functional analysis, or further purification by immunoisolation) once protein analysis has revealed the fusosome distribution across fractions.

In an embodiment, using this assay, the average density of the fusosomes will be 1.25 g/ml +/- 0.05 standard deviation. In an embodiment, the average density of the fusosomes will be in the range of 1-1.1, 1.05-1.15, 1.1-1.2, 1.15-1.25, 1.2-1.3, or 1.25-1.35. In an embodiment, the average density of the fusosomes will be less than 1 or more than 1.35.

Example 31: Measuring nuclear envelope content

This Example describes a measurement of the nuclear envelope content in enucleated fusosomes. The nuclear envelope isolates DNA from the cytoplasm of the cell.

In an embodiment, a purified fusosome composition comprises a mammalian cell, such as HEK-293Ts (293 [HEK-293] (ATCC® CRL-1573™), that has been enucleated as described herein. This Example describes the quantification of different nuclear membrane proteins as a proxy to measure the amount of intact nuclear membrane that remains after fusosome generation.

In this Example, $10x10^6$ HEK-293Ts and the equivalent amount of fusosomes prepared from $10x10^6$ HEK-293Ts are fixed for 15 min using 3.7% PFA, washed with 1xPBS buffer, pH 7.4 and permeabilized simultaneously, and then blocked for 15 min using 1xPBS buffer containing 1% Bovine Serum Albumin and 0.5% Triton® X-100, pH 7.4. After permeabilization, fusosomes and cells are incubated for 12 hours at 4°C with different primary antibodies, e.g. (anti-RanGAP1 antibody [EPR3295] (Abcam - ab92360), anti-NUP98 antibody [EPR6678] - nuclear pore marker (Abcam - ab124980), anti-nuclear pore complex proteins antibody [Mab414] - (Abcam- ab24609), anti-importin 7 antibody (Abcam - ab213670), at manufacturer suggested concentrations diluted in 1xPBS buffer containing 1% bovine serum albumin and 0.5% Triton® X-100, pH 7.4. Fusosomes and cells are then washed with 1xPBS buffer, pH 7.4, and incubated for 2hr at 21°C with an appropriate fluorescent secondary antibody that detects the previous specified primary antibody at manufacturer suggested concentrations diluted in 1xPBS buffer containing 1% bovine serum albumin and 0.5% detergent, pH 7.4. Fusosomes and cells are then washed with 1xPBS buffer, re-suspended in 300µL of 1xPBS buffer, pH 7.4 containing 1 µg/ml Hoechst 33342, filtered through a 20 µm FACS tube and analyzed by flow cytometry.

Negative controls are generated using the same staining procedure but with no primary antibody added. Flow cytometry is performed on a FACS cytometer (Becton Dickinson, San Jose, CA, USA) with 488nm argon laser excitation, and a 530+/-30nm emission spectrum is collected. FACS acquisition software is used for acquisition and analysis. The light scatter channels are set on linear gains, and the fluorescence channels on a logarithmic scale, with a minimum of 10,000 cells analyzed in each condition. The relative intact nuclear membrane

content is calculated based on the median intensity of fluorescence in each sample. All events are captured in the forward and side scatter channels.

The normalized fluorescence intensity value for the fusosomes is determined by subtracting from the median fluorescence intensity value of the fusosome the median fluorescence intensity value of the respective negative control sample. Then the normalized fluorescence for the fusosomes samples is normalized to the respective nucleated cell samples in order to generate quantitative measurements of intact nuclear membrane content.

In an embodiment, enucleated fusosomes will comprise less than 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% fluorescence intensity or nuclear envelope content compared to the nucleated parental cells.

Example 32: Measuring chromatin levels

This Example describes measurement of chromatin in enucleated fusosomes.

DNA can be condensed into chromatin to allow it to fit inside the nucleus. In an embodiment, a purified fusosome composition as produced by any one of the methods described herein will comprise low levels of chromatin.

Enucleated fusosomes prepared by any of the methods previously described and positive control cells (e.g., parental cells) are assayed for chromatin content using an ELISA with antibodies that are specific to histone protein H3 or histone protein H4. Histones are the chief protein component of chromatin, with H3 and H4 the predominant histone proteins.

Histones are extracted from the fusosome preparation and cell preparation using a commercial kit (e.g. Abcam Histone Extraction Kit (ab113476)) or other methods known in the art. These aliquots are stored at -80 C until use. A serial dilution of standard is prepared by diluting purified histone protein (either H3 or H4) from 1 to 50 ng/µl in a solution of the assay buffer. The assay buffer may be derived from a kit supplied by a manufacturer (e.g. Abcam Histone H4 Total Quantification Kit (ab156909) or Abcam Histone H3 total Quantification Kit (ab115091)). The assay buffer is added to each well of a 48- or 96-well plate, which is coated with an anti-histone H3 or anti-H4 antibody and sample or standard control is added to the well to bring the total volume of each well to 50 µl. The plate is then covered and incubated at 37 degrees for 90 to 120 minutes.

After incubation, any histone bound to the anti-histone antibody attached to the plate is prepared for detection. The supernatant is aspirated and the plate is washed with 150 µl of wash buffer. The capture buffer, which includes an anti-histone H3 or anti-H4 capture antibody, is then added to the plate in a volume of 50 µl and at a concentration of 1 µg/mL. The plate is then incubated at room temperature on an orbital shaker for 60 minutes.

Next, the plate is aspirated and washed 6 times using wash buffer. Signal reporter molecule activatable by the capture antibody is then added to each well. The plate is covered and incubated at room temperature for 30 minutes. The plate is then aspirated and washed 4 times using wash buffer. The reaction is stopped by adding stop solution. The absorbance of each well in the plate is read at 450 nm, and the concentration of histones in each sample is calculated according to the standard curve of absorbance at 450 nm vs. concentration of histone in standard samples.

In an embodiment, fusosome samples will comprise less than 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% the histone concentration of the nucleated parental cells.

Example 33: Measuring miRNA content in fusosomes

This example describes quantification of microRNAs (miRNAs) in fusosomes. In an embodiment, a fusosome comprises miRNAs.

MiRNAs are regulatory elements that, among other activities, control the rate by which messenger RNAs (mRNAs) are translated into proteins. In an embodiment, fusosomes carrying miRNA may be used to deliver the miRNA to target sites.

Fusosomes are prepared by any one of the methods described in previous Examples. RNA from fusosomes or parental cells is prepared as described previously. At least one miRNA gene is selected from the Sanger Center miRNA Registry at www.sanger.ac.uk/Software/Rfam/mirna/index.shtml. miRNA is prepared as described in Chen *et al*, *Nucleic Acids Research*, 33(20), 2005. All TaqMan miRNA assays are available through Thermo Fisher (A25576, Waltham, MA).

qPCR is carried out according to manufacturer's specifications on miRNA cDNA, and C_T values are generated and analyzed using a real-time PCR system as described herein.

In an embodiment, the miRNA content of fusosomes will be at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater than that of their parental cells.

Example 34: Quantifying expression of an endogenous RNA or synthetic RNA in fusosomes

This example describes quantification of levels of endogenous RNA with altered expression, or a synthetic RNA that is expressed in a fusosome.

The fusosome or parental cell is engineered to alter the expression of an endogenous or synthetic RNA that mediates a cellular function to the fusosomes.

Transposase vectors (System Biosciences, Inc.) includes the open reading frame of the Puromycin resistance gene together with an open reading frame of a cloned fragment of a protein agent. The vectors are electroporated into 293Ts using an electroporator (Amaxa) and a 293T cell line specific nuclear transfection kit (Lonza).

Following selection with puromycin for 3-5 days in DMEM containing 20% Fetal Bovine Serum and 1x Penicillin/Streptomycin, fusosomes are prepared from the stably expressing cell line by any one of the methods described in previous Examples.

Individual fusosomes are isolated and protein agent or RNA per fusosome is quantified as described in a previous Example.

In an embodiment, the fusosomes will have at least 1, 2, 3, 4, 5, 10, 20, 50, 100, 500, 10^3, 5.0 x 10^3, 10^4, 5.0 x 10^4, 10^5, 5.0 x 10^5, 10^6, 5.0 x 10^6, or more of the RNA per fusosome.

Example 35: Measuring proteomic composition in fusosomes

This Example describes quantification of the protein composition of fusosomes. In an embodiment, the protein composition of fusosomes will be similar to the cells that they are derived from.

Fusosomes are prepared by any one of the methods described in previous Examples. Fusosomes are resuspended in lysis buffer (7M Urea, 2M Thiourea, 4% (w/v) Chaps in 50 mM Tris pH 8.0) and incubated for 15 minutes at room temperature with occasional vortexing. Mixtures are then lysed by sonication for 5 minutes in an ice bath and spun down for 5 minutes

at 13,000 RPM. Protein content is determined by a colorimetric assay (Pierce) and protein of each sample is transferred to a new tube and the volume is equalized with 50 mM Tris pH 8.

Proteins are reduced for 15 minutes at 65 Celsius with 10 mM DTT and alkylated with 15 mM iodoacetamide for 30 minutes at room temperature in the dark. Proteins are precipitated with gradual addition of 6 volumes of cold (-20 Celsius) acetone and incubated overnight at -80 Celsius. Protein pellets are washed 3 times with cold (-20 Celsius) methanol. Proteins are resuspended in 50 mM Tris pH 8.3.

Next, trypsin/lysC is added to the proteins for the first 4h of digestion at 37 Celsius with agitation. Samples are diluted with 50mM Tris pH 8 and 0.1% sodium deoxycholate is added with more trypsin/lysC for digestion overnight at 37 Celsius with agitation. Digestion is stopped and sodium deoxycholate is removed by the addition of 2% v/v formic acid. Samples are vortexed and cleared by centrifugation for 1 minute at 13,000 RPM. Peptides are purified by reversed phase solid phase extraction (SPE) and dried down. Samples are reconstituted in 20 µl of 3% DMSO, 0.2% formic acid in water and analyzed by LC-MS.

To have quantitative measurements, a protein standard is also run on the instrument. Standard peptides (Pierce, equimolar, LC-MS grade, #88342) are diluted to 4, 8, 20, 40 and 100 fmol/ul and are analyzed by LC-MS/MS. The average AUC (area under the curve) of the 5 best peptides per protein (3 MS/MS transition/peptide) is calculated for each concentration to generate a standard curve.

Acquisition is performed with a high resolution mass spectrometer (ABSciex, Foster City, CA, USA) equipped with an electrospray interface with a 25 µm iD capillary and coupled with micro-ultrahigh performance liquid chromatography (µUHPLC) (Eksigent, Redwood City, CA, USA). Analysis software is used to control the instrument and for data processing and acquisition. The source voltage is set to 5.2 kV and maintained at 225°C, curtain gas is set at 27 psi, gas one at 12 psi and gas two at 10 psi. Acquisition is performed in Information Dependent Acquisition (IDA) mode for the protein database and in SWATH acquisition mode for the samples. Separation is performed on a reversed phase column 0.3 µm i.d., 2.7 µm particles, 150mm long (Advance Materials Technology, Wilmington, DE) which is maintained at 60°C. Samples are injected by loop overfilling into a 5µL loop. For the 120 minute (samples) LC gradient, the mobile phase includes the following: solvent A (0.2% v/v formic acid and 3%

DMSO v/v in water) and solvent B (0.2% v/v formic acid and 3% DMSO in EtOH) at a flow rate of 3 µL/min.

For the absolute quantification of the proteins, a standard curve (5 points, R2>0.99) is generated using the sum of the AUC of the 5 best peptides (3 MS/MS ion per peptide) per protein. To generate a database for the analysis of the samples, the DIAUmpire algorithm is run on each of the 12 samples and combined with the output MGF files into one database. This database is used with software (ABSciex) to quantify the proteins in each of the samples, using 5 transition/peptide and 5 peptide/protein maximum. A peptide is considered as adequately measured if the score computed is superior to 1.5 or had a FDR < 1%. The sum of the AUC of each of the adequately measured peptides is mapped on the standard curve, and is reported as fmol.

The resulting protein quantification data is then analyzed to determine protein levels and proportions of known classes of proteins as follows: enzymes are identified as proteins that are annotated with an Enzyme Commission (EC) number; ER associated proteins are identified as proteins that had a Gene Ontology (GO; http://www.geneontology.org) cellular compartment classification of ER and not mitochondria; exosome associated proteins are identified as proteins that have a Gene Ontology cellular compartment classification of exosomes and not mitochondria; and mitochondrial proteins are identified as proteins that are identified as mitochondrial in the MitoCarta database (Calvo et al., NAR 2015l doi:10.1093/nar/gkv1003). The molar ratios of each of these categories are determined as the sum of the molar quantities of all the proteins in each class divided by the sum of the molar quantities of all identified proteins in each sample.

Fusosome proteomic composition is compared to parental cell proteomic composition. In an embodiment, a similar proteomic compositions between fusosomes and parental cells will be observed when >50% of the identified proteins are present in the fusosome, and of those identified proteins the level is >25% of the corresponding protein level in the parental cell.

Example 36: Measuring GAPDH in fusosomes

This assay describes quantification of the level of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) in the fusosomes, and the relative level of GAPDH in the fusosomes compared to the parental cells.

GAPDH is measured in the parental cells and the fusosomes using a standard commercially available ELISA for GAPDH (ab176642, Abcam) per the manufacturer's directions.

Total protein levels are similarly measured via bicinchoninic acid assay as previously described in the same volume of sample used to measure GAPDH. In embodiments, using this assay, the level of GAPDH per total protein in the fusosomes will be <100ng GAPDH / µg total protein. Similarly, in embodiments, the decrease in GAPDH levels relative to total protein from the parental cells to the fusosomes will be greater than a 10% decrease.

In an embodiment, GAPDH content in the preparation in ng GAPDH / µg total protein will be less than 500, less than 250, less than 100, less than 50, less than 20, less than 10, less than 5, or less than1.

In an embodiment, the decrease in GAPDH per total protein in ng/µg from the parent cell to the preparation will be more than 1%, more than 2.5%, more than 5%, more than 10%, more than 15%, more than 20%, more than 30%, more than 40%, more than 50%, more than 60%, more than 70%, more than 80%, or more than 90%.

Example 37: Measuring calnexin in fusosomes

This assay describes quantification of the level of calnexin (CNX) in the fusosomes, and the relative level of CNX in the fusosomes compared to the parental cells.

Calnexin is measured in the starting cells and the preparation using a standard commercially available ELISA for calnexin (MBS721668, MyBioSource) per the manufacturer's directions.

Total protein levels are similarly measured via bicinchoninic acid assay as previously described in the same volume of sample used to measure calnexin. In embodiments, using this assay, the level of calnexin per total protein in the fusosomes will be <100ng calnexin / µg total protein. Similarly, in embodiments, the increase in calnexin levels relative to total protein from the parental cell to the fusosomes will be greater than a 10% increase.

In an embodiment, calnexin content in the preparation in ng calnexin / µg total protein will be less than 500, 250, 100, 50, 20, 10, 5, or 1.

In an embodiment, the decrease in calnexin per total protein in ng/µg from the parent cell to the preparation will be more than 1%, 2.5%, 5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%.

Example 38: Comparison of soluble to insoluble protein mass

This Example describes quantification of the soluble:insoluble ratio of protein mass in fusosomes. In an embodiment, the soluble:insoluble ratio of protein mass in fusosomes will be similar to nucleated cells.

Fusosomes are prepared by any one of the methods described in previous Examples. The fusosome preparation is tested to determine the soluble : insoluble protein ratio using a standard bicinchoninic acid assay (BCA) (e.g. using the commercially available Pierce™ BCA Protein Assay Kit, Thermo Fischer product# 23225). Soluble protein samples are prepared by suspending the prepared fusosomes or parental cells at a concentration of 1x10^7 cells or fusosomes/mL in PBS and centrifuging at 1600g to pellet the fusosomes or cells. The supernatant is collected as the soluble protein fraction.

The fusosomes or cells in the pellet are lysed by vigorous pipetting and vortexing in PBS with 2% Triton-X-100. The lysed fraction represents the insoluble protein fraction.

A standard curve is generated using the supplied BSA, from 0 to 20 µg of BSA per well (in triplicate). The fusosome or cell preparation is diluted such that the quantity measured is within the range of the standards. The fusosome preparation is analyzed in triplicate and the mean value is used. The soluble protein concentration is divided by the insoluble protein concentration to yield the soluble:insoluble protein ratio.

In an embodiment, the fusosome soluble:insoluble protein ratio will be within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or greater compared to the parental cells.

Example 39: Measuring LPS in fusosomes

This example describes quantification of levels of lipopolysaccharides (LPS) in fusosomes as compared to parental cells. In an embodiment, fusosomes will have lower levels of LPS compared to parental cells.

LPS are a component of bacterial membranes and potent inducer of innate immune responses.

The LPS measurements are based on mass spectrometry as described in the previous Examples.

In an embodiment, less than 5%, 1%, 0.5%, 0.01%, 0.005%, 0.0001%, 0.00001% or less of the lipid content of fusosomes will be LPS.

Example 40: Ratio of lipids to proteins in fusosomes

This Example describes quantification of the ratio of lipid mass to protein mass in fusosomes. In an embodiment, fusosomes will have a ratio of lipid mass to protein mass that is similar to nucleated cells.

Total lipid content is calculated as the sum of the molar content of all lipids identified in the lipidomics data set outlined in a previous Example. Total protein content of the fusosomes is measured via bicinchoninic acid assay as described herein.

Alternatively, the ratio of lipids to proteins can be described as a ratio of a particular lipid species to a specific protein. The particular lipid species is selected from the lipidomics data produced in a previous Example. The specific protein is selected from the proteomics data produced in a previous Example. Different combinations of selected lipid species and proteins are used to define specific lipid:protein ratios.

Example 41: Ratio of proteins to DNA in fusosomes

This Example describes quantification of the ratio of protein mass to DNA mass in fusosomes. In an embodiment, fusosomes will have a ratio of protein mass to DNA mass that is much greater than cells.

Total protein content of the fusosomes and cells is measured as described in in a previous Example. The DNA mass of fusosomes and cells is measured as described in a previous Example. The ratio of proteins to total nucleic acids is then determined by dividing the total protein content by the total DNA content to yield a ratio within a given range for a typical fusosome preparation.

Alternatively, the ratio of proteins to nucleic acids is determined by defining nucleic acid levels as the level of a specific house-keeping gene, such as GAPDH, using semi-quantitative real-time PCR (RT-PCR).

The ratio of proteins to GAPDH nucleic acids is then determined by dividing the total protein content by the total GAPDH DNA content to define a specific range of protein:nucleic acid ratio for a typical fusosome preparation.

Example 42: Measuring fusion with a target cell

This example describes quantification of fusosome fusion with a target cell compared to a non-target cell.

In an embodiment, fusosome fusion with a target cell allows the cell-specific delivery of a cargo, carried within the lumen of the fusosome, to the cytosol of the recipient cell. Fusosomes produced by the herein described methods are assayed for fusion rate with a target cell as follows.

In this example, the fusosome comprises a HEK293T cell expressing Myomaker on its plasma membrane. In addition, the fusosome expresses mTagBFP2 fluorescent protein and Cre recombinase. The target cell is a myoblast cell, which expresses both Myomaker and Myomixer, and the non-target cell is a fibroblast cell, which expresses neither Myomaker nor Myomixer. A Myomaker-expressing fusosome is predicted to fuse with the target cell that expresses both Myomaker and Myomixer but not the non-target cell (Quinn et al., 2017, Nature Communications, 8, 15665. doi.org/10.1038/ncomms15665) (Millay et al., 2013, Nature, 499(7458), 301–305. doi.org/10.1038/nature12343). Both the target and non-target cell types are isolated from mice and stably-express "LoxP-stop-Loxp-tdTomato" cassette under a CMV promoter, which upon recombination by Cre turns on tdTomato expression, indicating fusion.

The target or non-target recipient cells are plated into a black, clear-bottom 96-well plate. Both target and non-target cells are plated for the different fusion groups. Next, 24 hours after plating the recipient cells, the fusosomes expressing Cre recombinase protein and Myomaker are applied to the target or non-target recipient cells in DMEM media. The dose of fusosomes is correlated to the number of recipient cells plated in the well. After applying the fusosomes, the cell plate is centrifuged at 400g for 5 minutes to help initiate contact between the fusosomes and the recipient cells.

Starting at four hours after fusosome application, the cell wells are imaged to positively identify RFP-positive cells versus GFP-positive cells in the field or well.

In this example, cell plates are imaged using an automated microscope (www.biotek.com/products/imaging-microscopy-automated-cell-imagers/lionheart-fx-automated-live-cell-imager/). The total cell population in a given well is determined by first staining the cells with Hoechst 33342 in DMEM media for 10 minutes. Hoechst 33342 stains cell nuclei by intercalating into DNA and therefore is used to identify individual cells. After staining, the Hoechst media is replaced with regular DMEM media.

The Hoechst is imaged using the 405 nm LED and DAPI filter cube. GFP is imaged using the 465 nm LED and GFP filter cube, while RFP is imaged using 523 nm LED and RFP filter cube. Images of target and non-target cell wells are acquired by first establishing the LED intensity and integration times on a positive-control well; i.e., recipient cells treated with adenovirus coding for Cre recombinase instead of fusosomes.

Acquisition settings are set so that RFP and GFP intensities are at the maximum pixel intensity values but not saturated. The wells of interest are then imaged using the established settings. Wells are imaged every 4 hours to acquire time-course data for rates of fusion activity.

Analysis of GFP and RFP-positive wells is performed with software provided with the fluorescent microscope or other software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, rsb.info.nih.gov/ij/, 1997–2007).

The images are pre-processed using a rolling ball background subtraction algorithm with a 60 μm width. The total cell mask is set on the Hoechst-positive cells. Cells with Hoechst intensity significantly above background intensities are thresholded and areas too small or large to be Hoechst-positive cells are excluded.

Within the total cell mask, GFP and RFP-positive cells are identified by again thresholding for cells significantly above background and extending the Hoechst (nuclei) masks for the entire cell area to include the entire GFP and RFP cellular fluorescence. The number of RFP-positive cells identified in control wells containing target or non-target recipient cells is used to subtract from the number of RFP-positive cells in the wells containing fusosome (to subtract for non-specific Loxp recombination). The number of RFP-positive cells (fused recipient cells) is then divided by the sum of the GFP-positive cells (recipient cells that have not fused) and RFP-positive cells at each time point to quantify the rate of fusosome fusion within

the recipient cell population. The rate is normalized to the given dose of fusosome applied to the recipient cells. For rates of targeted fusion (fusosome fusion to targeted cells), the rate of fusion to the non-target cell is subtracted from the rate of fusion to the target cell in order to quantify rates of targeted fusion.

In an embodiment, the average rate of fusion for the fusosomes with the target cells will be in the range of 0.01-4.0 RFP/GFP cells per hour for target cell fusion or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than non-target recipient cells with fusosomes. In an embodiment, groups with no fusosome applied will show a background rate of <0.01 RFP/GFP cells per hour.

Example 43: In vitro fusion to deliver a membrane protein

This example describes fusosome fusion with a cell *in vitro*. In an embodiment, fusosome fusion with a cell in vitro results in delivery of an active membrane protein to the recipient cell.

In this example, the fusosomes are generated from a HEK293T cell expressing the Sendai virus HVJ-E protein (Tanaka et al., 2015, Gene Therapy, 22(October 2014), 1–8. doi.org/10.1038/gt.2014.12). In an embodiment, the fusosomes are generated to express the membrane protein, GLUT4, which is found primarily in muscle and fat tissues and is responsible for the insulin-regulated transport of glucose into cells. Fusosomes with and without GLUT4 are prepared from HEK293T cells as described by any of the methods described in a previous Example.

Muscles cells, such as, C2C12 cells, are then treated with fusosomes expressing GLUT4, fusosomes that do not express GLUT4, PBS (negative control), or insulin (positive control). The activity of GLUT4 on C2C12 cells is measured by the uptake of the fluorescent 2-deoxyglucose analog, 2-[N-(7-nitrobenz-2-oxa-1,3-diaxol-4-yl)amino]-2-deoxyglucose (2-NBDG). The fluorescence of C2C12 cells is assessed via microscopy using methods described in previous Examples.

In an embodiment, C2C12 cells that are treated with fusosomes that express GLUT4 and insulin are expected to demonstrate increased fluorescence compared to C2C12 cells treated with PBS or fusosomes not expressing GLUT4.

See, also, Yang et al., *Advanced Materials* 29, 1605604, 2017.

Example 44: Measuring extravasation from blood vessels

This Example describes quantification of fusosome extravasation across an endothelial monolayer as tested with an in vitro microfluidic system (J.S Joen et al. 2013, journals.plos.org/plosone/article?id=10.1371/journal.pone.0056910).

Cells extravasate from the vasculature into surrounding tissue. Without wishing to be bound by theory, extravasation is one way for fusosomes to reach extravascular tissues.

The system includes three independently addressable media channels, separated by chambers into which an ECM-mimicking gel can be injected. In brief, the microfluidics system has molded PDMS (poly-dimethyl siloxane; Silgard 184; Dow Chemical, MI) through which access ports are bored and bonded to a cover glass to form microfluidic channels. Channel cross-sectional dimensions are 1 mm (width) by 120 μm (height). To enhance matrix adhesion, the PDMS channels are coated with a PDL (poly-D-lysine hydrobromide; 1 mg/ml; Sigma-Aldrich, St. Louis, MO) solution.

Next, collagen type I (BD Biosciences, San Jose, CA, USA) solution (2.0 mg/ml) with phosphate-buffered saline (PBS; Gibco) and NaOH is injected into the gel regions of the device via four separate filling ports and incubated for 30 min to form a hydrogel. When the gel is polymerized, endothelial cell medium (acquired from suppliers such as Lonza or Sigma) is immediately pipetted into the channels to prevent dehydration of the gel. Upon aspirating the medium, diluted hydrogel (BD science) solution (3.0 mg/ml) is introduced into the cell channel and the excess hydrogel solution is washed away using cold medium.

Endothelial cells are introduced into the middle channel and allowed to settle to form an endothelium. Two days after endothelial cell seeding, fusosomes or macrophage cells (positive control) are introduced into the same channel where endothelial cells had formed a complete monolayer. The fusosomes are introduced so they adhere to and transmigrate across the monolayer into the gel region. Cultures are kept in a humidified incubator at 37°C and 5% CO_2. A GFP-expressing version of the fusosome is used to enable live-cell imaging via fluorescent microscopy. On the following day, cells are fixed and stained for nuclei using DAPI staining in the chamber, and multiple regions of interest are imaged using confocal microscope to determine how many fusosomes passed through the endothelial monolayer.

In an embodiment, DAPI staining will indicate that fusosomes and positive control cells are able to pass through the endothelial barrier after seeding.

Example 45: Measuring chemotactic cell mobility

This Example describes quantification of fusosome chemotaxis. Cells can move towards or away from a chemical gradient via chemotaxis. In an embodiment, chemotaxis will allow fusosomes to home to a site of injury, or track a pathogen. A purified fusosome composition as produced by any one of the methods described in previous Examples is assayed for its chemotactic abilities as follows.

A sufficient number of fusosomes or macrophage cells (positive control) are loaded in a micro-slide well according to the manufacturer's provided protocol in DMEM media (ibidi.com/img/cms/products/labware/channel_slides/S_8032X_Chemotaxis/IN_8032X_Chemotaxis.pdf). Fusosomes are left at 37°C and 5% CO2 for 1h to attach. Following cell attachment, DMEM (negative control) or DMEM containing MCP1 chemoattractant is loaded into adjacent reservoirs of the central channel and the fusosomes are imaged continuously for 2 hours using a Zeiss inverted widefield microscope. Images are analyzed using ImageJ software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://rsb.info.nih.gov/ij/, 1997–2007). Migration co-ordination data for each observed fusosome or cell is acquired with the manual tracking plugin (Fabrice Cordelières, Institut Curie, Orsay, France). Chemotaxis plots and migration velocities is determined with the Chemotaxis and Migration Tool (ibidi).

In an embodiment, the average accumulated distance and migration velocity of fusosomes will be within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than the response of the positive control cells to chemokine. The response of cells to a chemokine is described, e.g., in Howard E. Gendelman et al., *Journal of Neuroimmune Pharmacology*, 4(1): 47-59, 2009.

Example 46: Measuring homing potential

This Example describes homing of fusosomes to a site of injury. Cells can migrate from a distal site and/or accumulate at a specific site, e.g., home to a site. Typically, the site is a site of injury. In an embodiment, fusosomes will home to, e.g., migrate to or accumulate at, a site of injury.

Eight week old C57BL/6J mice (Jackson Laboratories) are dosed with notexin (NTX) (Accurate Chemical & Scientific Corp), a myotoxin, in sterile saline by intramuscular (IM) injection using a 30G needle into the right tibialis anterior (TA) muscle at a concentration of 2 μg/mL. The skin over the tibialis anterior (TA) muscle is prepared by depilating the area using a chemical hair remover for 45 seconds, followed by 3 rinses with water. This concentration is chosen to ensure maximum degeneration of the myofibers, as well as minimal damage to their satellite cells, the motor axons and the blood vessels.

On day 1 after NTX injection, mice receive an IV injection of fusosomes or cells that express firefly luciferase. Fusosomes are produced from cells that stably express firefly luciferase by any one of the methods described in previous Examples. A bioluminescent imaging system (Perkin Elmer) is used to obtain whole animal images of bioluminescence at 0, 1, 3, 7, 21, and 28 post injection.

Five minutes before imaging, mice receive an intraperitoneal injection of bioluminescent substrate (Perkin Elmer) at a dose of 150mg/kg in order to visualize luciferase. The imaging system is calibrated to compensate for all device settings. The bioluminescent signal is measured using Radiance Photons, with Total Flux used as a measured value. The region of interest (ROI) is generated by surrounding the signal of the ROI in order to give a value in photons/second. An ROI is assessed on both the TA muscle treated with NTX and on the contralateral TA muscle, and the ratio of photons/second between NTX-treated and NTX-untreated TA muscles is calculated as a measure of homing to the NTX-treated muscle.

In an embodiment, the ratio of photons/second between NTX-treated and NTX-untreated TA muscles in fusosomes and cells will be greater than 1 indicating site specific accumulation of luciferase-expressing fusosomes at the injury.

See, for example, Plant et al., *Muscle Nerve* 34(5)L 577-85, 2006.

Example 47: Measuring phagocytic activity

This Example demonstrates phagocytic activity of fusosomes. In an embodiment, fusosomes have phagocytic activity, e.g., are capable of phagocytosis. Cells engage in phagocytosis, engulfing particles, enabling the sequestration and destruction of foreign invaders, like bacteria or dead cells.

A purified fusosome composition as produced by any one of the methods described in previous Examples comprising a fusosome from a mammalian macrophage having partial or complete nuclear inactivation was capable of phagocytosis assayed via pathogen bioparticles. This estimation was made by using a fluorescent phagocytosis assay according to the following protocol.

Macrophages (positive control) and fusosomes were plated immediately after harvest in separate confocal glass bottom dishes. The macrophages and fusosomes were incubated in DMEM+10%FBS+1%P/S for 1h to attach. Fluorescein-labeled *E. coli* K12 and non-fluorescein-labeled *Escherichia coli* K-12 (negative control) were added to the macrophages/fusosomes as indicated in the manufacturer's protocol, and were incubated for 2h, tools.thermofisher.com/content/sfs/manuals/mp06694.pdf. After 2h, free fluorescent particles were quenched by adding Trypan blue. Intracellular fluorescence emitted by engulfed particles was imaged by confocal microscopy at 488 excitation. The number of phagocytotic positive fusosome were quantified using image J software.

The average number of phagocytotic fusosomes was at least 30% 2h after bioparticle introduction, and was greater than 30% in the positive control macrophages.

Example 48: Measuring potential for protein secretion

This Example describes quantification of secretion by fusosomes. In an embodiment, fusosomes will be capable of secretion, e.g., protein secretion. Cells can dispose or discharge of material via secretion. In an embodiment, fusosomes will chemically interact and communicate in their environment via secretion.

The capacity of fusosomes to secrete a protein at a given rate is determined using the Gaussia luciferase flash assay from ThermoFisher Scientific (catalog #16158). Mouse embryonic fibroblast cells (positive control) or fusosomes as produced by any one of the methods described in previous Examples are incubated in growth media and samples of the media are collected every 15 minutes by first pelleting the fusosomes at 1600g for 5min and then collecting the supernatant. The collected samples are pipetted into a clear-bottom 96-well plate. A working solution of assay buffer is then prepared according to the manufacturer's instructions.

Briefly, colenterazine, a luciferin or light-emitting molecule, is mixed with flash assay buffer and the mixture is pipetted into each well of the 96 well plate containing samples.

Negative control wells that lack cells or fusosomes include growth media or assay buffer to determine background Gaussia luciferase signal. In addition, a standard curve of purified Gaussia luciferase (Athena Enzyme Systems, catalog #0308) is prepared in order to convert the luminescence signal to molecules of Gaussia luciferase secretion per hour.

The plate is assayed for luminescence, using 500 msec integration. Background Gaussia luciferase signal is subtracted from all samples and then a linear best-fit curve is calculated for the Gaussia luciferase standard curve. If sample readings do not fit within the standard curve, they are diluted appropriately and re-assayed. Using this assay, the capacity for fusosomes to secrete Gaussia luciferase at a rate (molecules/hour) within a given range is determined.

In an embodiment, fusosomes will be capable of secreting proteins at a rate that is 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than the positive control cells.

Example 49: Measuring signal transduction potential

This Example describes quantification of signal transduction in fusosomes. In an embodiment, fusosomes are capable of signal transduction. Cells can send and receive molecular signals from the extracellular environment through signaling cascades, such as phosphorylation, in a process known as signal transduction. A purified fusosome composition as produced by any one of the methods described in previous Examples comprising a fusosome from a mammalian cell having partial or complete nuclear inactivation is capable of signal transduction induced by insulin. Signal transduction induced by insulin is assessed by measuring AKT phosphorylation levels, a key pathway in the insulin receptor signaling cascade, and glucose uptake in response to insulin.

To measure AKT phosphorylation, cells, e.g., Mouse Embryonic Fibroblasts (MEFs) (positive control), and fusosomes are plated in 48-well plates and left for 2 hours in a humidified incubator at 37°C and 5% CO_2. Following cell adherence, insulin (e.g. at 10 nM), or a negative control solution without insulin, is add to the well containing cells or fusosomes for 30 min. After 30 minutes, protein lysate is made from the fusosomes or cells, and phospho-AKT levels are measured by western blotting in insulin stimulated and control unstimulated samples.

Glucose uptake in response to insulin or negative control solution is measured as it is explained in the glucose uptake section by using labeled glucose (2-NBDG). (S. Galic et al., *Molecular Cell Biology* 25(2): 819-829, 2005).

In an embodiment, fusosomes will enhance AKT phosphorylation and glucose uptake in response to insulin over the negative controls by at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater.

Example 50: Measuring ability to transport glucose across cell membrane

This Example describes quantification of the levels of a 2-NBDG (2-(*N*-(7-Nitrobenz-2-oxa-1,3-diazol-4-yl)Amino)-2-Deoxyglucose) a fluorescent glucose analog that can be used to monitor glucose uptake in live cells, and thus measure active transport across the lipid bilayer. In an embodiment, this assay can be used to measure the level of glucose uptake and active transport across the lipid bilayer of the fusosome.

A fusosome composition is produced by any one of the methods described in previous Examples. A sufficient number of fusosomes are then incubated in DMEM with no glucose, 20% Fetal Bovine Serum and 1x Penicillin/Streptomycin for 2hr at 37°C and 5% CO_2. After a 2hr glucose starvation period, the medium is changed such that it includes DMEM with no glucose, 20% Fetal Bovine Serum, 1x Penicillin/Streptomycin and 20 uM 2-NBDG (ThermoFisher) and incubated for an additional 2hr at 37°C and 5% CO_2.

Negative control fusosomes are treated the same, except an equal amount of DMSO is added in place of 2-NBDG.

The fusosomes are then washed thrice with 1xPBS and re-suspended in an appropriate buffer, and transferred to a 96 well imaging plate. 2-NBDG fluorescence is then measured in a fluorimeter using a GFP light cube (469/35 excitation filter and a 525/39 emission filter) to quantify the amount of 2-NBDG that has been transported across the fusosome membrane and accumulated in the fusosome in the 1hr loading period.

In an embodiment, 2-NBDG fluorescence will be higher in the fusosome with 2-NBDG treatment as compared to the negative (DMSO) control. Fluorescence measure with a 525/39 emission filter will correlate with to the number of 2-NBDG molecules present.

Example 51: Measuring esterase activity in the cytosol

This Example describes quantification of esterase activity, as a surrogate for metabolic activity, in fusosomes. The cytosolic esterase activity in fusosomes is determined by quantitative assessment of calcein-AM staining (Bratosin et al., Cytometry 66(1): 78-84, 2005).

The membrane-permeable dye, calcein-AM (Molecular Probes, Eugene OR USA), is prepared as a stock solution of 10 mM in dimethylsulfoxide and as a working solution of 100 mM in PBS buffer, pH 7.4. Fusosomes as produced by any one of the methods described in previous Examples or positive control parental Mouse Embryonic Fibroblast cells are suspended in PBS buffer and incubated for 30 minutes with calcein-AM working solution (final concentration in calcein-AM: 5 mM) at 37°C in the dark and then diluted in PBS buffer for immediate flow cytometric analysis of calcein fluorescence retention.

Fusosomes and control parental Mouse Embryonic Fibroblast cells are experimental permeabilized as a negative control for zero esterase activity with saponin as described in (Jacob et al., *Cytometry* 12(6): 550-558, 1991). Fusosomes and cells are incubated for 15 min in 1% saponin solution in PBS buffer, pH 7.4, containing 0.05% sodium azide. Due to the reversible nature of plasma membrane permeabilization, saponin is included in all buffers used for further staining and washing steps. After saponin permeabilization, fusosomes and cells are suspended in PBS buffer containing 0.1% saponin and 0.05% sodium azide and incubated (37C in the dark for 45 min) with calcein-AM to a final concentration of 5 mM, washed three times with the same PBS buffer containing 0.1% saponin and 0.05% sodium azide, and analyzed by flow cytometry. Flow cytometric analyses are performed on a FACS cytometer (Becton Dickinson, San Jose, CA, USA) with 488nm argon laser excitation and emission is collected at 530+/-30nm. FACS software is used for acquisition and analysis. The light scatter channels are set on linear gains, and the fluorescence channels are set on a logarithmic scale, with a minimum of 10,000 cells analyzed in each condition. Relative esterase activities are calculated based on the intensity of calcein-AM in each sample. All events are captured in the forward and side scatter channels (alternatively, a gate can be applied to select only the fusosome population). The fluorescence intensity (FI) value for the fusosomes is determined by subtracting the FI value of the respective negative control saponin-treated sample. The normalized esterase activity for the fusosomes samples are normalized to the respective positive control cell samples in order to generate quantitative measurements for cytosolic esterase activities.

In an embodiment, a fusosome preparation will have within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater esterase activity compared to the positive control cell.

See also, Bratosin D, Mitrofan L, Palii C, Estaquier J, Montreuil J. Novel fluorescence assay using calcein-AM for the determination of human erythrocyte viability and aging. Cytometry A. 2005 Jul;66(1):78-84; and Jacob BC, Favre M, Bensa JC. Membrane cell permeabilisation with saponin and multiparametric analysis by flow cytometry. Cytometry 1991;12:550–558.

Example 52: Measuring acetylcholinesterase activity in fusosomes

Acetylcholinesterase activity is measured using a kit (MAK119, SIGMA) that follows a procedure described previously (Ellman, et al., Biochem. Pharmacol. 7, 88, 1961) and following the manufacturer's recommendations.

Briefly, fusosomes are suspended in 1.25 mM acetylthiocholine in PBS, pH 8, mixed with 0.1 mM 5,5-dithio-bis(2-nitrobenzoic acid) in PBS, pH 7. The incubation is performed at room temperature but the fusosomes and the substrate solution are pre-warmed at 37 °C for 10 min before starting the optical density readings.

Changes in absorption are monitored at 450 nm for 10 min with a plate reader spectrophotometer (ELX808, BIO-TEK instruments, Winooski, VT, USA). Separately, a sample is used for determining the protein content of the fusosomes via bicinchoninic acid assay for normalization. Using this assay, the fusosomes are determined to have <100 AChE activity units/μg of protein.

In an embodiment, AChE activity units/μg of protein values will be less than 0.001, 0.01, 0.1, 1, 10, 100, or 1000.

Example 53: Measuring metabolic activity level

This Example describes quantification of the measurement of citrate synthase activity in fusosomes.

Citrate synthase is an enzyme within the tricarboxylic acid (TCA) cycle that catalyzes the reaction between oxaloacetate (OAA) and acetyl-CoA to generate citrate. Upon hydrolysis of acetyl-CoA, there is a release of CoA with a thiol group (CoA-SH). The thiol group reacts with

a chemical reagent, 5,5-Dithiobis-(2-nitrobenzoic acid) (DTNB), to form 5-thio-2-nitrobenzoic acid (TNB), which is a yellow product that can be measured spectrophotometrically at 412 nm (Green 2008). Commercially-available kits, such as the Abcam Human Citrate Synthase Activity Assay Kit (Product #ab119692) provide all the necessary reagents to perform this measurement.

The assay is performed as per the manufacturer's recommendations. Fusosome sample lysates are prepared by collecting the fusosomes as produced by any one of the methods described in previous Examples and solubilizing them in Extraction Buffer (Abcam) for 20 minutes on ice. Supernatants are collected after centrifugation and protein content is assessed by bicinchoninic acid assay (BCA, ThermoFisher Scientific) and the preparation remains on ice until the following quantification protocol is initiated.

Briefly, fusosome lysate samples are diluted in 1X Incubation buffer (Abcam) in the provided microplate wells, with one set of wells receiving only 1X Incubation buffer. The plate is sealed and incubated for 4 hours at room temperature with shaking at 300rpm. The buffer is then aspirated from the wells and 1X Wash buffer is added. This washing step is repeated once more. Then, 1X Activity solution is added to each well, and the plate is analyzed on a microplate reader by measuring absorbance at 412nm every 20 seconds for 30 minutes, with shaking between readings.

Background values (wells with only 1X Incubation buffer) are subtracted from all wells, and the citrate synthase activity is expressed as the change in absorbance per minute per µg of fusosome lysate sample loaded (ΔmOD@412nm/min/ug protein). Only the linear portion from 100-400 seconds of the kinetic measurement is used to calculate the activity.

In an embodiment, a fusosome preparation will have within 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater synthase activity compared to the control cell.

See, for example, Green HJ et al. Metabolic, enzymatic, and transporter response in human muscle during three consecutive days of exercise and recovery. Am J Physiol Regul Integr Comp Physiol 295: R1238-R1250, 2008.

Example 54: Measuring respiration levels

This Example describes quantification of the measurement of respiration level in fusosomes. Respiration level in cells can be a measure of oxygen consumption, which powers metabolism. Fusosome respiration is measured for oxygen consumption rates by a Seahorse extracellular flux analyzer (Agilent) (Zhang 2012).

Fusosomes as produced by any one of the methods described in previous Examples or cells are seeded in a 96-well Seahorse microplate (Agilent). The microplate is centrifuged briefly to pellet the fusosomes and cells at the bottom of the wells. Oxygen consumption assays are initiated by removing growth medium, replacing with a low-buffered DMEM minimal medium containing 25mM glucose and 2mM glutamine (Agilent) and incubating the microplate at 37°C for 60 minutes to allow for temperature and pH equilibrium.

The microplate is then assayed in an extracellular flux analyzer (Agilent) that measures changes in extracellular oxygen and pH in the media immediately surrounding adherent fusosomes and cells. After obtaining steady state oxygen consumption (basal respiration rate) and extracellular acidification rates, oligomycin (5μM), which inhibits ATP synthase, and proton ionophore FCCP (carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone; 2μM), which uncouples mitochondria, are added to each well in the microplate to obtain values for maximal oxygen consumption rates.

Finally, 5 μM antimycin A (inhibitor of mitochondria complex III) is added to confirm that respiration changes are due mainly to mitochondrial respiration. The minimum rate of oxygen consumption after antimycin A addition is subtracted from all oxygen consumption measurements to remove the non-mitochondrial respiration component. Cell samples that do not appropriately respond to oligomycin (at least a 25% decrease in oxygen consumption rate from basal) or FCCP (at least a 50% increase in oxygen consumption rate after oligomycin) are excluded from the analysis. Fusosomes respiration level is then measured as pmol O2/min/1e4 fusosomes.

This respiration level is then normalized to the respective cell respiration level. In an embodiment, fusosomes will have at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater respiration level compared to the respective cell samples.

See, for example, Zhang J, Nuebel E, Wisidagama DRR, et al. Measuring energy metabolism in cultured cells, including human pluripotent stem cells and differentiated cells. Nature protocols. 2012;7(6):10.1038/nprot.2012.048. doi:10.1038/nprot.2012.048.

Example 55: Measuring phosphatidylserine levels of fusosomes

This Example describes quantification of the level of annexin-V binding to the surface of fusosomes.

Dying cells can display phosphatidylserine on the cell surface which is a marker of apoptosis in the programmed cell death pathway. Annexin-V binds to phosphatidylserine, and thus, annexin-V binding is a proxy for viability in cells.

Fusosomes were produced as described herein. For detection of apoptosis signals, fusosomes or positive control cells were stained with 5% annexin V fluor 594 (A13203, Thermo Fisher, Waltham, MA). Each group (detailed in the table below) included an experimental arm that was treated with an apoptosis-inducer, menadione. Menadione was added at 100 μM menadione for 4 h. All samples were run on a flow cytometer (Thermo Fisher, Waltham, MA) and fluorescence intensity was measured with the YL1 laser at a wavelength of 561 nm and an emission filter of 585/16 nm. The presence of extracellular phophatidyl serine was quantified by comparing fluorescence intensity of annexin V in all groups.

The negative control unstained fusosomes were not positive for annexin V staining.

In an embodiment, fusosomes were capable of upregulating phosphatidylserine display on the cell surface in response to menadione, indicating that non-menadione stimulated fusosomes are not undergoing apoptosis. In an embodiment, positive control cells that were stimulated with menadione demonstrated higher-levels of annexin V staining than fusosomes not stimulated with menadione.

Table 8: Annexin V staining parameter

Experimental Arm	Mean Fluorescence Intensity of Annexin V Signal (and standard deviation)
Unstained Fusosomes (negative control)	941 (937)
Stained Fusosomes	11257 (15826)
Stained Fusosomes + Menadione	18733 (17146)

Stained Macrophages + Menadione (positive control)	14301 (18142)

Example 56: Measuring juxtacrine-signaling levels

This Example describes quantification of juxtacrine-signaling in fusosomes.

Cells can form cell-contact dependent signaling via juxtacrine signaling. In an embodiment, presence of juxtacrine signaling in fusosomes will demonstrate that fusosomes can stimulate, repress, and generally communicate with cells in their immediate vicinity.

Fusosomes produced by any one of the methods described in previous Examples from mammalian bone marrow stromal cells (BMSCs) having partial or complete nuclear inactivation trigger IL-6 secretion via juxtacrine signaling in macrophages. Primary macrophages and BMSCs are co-cultured. Bone marrow-derived macrophages are seeded first into 6-well plates, and incubated for 24h, then primary mouse BMSC-derived fusosomes or BMSC cells (positive control parental cells) are placed on the macrophages in a DMEM medium with 10% FBS. The supernatant is collected at different time points (2, 4, 6, 24 hours) and analyzed for IL-6 secretion by ELISA assay. (Chang J. et al., 2015).

In an embodiment, the level of juxtacrine signaling induced by BMSC fusosomes is measured by an increase in macrophage-secreted IL-6 levels in the media. In an embodiment, the level of juxtacrine signaling will be at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than the levels induced by the positive control bone marrow stromal cells (BMSCs).

Example 57: Measuring paracrine-signaling levels

This Example describes quantification of paracrine signaling in fusosomes.

Cells can communicate with other cells in the local microenvironment via paracrine signaling. In an embodiment, fusosomes will be capable of paracrine signaling, e.g., to communicate with cells in their local environment. In an embodiment, the ability of fusosomes to trigger Ca^{2+} signaling in endothelial cells via paracrine-derived secretion with the following protocol will measure Ca^{2+} signaling via the calcium indicator, fluo-4 AM.

To prepare the experimental plate, murine pulmonary microvascular endothelial cells (MPMVECs) are plated on a 0.2% gelatin coated 25mm glass bottom confocal dish (80% confluence). MPMVECs are incubated at room temperature for 30 min in ECM containing 2%

BSA and 0.003% pluronic acid with 5 μM fluo-4 AM (Invitrogen) final concentration to allow loading of fluo-4 AM. After loading, MPMVECs are washed with experimental imaging solution (ECM containing 0.25% BSA) containing sulfinpyrazone to minimize dye loss. After loading fluo-4, 500μl of pre-warmed experimental imaging solution is added to the plate, and the plate is imaged by a Zeiss confocal imaging system.

In a separate tube, freshly isolated murine macrophages are either treated with 1μg/ml LPS in culture media (DMEM+10% FBS) or not treated with LPS (negative control). After stimulation, fusosomes are generated from macrophages by any one of the methods described in previous Examples.

Fusosomes or parental macrophages (positive control) are then labeled with cell tracker red, CMTPX (Invitrogen), in ECM containing 2%BSA and 0.003% pluronic acid. Fusosomes and macrophages are then washed and resuspended in experimental imaging solution. Labeled fusosomes and macrophages are added onto the fluo-4 AM loaded MPMVECs in the confocal plate.

Green and red fluorescence signal is recorded every 3s for 10-20 min using Zeiss confocal imaging system with argon ion laser source with excitation at 488 and 561 nm for fluo-4 AM and cell tracker red fluorescence respectively. Fluo-4 fluorescence intensity changes are analyzed using imaging software (Mallilankaraman, K. et al., *J Vis Exp.* (58): 3511, 2011). The level of Fluo-4 intensity measured in negative control fusosome and cell groups is subtracted from LPS-stimulated fusosome and cell groups.

In an embodiment, fusosomes, e.g., activated fusosomes, will induce an increase in Fluo-4 fluorescence intensity that is at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than the positive control cell groups.

Example 58: Measuring ability to polymerize actin for mobility

This Example describes quantification of cytoskeletal components, such as actin, in fusosomes. In an embodiment, fusosomes comprise cytoskeletal components such as actin, and are capable of actin polymerization.

Cells use actin, which is a cytoskeletal component, for motility and other cytoplasmic processes. The cytoskeleton is essential to creating motility driven forces and coordinating the process of movement

C2C12 cells were enucleated as described herein. Fusosomes obtained from the 12.5% and 15% Ficoll layers were pooled and labeled 'Light', while fusosomes from the 16 – 17% layers were pooled and labeled 'Medium'. Fusosomes or cells (parental C2C12 cells, positive control) were resuspended in DMEM + Glutamax + 10% Fetal Bovine Serum (FBS), plated in 24-well ultra-low attachment plates (#3473, Corning Inc, Corning, NY) and incubated at 37 °C + 5% CO_2. Samples were taken periodically (5.25 hr, 8.75 hr, 26.5 hr) and stained with 165 μM rhodamine phalloidin (negative control was not stained) and measured on a flow cytometer (#A24858, Thermo Fisher, Waltham, MA) with a FC laser YL1 (561 nm with 585/16 filter) to measure F-actin cytoskeleton content. The fluorescence intensity of rhodamine phalloidin in fusosomes was measured along with unstained fusosomes and stained parental C2C12 cells.

Fusosome fluorescence intensity was greater (FIG. 1) than the negative control at all timepoints, and fusosomes were capable of polymerizing actin at a similar rate to the parental C2C12 cells.

Additional cytoskeletal components, such as those listed in the table below, are measured via a commercially available ELISA systems (Cell Signaling Technology and MyBioSource), according to manufacturer's instructions.

Table 9: Cytoskeletal components

Cytoskeletal protein measured	Commercial Kit Type	Kit ID
Actin	Path Scan Total B-Actin Sandwich ELISA Kit	Cell Signaling, 7880
Arp2/3	Human Actin Related protein 2/3 complex subunit(APRC2) ELISA KIT	MyBioSource, MBS7224740
Formin	**Formin Binding Protein 1 (FNBP1), ELISA Kit**	MyBioSource, MBS9308864
Coronin	Human Coronin 1A ELISA Kit	MyBioSource, MBS073640
Dystrophin	Human dystrophin ELISA Kit	MyBioSource MBS722223
Keratin	Human Keratin 5 ELISA Kit	MyBioSource, MBS081200
Myosin	Human Myosin IG (MYO1G) ELISA Kit	MyBioSource, MBS9312965
Tubulin	Human Tubulin Beta 3 ELISA Kit	MyBioSource, MBS097321

Then 100uL of appropriately-diluted lysate is added to the appropriate well from the microwell strips. The microwells are sealed with tape and incubated for 2 hrs at 37C. After incubation, the sealing tape is removed and the contents are discarded. Each microwell is washed four times with 200uL of 1X Wash Buffer. After each individual wash, plates are struck onto an absorbent cloth so that the residual wash solution is removed from each well. However, wells are not completely dry at any time during the experiment.

Next, 100ul of the reconstituted Detection Antibody (green) is added each individual well, except for negative control wells. Then wells are sealed and incubated for 1 hour at 37°C. The washing procedure is repeated after incubation is complete. 100uL of reconstituted HRP-Linked secondary antibody (red) is added to each of the wells. The wells are sealed with tape and incubated for 30 minutes at 37°C. The sealing tape is then removed and the washing procedure is repeated. 100uL of TMB Substrate is then added to each well. The wells are sealed with tape, then incubated for 10 minutes at 37°C. Once this final incubation is complete, 100uL of STOP solution is added to each of the wells and the plate is shaken gently for several seconds.

Spectrophotometric analysis of the assay is conducted within 30 minutes of adding the STOP solution. The underside of the wells is wiped with lint-free tissue and then absorbance is read at 450nm. In an embodiment, fusosome samples that have been stained with the detection antibody will absorb more light at 450 nm that negative control fusosome samples, and absorb less light than cell samples that have been stained with the detection antibody.

Example 59: Measuring average membrane potential

This Example describes quantification of the mitochondrial membrane potential of fusosomes. In an embodiment, fusosomes comprising a mitochondrial membrane will maintain mitochondrial membrane potential.

Mitochondrial metabolic activity can be measured by mitochondrial membrane potential. The membrane potential of the fusosome preparation is quantified using a commercially available dye, TMRE, for assessing mitochondrial membrane potential (TMRE: tetramethyl rhodamine, ethyl ester, perchlorate, Abcam, Cat# T669).

Fusosomes are generated by any one of the methods described in previous Examples. Fusosomes or parental cells are diluted in growth medium (phenol-red free DMEM with 10% fetal bovine serum) in 6 aliquots (untreated and FCCP-treated triplicates). One aliquot of the

samples is incubated with FCCP, an uncoupler that eliminates mitochondrial membrane potential and prevents TMRE staining. For FCCP-treated samples, 2μM FCCP is added to the samples and incubated for 5 minutes prior to analysis. Fusosomes and parental cells are then stained with 30nM TMRE. For each sample, an unstained (no TMRE) sample is also prepared in parallel. Samples are incubated at 37°C for 30 minutes. The samples are then analyzed on a flow cytometer with 488nm argon laser, and excitation and emission is collected at 530+/-30nm.

Membrane potential values (in millivolts, mV) are calculated based on the intensity of TMRE. All events are captured in the forward and side scatter channels (alternatively, a gate can be applied to exclude small debris). The fluorescence intensity (FI) value for both the untreated and FCCP-treated samples are normalized by subtracting the geometric mean of the fluorescence intensity of the unstained sample from the geometric mean of the untreated and FCCP-treated sample. The membrane potential state for each preparation is calculated using the normalized fluorescent intensity values with a modified Nernst equation (see below) that can be used to determine mitochondrial membrane potential of the fusosomes or cells based on TMRE fluorescence (as TMRE accumulates in mitochondria in a Nernstian fashion).

Fusosome or cell membrane potential is calculated with the following formula: (mV) = -61.5 * log(FIuntreated-normalized/FIFCCP-treated-normalized). In an embodiment, using this assay on fusosome preparations from C2C12 mouse myoblast cells, the membrane potential state of the fusosome preparation will be within about 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than the parental cells. In an embodiment, the range of membrane potential is about -20 to -150mV.

Example 60: Measuring persistence half-life in a subject

This Example describes the measurement of fusosome half-life.

Fusosomes are derived from cells that express Gaussia luciferase produced by any one of the methods described in previous Examples, and pure, 1:2, 1:5, and 1:10 dilutions in buffered solution are made. A buffered solution lacking fusosomes is used as a negative control.

Each dose is administered to three eight week old male C57BL/6J mice (Jackson Laboratories) intravenously. Blood is collected from the retro-orbital vein at 1, 2, 3, 4, 5, 6, 12, 24, 48, and 72 hours after intravenous administration of the fusosomes. The animals are sacrificed at the end of the experiment by CO_2 inhalation.

Blood is centrifuged for 20 min at room temperature. The serum samples are immediately frozen at −80°C until bioanalysis. Then, each blood sample is used to carry out a Gaussia luciferase activity assay after mixing the samples with Gaussia luciferase substrate (Nanolight, Pinetop, AZ). Briefly, colenterazine, a luciferin or light-emitting molecule, is mixed with flash assay buffer and the mixture is pipetted into wells containing blood samples in a 96 well plate. Negative control wells that lack blood contain assay buffer to determine background Gaussia luciferase signal.

In addition, a standard curve of positive-control purified Gaussia luciferase (Athena Enzyme Systems, catalog #0308) is prepared in order to convert the luminescence signal to molecules of Gaussia luciferase secretion per hour. The plate is assayed for luminescence, using 500 msec integration. Background Gaussia luciferase signal is subtracted from all samples and then a linear best-fit curve is calculated for the Gaussia luciferase standard curve. If sample readings do not fit within the standard curve, they are diluted appropriately and re-assayed. The luciferase signal from samples taken at 1, 2, 3, 4, 5, 6, 12, 24, 48, and 72 hours is interpolated to the standard curve. The elimination rate constant k_e (h^{-1}) is calculated using the following equation of a one-compartment model: $C(t) = C_0 \, x \, e^{-kext}$, in which $C(t)$ (ng/mL) is the concentration of fusosomes at time t (h) and C_0 the concentration of fusosomes at time = 0 (ng/mL). The elimination half-life $t_{1/2,e}$ (h) is calculated as $\ln(2)/k_e$.

In an embodiment, fusosomes will have a half-life of at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or greater than the negative control cells.

Example 61: Delivery of fusosomes via non-endocytic pathway

This example describes quantification of fusosome delivery of Cre to a recipient cell via a non-endocytic pathway.

In an embodiment, fusosomes will deliver agents via a fusosome-mediated, non-endocytic pathway. Without wishing to be bound by theory, delivery of an agent, e.g., Cre, which is carried within the lumen of the fusosomes, directly to the cytosol of the recipient cells without any requirement for endocytosis-mediated uptake of the fusosomes, will occur through a fusosome-mediated, non-endocytic pathway delivery.

In this example, the fusosome comprises a HEK293T cell expressing the Sendai virus H and F protein on its plasma membrane (Tanaka et al., 2015, Gene Therapy, 22(October 2014), 1–

8. https://doi.org/10.1038/gt.2014.123). In addition, the fusosome expresses mTagBFP2 fluorescent protein and Cre recombinase. The target cell is a RPMI8226 cell which stably-expresses "LoxP-GFP-stop-LoxP-RFP" cassette under a CMV promoter, which upon recombination by Cre switches from GFP to RFP expression, indicating fusion and Cre, as a marker, delivery.

Fusosomes produced by the herein described methods are assayed for delivery of Cre via a non-endocytic pathway as follows. The recipient cells are plated into a black, clear-bottom 96-well plate. Next, 24 hours after plating the recipient cells, the fusosomes expressing Cre recombinase protein and possessing the particular fusogen protein are applied to the recipient cells in DMEM media. To determine the level of Cre delivery via a non-endocytic pathway, a parallel group of recipient cells receiving fusosomes is treated with an inhibitor of endosomal acidification, chloroquine (30 μg/mL). The dose of fusosomes is correlated to the number of recipient cells plated in the well. After applying the fusosomes, the cell plate is centrifuged at 400g for 5 minutes to help initiate contact between the fusosomes and the recipient cells. The cells are then incubated for 16 hours and agent delivery, Cre, is assessed via imaging.

The cells are imaged to positively identify RFP-positive cells versus GFP-positive cells in the field or well. In this example cell plates are imaged using an automated fluorescence microscope. The total cell population in a given well is determined by first staining the cells with Hoechst 33342 in DMEM media for 10 minutes. Hoechst 33342 stains cell nuclei by intercalating into DNA and therefore is used to identify individual cells. After staining, the Hoechst media is replaced with regular DMEM media.

The Hoechst is imaged using the 405 nm LED and DAPI filter cube. GFP is imaged using the 465 nm LED and GFP filter cube, while RFP is imaged using 523 nm LED and RFP filter cube. Images of the different cell groups are acquired by first establishing the LED intensity and integration times on a positive-control well; i.e., recipient cells treated with adenovirus coding for Cre recombinase instead of fusosomes.

Acquisition settings are set so that RFP and GFP intensities are at the maximum pixel intensity values but not saturated. The wells of interest are then imaged using the established settings.

Analysis of GFP and RFP-positive wells is performed with software provided with the fluorescence microscope or other software (Rasband, W.S., ImageJ, U. S. National Institutes of

Health, Bethesda, Maryland, USA, 1997–2007). The images are pre-processed using a rolling ball background subtraction algorithm with a 60 μm width. The total cell mask is set on the Hoechst-positive cells. Cells with Hoechst intensity significantly above background intensities are used to set a threshold, and areas too small or large to be Hoechst-positive cells are excluded.

Within the total cell mask, GFP and RFP-positive cells are identified by again setting a threshold for cells significantly above background and extending the Hoechst (nuclei) masks for the entire cell area to include the entire GFP and RFP cellular fluorescence.

The number of RFP-positive cells identified in control wells containing recipient cells is used to subtract from the number of RFP-positive cells in the wells containing fusosomes (to subtract for non-specific Loxp recombination). The number of RFP-positive cells (recipient cells that received Cre) is then divided by the sum of GFP-positive cells (recipient cells that have not received Cre) and RFP-positive cells to quantify the fraction of fusosome Cre delivered to the recipient cell population. The level is normalized to the given dose of fusosomes applied to the recipient cells. To calculate the value of fusosome Cre delivered via a non-endocytic pathway, the level of fusosome Cre delivery in the presence of chloroquine (FusL+CQ) is determined as well as the level of fusosome Cre delivery in the absence of chloroquine (FusL-CQ). To determine the normalized value of fusosome Cre delivered via a non-endocytic pathway, the following equation is used: [(FusL-CQ)-(FusL+CQ)]/(FusL-CQ).

In an embodiment, the average level of fusosome Cre delivered via a non-endocytic pathway for a given fusosome will be in the range of 0.1-0.95, or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than chloroquine treated recipient cells.

Example 62: Delivery of fusosomes via endocytic pathway

This example describes fusosome delivery of Cre to a recipient cell via an endocytic pathway.

In an embodiment, fusosomes will deliver agents via a fusosome-mediated, endocytic pathway. Without wishing to be bound by theory, delivery of an agent, e.g., a cargo, carried in the lumen of the fusosomes, to the recipient cells with the route of uptake being endocytosis-dependent will occur through a fusosome-mediated, endocytic pathway delivery.

In this example the fusosome comprises microvesicles that were produced by extruding a HEK293T cell expressing a fusogen protein on its plasma membrane through a 2 μm filter (Lin et al., 2016, Biomedical Microdevices, 18(3). doi.org/10.1007/s10544-016-0066-y)(Riedel, Kondor-Koch, & Garoff, 1984, The EMBO Journal, 3(7), 1477–83. Retrieved from www.ncbi.nlm.nih.gov/pubmed/6086326). In addition, the fusosome expresses mTagBFP2 fluorescent protein and Cre recombinase. The target cell is a PC3 cell which stably-expresses "LoxP-GFP-stop-LoxP-RFP" cassette under a CMV promoter, which upon recombination by Cre switches from GFP to RFP expression, indicating fusion and Cre, as a marker, delivery.

Fusosomes produced by the herein described methods are assayed for delivery of Cre via an endocytic pathway as follows. The recipient cells are plated into a cell culture multi-well plate compatible with the imaging system to be used (in this example cells are plated in a black, clear-bottom 96-well plate). Next, 24 hours after plating the recipient cells, the fusosomes expressing Cre recombinase protein and possessing the particular fusogen protein are applied to the recipient cells in DMEM media. To determine the level of Cre delivery via an endocytic pathway, a parallel group of recipient cells receiving fusosomes is treated with an inhibitor of endosomal acidification, chloroquine (30 μg/mL). The dose of fusosomes is correlated to the number of recipient cells plated in the well. After applying the fusosomes, the cell plate is centrifuged at 400g for 5 minutes to help initiate contact between the fusosomes and the recipient cells. The cells are then incubated for 16 hours and agent delivery, Cre, is assessed via imaging.

The cells are imaged to positively identify RFP-positive cells versus GFP-positive cells in the field or well. In this example cell plates are imaged using an automated fluorescent microscope. The total cell population in a given well is determined by first staining the cells with Hoechst 33342 in DMEM media for 10 minutes. Hoechst 33342 stains cell nuclei by intercalating into DNA and therefore is used to identify individual cells. After staining the Hoechst media is replaced with regular DMEM media.

The Hoechst is imaged using the 405 nm LED and DAPI filter cube. GFP is imaged using the 465 nm LED and GFP filter cube, while RFP is imaged using 523 nm LED and RFP filter cube. Images of the different cell groups are acquired by first establishing the LED intensity and integration times on a positive-control well; i.e., recipient cells treated with adenovirus coding for Cre recombinase instead of fusosomes.

Acquisition settings are set so that RFP and GFP intensities are at the maximum pixel intensity values but not saturated. The wells of interest are then imaged using the established settings.

Analysis of GFP and RFP-positive wells is performed with software provided with the fluorescent microscope or other software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, 1997–2007). The images are pre-processed using a rolling ball background subtraction algorithm with a 60 μm width. The total cell mask is set on the Hoechst-positive cells. Cells with Hoechst intensity significantly above background intensities are thresholded and areas too small or large to be Hoechst-positive cells are excluded.

Within the total cell mask, GFP and RFP-positive cells are identified by again thresholding for cells significantly above background and extending the Hoechst (nuclei) masks for the entire cell area to include the entire GFP and RFP cellular fluorescence.

The number of RFP-positive cells identified in control wells containing recipient cells is used to subtract from the number of RFP-positive cells in the wells containing fusosomes (to subtract for non-specific Loxp recombination). The number of RFP-positive cells (recipient cells that received Cre) is then divided by the sum of the GFP-positive cells (recipient cells that have not received Cre) and RFP-positive cells to quantify the fraction of fusosome Cre delivered to the recipient cell population. The level is normalized to the given dose of fusosomes applied to the recipient cells. To calculate the value of fusosome Cre delivered via an endocytic pathway, the level of fusosome Cre delivery in the presence of chloroquine (FusL+CQ) is determined as well as the level of fusosome Cre delivery in the absence of chloroquine (FusL-CQ). To determine the normalized value of fusosome Cre delivered via an endocytic pathway, the following equation is used: (FusL+CQ)/(FusL-CQ).

In an embodiment, the average level of fusosome Cre delivered via an endocytic pathway for a given fusosome will be in the range of 0.01-0.6, or at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater than chloroquine treated recipient cells.

Example 63: Delivery of fusosomes via a dynamin mediated pathway, a macropinocytosis pathway, or an actin mediated pathway

This example describes fusosome delivery of Cre to a recipient cell via a dynamin mediated pathway. A fusosome comprising a microvesicle may be produced as described in the

preceding example. Fusosomes are assayed for delivery of Cre via a dynamin-mediated pathway according to the preceding example, except that a group of recipient cells receiving fusosomes is treated with an inhibitor of dynamin, Dynasore (120 μM). To calculate the value of fusosome Cre delivered via a dynamin-mediated pathway, the level of fusosome Cre delivery in the presence of Dynasore (FusL+DS) is determined as well as the level of fusosome Cre delivery in the absence of Dynasore (FusL-DS). The normalized value of fusosome Cre delivered may be calculated as described in the preceding example.

This example also describes delivery of Cre to a recipient cell via macropinocytosis. A fusosome comprising a microvesicle may be produced as described in the preceding example. Fusosomes are assayed for delivery of Cre via macropinocytosis according to the preceding example, except that a group of recipient cells receiving fusosomes is treated with an inhibitor of macropinocytosis, 5-(*N*-ethyl-*N*-isopropyl)amiloride (EIPA) (25 μM). To calculate the value of fusosome Cre delivered via macropinocytosis, the level of fusosome Cre delivery in the presence of EIPA (FusL+EPIA) is determined as well as the level of fusosome Cre delivery in the absence of EPIA (FusL-EIPA). The normalized value of fusosome Cre delivered may be calculated as described in the preceding example.

This example also describes fusosome delivery of Cre to a recipient cell via an actin mediated pathway. A fusosome comprising a microvesicle may be produced as described in the preceding example. Fusosomes are assayed for delivery of Cre via macropinocytosis according to the preceding example, except that a group of recipient cells receiving fusosomes is treated with an inhibitor of actin polymerization, Latrunculin B (6 μM). To calculate the value of fusosome Cre delivered via an actin-mediated pathway, the level of fusosome Cre delivery in the presence of Latrunculin B (FusL+ LatB) is determined as well as the level of fusosome Cre delivery in the absence of Latrunculin B (FusL- LatB). The normalized value of fusosome Cre delivered may be calculated as described in the preceding example.

Example 64: In vivo delivery of protein

This example describes the delivery of therapeutic agents to the eye by fusosomes.

Fusosomes are derived from hematopoietic stem and progenitor cells using any of the methods described in previous Examples and are loaded with a protein that is deficient in a mouse knock-out.

Fusosomes are injected subretinally into the right eye of a mouse that is deficient for the protein and vehicle control is injected into the left eye of the mice. A subset of the mice is euthanized when they reach 2 months of age.

Histology and H&E staining of the harvested retinal tissue is conducted to count the number of cells rescued in each retina of the mice (described in Sanges et al., The Journal of Clinical Investigation, 126(8): 3104-3116, 2016).

The level of the injected protein is measured in retinas harvested from mice euthanized at 2 months of age via a western blot with an antibody specific to the PDE6B protein.

In an embodiment, the left eyes of mice, which are administered fusosomes, will have an increased number of nuclei present in the outer nuclear level of the retina compared to the right eyes of mice, which are treated with vehicle. The increased protein is suggestive of complementation of the mutated PBE6B protein.

Example 65: Assessment of teratoma formation after administration of fusosome

This Example describes the absence of teratoma formation with a fusosome. In an embodiment, a fusosome will not result in teratoma formation when administered to a subject.

The fusosomes are produced by any one of the methods described in a previous Example. Fusosomes, tumor cells (positive control) or vehicle (negative control) are subcutaneously injected in PBS into the left flank of mice (12-20 weeks old). Teratoma, e.g., tumor, growth is analyzed 2-3 times a week by determination of tumor volume by caliper measurements for eight weeks after fusosome, tumor cell, or vehicle injection.

In an embodiment, mice administered fusosomes or vehicle will not have a measurable tumor formation, e.g., teratoma, via caliper measurements. In an embodiment, positive control animals treated with tumor cells will demonstrate an appreciable tumor, e.g., teratoma, size as measured by calipers over the eight weeks of observation.

Example 66: Measuring total RNA in a fusosome and source cell

This Example describes a method to quantify the amount of RNA in a fusosome relative to a source cell. In an embodiment, a fusosome will have similar RNA levels to the source cell. In this assay, RNA levels are determined by measuring total RNA.

Fusosomes are prepared by any one of the methods described in previous Examples. Preparations of the same mass as measured by protein of fusosomes and source cells are used to isolate total RNA (e.g., using a kit such as Qiagen RNeasy catalog #74104), followed by determination of RNA concentration using standard spectroscopic methods to assess light absorbance by RNA (e.g. with Thermo Scientific NanoDrop).

In an embodiment, the concentration of RNA in fusosomes will be 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% of that of source cells per mass of protein.

Example 67: Isolating fusogenic microvesicles freely released from cells

This example describes the isolation of fusogenic microvesicles freely released from cells. Fusogenic microvesicles were isolated as follows. 9.2 x 10^6 HEK-293T (ATCC, Cat# CRL-3216) were reverse transfected using Xfect transfection reagent (Takara, Cat# 631317) with 10 μg of the pcDNA3.1 expression plasmid containing the open reading frame for VSVg and 15ug of the pcDNA3.1 expression plasmid containing the open reading frame for bacteriophage P1 Cre Recombinase with a SV40 Nuclear localization sequence in 7.5mL of complete media (Dulbecco's Modified Eagle Medium (DMEM) supplemented with GlutaMAX (ThermoFisher), 10% fetal calf serum (ThermoFisher), and penicillin/streptomycin antibiotics (ThermoFisher)) in a 100mm collagen coated dish (Corning). Twelve hours after seeding, an additional 7.5mL of complete medium was carefully added. The cells were separated from culture media by centrifugation at 200 *x g* for 10 minutes. Supernatants were collected and centrifuged sequentially twice at 500 x g for 10 minutes, once at 2,000 x g for 15 minutes, once at 10,000 x g for 30 min, and once at 70,000 x g for 60 minutes. Freely released fusosomes were pelleted during the final centrifugation step, resuspended in PBS and repelleted at 70,000 x g. The final pellet was resuspended in PBS.

See also, Wubbolts R *et al.* Proteomic and Biochemical Analyses of Human B Cell-derived Exosomes: Potential Implications for their Function and Multivesicular Body Formation. *J. Biol. Chem.* 278:10963-10972 2003.

Example 68: Measuring the average size distribution of fusosomes

This Example describes measurement of the size distribution of fusosomes.

Fusosomes were prepared as described herein by transient transfection of HEK293T with VSV-G, enucleation and subsequent fractionation with Ficoll. The fusosomes were measured to determine the size distribution using the method of Example 27, as shown in FIG. 3. It is contemplated that the fusosomes can have less than about 50%, 40%, 30%, 20%, 10%, 5%, or less of the parental cell's variability in size distribution within 90% of the sample. It is contemplated that the fusosomes can have 58% less of the parental cell's variability in size distribution within 90% of the sample.

Example 69: Average volume of fusosomes

This example describes measurement of the average volume of fusosomes. Varying the size (e.g., volume) of fusosomes can make them versatile for distinct cargo loading, therapeutic design or application.

Fusosomes were prepared as described herein by transient transfection of HEK293T with VSV-G, enucleation and subsequent fractionation with Ficoll. The positive control was HEK293T cells.

Analysis with a combination of NTA and confocal microscopy as described in Example 27 was used to determine the size of the fusosomes. The diameter of the fusosomes were measured and the volume calculated, as shown in FIG. 4. It is contemplated that fusosomes can have an average size of greater than 50 nm in diameter. It is contemplated that fusosomes can have an average size of 129 nm in diameter.

Example 70: Comparison of soluble to insoluble protein mass

This Example describes quantification of the soluble:insoluble ratio of protein mass in fusosomes. The soluble:insoluble ratio of protein mass in fusosomes can, in some instances, be similar to that of nucleated cells.

Fusosomes were prepared as described herein by transient transfection of HEK293T with VSV-G, enucleation and subsequent fractionation with Ficoll. The fusosome preparation was tested to determine the soluble:insoluble protein ratio using a standard bicinchoninic acid assay (BCA) (Pierce™ BCA Protein Assay Kit, Thermo Fischer product# 23225). Soluble protein samples were prepared by suspending the prepared fusosomes or parental cells at a concentration of $1x10^7$ cells or ~1 mg/mL total fusosomes in PBS and centrifuging at 1,500 x g to pellet the

cells or 16,000 x g to pellet the fusosomes. The supernatant was collected as the soluble protein fraction.

The fusosomes or cells were then resuspended in PBS. This suspension represents the insoluble protein fraction.

A standard curve was generated using the supplied BSA, from 0 to 15 µg of BSA per well (in duplicate). The fusosome or cell preparation was diluted such that the quantity measured is within the range of the standards. The fusosome preparation was analyzed in duplicate and the mean value was used. The soluble protein concentration was divided by the insoluble protein concentration to yield the soluble:insoluble protein ratio (FIG. 5).

Example 71: Measuring fusion with a target cell

Fusosomes derived from HEK-293T cells expressing the engineered hemagglutinin glycoprotein of measles virus (MvH) and the fusion protein (F) on the cell surface and containing Cre recombinase protein were generated, as described herein. The MvH was engineered so that its natural receptor binding is ablated and target cell specificity is provided through a single-chain anbitody (scFv) that recognizes the cell surface antigen, in this case the scFv is designed to target CD8, a co-receptor for the T cell receptor. A control fusosome was used which was derived from HEK-293T cells expressing the fusogen VSV-G on its surface and containing Cre recombinase protein. The target cell was a HEK-293T cell engineered to express a "Loxp-GFP-stop-Loxp-RFP" cassette under CMV promoter, as well as engineered to over-express the co-receptors CD8a and CD8b. The non-target cell was the same HEK-293T cell expressing "Loxp-GFP-stop-Loxp-RFP" cassette but without CD8a/b over-expression. The target or non-target recipient cells were plated 30,000 cells/well into a black, clear-bottom 96-well plate and cultured in DMEM media with 10% fetal bovine serum at 37ºC and 5% CO_2. Four to six hours after plating the recipient cells, the fusosomes expressing Cre recombinase protein and MvH+F were applied to the target or non-target recipient cells in DMEM media. Recipient cells were treated with 10 µg of fusosomes and incubated for 24 hours at 37°C and 5% CO_2.

Cell plates were imaged using an automated microscope (www.biotek.com/products/imaging-microscopy-automated-cell-imagers/lionheart-fx-automated-live-cell-imager/). The total cell population in a given well was determined by staining the cells with Hoechst 33342 in DMEM media for 10 minutes. Hoechst 33342 stains cell

nuclei by intercalating into DNA and therefore is used to identify individual cells. The Hoechst was imaged using the 405 nm LED and DAPI filter cube. GFP was imaged using the 465 nm LED and GFP filter cube, while RFP was imaged using 523 nm LED and RFP filter cube. Images of target and non-target cell wells were acquired by first establishing the LED intensity and integration times on a positive-control well; i.e., recipient cells treated with adenovirus coding for Cre recombinase instead of fusosomes.

Acquisition settings were set so that Hoescht, RFP, and GFP intensities are at the maximum pixel intensity values but not saturated. The wells of interest were then imaged using the established settings. Focus was set on each well by autofocusing on the Hoescht channel and then using the established focal plane for the GFP and RFP channels. Analysis of GFP and RFP-positive cells was performed with Gen5 software provided with automated fluorescent microscope (https://www.biotek.com/products/software-robotics-software/gen5-microplate-reader-and-imager-software/).

The images were pre-processed using a rolling ball background subtraction algorithm with a 60 μm width. Cells with GFP intensity significantly above background intensities were thresholded and areas too small or large to be GFP-positive cells were excluded. The same analysis steps were applied to the RFP channel. The number of RFP-positive cells (recipient cells receiving Cre) was then divided by the sum of the GFP-positive cells (recipient cells that did not show delivery) and RFP-positive cells to quantify the percent RFP conversion, which describes the amount of fusosome fusion within the target and non-target recipient cell population. For amounts of targeted fusion (fusosome fusion to targeted recipient cells), the percent RFP conversion value is normalized to the percentage of recipient cells that are target recipient cells (i.e., expressing CD8), which was assessed by staining with anti-CD8 antibody conjugated to phycoerythrin (PE) and analyzed by flow cytometry. Finally, the absolute amount of targeted fusion was determined by subtracting the amount of non-target cell fusion from the target cell fusion amount (any value <0 was considered to be 0).

With this assay, the fusosome derived from a HEK-293T cell expressing the engineered MvH(CD8)+F on its surface and containing Cre recombinase protein showed a percentage RFP conversion of 25.2 +/- 6.4% when the recipient cell was the target HEK-293T cell expressing the "Loxp-GFP-stop-Loxp-RFP" cassette, and 51.1% of these recipient cells were observed to be CD8-positive. From these results, the normalized percentage RFP conversion or amount of

targeted fusion was determined to be 49.3 +/- 12.7% for targeted fusion. The same fusosome showed a percentage RFP conversion of 0.5 +/- 0.1% when the recipient was the non-target HEK-293T cell expressing "Loxp-GFP-stop-Loxp-RFP" but with no expression of CD8. Based on the above, the absolute amount of targeted fusion for the MvH(CD8)+F fusosome determined to be 48.8% and the absolute amount of targeted fusion for the control VSV-G fusosome was determined to be 0% (FIG. 6).

Example 72: Measuring ability to transport glucose across cell membrane

Fusosomes from HEK-293T cells expressing the envelope glycoprotein G from vesicular stomatitis virus (VSV-G) on the cell surface and expressing Cre recombinase protein were generated according by the standard procedure of ultracentrifugation through a Ficoll gradient to obtain small particle fusosomes as described herein. To measure the ability of the fusosomes to transport glucose across the cell membrane, the levels of a 2-NBDG (2-(*N*-(7-Nitrobenz-2-oxa-1,3-diazol-4-yl)Amino)-2-Deoxyglucose) fluorescent glucose analog, that can be used to monitor glucose uptake in live cells, was quantified to assess active transport across the lipid bilayer. A commercially-available kit from Biovision Inc. (Cat #K682) was used for the assay according to manufacturer's instructions.

Briefly, the fusosome sample was measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Next 40ug of fusosome total protein was pelleted by centrifugation at 3000g for 5 minutes in a table-top centrifuge, followed by resuspension in 400uL of DMEM supplemented with 0.5% fetal bovine serum. This was done in duplicate for each sample, and one of the duplicates was treated with 4uL of phloretin (provided with the kit), a natural phenol that inhibits glucose uptake, as a control for glucose uptake inhibition. The samples were then incubated for 1 hour at room temperature. After the incubation, the fusosome sample was pelleted and resuspended in 400uL of glucose uptake mix prepared previously (see Table 10 below for formulation). Samples pre-treated with phloretin were resuspended in glucose uptake mix with phloretin; samples not pre-treated were resuspended in glucose uptake mix with 20uL of PBS instead of phloretin. Also a parallel set of fusosome samples were resuspended in DMEM media with 0.5% FBS only as a negative control for flow cytometry analysis.

Table 10: Glucose uptake mix formulation

Reagent	Volume (uL)
DMEM media with 0.5% FBS	1880
2-NBDG reagent	20
Glucose Uptake Enhancer	100
Optional: Phloretin	20

The samples were then incubated at 37°C with 5% CO_2 for 30 minutes. After the incubation cells were pelleted, washed once with 1mL of 1X Analysis Buffer (provided with kit), pelleted again, and resuspended in 400uL of 1X Analysis Buffer.

The samples were then measured for 2-NBDG uptake by flow cytometry analysis using an Invitrogen Attune NxT acoustic focusing cytometer. 2-NBDG was excited with a 488nm laser and emission captured at 513±26nm. Forward and side scatter gating was initially used to capture fusosome-sized events and discard small debris. Events positive for 2-NBDG were determined by gating at the minimum level for which the 2-NBDG negative control sample showed <0.5% of events positive for 2-NBDG staining. The gated cells positive for 2-NBDG fluorescence were then assessed for the mean fluorescence intensity (F.I.) of 2-NBDG in order to calculate a value for glucose uptake for the fusosomes with and without phloretin treatment.

With this assay, the fusosome derived from a HEK-293T cell expressing the VSV-G and Cre showed a 2-NBDG mean F.I. of 631.0 +/- 1.4 without phloretin treatment and a mean F.I. of 565.5 +/- 4.9 with phloretin treatment (FIG. 7).

Example 73: Measuring esterase activity in the cytosol

Fusosomes from C2C12 cells were generated according to the standard procedure of ultracentrifugation through a Ficoll gradient to obtain small particle fusosomes as described herein. To measure the esterase activity in the cytosol of the fusosomes, samples were stained with Calcein AM (BD Pharmigen, Cat #564061), a fluorescein derivative and nonfluorescent vital dye that passively crosses the cell membrane of viable cells and is converted by cytosolic esterases into green fluorescent calcein, which is retained by cells with intact membranes and inactive multidrug resistance protein.

Briefly, the fusosome sample was measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Next 20ug of fusosome total protein was pelleted by centrifugation at 3000g for 5 minutes in a table-

top centrifuge, followed by resuspension in 400uL of DMEM supplemented with 0.5% fetal bovine serum. The membrane-permeable dye, calcein-AM was prepared as a stock solution of 10 mM in dimethylsulfoxide and as a working solution of 1 mM in PBS buffer, pH 7.4. VSV-G fusosomes were stained with 1 μM solution of calcein-AM diluted in DMEM media. Samples were incubated at 37°C in the dark for 30 minutes and then pelleted by centrifugation. After washing twice with PBS buffer, fusosomes were resuspended in PBS and analyzed by flow cytometry.

The samples were measured for calcein fluorescence retention using an Invitrogen Attune NxT acoustic focusing cytometer. Calcein AM was excited with a 488nm laser and emission captured at 513±26nm. Forward and side scatter gating was initially used to capture fusosome-sized events and discard small debris. Events positive for calcein were determined by gating at the minimum level for which the calcein negative control sample showed <0.5% of events positive for calcein staining. The gated cells positive for calcein fluorescence were then assessed for the mean fluorescence intensity (F.I.) of calcein in order to calculate a value for esterase activity in the cytosol of fusosomes.

With this assay the fusosome derived from a C2C12 cell showed an esterase activity (mean calcein F.I.) of 631.0 +/- 1.4 (FIG. 8).

Example 74: Measuring acetylcholinesterase activity in fusosomes

Fusosomes from HEK-293T cells expressing the placental cell-cell fusion protein syncytin-1 (Syn1) on the cell surface and expressing Cre recombinase protein were generated as described herein. Acetylcholinesterase activity was measured using the FluoroCet Quantitation Kit (System Biosciences, Cat #FCET96A-1) following the manufacturer's recommendations.

Briefly, fusosomes were pelleted via ultracentrifugation at 120,000 g for 90 minutes and resuspended carefully in phosphate-buffered saline (PBS). Next fusosomes were quantified for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. After BCA quantification of protein concentration, 1000ng of total fusosome protein was diluted with PBS to a volume of 60uL, followed by addition of 60uL of Lysis Buffer to lyse the particles. After a 30 minute incubation on ice the samples were ready to run in the FluoroCet assay.

In duplicate wells of a 96-well plate, 50uL of lysed fusosome sample was mixed with 50uL of Working stock of Buffer A and 50uL of Working stock of Buffer B. In parallel, a standard curve was prepared by pipetting 2uL of the provided standard in 126uL of 1X Reaction buffer. This standard solution was then serial diluted 5X to make a six-point standard curve consisting of 2.0E+08, 1.0E+08, 5.0E+07, 2.5E+07, 1.25E+07, and 6.25E+06 exosome equivalents of acetylcholinesterase activity. 50uL of each standard was then mixed with 50uL of Working stock of Buffer A and 50uL of Working stock of Buffer B in duplicate wells of the 96-well plate. 50uL of 1X Reaction buffer was used as a blank. The plate was mixed by tapping the sides followed by incubation in the dark for 20 minutes at room temperature. The plate was then measured immediately using a fluorescence plate reader set at Excitation: 530-570nm and Emission: 590- 600nm. The plate was shaken for 30 sec before reading.

The relative fluorescence units (RFU) were then plotted against the known exosome equivalents of acetylcholinesterase activity after subtracting the RFU values from the blank wells. A linear regression line was then calculated and the equation used to determine the acetylcholinesterase activity (in exosome equivalents) for the fusosome samples from the measured RFU values. The measured acetylcholinesterase activity for Syn1 fusosomes are shown in Table 11:

Table 11: Acetylcholinesterase activity in fusosomes and control particles

Sample	**Acetylcholinesterase activity (exosome equivalents)**
Syn1 fusosomes	6.83E+05 +/- 2.21E+05

Example 75: Measuring metabolic activity level

Fusosomes from HEK-293T cells expressing the envelope glycoprotein G from vesicular stomatitis virus (VSV-G) on the cell surface and expressing Cre recombinase protein were generated as described herein. To determine the metabolic activity level of the fusosome preparation, citrate synthase activity was assessed using a commercially available kit from Sigma (Cat #CS0720) which provides all of the necessary reagents. Citrate synthase is an enzyme within the tricarboxylic acid (TCA) cycle that catalyzes the reaction between oxaloacetate (OAA) and acetyl-CoA to generate citrate. Upon hydrolysis of acetyl-CoA, there is a release of CoA with a thiol group (CoA-SH). The thiol group reacts with a chemical reagent, 5,5-Dithiobis-(2-nitrobenzoic acid) (DTNB), to form 5-thio-2-nitrobenzoic acid (TNB), which has a yellow product that can be measured spectrophotometrically at 412 nm.

The assay was performed as per the manufacturer's recommendations. Briefly, fusosome sample was measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Next 400ug of fusosome total protein was pelleted by centrifugation at 3000g for 5 minutes in a table-top centrifuge. The fusosomes were washed once by pelleting again and resuspending in ice-cold PBS. Fusosomes were pelleted again and supernatant was removed. The pellet was lysed in 100uL of CellLytic M buffer with 1X protease inhibitors. After mixing by pipetting, the lysed sample was incubated for 15 minutes at room temperature to complete lysis. The sample was then centrifuged at 12,000g for 10 minutes and the supernatant was transferred to a new microcentrifuge tube and stored at -80°C until the subsequent assay was performed.

To initiate the citrate synthase activity assay, all assay solutions were warmed to room temperature prior to using. The lysed fusosome sample was mixed with assay solutions according to Table 12 below:

Table 12: Reaction Scheme for Citrate Synthase Activity Measurement in 96 Well Plate

Sample	Assay buffer	30mM Acetyl CoA solution	10mM DTNB solution	10mM OAA solution (added last)
4uL	182uL	2uL	2uL	10uL

The volumes in Table 12 represent volumes for a single well of a 96-well plate. Samples were measured in duplicates. All components of the reaction were mixed and pipetted into a single well of a 96-well plate. The absorbance at 412nm was then analyzed on a microplate reader for 1.5 minutes to measure the baseline reaction. Next, 10uL of the 10mM OAA solution was added to each well to initiate the reaction. The plate was shaken for 10 seconds in the microplate reader before reading the absorbance at 412nm for 1.5 minutes with a measurement every 10 seconds.

To calculate the citrate synthase activity, the absorbance at 412nm was plotted against time for each reaction. The change in absorbance per minute was calculated for the linear range of the plot for before (endogenous activity) and after (total activity) OAA addition. The net citrate synthase activity was then calculated by subtracting the endogenous activity from the total activity for the sample. This value was then used to calculate the citrate synthase activity based on the equation and constant values provided by the manufacturer. The measured citrate synthase activity for the VSV-G fusosomes was 1.57E-02 +/- 1.86E-03 umol/ug fusosome/min.

Example 76: Measuring respiration levels

Fusosomes from HEK-293T cells expressing the envelope glycoprotein G from vesicular stomatitis virus (VSV-G) on the cell surface were generated according by the standard procedure of ultracentrifugation through a Ficoll gradient to obtain small particle fusosomes as described herein. Respiration level in the fusosome preparation were determined by measuring mitochondrial oxygen consumption rates by a Seahorse extracellular flux analyzer (Agilent).

Briefly, the fusosome sample was measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Next 20 µg of fusosome total protein was pelleted by centrifugation at 3000g for 5 minutes in a table-top centrifuge, followed by resuspension (in quadruplicates) in 150 µL of XF Assay media (Agilent Cat # 103575-100) supplemented with 25 mM glucose and 2 mM glutamine (pH 7.4). The resuspended samples were then added to one well of a 96-well Seahorse plate (Agilent).

Oxygen consumption assays were initiated by incubating the 96-well Seahorse plate with samples at 37°C for 60 minutes to allow temperature and pH to reach equilibrium. The microplate was then assayed in the XF96 Extracellular Flux Analyzer (Agilent) to measure extracellular flux changes of oxygen and pH in the media immediately surrounding the fusosomes. After obtaining steady state oxygen consumption and extracellular acidification rates, oligomycin (5µM), which inhibits ATP synthase, and proton ionophore FCCP (carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone; 2µM), which uncouples mitochondria, were injected sequentially through reagent delivery chambers for each well in the microplate to obtain values for maximal oxygen consumption rates. Finally, 5 µM antimycin A (inhibitor of mitochondrial complex III) was injected to confirm that respiration changes were due mainly to mitochondrial respiration. The rates of antimycin A respiration were subtracted from the other three respiration rates in order to determine the basal, uncoupled (oligomycin-resistant), and maximal (FCCP-induced) mitochondrial respiration rates.

Using this assay it was determined that donor VSV-G fusosomes showed basal, uncoupled, and maximal oxygen consumption (respiration) rates according to Table 13 below.

Table 13: Respiration rates of VSV-G fusosomes

Respiration state	Mitochondrial oxygen consumption (respiration) rate

	(pmol/min/20µg fusosome) AVG ± SEM
Basal	11.3 ± 3.0
Uncoupled	10.1 ± 2.3
Maximal	20.0 ± 1.9

Example 77: Measuring phosphatidylserine levels of fusosomes

Fusosomes from HEK-293T cells expressing the envelope glycoprotein G from vesicular stomatitis virus (VSV-G) on the cell surface and expressing Cre recombinase protein were generated according by the standard procedure of ultracentrifugation through a Ficoll gradient to obtain small particle fusosomes as described herein. To measure the phosphatidylserine levels of the fusosomes, annexin V staining was performed using a commercially available annexin V conjugated with Alexa Fluor 647 dye (Cat #A23204) according to the manufacturer's instructions. Annexin V is a cellular protein that can bind phosphatidylserine when it is exposed on the outer leaflet of the plasma membrane; thus, the readout of annexin V binding to a sample can provide an assessment of phosphatidylserine levels in the sample.

Briefly, the fusosome sample was measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Next 40 µg of fusosome total protein was pelleted by centrifugation (in sample triplicates) at 3000g for 5 minutes in a table-top centrifuge, followed by resuspension in 400uL of DMEM supplemented with 2% fetal bovine serum. One sample was treated with 40 µM antimycin A. The samples were then incubated for 1 hour at 37C. After the incubation samples were then pelleted by centrifugation again and resuspended in 100 µL annexin-binding buffer (ABB; 10 mM HEPES, 140 mM NaCl, 2.5 mM $CaCl_2$, pH 7.4). Next 5 µL of annexin V conjugated with Alexa Fluor 647 was added to each sample (except for the negative control with no annexin V staining). The samples were incubated for 15 minutes at room temperature followed by addition of 400 µL ABB.

The samples were then measured for annexin V staining by flow cytometry analysis using an Invitrogen Attune NxT acoustic focusing cytometer. Annexin V conjugated with Alexa Fluor 647 was excited with a 638 nm laser and emission captured at 670 ± 14 nm. Forward and side scatter gating was initially used to capture fusosome-sized events and discard small debris. Events positive for Alexa Fluor 647 (annexin V) staining were determined by gating at the minimum level for which the unstained, annexin V-negative control sample showed <0.5% of

events positive for Alexa Fluor 647 staining. The gated events positive for Alexa Fluor 647 staining were then assessed for the percentage of annexin V-positive events of the total parent population (fusosome-sized events in the forward/side scatter gate) and this value was used as the quantification of phosphatidylserine levels in the fusosome sample.

With this assay the fusosome derived from a HEK-293T cell expressing the VSV-G and Cre showed a % annexin V-positive fusosomes of 63.3 ± 2.3% without antimycin A treatment and percentage of annexin V-positive fusosomes of 67.6 ± 5.7% with antimycin A treatment.

Example 78: Measuring average mitochondrial membrane potential

Fusosomes from HEK-293T cells expressing the envelope glycoprotein G from vesicular stomatitis virus (VSV-G) on the cell surface and expressing Cre recombinase protein were generated according by the standard procedure of ultracentrifugation through a Ficoll gradient to obtain small particle fusosomes as described herein. To measure the average mitochondrial membrane potential levels of the fusosomes, a commercially available dye that is mitochondrial membrane potential sensitive, tetramethyl rhodamine, ethyl ester, perchlorate (TMRE; Abcam, Cat# T669) was used for assessing mitochondrial membrane potential. To normalize TMRE fluorescence intensity (FI) to the amount of mitochondria in the sample, MitoTracker Green FM dye (MTG; ThermoFisher, Cat #M7514) was used to co-stain samples in order to normalize TMRE FI to the MTG FI and thus to the amount of mitochondria in the sample. In addition, carbonyl cyanide-p-trifluoromethoxyphenylhydrazone (FCCP; Sigma Cat #C2920) was used to treat a parallel set of samples in order to fully depolarize the mitochondrial membrane potential and thus allow quantification of mitochondrial membrane potential in millivolts based on the decrease in TMRE FI.

Briefly, the fusosome sample was measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Next 40 µg of fusosome total protein was pelleted by centrifugation (in sample quadruplicates for untreated and FCCP-treated duplicates) at 3000g for 5 minutes in a table-top centrifuge, followed by resuspension in 100uL of DMEM supplemented with 2% fetal bovine serum and containing TMRE and MTG dyes at a final concentration of 30 nM and 200 nM, respectively. A parallel set of fusosome samples was left unstained as a negative control. The samples were incubated at for 45 minutes at 37ºC. After incubation, samples were pelleted by centrifugation and resuspended

in 400 µL of phenol red-free DMEM media containing 30 nm TMRE. One set of duplicates was treated with 20 µM FCCP for 5 minutes before assessment by flow cytometry.

The samples were then measured for annexin V staining by flow cytometry analysis using an Invitrogen Attune NxT acoustic focusing cytometer. MTG was excited with a 488 nm laser and emission captured at 530 ± 30 nm. TMRE was excited with a 561 nm laser and emission captured at 585 ± 16 nm. Forward and side scatter gating was initially used to capture fusosome-sized events and discard small debris. Events positive for MTG and TMRE staining were determined by gating at the minimum level for which the unstained control sample showed <0.5% of events positive for MTG or TMRE staining. The gated events positive for MTG and TMRE staining were then assessed for the mean FI of MTG and TMRE.

Membrane potential values (in millivolts, mV) are calculated based on the intensity of TMRE after normalizing TMRE FI values to MTG FI values. This TMRE/MTG ratio value allows for normalization TMRE intensity to the amount of mitochondria in the sample. The TMRE/MTG ratio value for both the untreated and FCCP-treated samples are calculated and used to determine the membrane potential in millivolts using a modified Nernst equation (see below) that can determine mitochondrial membrane potential based on TMRE fluorescence (as TMRE accumulates in mitochondria in a Nernstian fashion). Fusosome membrane potential is calculated with the following formula: (mV) = -61.5 * log(FI(untreated)/FI(FCCP-treated)). Using this equation, the calculated mitochondrial membrane potential of the VSV-G fusosome sample was -29.6 ± 1.5 millivolts.

Example 79: Measuring targeting potential in a subject (BiVs- Cre Gesicles)

This example assesses the ability of a fusosome to target a specific body site. Fusosomes were derived using methods as described herein and were loaded with cre-recombinase protein.

Two doses of fusosomes (1x and 3x) were delivered into Loxp Luciferase (Jackson Laboratory, 005125) mice were injected intravenously (I.V.) via tail vein. Mice were placed underneath a heat lamp (utilizing a 250W(infrared) heat lamp bulb) for ~5 minutes (or until mice begin to groom their whiskers excessively) to dilate the tail vein. Mice were placed on a restrainer and tail was wiped down with 70% ethanol to better visualize the vein.

Using a tuberculin syringe, 200 µL of fusosome 1x solution (8.5e8±1.4e8 particles/ µL, mean(SEM)) or 3x solution (2.55e9±1.4e8 particles/ µL, mean(SEM)) was injected IV. Upon completion of injection, the syringe was removed, and pressure was applied to the injection site.

After fusion, CRE protein translocated to the nucleus to carry out recombination, which resulted in the constitutive expression of luciferase. Three days post-treatment, the ventral region of subjects was prepared by depilating the area (Nair Hair Remover cream for 45 seconds, followed by cleaning the area with 70% ethanol). Subjects were then treated with D-luciferin (Perkin Elmer, 150 mg/kg) via intraperitoneal administration. This enabled the detection of luciferase expression via in vivo bioluminescent imaging. The animal was placed into an *in vivo* bioluminescent imaging chamber (Perkin Elmer) which houses a cone anesthetizer (isoflurane) to prevent animal motion. Photon collection was carried out between 3-15 minutes post-injection to observe the maximum bioluminescent signal due to D-luciferin pharmacokinetic clearance. Maximum radiance was recorded in photons/sec/cm2/radians. Total flux, which integrates the radiance over the area, was quantified using a region of interest (ROI) tool within the Living Image Software (Perkin Elmer) and reported in photons/sec.

Evidence of protein (Cre recombinase) delivery by fusosomes was detected by bioluminescent imaging in the recipient tissue of the animal, as shown in FIGS. 9A-9B. Signal was seen primarily in the spleen and liver, with the 3x group showing the highest signal.

Following whole body imaging, mice were cervically dislocated and liver, heart, lungs, kidney, small intestines, pancreas, and spleen were collected and imaged within 5 minutes of euthanasia. Evidence of protein (Cre recombinase) delivery to the liver and spleen by fusosomes was detected by bioluminescent imaging in the extracted recipient tissue of the animals. This can be seen in FIGS. 10A-10B. Signal was highest in spleen and the lowest in heart, with the 3x group showing the highest significant signal ($p=0.0004$ as compared to heart).

Example 80: Delivery of fusosomes via a pathway that is independent of lysosome acidification

Often, entry of complex biological cargo into target cells is accomplished by endocytosis. Endocytosis requires the cargo to enter an endosome, which matures into an acidified lysosome. Disadvantageously, cargo that enters a cell through endocytosis may become trapped in an endosome or lysosome and be unable to reach the cytoplasm. The cargo

may also be damaged by acidic conditions in the lysosome. Some viruses are capable of non-endocytic entry into target cells; however this process is incompletely understood. This example demonstrates that a viral fusogen can be isolated from the rest of the virus and confer non-endocytic entry on a fusosome that lacks other viral proteins.

Fusosomes from HEK-293T cells expressing the Nipah virus receptor-binding G protein and fusion F protein (NivG+F) on the cell surface and expressing Cre recombinase protein were generated according by the standard procedure of ultracentrifugation through a Ficoll gradient to obtain small particle fusosomes, as described herein. To demonstrate delivery of the fusosome to a recipient cell via a non-endocytic pathway, the NivG+F fusosomes were used to treat recipient HEK-293T cells engineered to express a "Loxp-GFP-stop-Loxp-RFP" cassette under CMV promoter. NivF protein is a pH-independent envelope glycoprotein that has been shown to not require environmental acidification for activation and subsequent fusion activity (Tamin, 2002).

The recipient cells were plated 30,000 cells/well into a black, clear-bottom 96-well plate. Four to six hours after plating the recipient cells, the NivG+F fusosomes expressing Cre recombinase protein were applied to the target or non-target recipient cells in DMEM media. The fusosome sample was first measured for total protein content by bicinchoninic acid assay (BCA, ThermoFisher, Cat #23225) according to manufacturer's instructions. Recipient cells were treated with 10 µg of fusosomes and incubated for 24 hrs at 37ºC and 5% CO2. To demonstrate that Cre delivery via NivG+F fusosomes was through a non-endocytic pathway, a parallel wells of recipient cells receiving NivG+F fusosome treatment were co-treated with an inhibitor of endosome/lysosome acidification, bafilomycin A1 (Baf; 100 nM; Sigma, Cat #B1793).

Cell plates were imaged using an automated microscope (www.biotek.com/products/imaging-microscopy-automated-cell-imagers/lionheart-fx-automated-live-cell-imager/). The total cell population in a given well was determined by staining the cells with Hoechst 33342 in DMEM media for 10 minutes. Hoechst 33342 stains cell nuclei by intercalating into DNA and was therefore used to identify individual cells. Hoechst staining was imaged using the 405 nm LED and DAPI filter cube. GFP was imaged using the 465 nm LED and GFP filter cube, while RFP was imaged using the 523 nm LED and RFP filter cube. Images of target and non-target cell wells were acquired by first establishing the LED intensity and integration times on a positive control well containing recipient cells treated with adenovirus coding for Cre recombinase instead of fusosomes.

Acquisition settings were set so that Hoescht, RFP, and GFP intensities were at the maximum pixel intensity values but not saturated. The wells of interest were then imaged using the established settings. Focus was set on each well by autofocusing on the Hoescht channel and then using the established focal plane for the GFP and RFP channels. Analysis of GFP and RFP-positive cells was performed with Gen5 software provided with automated fluorescent microscope (https://www.biotek.com/products/software-robotics-software/gen5-microplate-reader-and-imager-software/).

The images were pre-processed using a rolling ball background subtraction algorithm with a 60 μm width. Cells with GFP intensity significantly above background intensities were thresholded and areas too small or large to be GFP-positive cells were excluded. The same analysis steps were applied to the RFP channel. The number of RFP-positive cells (recipient cells receiving Cre) was then divided by the sum of the GFP-positive cells (recipient cells that did not show delivery) and RFP-positive cells to quantify the percentage RFP conversion, which indicates the amount of fusosome fusion with the recipient cells.

With this assay, the fusosome derived from a HEK-293T cell expressing NivG+F on its surface and containing Cre recombinase protein showed significant delivery via a lysosome-independent pathway, which is consistent with entry via a non-endocytic pathway, as evidenced by a significant delivery of Cre cargo by NivG+F fusosomes even when recipient cells were co-treated with Baf to inhibit endocytosis-mediated uptake (FIG. 11). In this case, the inhibition of cargo delivery by Baf co-treatment was 23.4%.

Example 81: Measuring ability to polymerize actin for mobility

Fusosomes were generated by the standard procedure of harvesting and preparing fusosomes produced from HEK-293T cells expressing the envelope glycoprotein G from vesicular stomatitis virus (VSV-G) on the cell surface, as described herein. Control particles (non-fusogenic fusosomes) were produced from HEK-293T cells reverse transiently transfected with pcDNA3.1 empty vector. Fusosomes and parental cells were then assayed for their ability to polymerize actin (over time) using a rhodamine phalloidin-flow cytometry assay and Tubulin ELISA. Briefly, approximately $1x10^6$ fusosomes corresponding to 60 μL of a standard VSV-G fusosome preparation and $1x10^5$ parent cells used to generate the fusosomes were plated in 1 mL of complete media in a 96 well low-attachment multi-well plate in complete and incubated at

37°C and 5% CO_2. Samples were taken periodically, at 3 hr, 5 hr and 24 hr post plating. Samples were centrifuged at 21,000xg for 10mins, re-suspended in 200uL 4% (v/v) PFA in phosphate buffered saline for 10mins, washed with 1mL of phosphate buffered saline, centrifuged at 21,000xg for 10mins, washed again and stored at 4°C until further use.

For rhoamine-phalloidin staining, samples were centrifuged at 21,000xg for 10mins, and incubated in 100uL of 0.1% (v/v) Triton X-100 in phosphate buffered saline for 20 mins. Following the 20-min incubation, an additional 100uL of 0.1% (v/v) Triton X-100 in phosphate buffered saline containing 165 μM rhodamine-phalloidin was added to the sample and pipette mixed, negative control received and additional 100uL of 100uL of 0.1% (v/v) Triton X-100 in phosphate buffered saline only. Samples were incubated for 45 mins before being washed with 1mL of phosphate buffered saline, centrifuged at 21,000xg for 10mins, washed again and re-suspended in 300uL of phosphate buffered saline and analyzed by flow cytometry (Attune, ThermoFisher) using a 561 nm laser for excitation, and 585 +/- 16nm filter emission, as shown in the table below:

Flow cytometer settings

Dye	Attune Laser/Filter	Laser Wavelength	Emission Filter (nm)
AF47	YL1	585	585/16

Attune NxT software was used for acquisition and FlowJo used analysis. For data acquisition the FSC and SSC channels were set on linear axis to determine a population representative of the cells or fusosomes. This population was then gated and events only inside this gate were used to display events in the 585 +/- 16nm emission channel on a logarithmic scale. A minimum of 10,000 events within the cells or fusosomes gate was collected for in each condition. For data analysis, the FSC and SSC channels were set on linear axis to determine a population representative of the cells or fusosomes. This population was then gated and events only inside this gate were used to display events in the 585 +/- 16nm emission channel on a logarithmic scale. The negative control 585 +/- 16nm emission was used to determine where to place the gate on the histogram such that it was less the gate include less than 1% positive. Using analysis criteria listed above parent cells demonstrated 19.9%, 24.8% and 82.5% rhodamine-phalloidin positive events, at the 3hr, 5hr and 24hr time-points, respectively. The fusosomes

were 44.6%, 41.9% and 34.9% rhodamine-phalloidin at the 3hr, 5hr and 24hr time-points, respectively (FIG. 2). This example demonstrates that fusosomes do not increase in amount of actin over time, whereas the parent cells do.

Example 82: Measuring GAPDH in fusosomes

This example describes quantification of the level of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) in the fusosomes, and the relative level of GAPDH in the fusosomes compared to the parental cells. Fusosomes were prepared as described in Examples 67 and 86.

GAPDH was measured in the parental cells and the fusosomes using a standard commercially available ELISA for GAPDH (ab176642, Abcam) per the manufacturer's directions. Total protein levels were similarly measured via bicinchoninic acid assay. Measured GAPDH and protein levels are shown in the table below:

	[Protein] (mg/mL)	[GAPDH] (ng/mL)	GAPDH:Protein (µg/g)
Fusosomes	0.82	37.2	45.3
Cells	0.45	50.4	112.0

GAPDH: Total protein ratios are also shown in FIG. 12.

Example 83: Ratio of lipids to proteins in fusosomes

This Example describes quantification of the ratio of lipid mass to protein mass in fusosomes. It is contemplated that fusosomes can have a ratio of lipid mass to protein mass that is similar to that of nucleated cells. Fusosomes and parental cells were prepared as described herein in Examples 67 and 86.

The lipid content was calculated using choline-containing phospholipids as a subset of total lipids using a commercially available phospholipid assay kit (MAK122 Sigma St. Louis, MO) according to manufacturer's instructions. Total protein content of the fusosomes was measured via bicinchoninic acid assay as described herein. Measured phospholipid levels, protein levels, and the ratio of phospholipids to protein are shown in FIG. 13 and the table below:

	Phospholipids (µM)	Protein (g/L)	Phospholipids:Protein (µmol/g)
Fusosomes	115.6	0.82	141.0

Cells	47.9	0.45	106.4

Example 84: Ratio of proteins to DNA in fusosomes

This Example describes quantification of the ratio of protein mass to DNA mass in fusosomes. It is contemplated that fusosomes can have a ratio of protein mass to DNA mass that is much greater than that of cells. Fusosomes were prepared as described in Examples 67 and 86.

Total protein content of the fusosomes and cells was measured via bicinchoninic acid as described herein. The DNA mass of fusosomes and cells were measured by absorption at 280 nm after extraction of total DNA using a commercially available isolation kit (#69504 Qiagen Hilden, Germany) according to the manufacturer's instructions. The ratio of proteins to total nucleic acids was determined by dividing the total protein content by the total DNA content to yield a ratio within a given range for a typical fusosome preparation. Measured protein levels, DNA levels, and the ratio of protein to DNA are shown in FIG. 14 and the table below:

	[Protein] (mg/mL)	[DNA] (ng/μL)	Protein:DNA (g/g)
Fusosomes	0.82	29.5	27.8
Cells	0.45	15.9	28.3

Example 85: Ratio of lipids to DNA in fusosomes

This Example describes quantification of the ratio of lipids to DNA in fusosomes compared to parental cells. In an embodiment, fusosomes will have a greater ratio of lipids to DNA compared to parental cells. Fusosomes were prepared as described previously in Examples 67 and 86.

This ratio is defined as the lipid content outlined in Example 40, and nucleic acid content is determined as described in Example 41. Measured lipid levels, DNA levels, and the ratio of lipid to DNA are shown in FIG. 15 and the table below:

	[Lipids] (μM)	[DNA] (ng/μL)	Lipids:DNA (μmol/mg)
Fusosomes	115.6	29.5	3.92
Cells	47.9	15.9	3.01

Example 86: Measuring lipid composition in fusosomes

This Example describes quantification of the lipid composition of fusosomes. It is contemplated that the lipid composition of fusosomes can be similar to the cells from which they

are derived. Lipid composition affects important biophysical parameters of fusosomes and cells, such as size, electrostatic interactions, and colloidal behavior.

The lipid measurements were based on mass spectrometry. Fusosomes were prepared as described herein by transient transfection of VSV-G and GFP in 10 cm dishes, followed by filtration and ultracentrifugation of the conditioned media 48 h after transfection to obtain fusosomes. Transfected cells were harvested in parallel to the conditioned media and submitted for analysis. Exosomes were also harvested from cells that were not transfected with VSV-G or GFP.

Mass spectrometry-based lipid analysis was performed by Lipotype GmbH (Dresden, Germany) as described (Sampaio et al. 2011). Lipids were extracted using a two-step chloroform/methanol procedure (Ejsing et al. 2009). Samples were spiked with internal lipid standard mixture containing: cardiolipin 16:1/15:0/15:0/15:0 (CL), ceramide 18:1;2/17:0 (Cer), diacylglycerol 17:0/17:0 (DAG), hexosylceramide 18:1;2/12:0 (HexCer), lyso-phosphatidate 17:0 (LPA), lyso-phosphatidylcholine 12:0 (LPC), lyso-phosphatidylethanolamine 17:1 (LPE), lyso-phosphatidylglycerol 17:1 (LPG), lyso-phosphatidylinositol 17:1 (LPI), lyso-phosphatidylserine 17:1 (LPS), phosphatidate 17:0/17:0 (PA), phosphatidylcholine 17:0/17:0 (PC), phosphatidylethanolamine 17:0/17:0 (PE), phosphatidylglycerol 17:0/17:0 (PG), phosphatidylinositol 16:0/16:0 (PI), phosphatidylserine 17:0/17:0 (PS), cholesterol ester 20:0 (CE), sphingomyelin 18:1;2/12:0;0 (SM), triacylglycerol 17:0/17:0/17:0 (TAG) and cholesterol D6 (Chol).

After extraction, the organic phase was transferred to an infusion plate and dried in a speed vacuum concentrator. 1st step dry extract was re-suspended in 7.5 mM ammonium acetate in chloroform/methanol/propanol (1:2:4, V:V:V) and 2nd step dry extract in 33% ethanol solution of methylamine in chloroform/methanol (0.003:5:1; V:V:V). All liquid handling steps were performed using Hamilton Robotics STARlet robotic platform with the Anti Droplet Control feature for organic solvents pipetting.

Samples were analyzed by direct infusion on a QExactive mass spectrometer (Thermo Scientific) equipped with a TriVersa NanoMate ion source (Advion Biosciences). Samples were analyzed in both positive and negative ion modes with a resolution of $R_{m/z=200}=280000$ for MS and $R_{m/z=200}=17500$ for MSMS experiments, in a single acquisition. MSMS was triggered by an inclusion list encompassing corresponding MS mass ranges scanned in 1 Da increments

(Surma et al. 2015). Both MS and MSMS data were combined to monitor CE, DAG and TAG ions as ammonium adducts; PC, PC O-, as acetate adducts; and CL, PA, PE, PE O-, PG, PI and PS as deprotonated anions. MS only was used to monitor LPA, LPE, LPE O-, LPI and LPS as deprotonated anions; Cer, HexCer, SM, LPC and LPC O- as acetate adducts and cholesterol as ammonium adduct of an acetylated derivative (Liebisch et al. 2006).

Data were analyzed with in-house developed lipid identification software based on LipidXplorer (Herzog et al. 2011; Herzog et al. 2012). Data post-processing and normalization were performed using an in-house developed data management system. Only lipid identifications with a signal-to-noise ratio >5, and a signal intensity 5-fold higher than in corresponding blank samples were considered for further data analysis.

Fusosome lipid composition was compared to lipid compositions of parental cells, with undetected lipid species assigned a value of zero. The lipid species identified in fusosomes and parental cells are shown in the table below:

	Total Lipid Species Identified	Shared Lipid Species (identified in both parental cells and fusomes)	Shared Lipid Species with 25% of parental expression in fusosomes	Fraction of Shared Lipid Species to Total Lipids
Fusosomes	679	569	548	0.700
Parental Cells	783			

It is contemplated that fusosomes and parental cells can have a similar lipid composition if ≥70% of the lipid species identified in any replicate sample of the parental cells are present in any replicate sample of the fusosomes, and of those identified lipids, the average level in the fusosome can be >25% of the corresponding average lipid species level in the parental cell.

Example 87: Measuring proteomic composition in fusosomes

This Example describes quantification of the protein composition of fusosomes. It is contemplated that the protein composition of fusosomes can be similar to the parental cells from which they are derived.

Fusosomes and parental cells were prepared as described herein by the method of Examples 67 and 86.

Each sample was resuspended in lysis buffer (6 M urea, 2 M thiourea, 4% CHAPS, 50 mM Tris pH 8.0), sonicated on an ice bath and ran through a small gauge syringe. Proteins were

reduced with 10 mM DTT for 15 minutes at 65°C and alkylated with 15 mM iodoacetamide (IAA) for 30 minutes in the dark at room temperature. Excess IAA was quenched with an additional 10 mM DTT. Proteins were then precipitated with the addition of 8 volumes of ice cold acetone + 1 volume of ice cold methanol and placed at -80°C overnight. The precipitated proteins were pelleted by centrifugation. Remaining lysis buffer was washed with 200 µl of ice cold methanol 3 times. Proteins were resuspended in 0.75 M urea + 50 mM Tris pH 8.0 + 1 µg Trypsin/LysC and pre-digested for 4 hours at 37°C with agitation. An additional 1 µg of trypsin/LysC was added to the proteins and the digestion was continued overnight. Peptides were purified by reversed phase SPE and analyzed by LC-MS.

A replicate sample for each condition was lysed and combined in one tube. This pool was then either subjected to the same preparation protocol as the samples and analyzed by LC-MS in information dependent acquisition or separated on a gel as described below.

A total of 100 µg of pooled proteins was placed in 2x Laemmli loading buffer and separated on a 12.5% SDS PAGE. Proteins were briefly stained with Coomassie blue and the protein lanes were separated into 12 fractions. Each fraction was then dehydrated with 50% acetonitrile and rehydrated with 10 mM DTT for the reduction. Gel pieces were placed at 65°C for 15 minutes and alkylated for 30 minutes at room temperature with 15 mM IAA in the dark. Gels were further dehydrated with 50% acetonitrile and rehydrated in 50 mM Tris pH 8 with 1 µg of trypsin/LysC overnight at 37°C. Peptides were extracted from the gel by dehydration and sonication. Peptides were purified by reversed phase SPE and analyzed by LC-MS/MS (1xIDA per fraction).

Acquisition was performed with an ABSciex TripleTOF 5600 (ABSciex, Foster City, CA, USA) equipped with an electrospray interface with a 25 µm iD capillary and coupled to an Eksigent µUHPLC (Eksigent, Redwood City, CA, USA). Analyst TF 1.7 software was used to control the instrument and for data processing and acquisition. Acquisition was performed in Information Dependent Acquisition (IDA) mode for the 12 fractions from the gel or the unfractionated pool. The samples were analyzed in SWATH acquisition mode. For the IDA mode, the source voltage was set to 5.2 kV and maintained at 225°C, curtain gas was set at 27 psi, gas one at 12 psi and gas two at 10 psi. For the SWATH mode, the source voltage was set to 5.5 kV and maintained at 225°C, curtain gas was set at 25 psi, gas one at 16 psi and gas two at 15 psi. Separation was performed on a reversed phase HALO C18-ES column 0.3 mm i.d., 2.7 µm

particles, 150 mm long (Advance Materials Technology, Wilmington, DE) which was maintained at 60°C. Samples were injected by loop overfilling into a 5 µL loop. For the 60 minutes LC gradient, the mobile phase consisted of the following solvent A (0.2% v/v formic acid and 3% DMSO v/v in water) and solvent B (0.2% v/v formic acid and 3% DMSO in EtOH) at a flow rate of 3 µL/min.

To generate the ion library for the analysis of the samples, the ProteinPilot software was run on the wiff files that were generated by the IDA runs. This database was used in the Peakview software (ABSciex) to quantify the proteins in each of the samples, using 3 transition/peptide and 15 peptide/protein. To maximize the number of quantified proteins, the samples were quantified on a publicly available human SWATH database (Atlas) with the same parameters. A peptide was considered as adequately measured if the score computed by Peakview was superior to 1.5 and had an FDR < 1%. The quantification from each of the database was combined into one final quantification using the protein name from both databases. A correction factor was computed for every sample by taking into account the total signal of every protein in that sample when compared to the average of the total signal for every sample.

The fusosome proteomic composition was compared to the parental cell proteomic composition. A similar proteomic composition between fusosomes and parental cells was observed when >33% of the identified proteins were present in the fusosome, and of those identified proteins, the level was >25% of the corresponding protein level in the parental cell, as shown in the table below.

	Total Proteins Identified	Shared proteins (identified in both parental cells and fusomes)	Shared proteins with 25% of parental expression in fusosomes	Fraction of shared proteins to total proteins
Fusosomes	1926	1487	957	0.333
Cells	2870			

Example 88: Quantifying an endogenous or synthetic protein level per fusosome

This example describes quantification of an endogenous or synthetic protein cargo in fusosomes. Fusosomes can, in some instances, comprise an endogenous or synthetic protein cargo. The fusosome or parental cell described in this Example was engineered to alter the expression of an endogenous protein or express a synthetic cargo that mediates a therapeutic or novel cellular function.

Fusosomes and parental cells expressing GFP were prepared as described herein by the method of Examples 67 and 86. Quantification of GFP in fusosomes was accomplished using a commercially available ELISA kit (ab171581 Abcam Cambridge, United Kingdom) according to the manufacturer's instructions. Fusosome quantification was performed by Nanoparticle Tracking Analysis using a NanoSight NS300 (Malvern Instruments, Malvern, Worcestershire, United Kingdom). Results are shown in the table below.

	Concentration (#/mL)
GFP Protein	4.41×10^{13}
Fusosomes	2.66×10^{11}
GFP:Fusosome	165.8

It is contemplated that the fusosomes can have at least 1, 2, 3, 4, 5, 10, 20, 50, 100, or more protein agent molecules per fusosome. In an embodiment, the fusosomes will have 166 protein agent molecules per fusosome.

Example 89: Measuring markers of exosomal proteins in fusosomes

This assay describes quantification of the proportion of proteins that are known to be specific markers of exosomes.

Fusosomes were prepared as described herein by the method of Examples 67 and 86. Exosomes were prepared as described herein for fusosomes by the method of Examples 67 and 86 with the exception that the parental cells were not transfected with VSV-G or GFP. Protein quantification by mass spectrometry for fusosomes and exosomes was performed as described herein in Example 35.

The resulting protein quantification data was analyzed to determine protein levels and proportions of the known exosomal marker CD63. Average log intensities per group were calculated by adding 1 to intensity values from mass spectrometry, transforming by log10, and computing the mean across replicates. The results are shown in FIG. 16.

Example 90: Measuring calnexin in fusosomes

This assay describes quantification of the level of calnexin (CNX) in the fusosomes, and the relative level of CNX in the fusosomes compared to the parental cells.

Fusosomes and parental cells were prepared as described herein in Examples 67 and 86. Calnexin and total protein was measured using mass spectrometry conducted according to the method of Example 35. The calnexin signal intensity determined for parental cells and fusosomes is shown in FIG. 17.

In embodiments, using this assay, the average fractional content (calculated as described herein in Example 35) of CNX in the fusosomes will be $< 2.43 \times 10^{-4}$.

In an embodiment, the decrease in calnexin per total protein in ng/µg from the parent cell to the preparation will be more than 88%.

Example 91: Ratio of lipids to DNA in fusosomes

This Example describes quantification of the ratio of lipids to DNA in fusosomes compared to parental cells. In an embodiment, fusosomes will have a greater ratio of lipids to DNA compared to parental cells. Fusosomes were prepared as described previously in Examples 67 and 86.

This ratio is defined as the lipid content outlined in Example 40, and nucleic acid content is determined as described in Example 41. As shown in FIG. 18 and in the table below, fusosomes were found to exhibit a greater lipid:DNA ratio than parental cells.

	[Lipids] (µM)	[DNA] (ng/µL)	Lipids:DNA (µmol/mg)
Fusosomes	115.6	29.5	3.92
Cells	47.9	15.9	3.01

Example 92: Analyzing surface markers on fusosomes

This assay describes identification of surface markers on the fusosomes.

Fusosomes were prepared as described herein in Examples 67 and 86. Phosphatidylserine was measured by mass spectrometry as described herein in Examples 67 and 86. The quantity of phosphatidylserine relative to total lipids in fusosomes was determined to be 121% greater than the quantity of phosphatidylserine relative to total lipid in parental cells, as shown in the table below.

	Phosphatidylserine (molar %)	Phosphatidylserine Percent change
Fusosomes	14.6	121%
Parental Cells	6.6	

Example 93: Analysis of viral capsid proteins in fusosomes

In this example, the makeup of the sample preparation was analyzed and the proportion of proteins that are derived from viral capsid sources was assessed.

Fusosomes were prepared as described herein by the method of Examples 67 and 86. Protein quantification by mass spectrometry for fusosomes was performed as described herein in Example 35. The fractional content of the viral capsid proteins was calculated as described herein in Example 35, averaged over fusosome samples, and expressed as a percent.

Using this approach, the sample was found to contain 0.05% viral capsid protein, as shown in the table below. The only viral capsid protein detected was Complex of Rabbit Endogenous Lentivirus (RELIK) Capsid with Cyclophilin A (PDB 2XGY|B).

	Raw MS Intensity	Viral:Total Protein (%)
Viral Capsid Proteins	5.10×10^5	0.05
Total Proteins	9.46×10^8	

Example 94: Quantification of fusogen protein ratios in fusosomes

This example describes quantification of the ratio of fusogen protein to total protein or other proteins of interest in fusosomes. Other proteins of interest may include, but are not limited to, EGFP, CD63, ARRDC1, GAPDH, Calnexin (CNX), and TSG101. Fusosomes were prepared as described herein by the method of Examples 67 and 86. Protein quantification by mass spectrometry for fusosomes was performed as described herein in Example 35. The quantification of all proteins was calculated as described herein in Example 35, averaged over fusosome samples, and expressed as a fraction.

As shown in the table below, the fusogen was found to have a ratio to EGFP of 156.9, a ratio to CD63 of 2912.0, a ratio to ARRDC1 of 664.9, a ratio to GAPDH of 69.0, a ratio to CNX of 558.4, and a ratio to TSG101 of 3064.1.

Proteins	Raw MS Intensity	Fusogen:Protein(s) Ratio
VSV-G	1.29×10^8	N/A
Total Proteins	9.46×10^8	0.136
EGFP	8.22×10^5	156.9
CD63	4.43×10^4	2912.0
ARRDC1	1.94×10^5	664.9
GAPDH	1.87×10^6	69.0
CNX	2.31×10^5	558.4
TSG101	4.21×10^4	3064.1

Example 95: Quantification of endogenous and synthetic protein ratios in fusosomes

This example describes the quantification of an endogenous or synthetic protein cargo relative to total protein or other proteins of interest in fusosomes. Other proteins of interest may include, but are not limited to, VSV-G, CD63, ARRDC1, GAPDH, Calnexin (CNX), or TSG101. Fusosomes were prepared as described herein by the method of Examples 67 and 86. Protein quantification by mass spectrometry for fusosomes was performed as described herein in Example 35. The quantification of all proteins was calculated as described herein in Example 35, averaged over fusosome samples, and expressed as a fraction.

As shown in the table below, the synthetic protein cargo was found to have a ratio to VSV-G of 6.37×10^{-3}, a ratio to CD63 of 18.6, a ratio to ARRDC1 of 4.24, a ratio to GAPDH of 0.44, a ratio to CNX of 3.56, and a ratio to TSG101 of 19.52.

Proteins	Raw MS Intensity	Protein Cargo:Protein(s) Ratio
EGFP	8.22×10^{5}	N/A
Total Proteins	9.46×10^{8}	8.69×10^{-4}
VSV-G	1.29×10^{8}	6.37×10^{-3}
CD63	4.43×10^{4}	18.6
ARRDC1	1.94×10^{5}	4.24
GAPDH	1.87×10^{6}	0.44
CNX	2.31×10^{5}	3.56
TSG101	4.21×10^{4}	19.52

Example 96: Enriched lipid composition in fusosomes

This Example describes quantification of the lipid composition of fusosomes, parental cells, and exosomes. It is contemplated that the lipid composition of fusosomes can be enriched and/or depleted for specific lipids relative to the cells from which they are derived. Lipid composition affects important biophysical parameters of fusosomes and cells, such as size, electrostatic interactions, and colloidal behavior.

The lipid composition was measured as described in Examples 67 and 86. Fusosomes were prepared as described herein by transient transfection of VSV-G and GFP in 10 cm dishes, followed by filtration and ultracentrifugation of the conditioned media 48 hours after transfection to obtain fusosomes. Transfected cells were harvested in parallel to the conditioned media and submitted for analysis. Exosomes were prepared as described herein for fusosomes with the exception that the parental cells were not transfected with VSV-G or GFP.

The lipid composition for fusosomes, exosomes, and parental cells is shown in FIGS. 19A-19B. Compared to parental cells, fusosomes were enriched for cholesteryl ester, free cholesterol, ether-linked lyso-phosphatidylethanolamine, lyso-phosphatidylserine, phosphatidate, ether-linked phosphatidylethanolamine, phosphatidylserine, and sphingomyelin. Compared to parental cells, fusosomes are depleted for ceramide, cardiolipin, lyso-phosphatidylcholine, lyso-phosphatidylethanolamine, lyso-phosphatidylglycerol, lyso-phosphatidylinositol, ether-linked phosphatidylcholine, phosphatidylethanolamine, phosphatidylglycerol, phosphatidylinositol, and triacylglycerol. Compared to exosomes, fusosomes were enriched for cholesteryl ester, ceramide, diacylglycerol, lyso-phosphatidate, and phosphatidylethanolamine, triacylglycerol. Compared to exosomes, fusosomes are depleted for free cholesterol, hexosyl ceramide, lyso-phosphatidylcholine, ether-linked lyso-phosphatidylcholine, lyso-phosphatidylethanolamine, ether-linked lyso-phosphatidylethanolamine, and lyso-phosphatidylserine,

Example 97: Measuring compartment-specific proteomic content of fusosomes

This Example describes quantification of the proportion of proteins that are known to be derived from specific cellular compartments in fusosomes, fusosome parental cells, and exosomes.

Fusosomes and parental cells were prepared as described herein by the method of Examples 67 and 86. Exosomes were prepared as described herein for fusosomes by the method of Examples 67 and 86 with the exception that the parental cells were not transfected with VSV-G or GFP. Protein quantification by mass spectrometry for fusosomes and exosomes was performed as described herein in Example 35. The resulting protein quantification data was analyzed to determine protein levels and proportions of known exosomal, , endoplasmic reticulum, ribosome, nuclear, and mitochondrial proteins as annotated by Gene Ontology Cellular Compartment annotation terms (exosome: GO:0070062, endoplasmic reticulum: GO:0005783, ribosome: GO:0005840, GO:0022625, GO:0022626, GO:0022627, GO:0044391, GO:0042788, GO:0000313) with evidence code IDA (inferred by direct assay). The fraction of compartment-specific proteins relative to total protein in each sample was determined for fusosome samples, exosome samples, and parental cells.

As shown in FIG. 20, fusosomes were found to be depleted with endoplasmic reticulum protein compared to parental cells and exosomes. Fusosomes were also found to be depleted for

exosomal protein compared to exosomes. Fusosomes were depleted for mitochondrial protein compared to parental cells. Fusosomes were enriched for nuclear protein compared to parental cells. Fusosomes were enriched for ribosomal proteins compared to parental cells and exosomes.

Example 98: Measuring TSG101 and ARRDC1 content in fusosomes

This Example describes quantification of the proportion of proteins that are known to be important in fusosome release from cells.

Fusosomes and parental cells were prepared as described herein by the method of Examples 67 and 86. Exosomes were prepared as described herein for fusosomes by the method of Examples 67 and 86 with the exception that the parental cells were not transfected with VSV-G or GFP. Protein quantification by mass spectrometry for fusosomes and exosomes was performed as described herein in Example 35. The resulting protein quantification data was analyzed to determine protein levels and proportions of the protein TSG101 and ARRDC1. Average log intensities per group were calculated by adding 1 to intensity values from mass spectrometry, transforming by log10, and computing the mean across replicates. The percentage of total protein content of TSG101 or ARRDC1 in fusosomes relative to exosomes or parental cells was determined as the average log intensity of TSG101 or ARRDC1 for each sample, divided by the sum of intensities of all proteins in the same sample, averaged over replicates and expressed as a percent.

As shown in FIG. 21, ARRDC1 was found to be present at greater levels as a percentage of total protein content in fusosomes than in parental cells or exosomes. The level of ARRDC1 as a percentage of total protein content was at least 0.02% in fusosomes. TSG101 was found to be present at greater levels as a percentage of total protein content in fusosomes than in parental cells or exosomes. The level of TSG101 as a percentage of total protein content was at least 0.004% in fusosomes.

Example 99: Measuring serum inactivation of fusosomes after multiple administrations

This Example describes quantification of serum inactivation of fusosomes using an *in vitro* delivery assay following multiple administrations of the fusosome. It is contemplated that a modified fusosome, e.g., modified by a method described herein, can have a reduced (e.g., reduced compared to administration of an unmodified fusosome) serum inactivation following

multiple (e.g., more than one, e.g., 2 or more), administrations of the modified fusosome. In some instances, a fusosome described herein will not be inactivated by serum following multiple administrations.

A measure of immunogenicity for fusosomes is serum inactivation. In an embodiment, repeated injections of a fusosome can lead to the development of anti-fusosome antibodies, e.g., antibodies that recognize fusosomes. In an embodiment, antibodies that recognize fusosomes can bind in a manner that can limit fusosome activity or longevity and mediate complement degradation.

In this Example, serum inactivation is examined after one or more administrations of fusosomes. Fusosomes are produced by any one of the previous Examples. In this example, fusosomes are generated from: HEK293 cells modified with a lentiviral-mediated expression of HLA-G (hereafter HEK293-HLA-G), and HEK293 modified with a lentiviral-mediated expression of an empty vector (hereafter HEK293). In some embodiments, fusosomes are derived from cells that are expressing other immunoregulatory proteins.

Serum is drawn from the different cohorts: mice injected systemically and/or locally with 1, 2, 3, 5, 10 injections of vehicle (Fusosome naïve group), HEK293-HLA-G fusosomes, or HEK293 fusosomes. Sera are collected from mice by collecting fresh whole blood and allowing it to clot completely for several hours. Clots are pelleted by centrifugation and the serum supernatants are removed. A negative control is heat inactivated mouse serum. Negative control samples are heated at 56 degrees Celsius for 1 hour. Serum may be frozen in aliquots.

The fusosomes are tested for the dose at which 50% of cells in a recipient population receive the payload in the fusosomes. The fusosomes may be produced via any of the other examples described herein and may contain any of the payloads described herein. Many methods for assaying fusosome delivery of a payload to recipient cells are also described herein. In this particular example, the payload is Cre protein and the recipient cells are RPMI8226 cell which stably-expresses “LoxP-GFP-stop-LoxP-RFP” cassette under a CMV promoter, which upon recombination by Cre switches from GFP to RFP expression, indicating fusion and Cre, as a marker, of delivery. The identified dose at which 50% of the recipient cells are RFP positive is used for further experiments. In other embodiments, the identified dose at which 50% of the recipient cells receive the payload is used for further experiments.

To assess serum inactivation of fusosomes, fusosomes are diluted 1:5 into normal or heat-inactivated serum (or medium containing 10% heat-inactivated FBS as the no-serum control) and the mixture is incubated at 37 C for 1 h. Following the incubation, medium is added to the reaction for an additional 1:5 dilution and then serially diluted twice at a 1:10 ratio. Following this step, the fusosomes should be present at the previously identified dose at which 50% of the recipient cells have received the payload (e.g. are RFP positive). It is contemplated that the identified dose at which 50% of recipient cells receive the payload may be similar across fusosomes.

Fusosomes that have been exposed to serum are then incubated with recipient cells. The percent of cells which receive the payload, and thus are RFP positive, is calculated. The percent of cells which receive the payload may not be different between fusosome samples that have been incubated with serum and heat-inactivated serum from mice treated with HEK293-HLA-G fusosomes, indicating that there is not serum inactivation of fusosomes or an adaptive immune response. The percent of cells that receive the payload may not be different between fusosome samples that have been incubated from mice treated 1, 2, 3, 5 or 10 times with HEK293-HLA-G fusosomes, which would indicate that there was not serum inactivation of fusosomes or an adaptive immune response. In some instances, the percent of cells which receive the payload is not different between fusosome samples that have been incubated with serum from mice treated with vehicle and from mice treated with HEK293-HLA-G fusosomes, indicating that there is not serum inactivation of fusosomes or an adaptive immune response. In some instances, the percent of cells which receive the payload is less for fusosomes derived with HEK293 than for HEK293-HLA-G fusosomes, indicating that there is not serum inactivation of HEK293-HLA-G fusosomes or an adaptive immune response.

Example 100: Measuring complement targeting of fusosomes

This Example describes quantification of complement activity against fusosomes using an *in vitro* assay. It is contemplated that a modified fusosome described herein can induce reduced complement activity compared to a corresponding unmodified fusosome.

In this Example, serum from a mouse is assessed for complement activity against a fusosome. The example measures the level of complement C3a, which is a central node in all

complement pathways. Notably, the methods described herein may be equally applicable to humans, rats, monkeys with optimization to the protocol.

In this Example, fusosomes are produced by any one of the previous Examples. Fusosomes are generated from HEK293 cells modified with a lentiviral-mediated expression of a complement regulatory protein DAF (HEK293-DAF fusosomes) or HEK 293 cells not expressing a complementary regulatory protein (HEK293 fusosomes). Other complement regulatory proteins may also be used, such as proteins that bind decay-accelerating factor (DAF, CD55), e.g. factor H (FH)-like protein-1 (FHL-1), e.g. C4b-binding protein (C4BP), e.g. complement receptor 1 (CD35), e.g. Membrane cofactor protein (MCP, CD46), eg. Profectin (CD59), e.g. proteins that inhibit the classical and alternative complement pathway CD/C5 convertase enzymes, e.g. proteins that regulate MAC assembly

Serum is recovered from naïve mice, mice that are administered HEK293-DAF fusosomes, or mice that are administered HEK293 fusosomes. Sera are collected from mice by collecting fresh whole blood and allowing it to clot completely for several hours. Clots are pelleted by centrifugation and the serum supernatants are removed. A negative control is heat inactivated mouse serum. Negative control samples are heated at 56 degrees Celsius for 1 hour. Serum may be frozen in aliquots.

The different fusosomes are tested for the dose at which 50% of cells in a recipient population receive the payload in the fusosomes. The fusosomes may be produced via any of the other examples described herein and may contain any of the payloads described herein. Many methods for assaying fusosome delivery of a payload to recipient cells are also described herein. In this particular example, the payload is Cre protein and the recipient cells are RPMI8226 cell which stably-expresses "LoxP-GFP-stop-LoxP-RFP" cassette under a CMV promoter, which upon recombination by Cre switches from GFP to RFP expression, indicating fusion and Cre, as a marker, of delivery. The identified dose at which 50% of the recipient cells are RFP positive is used for further experiments. In other embodiments, the identified dose at which 50% of the recipient cells receive the payload is used for further experiments. In preferred embodiments, the identified dose at which 50% of recipient cells receive the payload is similar across fusosomes.

Two-fold dilutions of the fusosomes starting at the dose of fusosomes at which 50% of the recipient cells receive the payload in phosphate-buffered saline (PBS, pH 7.4) are mixed with a 1:10 dilution of the sera from mice treated with the same fusosomes or naïve mice (assay

volume, 20 μl) and incubated for 1 h at 37°C. The samples are further diluted 1:500 and used in an enzyme-linked immunosorbent assay (ELISA) specific for C3a. The ELISA is mouse complement C3a ELISA Kit product LS-F4210 sold by LifeSpan BioSciences Inc, which measures the concentration of C3a in a sample. The dose of fusosomes at which 200 pg/ml of C3a is present is compared across sera isolated from mice.

In some instances, the dose of fusosomes at which 200 pg/ml of C3a is present is greater for HEK293-DAF fusosomes incubated with HEK-293 DAF mouse sera than for HEK293 fusosomes incubated with HEK293 mouse sera, indicating that complement activity targeting fusosomes is greater in mice treated with HEK293 fusosomes than HEK293-DAF fusosomes. In some instances, the dose of fusosomes at which 200 pg/ml of C3a is present is greater for HEK293-DAF fusosomes incubated with naive mouse sera than for HEK293 fusosomes incubated with naive mouse sera, indicating that complement activity targeting fusosomes is greater in mice treated with HEK293 fusosomes than HEK293-DAF fusosomes.

WHAT IS CLAIMED IS:

1. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogenous agent; and

(ii) a positive target cell-specific regulatory element operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a CNS cell.

2. The fusosome of claim 1, wherein the nucleic acid further comprises a non-target cell-specific regulatory element (NTCSRE), operatively linked to the payload gene, wherein the NTCSRE decreases expression of the payload gene in a non-target cell relative to an otherwise similar fusosome lacking the NTCSRE, wherein the target cell is a first type of CNS cell and the non-target cell is a second, different type of CNS cell or a non-CNS cell optionally wherein:

the target cell is a neuron and the non-target cell is a glial cell, optionally wherein the glial cell is an oligodendrocyte, an astrocyte, or a microglia cell, or

the target cell is a glial cell, optionally wherein the glial cell isan oligodendrocyte, an astrocyte, or a microglia cell, and the non-target cell is a neuron.

3. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogenous agent; and

(ii) a promoter operatively linked to the payload gene, wherein the promoter is chosen from a SYN, NSE, CaMKII, aTubulin, PDGF, fSST, fNPY, GAD67, DLX5/6, VGLUT1, Dock10, ChAT, VAChT, Drd1a, TPH-2, GFAP, EAAT1, GS, CX3CR1, TMEM119, MBP, CNP, or CRFR2β promoter.

4. A fusosome comprising:

a) a lipid bilayer comprising a fusogen; and

b) a nucleic acid that comprises:

(i) a payload gene encoding an exogenous agent; and

(ii) a non-target cell-specific regulatory element (NTCSRE) operatively linked to the payload gene, wherein:

the NTCSRE decreases expression of the payload gene in a non-target cell or tissue relative to an otherwise similar fusosome lacking the NTCSRE, wherein the target cell is a first type of CNS cell and the non-target cell is a second, different type of CNS cell or a non-CNS cell; and

the NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site.

5. The fusosome of claim 4, wherein the nucleic acid further comprises a positive target cell-specific regulatory element operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a CNS cell.

6. The fusosome of any of claims 1-5, wherein the fusosome further comprises one or both of:

(i) a first exogenous or overexpressed immunosuppressive protein on the lipid bilayer; or

(ii) a first immunostimulatory protein that is absent or present at reduced levels, optionally wherein the reduced level is reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% compared to a fusosome generated from an otherwise similar, unmodified source cell.

7. The fusosome of any of claims 1-6, wherien the payload gene is a gene that treats a lysosomal storage disease or disorder or a CNS disease or disorder, optionally wherein the disease or disorder is a genetic deficiency.

8. A fusosome comprising:

a) a lipid bilayer comprising a fusogen;

b) a nucleic acid that comprises a payload gene encoding an exogenous agent for treating a lysososomal storage disease or disorder or a CNS disease or disorder; and

c) one or both of:

(i) a first exogenous or overexpressed immunosuppressive protein on the lipid bilayer; or

(ii) a first immunostimulatory protein that is absent or present at reduced levels (e.g., reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%) compared to a fusosome generated from an otherwise similar, unmodified source cell.

9. The fusosome of any of claims 6-8, which comprises (i) and (ii).

10. The fusosome of any of claims 6-8, which comprises (i) and further comprises a second exogenous or overexpressed immunosuppressive protein on the lipid bilayer.

11. The fusosome of any of claims 6-10, which comprises (ii) and further comprises a second immunostimulatory protein that is absent or present at reduced levels, optionally wherein the reduced level is reduced by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% compared to a fusosome generated from an otherwise similar, unmodified source cell.

12. The fusosome of any of claims 8-11, wherein the nucleic acid further comprises a positive target cell-specific regulatory element operatively linked to the payload gene, wherein the positive target cell-specific regulatory element increases expression of the payload gene in a target cell relative to an otherwise similar fusosome lacking the positive target cell-specific regulatory element, wherein the target cell is a CNS cell.

13. The fusosome of any of claims 8-12, wherein the nucleic acid further comprises a non-target cell-specific regulatory element (NTCSRE) operatively linked to the payload gene, wherein the NTCSRE decreases expression of the payload gene in a non-target cell or tissue relative to an otherwise similar fusosome lacking the NTCSRE, wherein the target cell is a first type of CNS cell and the non-target cell is a second, different type of CNS cell or a non-CNS cell, optionally wherein:

the target cell is a neuron and the non-target cell is a glial cell, optionally wherein the glial cell is an oligodendrocyte, an astrocyte, or a microglia cell, or

the target cell is a glial cell, optionally wherein the glial cell is an oligodendrocyte, an astrocyte, or a microglia cell, and the non-target cell is a neuron.

14. The fusosome of any of claims 6-13, wherein, when administered to a subject, one or more of:

i) the fusosome does not produce a detectable antibody response or antibodies against the fusosome are present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level;

ii) the fusosome does not produce a detectable cellular immune response, or a cellular immune response against the fusosome is present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level;

iii) the fusosome does not produce a detectable innate immune response), or the innate immune response against the fusosome is present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level,;

iv) less than 10%, 5%, 4%, 3%, 2%, or 1% of fusosomes are inactivated by serum;

v) a target cell that has received the exogenous agent from the fusosome does not produce a detectable antibody response, or antibodies against the target cell are present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level; or

vi) a target cell that has received the exogenous agent from the fusosome does not produce a detectable cellular immune response, or a cellular response against the target cell is present at a level of less than 10%, 5%, 4%, 3%, 2%, or 1% above a background level.

15. The fusosome of claim 14, wherein the background level is the corresponding level in the same subject prior to administration of the fusosome.

16. The fusosome of any of claims 6-15, wherein the immunosuppressive protein is a complement regulatory protein or CD47.

17. The fusosome of any of claims 6-16, wherein the immunostimulatory protein is an MHC I or MHC II protein.

18. The fusosome of any of claims 1-17, wherein one or more of:

i) the fusosome fuses at a higher rate with the CNS target cell than with a non-target cell, optionslly wherein the higher rate is by at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold;

ii) the fusosome fuses at a higher rate with the CNS target cell than with another fusosome, optionally wherein the higher rate is by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold;

iii) the fusosome fuses with CNS target cells at a rate such that the exogenous agent in the fusosome is delivered to at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, of CNS target cells after 24, 48, or 72 hours;

iv) the fusosome delivers the nucleic acid to the CNS target cell at a higher rate than to a non-target cell, optionally wherein the higher rate is by at least at least 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold;

v) the fusosome delivers the nucleic acid to the CNS target cell at a higher rate than to another fusosome, optionally wherein the higher rate is by at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 50-fold, or 100-fold; or

vi) the fusosome delivers the nucleic acid to athe CNS target cell at a rate such that the exogenous agent in the fusosome is delivered to at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90%, of target cells after 24, 48, or 72 hours.

19. The fusosome of any of claims 1-18, wherein the exogenous agent is chosen from: SYNE1, SETX, FMR1, SLC6A8, UBE3A, SOD1, TDP43, C9orf72, FXN, MECP2, ASPA, or ALDH7A1; or the exogenous agent is chosen from: TPP1, FUCA1, GALC, HEXA, HEXB, MANBA, ARSA, GNPTAB, or MCOLN1.

20. The fusosome of any of claims 1-19, wherein the payload gene is selected from among SYNE1, SETX, FMR1, SLC6A8, UBE3A, SOD1, TDP43, C9orf72, FXN, MECP2, ASPA and ALDH7A1.

21. The fusosome of any of claims 1-20, wherein the payload gene encodes an exogenous agent comprising the sequence set forth in any one of SEQ ID NOS: 134-145, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 134-145.

22. The fusosome of any of claims 1-19, wherein the payload gene is selected from TPP1, FUCA1, GALC, HEXA, HEXB, MANBA, ARSA, GNPTAB and MCOLN1.

23. The fusosome of any of claims 1-19 and 22, wherein the payload gene encodes an exogenous agent comprising the sequence set forth in any one of SEQ ID NOS: 146-154, a functional fragment thereof, or a functional variant thereof comprising an amino acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to an amino acid sequence set forth in any one of SEQ ID NOS: 146-154.

24. The fusosome of any of claims 1-23, wherein the fusogen targets a CNS cell, optionally wherein the CNS cell is a neuron or a glial cell, optionally wherein the CNS cell is a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell.

25. The fusosome of any of claims 1-24, wherein the fusogen is a viral envelope protein.

26. The fusosome of any of claims 1-25, wherein the fusogen comprises VSV-G.

27. The fusosome of any of claims 1-26, wherein the fusogen comprises a sequence chosen from Nipah virus F and G proteins, measles virus F and H proteins, tupaia paramyxovirus F and H proteins, paramyxovirus F and G proteins or F and H proteins or F and HN proteins, Hendra virus F and G proteins, Henipavirus F and G proteins, Morbilivirus F and H proteins, respirovirus F and HN protein, a Sendai virus F and HN protein, rubulavirus F and HN proteins, or avulavirus F and HN proteins, or a derivative thereof, or any combination thereof.

28. The fusosome of any of claims 1-24 and 27, wherein the fusogen comprises a domain of at least 100 amino acids in length having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to a wild-type paramyxovirus fusogen, optionally wherein the wild-type paramyxovirus fusogen is set forth in any one of SEQ ID NOS: 1-133.

29. The fusosome of claim 27, wherein the wild-type paramyxovirus is a Nipah virus, optionally wherein the Nipah virus is a henipavirus.

30. The fusosome of any of claims 1-29, wherein the fusogen is re-targeted for delivery to a CNS cell, optionally wherein the CNS cell is a neuron or a glial cell, optionally wherien the CNS cell is a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell.

31. The fusosome of any claims 1, 2, 5, 6, 7, and 12-30, wherein the positive target cell-specific regulatory element comprises a CNS cell-specific promoter, a CNS cell-specific enhancer, a CNS cell-specific splice site, a CNS cell-specific site extending half-life of an RNA

or protein, a CNS cell-specific mRNA nuclear export promoting site, a CNS cell-specific translational enhancing site, or a CNS cell-specific post-translational modification site.

32. The fusosome of any claims 1, 2, 5, 6, 7 and 12-31 wherein the positive target cell-specific regulatory element comprises a CNS cell-specific promoter.

33. The fusosome of claim 32, wherein the positive CNS cell-specific regulatory element comprises a promoter chosen from a SYN, NSE, CaMKII, aTubulin, PDGF, fSST, fNPY, GAD67, DLX5/6, VGLUT1, Dock10, ChAT, VAChT, Drd1a, TPH-2, GFAP, EAAT1, GS, CX3CR1, TMEM119, MBP, CNP, or CRFR2β promoter.

34. The fusosome of any of claims 2, 4-7, and 13-33, wherein the NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site.

35. The fusosome of any of claims 2, 4-7, and 13-34 , wherein the NTCSRE comprises a tissue-specific miRNA recognition sequence, tissue-specific protease recognition site, tissue-specific ubiquitin ligase site, tissue-specific transcriptional repression site, or tissue-specific epigenetic repression site.

36. The fusosome of any of claims 2, 4-7, and 13-35, wherein the NTCSRE comprises a non-target cell-specific miRNA recognition sequence, non-target cell-specific protease recognition site, non-target cell-specific ubiquitin ligase site, non-target cell-specific transcriptional repression site, or non-target cell-specific epigenetic repression site.

37. The fusosome of any of claims 2, 4-7, and 13-35, wherein the NTCSRE comprises a non-target cell-specific miRNA recognition sequence and the miRNA recognition sequence is able to be bound by one or more of miR-338-3p, miR-9, miR-125b-5p, miR-342-3p, or miR-124; optinally wherein the miRNA is or comprises the sequence set forth in any one of SEQ ID NOS: 156-162.

38. The fusosome of any of claims 34-37, wherein the NTCSRE is situated or encoded within a transcribed region encoding the exogenous agent, optionally wherein an RNA produced by the transcribed region comprises the miRNA recognition sequence within a UTR or coding region.

39. The fusosome of any of claims 1-38, wherein the nucleic acid comprises one or more insulator elements.

40. The fusosome of claim 39, wherein the nucleic acid comprises two insulator elements, optionally wherein the two insulator elements comprisea first insulator element upstream of the payload gene and a second insulator element downstream of the payload gene, optionally wherein the first insulator element and second insulator element comprise the same or different sequences.

41. The fusosome of any of claims 1-40, wherein the fusosome is a retroviral vector particle.

42. The fusosome of any of claims 1-41, wherein the nucleic acid is capable of integrating into the genome of a CNS cell.

43. The fusosome of any of claims 1-42, wherein the target cell is chosen from a CNS cell, optionally wherein the CNS cell is a neuron or a glial cell, optionally wherein the CNS cell is a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell.

44. A pharmaceutical composition comprising the fusosome of any of any of claims 1-43, and a pharmaceutically acceptable carrier, diluent, or excipient.

45. A method of delivering an exogenous agent to a subject comprising administering to the subject the fusosome of any of claims 1-43 or the pharmaceutical composition of claim 44, thereby delivering the exogenous agent to the subject.

46. A method of modulating a function, in a subject , CNS tissue , or a CNS cell, comprising contactingthe CNS tissue or the CNS cell of the subject with the fusosome of any of claims 1-43 or the pharmaceutical composition of claim 45.

47. The method of claim 46, wherein the CNS cell is neuron or a glial cell, optionally wherein the CNS cell is a pan-neuronal cell, a GABAergic neuron, a Glutamatergic neuron, a Cholinergic neuron, a Dopaminergic neuron, a Serotonergic neuron, a glial cell, an astrocyte, a microglial cell, an oligodendrocyte, or a choroid plexus cell.

48. The method of claim 46 or claim 47, wherein the CNS tissue or the CNS cell is present in a subject.

49. A method of treating a CNS disease or disorder or a lysosomal disease or disorder, comprising administering to the subject the fusosome of any of claims 1-43 or the pharmaceutical composition of claim 44.

50. The method of claim 49, wherein the CNS disease or disorder or the lysosomal disease or disorder is caused by a genetic deficiency.

51. A method of treating a genetic deficiency in a subject comprising administering to the subject the fusosome of any of claims 1-43 or the pharmaceutical composition of claim 44.

52. The method of claim 50 or claim 51, wherein the genetic deficiency is a genetic deficiency able to be treated by the payload gene encoding the exogenous agent.

53. The method of claim 49, claim 50 or claim 52, wherein the disease or disorder is selected from Spinocerebellar Ataxia; Autosomal Recessive, Type 1; Ataxia with Oculomotor Apraxia, Type 2; Fragile X Syndrome; Cerebral Creatine Deficiency Syndrome 1; Angelman Syndrome; Amyotrophic Lateral Sclerosis; Friedreich's Ataxia; Rett Syndrome; Canavan Disease; Pyridoxine-Dependent Epilepsy; Batten Disease, Fucosidosis; Krabbe Disease; Tay

Sachs Disease; Sandhoff Disease; Beta-mannosidosis; Metachromatic Leukodystrophy; Mucolipidosis Type IIIa; Mucolipidosis Type IIIb; or Mucolipidosis Type IV.

54. The method of any of claims 49-53, wherein the subject is a human subject.

55. A fusosome of any of claims 1-43 or the pharamaceutical composition of claim 44 for use in treating a subect with a CNS disease or disorder or a lysosomal disease or disorder.

56. Use of a fusosome of any of claims 1-43 or the pharamaceutical composition of claim 44 for manufacture of a medicament for use in treating a subect with a CNS disease or disorder or a lysosomal disease or disorder.

57. The fusosome or pharmaceutical composition for use of claim 55 or the use of claim 56, wherein the CNS disease or disorder or a lysosomal disease or disorder is caused by a genetic deficiency.

58. The fusosome or pharmaceutical composition for use of claim 55 or claim 57 or the use of claim 56 or claim 57, wherein the disease or disorder is selected from Spinocerebellar Ataxia; Autosomal Recessive, Type 1; Ataxia with Oculomotor Apraxia, Type 2; Fragile X Syndrome; Cerebral Creatine Deficiency Syndrome 1; Angelman Syndrome; Amyotrophic Lateral Sclerosis; Friedreich's Ataxia; Rett Syndrome; Canavan Disease; Pyridoxine-Dependent Epilepsy; Batten Disease, Fucosidosis; Krabbe Disease; Tay Sachs Disease; Sandhoff Disease; Beta-mannosidosis; Metachromatic Leukodystrophy; Mucolipidosis Type IIIa; Mucolipidosis Type IIIb; or Mucolipidosis Type IV.

59. A fusosome of any of claims 1-43 or the pharamaceutical composition of claim 44 for use in treating a genetic deficiency.

60. Use of a fusosome of any of claims 1-43 or the pharamaceutical composition of claim 44 for manufacture of a medicament for use in treating a genetic deficiency.

61. The fusosome or pharmaceutical composition for use of any of claims 57-59, or the use of claim 57, 58 and 60, wherein the genetic deficiency is a genetic deficiency able to be treated by the payload gene encoding the exogenous agent.

62. A method of making the fusosome of any of claims 1-43, comprising:

a) providing a cell that comprises the nucleic acid and the fusogen;

b) culturing the cell under conditions that allow for production of the fusosome, and

c) separating, enriching, or purifying the fusosome from the cell, thereby making the fusosome.

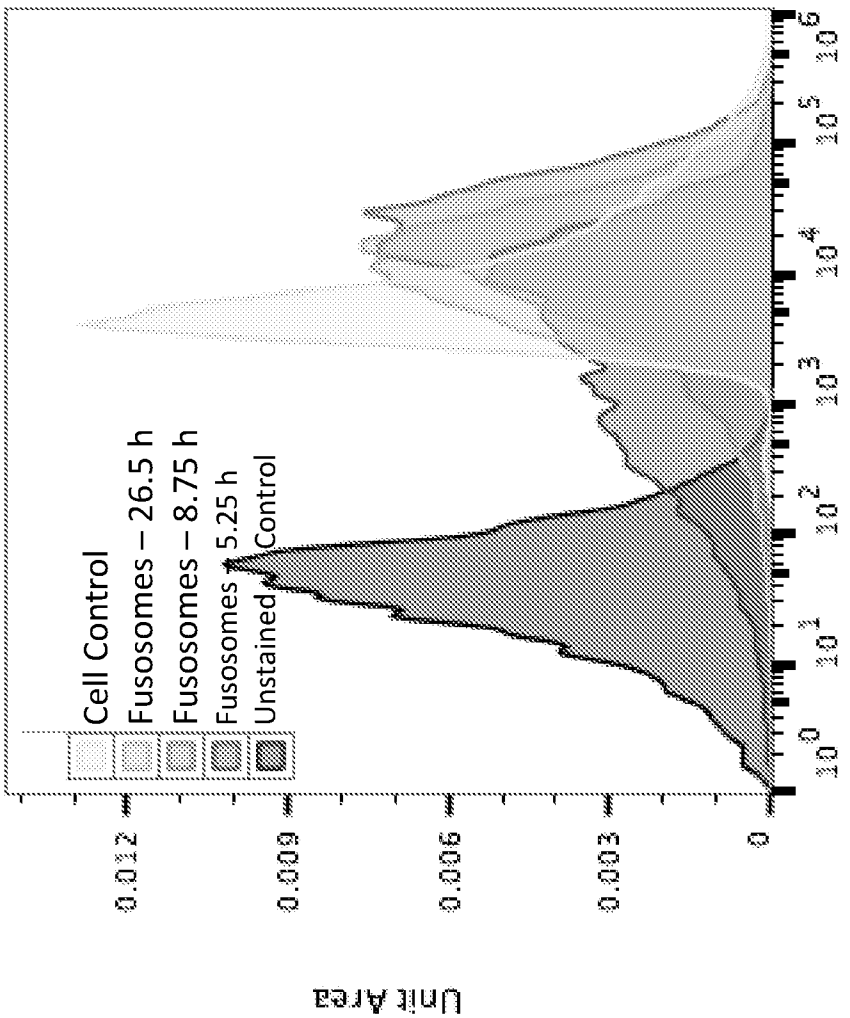

Cell Control
Fusosomes – 26.5 h
Fusosomes – 8.75 h
Fusosomes – 5.25 h
Unstained Control
Unit Area
0.012
0.009
0.006
0.003
0
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Como-YL1-A :: Phalloidin-A

Fig. 1

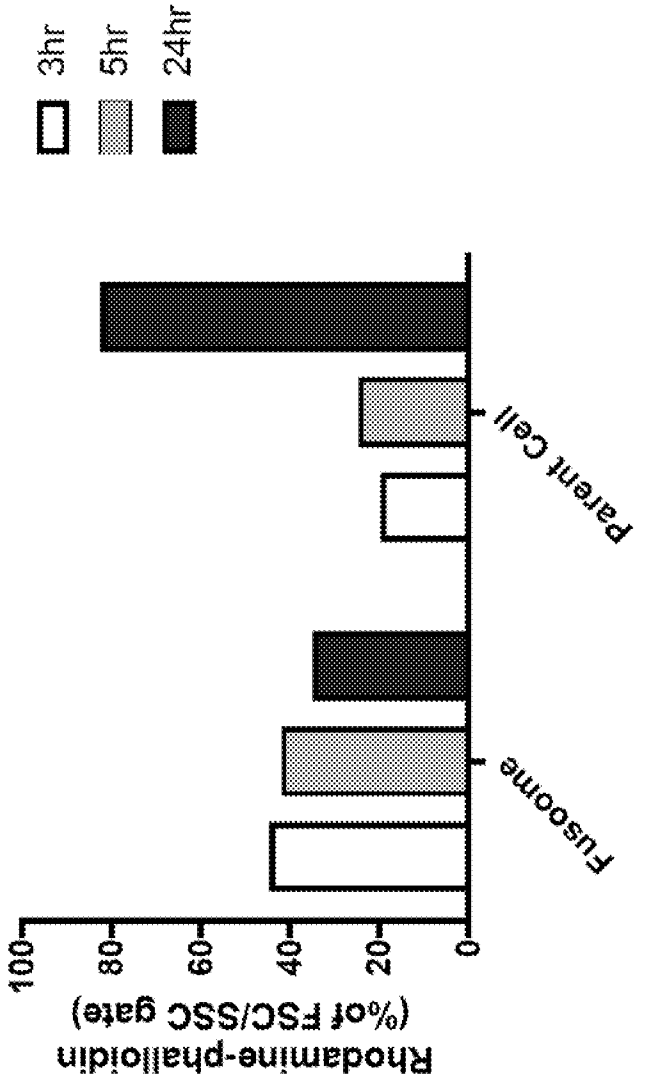

3hr
5hr
24hr
Rhodamine-phalloidin
(%of FSC/SSC gate)
100
80
60
40
20
0
Fusoome
Parent Cell

Fig. 2

	Fusosomes	Parental Cells
Average diameter (nm)	128.7	175.4
Min Diameter (nm)	14	29
Max Diameter (nm)	18720	19802
Median Diameter (nm)	134	99
10% Quantile (nm)	53	52
25% Quantile (nm)	88	66
75% Quantile (nm)	226	241
90% Quantile (nm)	4450	10649
Average volume (μm^3)	0.067	7.421

Fig. 3

	Fusosomes	Parental Cells
Average diameter (nm)	128.7	175.4
Average volume (μm^3)	0.067	7.421

Fig. 4

	[Protein] (mg/mL)		Ratio
	Insoluble	Soluble	
Cells	2.12	0.06	0.028
Fusosomes	1.28	0.72	0.563

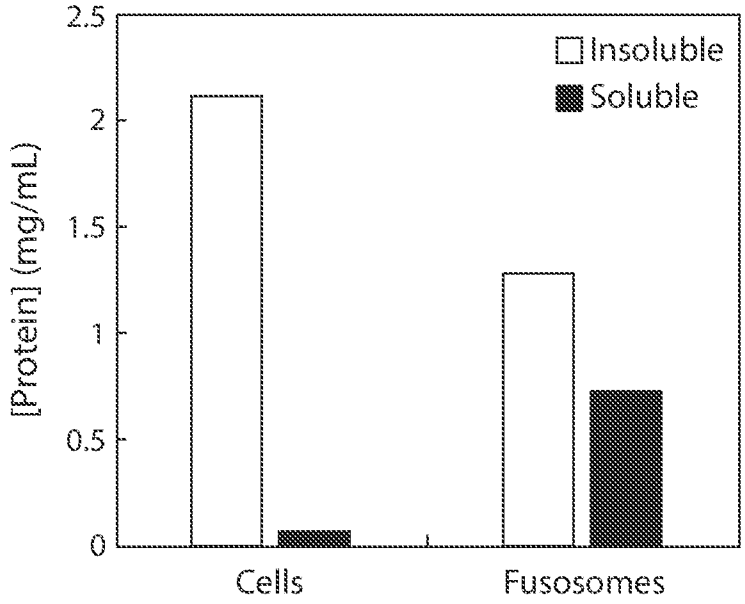
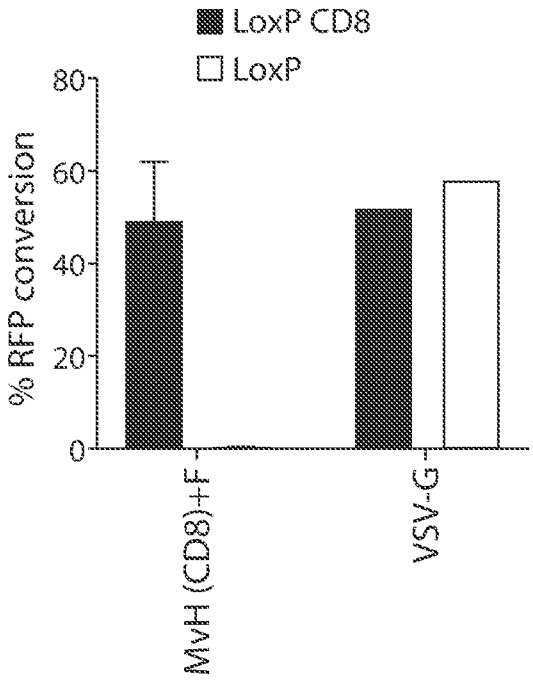


Fig. 5



Fusosome	Absolute amount of targeted fusion
MvH(CD8)+F	48.8%
VSV-G	0%

Fig. 6

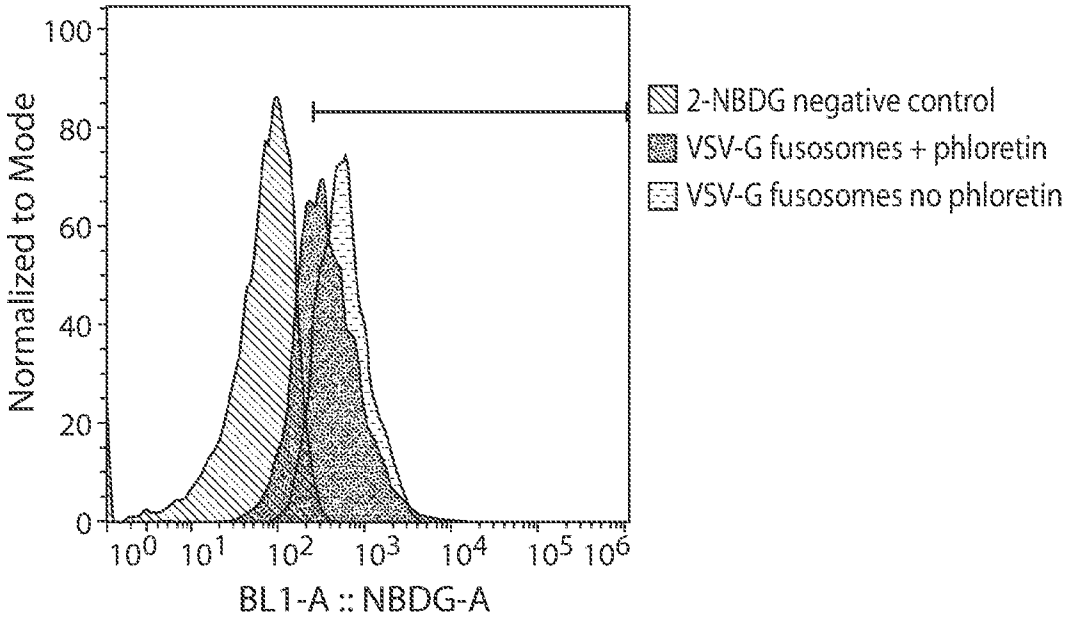

Normalized to Mode
100
80
60
40
20
0
10^0 10^1 10^2 10^3 10^4 10^5 10^6
BL1-A :: NBDG-A
2-NBDG negative control
VSV-G fusosomes + phloretin
VSV-G fusosomes no phloretin

Fig. 7

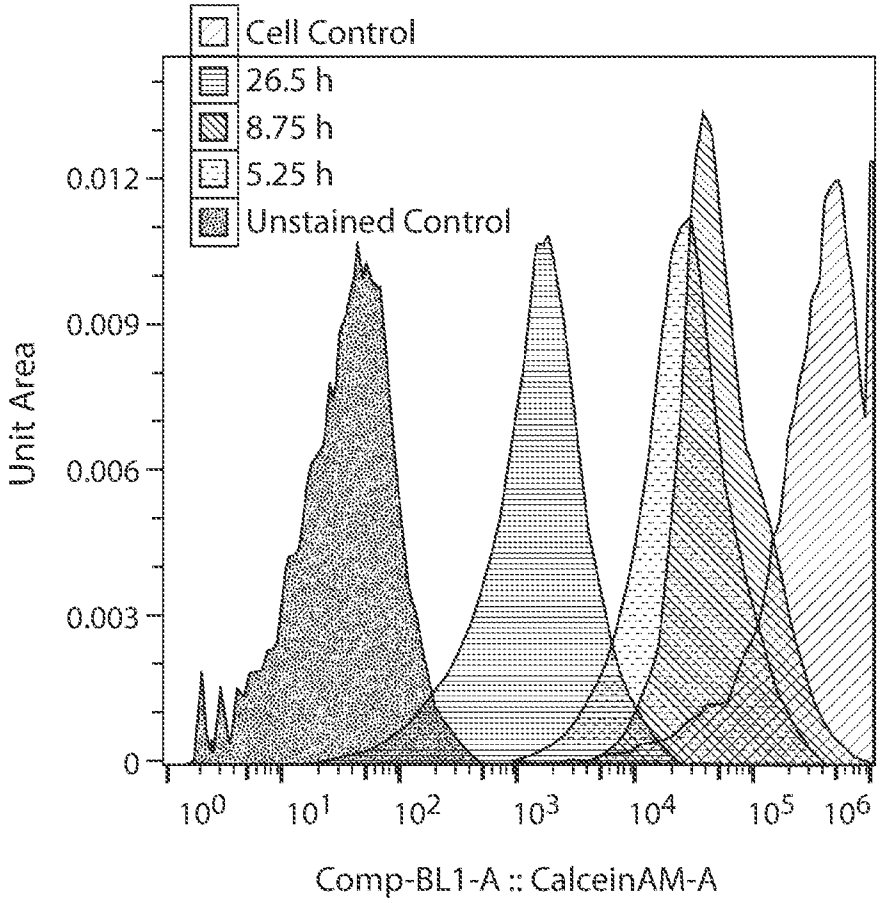

Cell Control
26.5 h
8.75 h
5.25 h
Unstained Control
Unit Area
0.012
0.009
0.006
0.003
0
10^0
10^1
10^2
10^3
10^4
10^5
10^6
Comp-BL1-A :: CalceinAM-A

Fig. 8

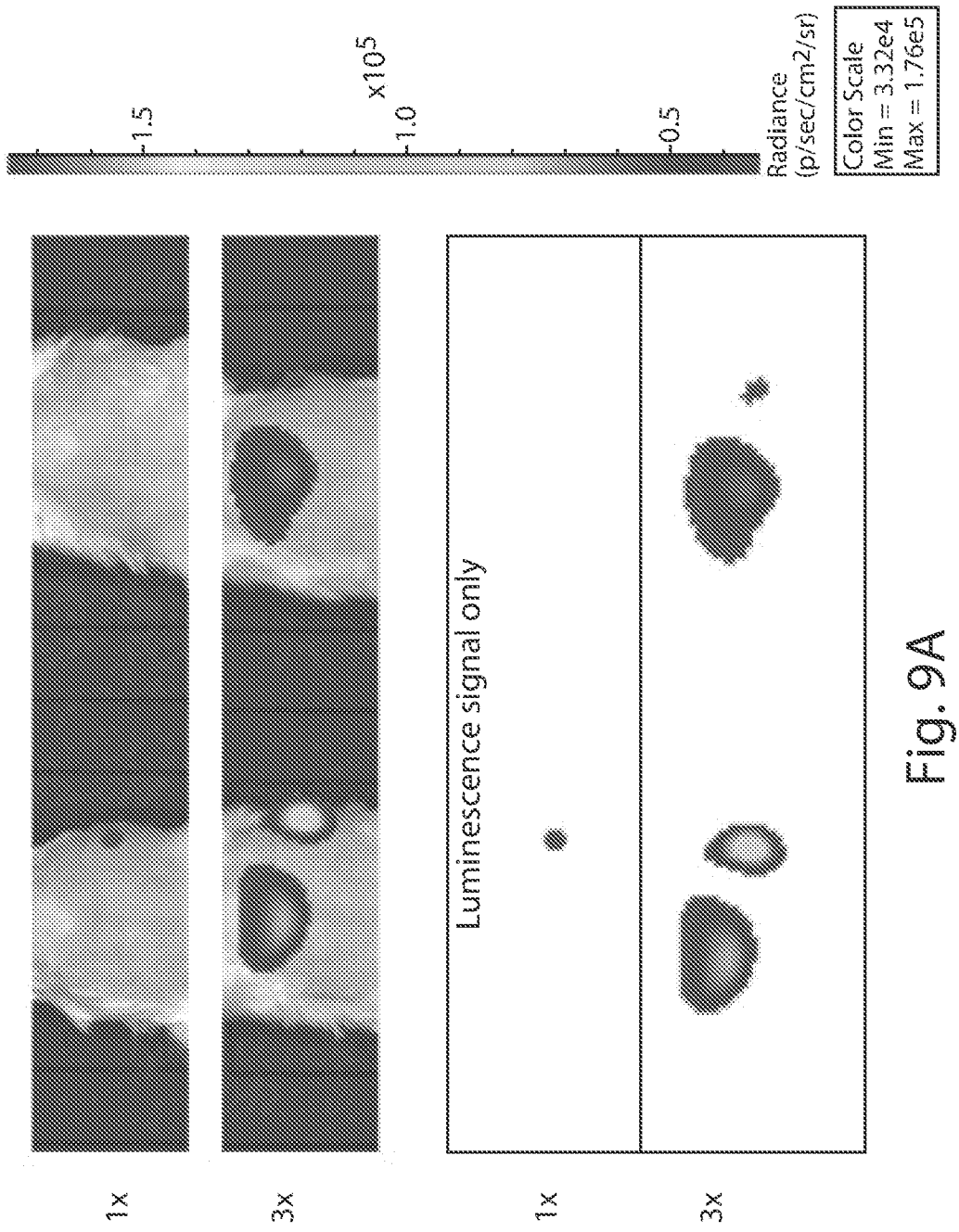

1.5
x10^5
1.0
0.5
Radiance (p/sec/cm2/sr)
Color Scale
Min = 3.32e4
Max = 1.76e5
Luminescence signal only
1x
3x
1x
3x

Fig. 9A

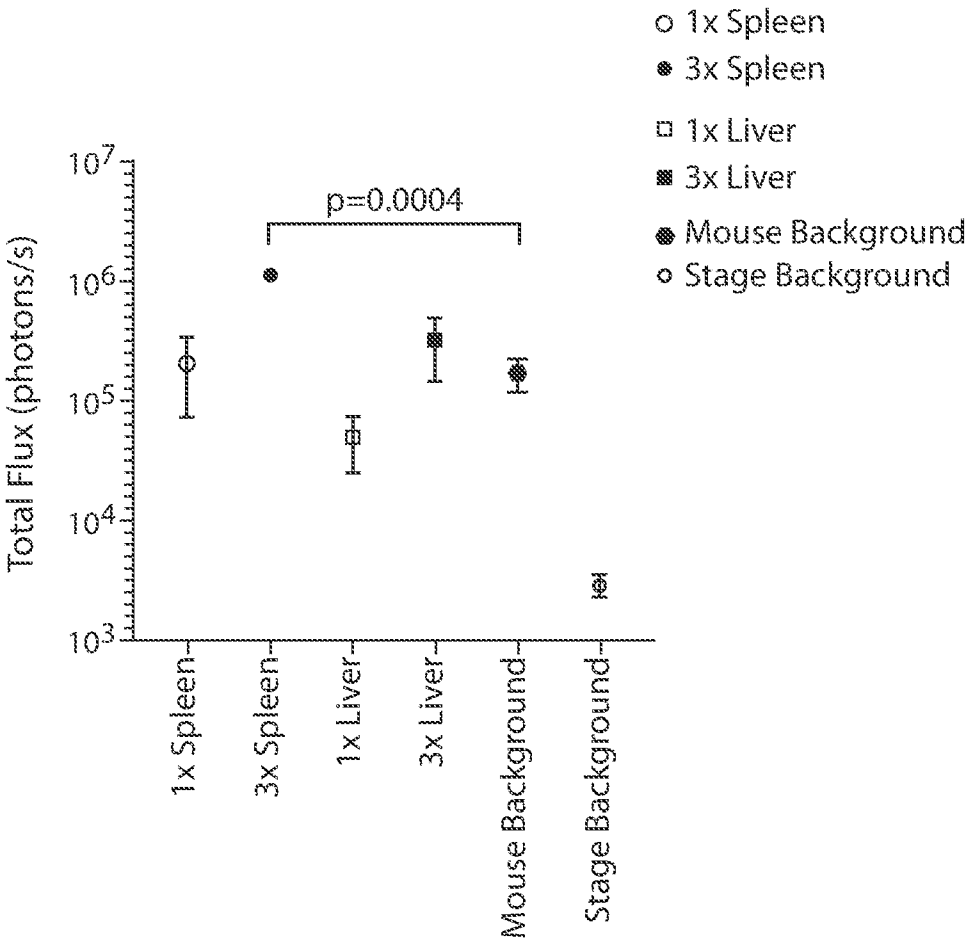

1x Spleen
3x Spleen
1x Liver
3x Liver
Mouse Background
Stage Background
p=0.0004
Total Flux (photons/s)
10^7
10^6
10^5
10^4
10^3
1x Spleen
3x Spleen
1x Liver
3x Liver
Mouse Background
Stage Background

Fig. 9B

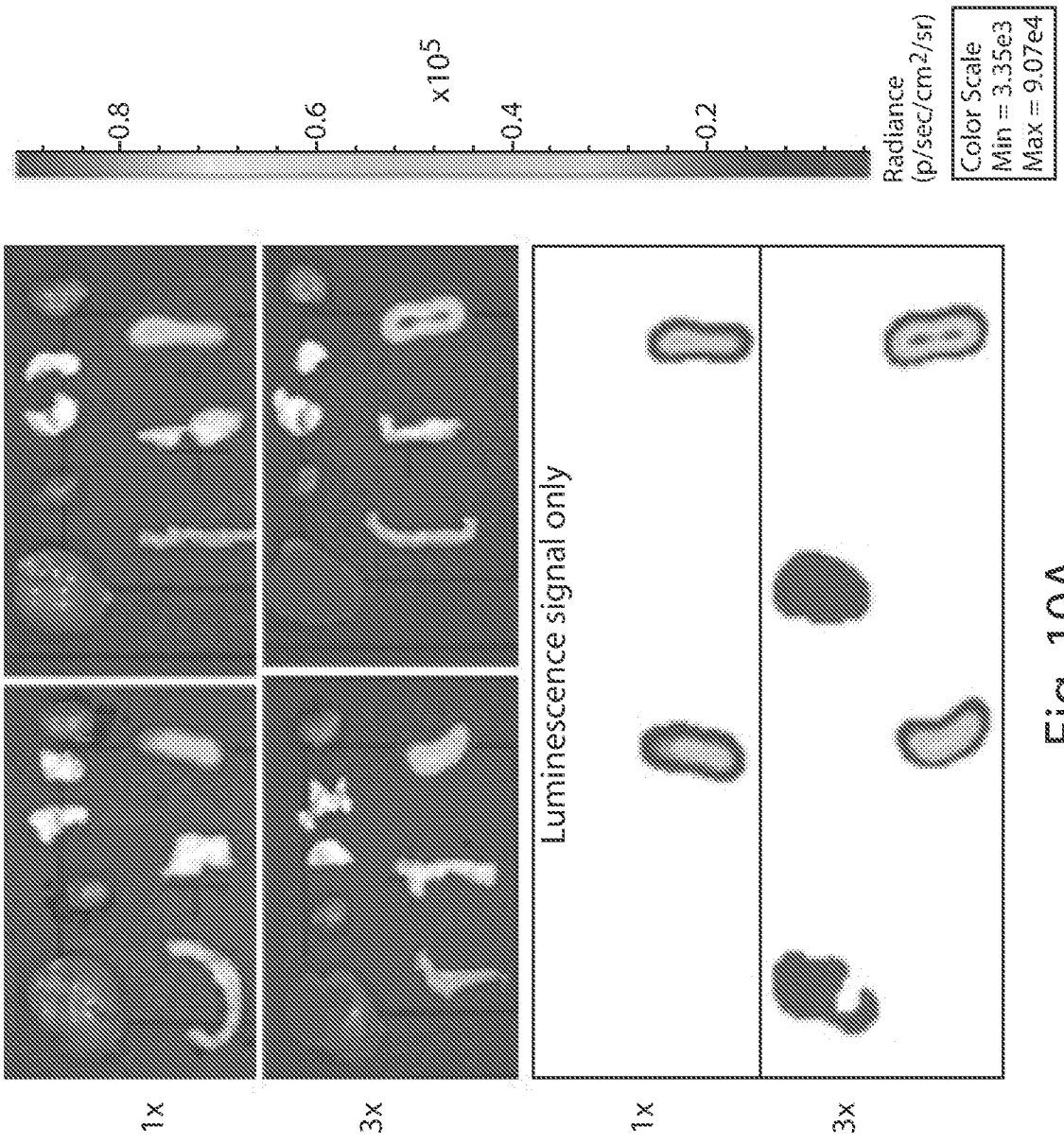

0.8
0.6
x10^5
0.4
0.2
Radiance (p/sec/cm²/sr)
Color Scale
Min = 3.35e3
Max = 9.07e4
Luminescence signal only
1x
3x
1x
3x

Fig. 10A

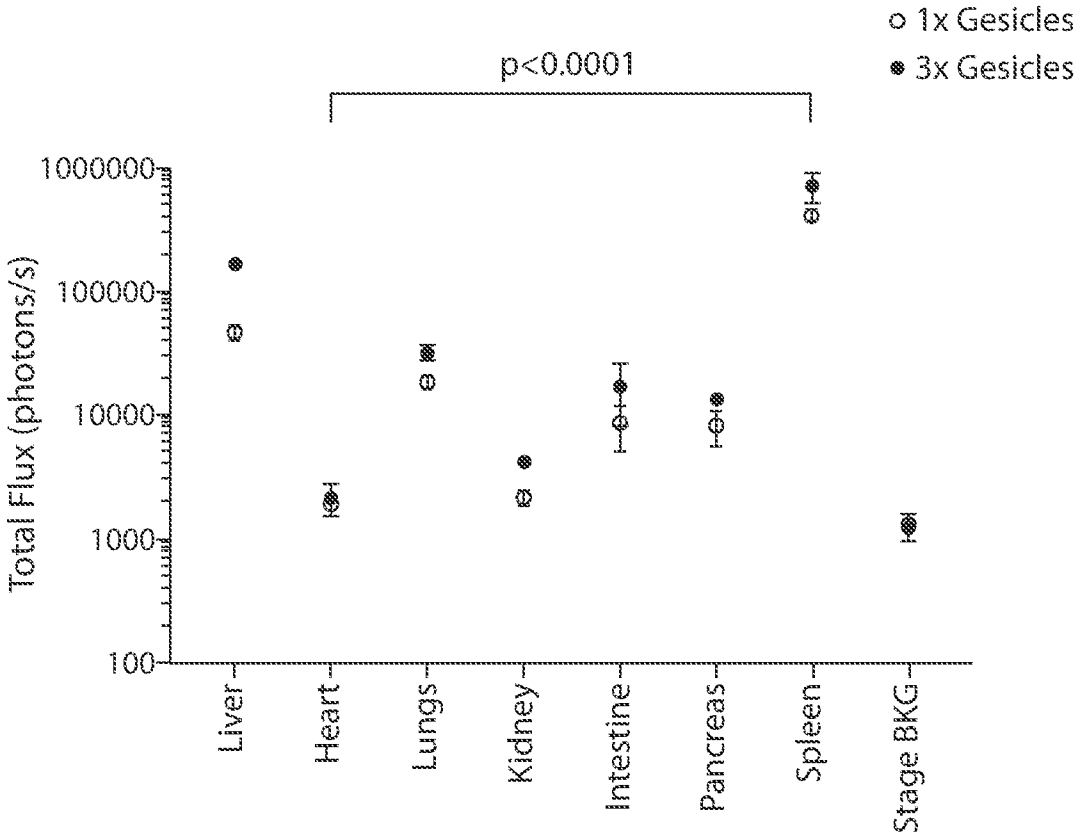

1x Gesicles
3x Gesicles
p<0.0001
Total Flux (photons/s)
1000000
100000
10000
1000
100
Liver
Heart
Lungs
Kidney
Intestine
Pancreas
Spleen
Stage BKG

Fig. 10B

Group	**% RFP conversion (± SD)**
Recipient + no fusosome	0.4 ± 0.2%
Recipient + NivG+F fusosome	88.9 ± 3.4%
Recipient + NivG+F fusosome + Baf	68.1 ± 2.7%

Fig. 11

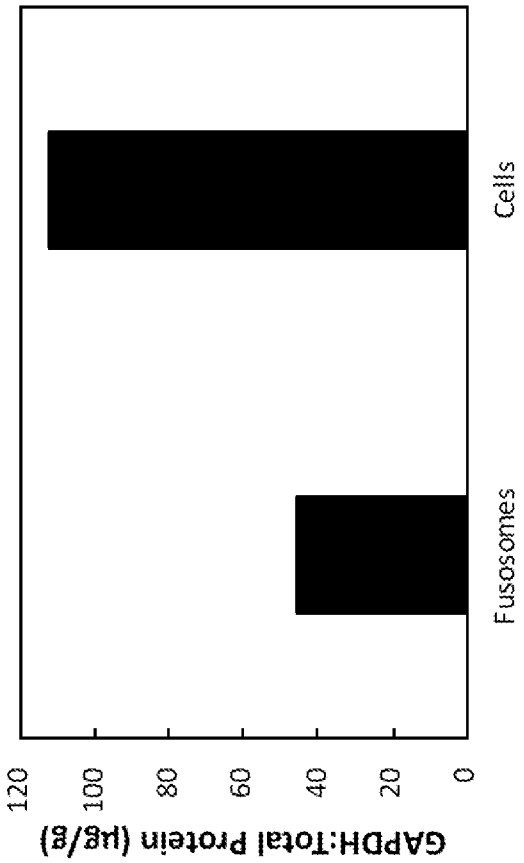


Fig. 12

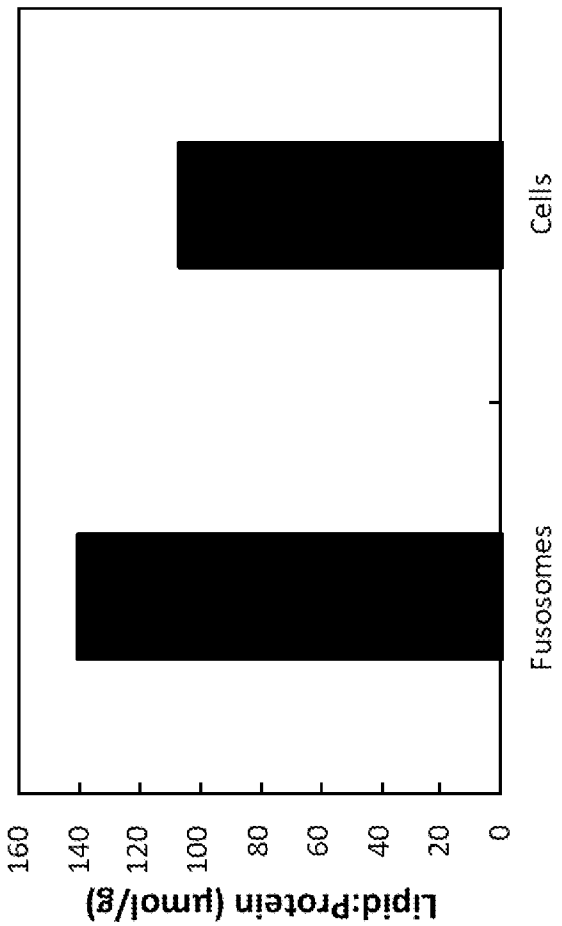

Lipid:Protein (μmol/g)
160
140
120
100
80
60
40
20
0
Fusosomes
Cells

Fig. 13

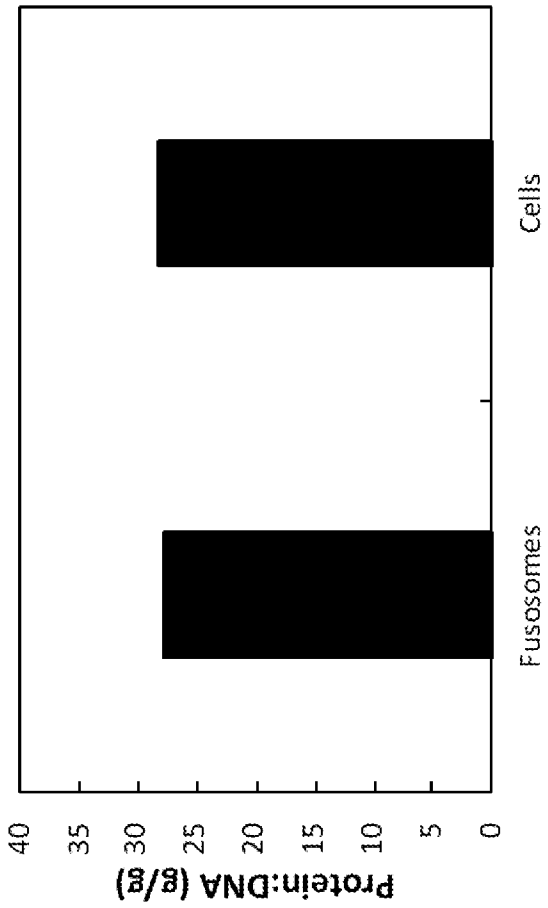


Fig. 14

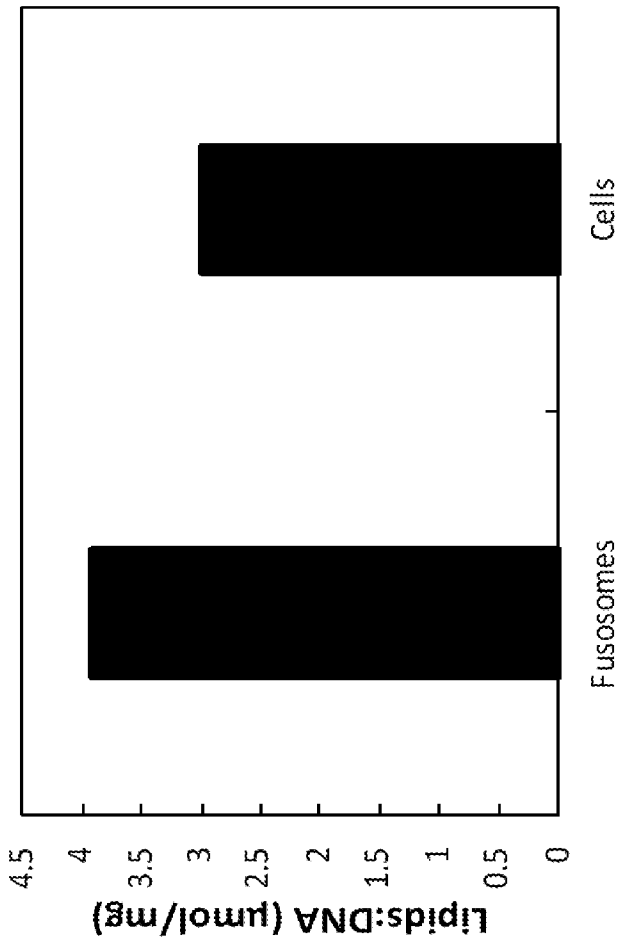


Fig. 15

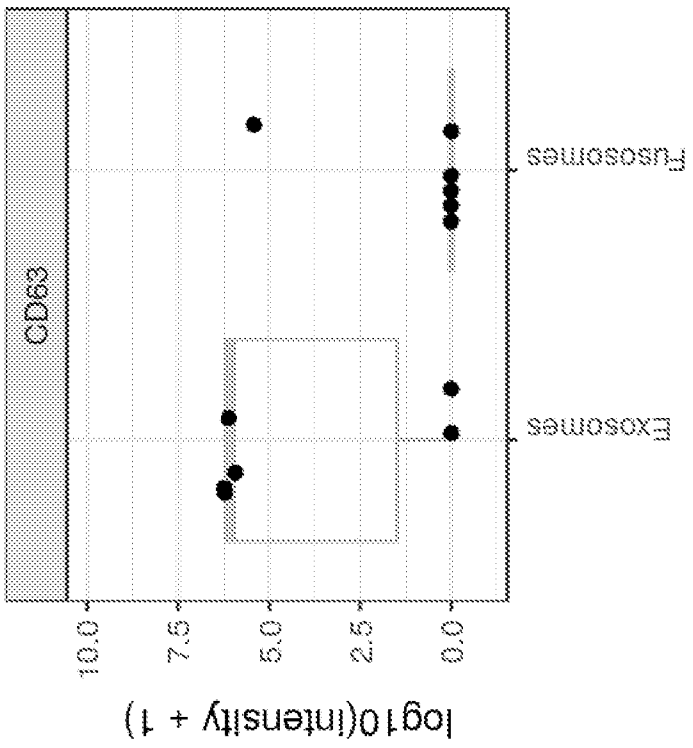

CD63
10.0
7.5
5.0
2.5
0.0
log10(intensity + 1)
Exosomes
Fusosomes
cohort

Fig. 16

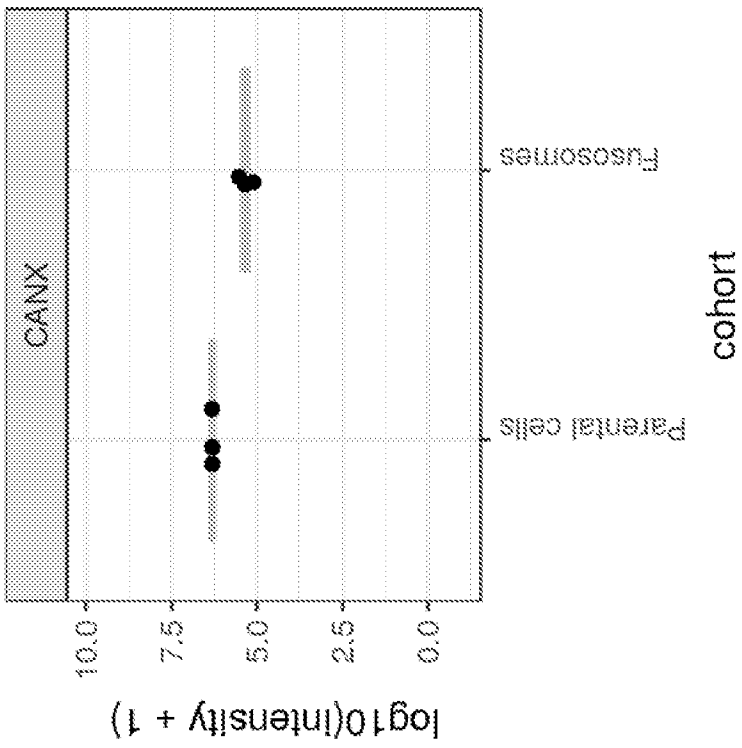

CANX
10.0
7.5
5.0
2.5
0.0
log10(Intensity + 1)
Parental cells
Fusosomes
cohort

Fig. 17

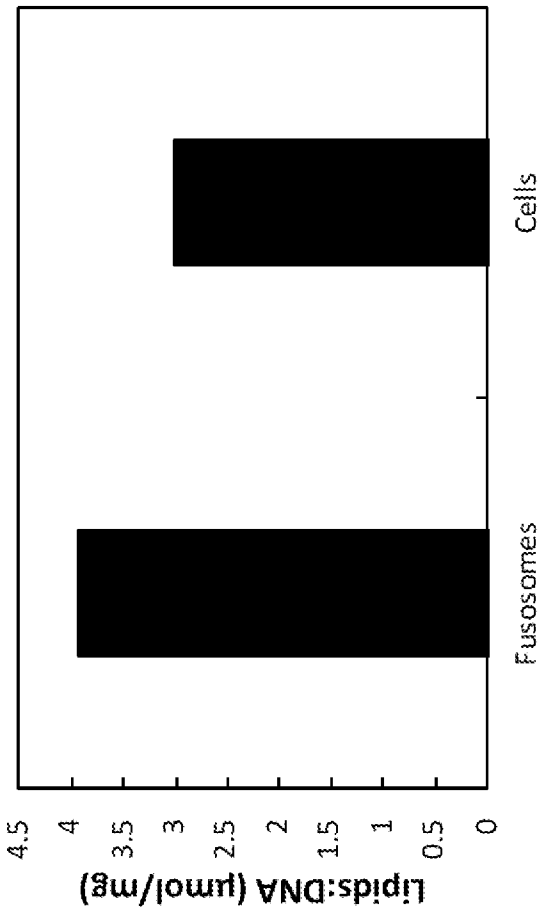


Fig. 18

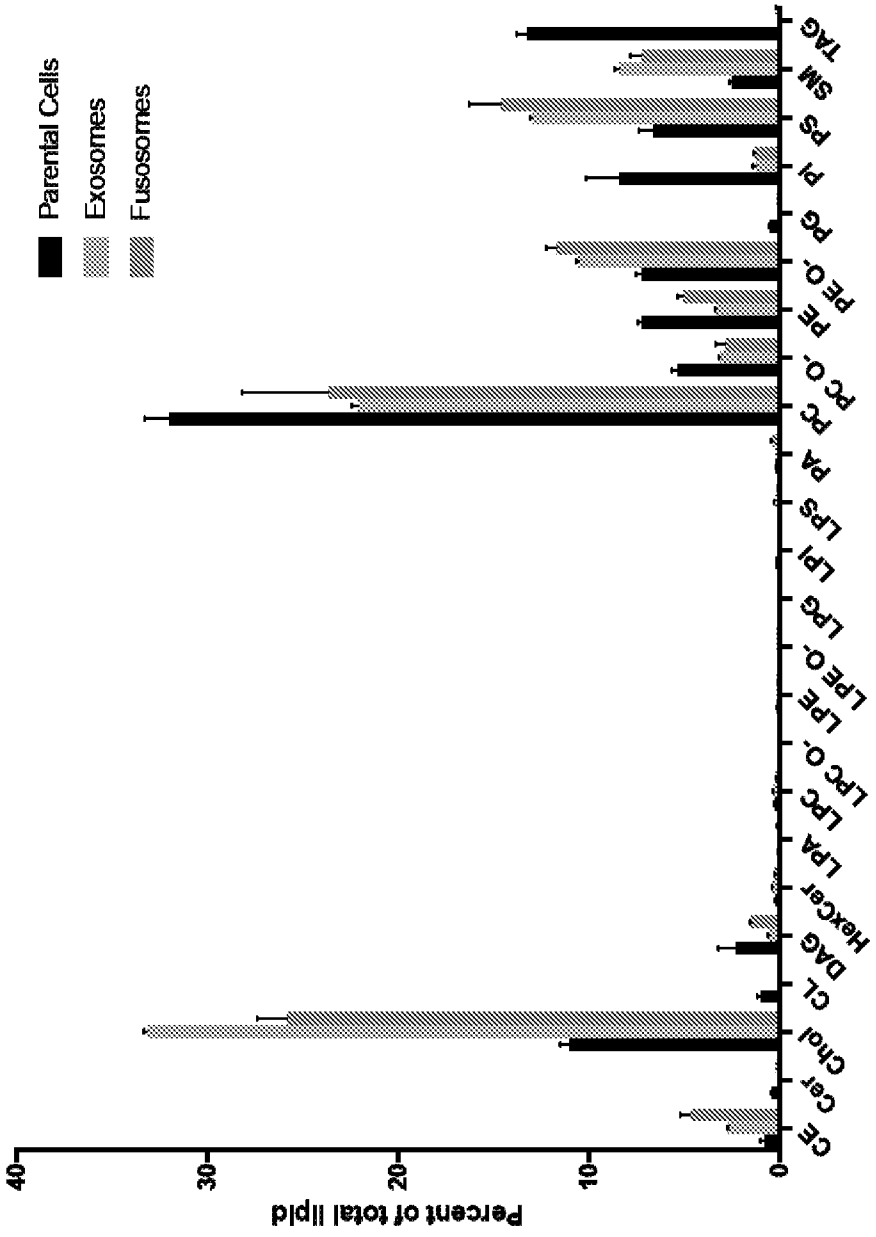

Parental Cells
Exosomes
Fusosomes
Percent of total lipid
0
10
20
30
40
CE
Cer
Chol
CL
DAG
HexCer
LPA
LPC
LPC O-
LPE
LPE O-
LPG
LPI
LPS
PA
PC
PC O-
PE
PE O-
PG
PI
PS
SM
TAG

Fig. 19A

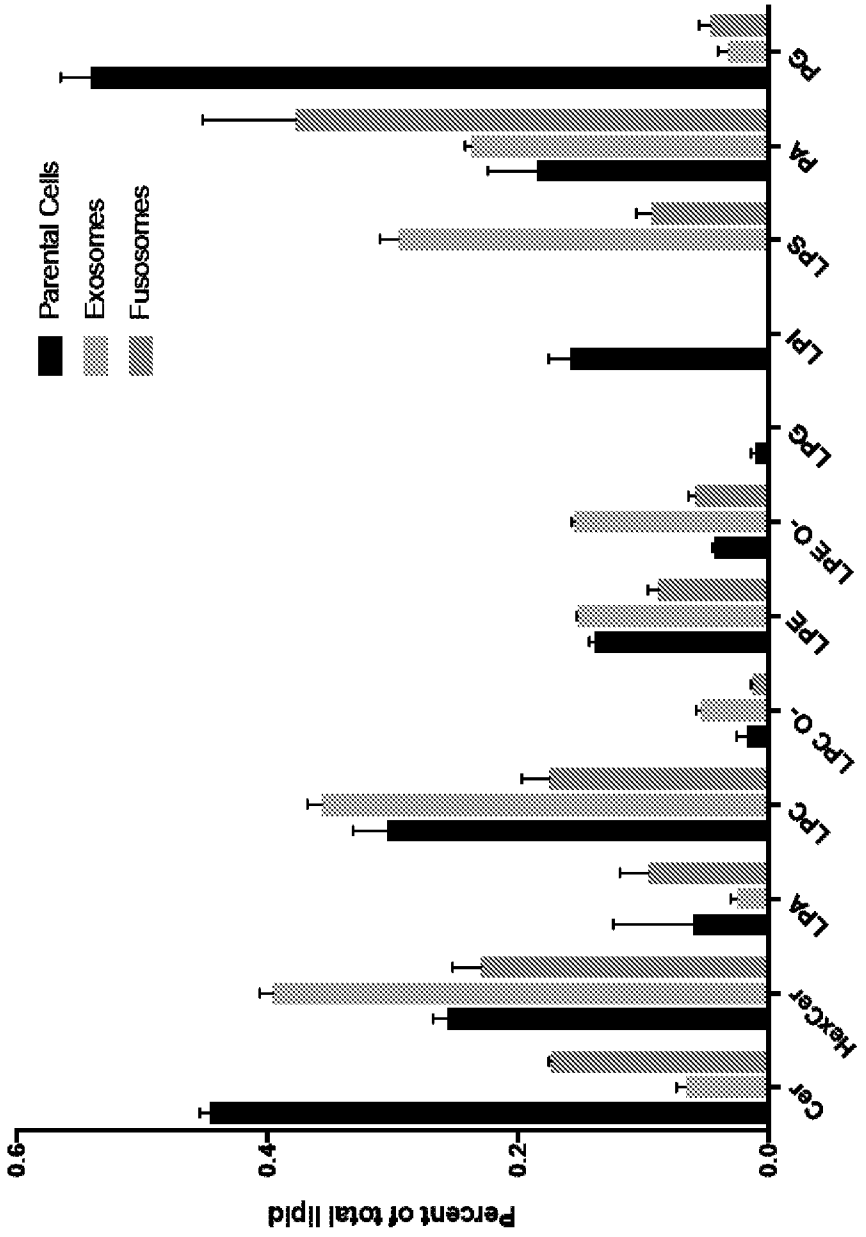

Parental Cells
Exosomes
Fusosomes
Percent of total lipid
0.0
0.2
0.4
0.6
Cer
Hexcer
LPA
LPC
LPC O-
LPE
LPE O-
LPG
LPI
LPS
PA
PG

Fig. 19B

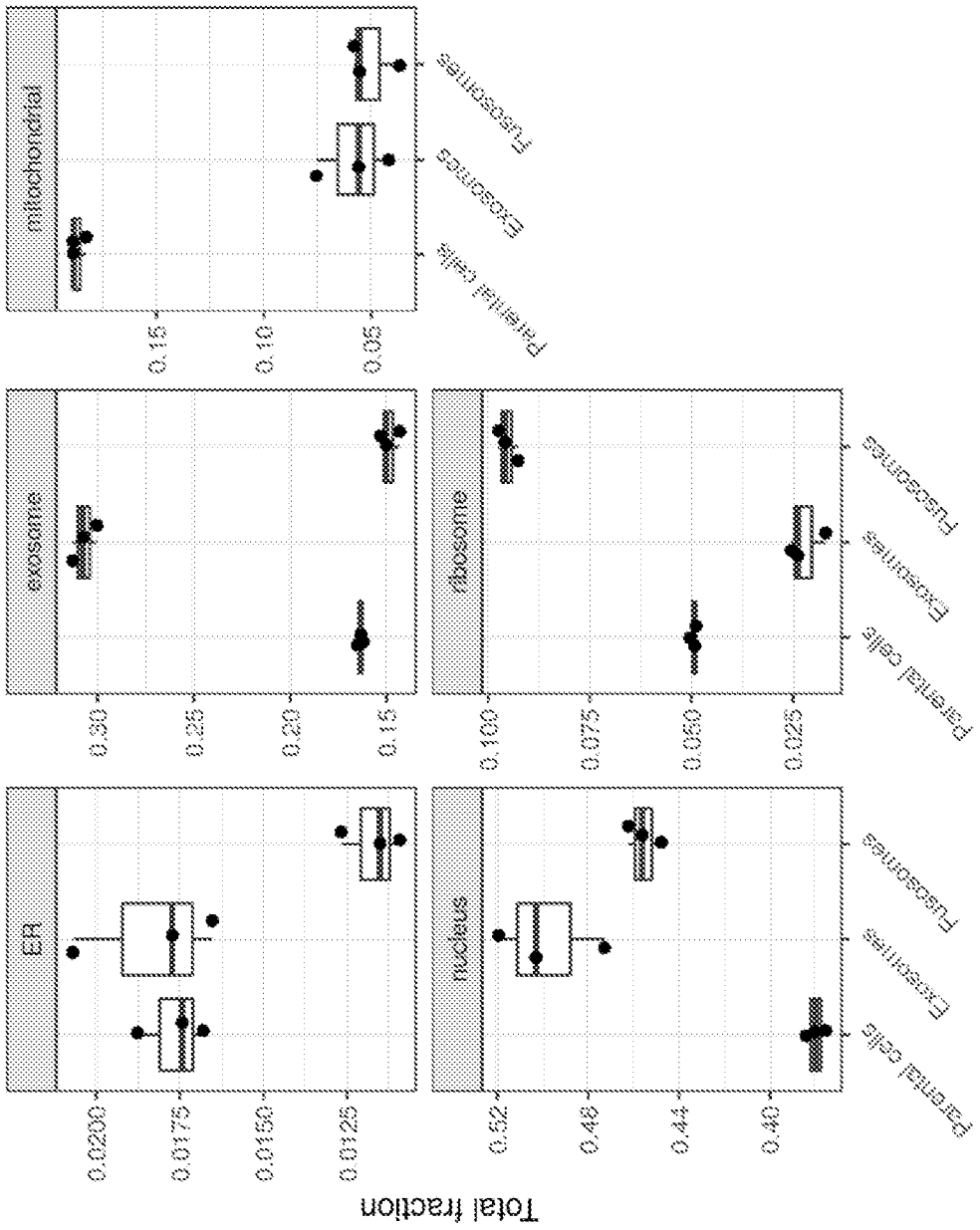

Total fraction
ER
exosome
mitochondrial
nucleus
ribosome
Parental cells
Exosomes
Fusosomes
0.0200
0.0175
0.0150
0.0125
0.30
0.25
0.20
0.15
0.15
0.10
0.05
0.52
0.48
0.44
0.40
0.100
0.075
0.050
0.025

Fig. 20

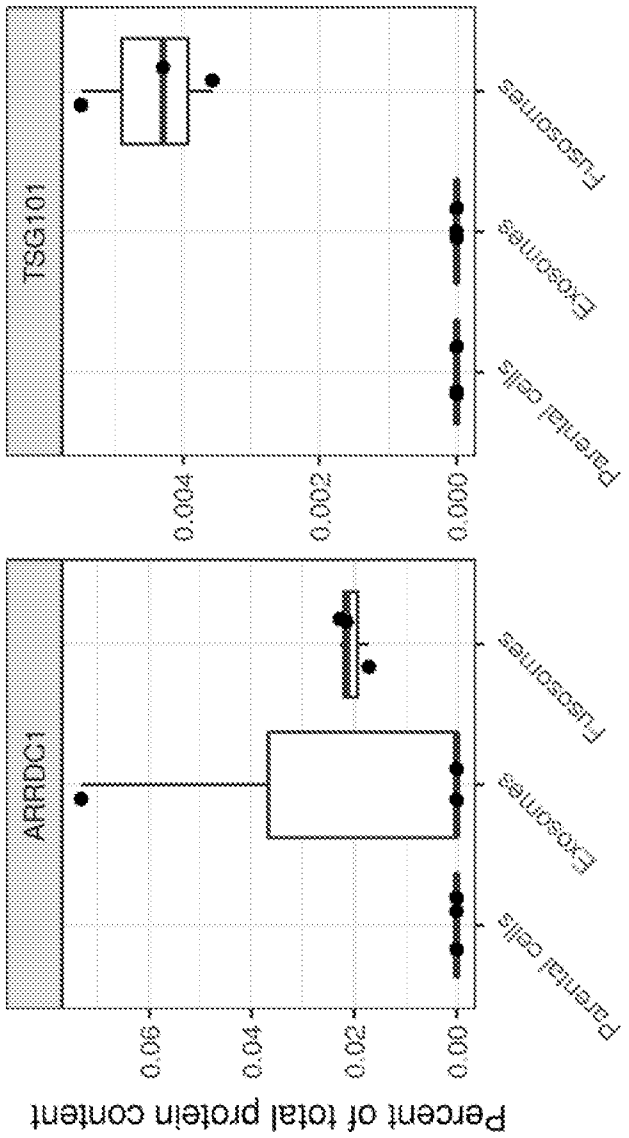

Percent of total protein content
ARRDC1
TSG101
0.00
0.02
0.04
0.06
0.000
0.002
0.004
Parental cells
Exosomes
Fusosomes

Fig. 21

SEQUENCE LISTING

<110> FLAGSHIP PIONEERING INNOVATIONS V, INC.

<120> FUSOSOME COMPOSITIONS FOR CNS DELIVERY

<130> V2050-7029WO_VL39023-03

<140> Not Yet Assigned
<141> Concurrently Herewith

<150> 62/767,358
<151> 2018-11-14

<150> 62/900,064
<151> 2019-09-13

<160> 162

<170> PatentIn version 3.5

<210> 1
<211> 589
<212> PRT
<213> human respiratory syncytal virus

<220>
<221> MISC_FEATURE
<223> fusion glycoprotein

<400> 1

Met Ile Pro Gln Ala Arg Thr Glu Leu Asn Leu Gly Gln Ile Thr Met
1 5 10 15

Glu Leu Leu Ile His Arg Ser Ser Ala Ile Phe Leu Thr Leu Ala Ile
20 25 30

Asn Ala Leu Tyr Leu Thr Ser Ser Gln Asn Ile Thr Glu Glu Phe Tyr
35 40 45

Gln Ser Thr Cys Ser Ala Val Ser Arg Gly Tyr Leu Ser Ala Leu Arg
50 55 60

Thr Gly Trp Tyr Thr Ser Val Ile Thr Ile Glu Leu Ser Asn Ile Lys
65 70 75 80

Glu Thr Lys Cys Asn Gly Thr Asp Thr Lys Val Lys Leu Ile Lys Gln
85 90 95

Glu Leu Asp Lys Tyr Lys Asn Ala Val Thr Glu Leu Gln Leu Leu Met
100 105 110

Gln Asn Thr Pro Ala Ala Asn Asn Arg Ala Arg Arg Glu Ala Pro Gln
115 120 125

Tyr Met Asn Tyr Thr Ile Asn Thr Thr Gly Ser Leu Asn Val Ser Ile
130 135 140

```
Ser Lys Lys Arg Lys Arg Arg Phe Leu Gly Phe Leu Leu Gly Val Gly
145                 150                 155                 160

Ser Ala Ile Ala Ser Gly Ile Ala Val Ser Lys Val Leu His Leu Glu
                165                 170                 175

Gly Glu Val Asn Lys Ile Lys Asn Ala Leu Leu Ser Thr Asn Lys Ala
            180                 185                 190

Val Val Ser Leu Ser Asn Gly Val Ser Val Leu Thr Ser Lys Val Leu
        195                 200                 205

Asp Leu Lys Asn Tyr Ile Asn Asn Gln Leu Leu Pro Ile Val Asn Gln
    210                 215                 220

Gln Ser Cys Arg Ile Ser Asn Ile Glu Thr Val Ile Glu Phe Gln Gln
225                 230                 235                 240

Lys Asn Ser Arg Leu Leu Glu Ile Thr Arg Glu Phe Ser Val Asn Ala
                245                 250                 255

Gly Val Thr Thr Pro Leu Ser Thr Tyr Met Leu Thr Asn Ser Glu Leu
            260                 265                 270

Leu Ser Leu Ile Asn Asp Met Pro Ile Thr Asn Asp Gln Lys Lys Leu
        275                 280                 285

Met Ser Ser Asn Val Gln Ile Val Arg Gln Gln Ser Tyr Ser Ile Met
    290                 295                 300

Ser Ile Ile Lys Glu Glu Val Leu Ala Tyr Val Val Gln Leu Pro Ile
305                 310                 315                 320

Tyr Gly Val Ile Asp Thr Pro Cys Trp Lys Leu His Thr Ser Pro Leu
                325                 330                 335

Cys Thr Thr Asn Ile Lys Glu Gly Ser Asn Ile Cys Leu Thr Arg Thr
            340                 345                 350

Asp Arg Gly Trp Tyr Cys Asp Asn Ala Gly Ser Val Ser Phe Phe Pro
        355                 360                 365

Gln Ala Asp Thr Cys Lys Val Gln Ser Asn Arg Val Phe Cys Asp Thr
    370                 375                 380

Met Asn Ser Leu Thr Leu Pro Ser Glu Val Ser Leu Cys Asn Thr Asp
385                 390                 395                 400

Ile Phe Asn Ser Lys Tyr Asp Cys Lys Ile Met Thr Ser Lys Thr Asp
                405                 410                 415
```

Ile Ser Ser Ser Val Ile Thr Ser Leu Gly Ala Ile Val Ser Cys Tyr
420 425 430

Gly Lys Thr Lys Cys Thr Ala Ser Asn Lys Asn Arg Gly Ile Ile Lys
435 440 445

Thr Phe Ser Asn Gly Cys Asp Tyr Val Ser Asn Lys Gly Val Asp Thr
450 455 460

Val Ser Val Gly Asn Thr Leu Tyr Tyr Val Asn Lys Leu Glu Gly Lys
465 470 475 480

Asn Leu Tyr Val Lys Gly Glu Pro Ile Ile Asn Tyr Tyr Asp Pro Leu
485 490 495

Val Phe Pro Ser Asp Glu Phe Asp Ala Ser Ile Ser Gln Val Asn Glu
500 505 510

Lys Ile Asn Gln Ser Leu Ala Phe Ile Arg Arg Ser Asp Glu Leu Leu
515 520 525

His Asn Val Asn Thr Gly Lys Ser Thr Thr Asn Ile Met Ile Thr Ala
530 535 540

Ile Ile Ile Val Ile Ile Val Val Leu Leu Ser Leu Ile Ala Ile Gly
545 550 555 560

Leu Leu Leu Tyr Cys Lys Ala Lys Asn Thr Pro Val Thr Leu Ser Lys
565 570 575

Asp Gln Leu Ser Gly Ile Asn Asn Ile Ala Phe Ser Lys
580 585

<210> 2
<211> 553
<212> PRT
<213> Newcastle disease virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 2

Met Asp Pro Lys Pro Ser Thr Ser Tyr Leu His Ala Phe Pro Leu Ile
1 5 10 15

Phe Val Ala Ile Ser Leu Val Phe Met Ala Gly Arg Ala Ser Ala Leu
20 25 30

Asp Gly Arg Pro Leu Ala Ala Ala Gly Ile Val Val Thr Gly Asp Lys
35 40 45

Ala Val Asn Ile Tyr Thr Ser Ser Gln Thr Gly Thr Ile Ile Ile Lys

50 55 60

Leu Leu Pro Asn Met Pro Lys Asp Lys Glu Gln Cys Ala Lys Ser Pro
65 70 75 80

Leu Asp Ala Tyr Asn Arg Thr Leu Thr Thr Leu Leu Ala Pro Leu Gly
85 90 95

Asp Ser Ile Arg Arg Ile Gln Glu Ser Val Thr Thr Ser Gly Gly Glu
100 105 110

Arg Gln Glu Arg Leu Val Gly Ala Ile Ile Gly Gly Val Ala Leu Gly
115 120 125

Val Ala Thr Ala Ala Gln Ile Thr Ala Ala Ser Ala Leu Ile Gln Ala
130 135 140

Asn Gln Asn Ala Ala Asn Ile Leu Lys Leu Lys Glu Ser Ile Ala Ala
145 150 155 160

Thr Asn Glu Ala Val His Glu Val Thr Ser Gly Leu Ser Gln Leu Ala
165 170 175

Val Ala Val Gly Lys Met Gln Gln Phe Val Asn Asp Gln Phe Asn Lys
180 185 190

Thr Ala Gln Glu Ile Asp Cys Ile Lys Ile Thr Gln Gln Val Gly Val
195 200 205

Glu Leu Asn Leu Tyr Leu Thr Glu Leu Thr Thr Val Phe Gly Pro Gln
210 215 220

Ile Thr Ser Pro Ala Leu Thr Gln Leu Thr Ile Gln Ala Leu Tyr Asn
225 230 235 240

Leu Ala Gly Gly Asn Met Asp Tyr Met Leu Thr Lys Leu Gly Val Gly
245 250 255

Asn Asn Gln Leu Ser Ser Leu Ile Ser Ser Gly Leu Ile Ser Gly Asn
260 265 270

Pro Ile Leu Tyr Asp Ser Gln Thr Gln Leu Leu Gly Ile Gln Val Thr
275 280 285

Leu Pro Ser Val Gly Asn Leu Asn Asn Met Arg Ala Thr Tyr Leu Glu
290 295 300

Thr Leu Ser Val Ser Thr Asn Lys Gly Phe Ala Ser Ala Leu Val Pro
305 310 315 320

Lys Val Val Thr Gln Val Gly Ser Val Ile Glu Glu Leu Asp Thr Ser
325 330 335

```
Tyr Cys Ile Glu Thr Asp Leu Asp Leu Tyr Cys Thr Arg Ile Val Thr
            340                 345                 350

Phe Pro Met Ser Pro Gly Ile Phe Ser Cys Leu Gly Gly Asn Thr Ser
        355                 360                 365

Ala Cys Met Tyr Ser Lys Thr Glu Gly Ala Leu Thr Thr Pro Tyr Met
    370                 375                 380

Thr Leu Lys Gly Ser Val Ile Ala Asn Cys Lys Met Thr Thr Cys Arg
385                 390                 395                 400

Cys Ala Asp Pro Pro Gly Ile Ile Ser Gln Asn Tyr Gly Glu Ala Val
                405                 410                 415

Ser Leu Ile Asp Lys Lys Val Cys Asn Ile Leu Thr Leu Asp Gly Ile
            420                 425                 430

Thr Leu Arg Leu Ser Gly Glu Phe Asp Ala Thr Tyr Gln Lys Asn Ile
        435                 440                 445

Ser Ile Gln Asp Ser Gln Val Val Ile Thr Gly Asn Leu Asp Ile Ser
    450                 455                 460

Thr Glu Leu Gly Asn Val Asn Asn Ser Ile Ser Asn Ala Leu Asp Lys
465                 470                 475                 480

Leu Glu Glu Ser Asn Ser Lys Leu Asp Lys Val Asn Val Arg Leu Thr
                485                 490                 495

Ser Thr Ser Ala Leu Ile Thr Tyr Ile Val Leu Thr Thr Ile Ala Leu
            500                 505                 510

Ile Cys Gly Ile Val Ser Leu Val Leu Ala Cys Tyr Ile Met Tyr Lys
        515                 520                 525

Gln Lys Ala Gln Gln Lys Thr Leu Leu Trp Leu Gly Asn Asn Thr Leu
    530                 535                 540

Asp Gln Met Arg Ala Thr Thr Lys Met
545                 550


<210>  3
<211>  553
<212>  PRT
<213>  measels virus strain AIK-C


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  3
```

```
Met Ser Ile Met Gly Leu Lys Val Asn Val Ser Ala Ile Phe Met Ala
1               5                   10                  15

Val Leu Leu Thr Leu Gln Thr Pro Thr Gly Gln Ile His Trp Gly Asn
            20                  25                  30

Leu Ser Lys Ile Gly Val Val Gly Ile Gly Ser Ala Ser Tyr Lys Val
        35                  40                  45

Met Thr Arg Ser Ser His Gln Ser Leu Val Ile Lys Leu Met Pro Asn
    50                  55                  60

Ile Thr Leu Leu Asn Asn Cys Thr Arg Val Glu Ile Ala Glu Tyr Arg
65                  70                  75                  80

Arg Leu Leu Arg Thr Val Leu Glu Pro Ile Arg Asp Ala Leu Asn Ala
                85                  90                  95

Met Thr Gln Asn Ile Arg Pro Val Gln Ser Val Ala Ser Ser Arg Arg
            100                 105                 110

His Lys Arg Phe Ala Gly Val Val Leu Ala Gly Ala Ala Leu Gly Val
        115                 120                 125

Ala Thr Ala Ala Gln Ile Thr Ala Gly Ile Ala Leu His Gln Ser Met
    130                 135                 140

Leu Asn Ser Gln Ala Ile Asp Asn Leu Arg Ala Ser Leu Glu Thr Thr
145                 150                 155                 160

Asn Gln Ala Ile Glu Ala Ile Arg Gln Ala Gly Gln Glu Met Ile Leu
                165                 170                 175

Ala Val Gln Gly Val Gln Asp Tyr Ile Asn Asn Glu Leu Ile Pro Ser
            180                 185                 190

Met Asn Gln Leu Ser Cys Asp Leu Ile Gly Gln Lys Leu Gly Leu Lys
        195                 200                 205

Leu Leu Arg Tyr Tyr Thr Glu Ile Leu Ser Leu Phe Gly Pro Ser Leu
    210                 215                 220

Arg Asp Pro Ile Ser Ala Glu Ile Ser Ile Gln Ala Leu Ser Tyr Ala
225                 230                 235                 240

Leu Gly Gly Asp Ile Asn Lys Val Leu Glu Lys Leu Gly Tyr Ser Gly
                245                 250                 255

Gly Asp Leu Leu Gly Ile Leu Glu Ser Arg Gly Ile Lys Ala Arg Ile
            260                 265                 270
```

```
Thr His Val Asp Thr Glu Ser Tyr Phe Ile Val Leu Ser Ile Ala Tyr
        275                 280                 285

Pro Thr Leu Ser Glu Ile Lys Gly Val Ile Val His Arg Leu Glu Gly
    290                 295                 300

Val Ser Tyr Asn Ile Gly Ser Gln Glu Trp Tyr Thr Thr Val Pro Lys
305                 310                 315                 320

Tyr Val Ala Thr Gln Gly Tyr Leu Ile Ser Asn Phe Asp Glu Ser Ser
                325                 330                 335

Cys Thr Phe Met Pro Glu Gly Thr Val Cys Ser Gln Asn Ala Leu Tyr
            340                 345                 350

Pro Met Ser Pro Leu Leu Gln Glu Cys Leu Arg Gly Tyr Thr Lys Ser
        355                 360                 365

Cys Ala Arg Thr Leu Val Ser Gly Ser Phe Gly Asn Arg Phe Ile Leu
    370                 375                 380

Ser Gln Gly Asn Leu Ile Ala Asn Cys Ala Ser Ile Leu Cys Lys Cys
385                 390                 395                 400

Tyr Thr Thr Gly Thr Ile Ile Asn Gln Asp Pro Asp Lys Ile Leu Thr
                405                 410                 415

Tyr Ile Ala Ala Asp Asn Cys Pro Val Val Glu Val Asn Gly Val Thr
            420                 425                 430

Ile Gln Val Gly Ser Arg Arg Tyr Pro Asp Ala Val Tyr Leu His Arg
        435                 440                 445

Ile Asp Leu Gly Pro Pro Ile Leu Leu Glu Arg Leu Asp Val Gly Thr
    450                 455                 460

Asn Leu Gly Asn Ala Ile Ala Lys Leu Glu Asp Ala Lys Glu Leu Leu
465                 470                 475                 480

Glu Ser Ser Asp Gln Ile Leu Arg Ser Met Lys Gly Leu Ser Ser Thr
                485                 490                 495

Cys Ile Val Tyr Ile Leu Ile Ala Val Cys Leu Gly Gly Leu Ile Gly
            500                 505                 510

Ile Pro Ala Leu Ile Cys Cys Cys Arg Gly Arg Cys Asn Lys Lys Gly
        515                 520                 525

Glu Gln Val Gly Met Ser Arg Pro Gly Leu Lys Pro Asp Leu Thr Gly
    530                 535                 540

Thr Ser Lys Ser Tyr Val Arg Ser Leu
```

545 550

<210> 4
<211> 539
<212> PRT
<213> Human metapneumovirus

<220>
<221> MISC_FEATURE
<223> fusion glycoprotein precursor

<400> 4

Met Ser Trp Lys Val Val Ile Ile Phe Ser Leu Leu Ile Thr Pro Gln
1 5 10 15

His Gly Leu Lys Glu Ser Tyr Leu Glu Glu Ser Cys Ser Thr Ile Thr
20 25 30

Glu Gly Tyr Leu Ser Val Leu Arg Thr Gly Trp Tyr Thr Asn Val Phe
35 40 45

Thr Leu Glu Val Gly Asp Val Glu Asn Leu Thr Cys Ser Asp Gly Pro
50 55 60

Ser Leu Ile Lys Thr Glu Leu Asp Leu Thr Lys Ser Ala Leu Arg Glu
65 70 75 80

Leu Lys Thr Val Ser Ala Asp Gln Leu Ala Arg Glu Glu Gln Ile Glu
85 90 95

Lys Pro Arg Gln Ser Arg Phe Val Leu Gly Ala Ile Ala Leu Gly Val
100 105 110

Ala Thr Ala Ala Ala Val Thr Ala Gly Val Ala Ile Ala Lys Thr Ile
115 120 125

Arg Leu Glu Ser Glu Val Thr Ala Ile Lys Asn Ala Leu Lys Thr Thr
130 135 140

Asn Glu Ala Val Ser Thr Leu Gly Asn Gly Val Arg Val Leu Ala Thr
145 150 155 160

Ala Val Arg Glu Leu Lys Asp Phe Val Ser Lys Asn Leu Thr Arg Ala
165 170 175

Ile Asn Lys Asn Lys Cys Asp Ile Asp Asp Leu Lys Met Ala Val Ser
180 185 190

Phe Ser Gln Phe Asn Arg Arg Phe Leu Asn Val Val Arg Gln Phe Ser
195 200 205

Asp Asn Ala Gly Ile Thr Pro Ala Ile Ser Leu Asp Leu Met Thr Asp
210 215 220

```
Ala Glu Leu Ala Arg Ala Val Ser Asn Met Pro Thr Ser Ala Gly Gln
225                 230                 235                 240

Ile Lys Leu Met Leu Glu Asn Arg Ala Met Val Arg Arg Lys Gly Phe
                245                 250                 255

Gly Ile Leu Ile Gly Val Tyr Gly Ser Ser Val Ile Tyr Met Val Gln
            260                 265                 270

Leu Pro Ile Phe Gly Val Ile Asp Thr Pro Cys Trp Ile Val Lys Ala
        275                 280                 285

Ala Pro Ser Cys Ser Glu Lys Lys Gly Asn Tyr Ala Cys Leu Leu Arg
    290                 295                 300

Glu Asp Gln Gly Trp Tyr Cys Gln Asn Ala Gly Ser Thr Val Tyr Tyr
305                 310                 315                 320

Pro Asn Glu Lys Asp Cys Glu Thr Arg Gly Asp His Val Phe Cys Asp
                325                 330                 335

Thr Ala Ala Gly Ile Asn Val Ala Glu Gln Ser Lys Glu Cys Asn Ile
            340                 345                 350

Asn Ile Ser Thr Thr Asn Tyr Pro Cys Lys Val Ser Thr Gly Arg His
        355                 360                 365

Pro Ile Ser Met Val Ala Leu Ser Pro Leu Gly Ala Leu Val Ala Cys
    370                 375                 380

Tyr Lys Gly Val Ser Cys Ser Ile Gly Ser Asn Arg Val Gly Ile Ile
385                 390                 395                 400

Lys Gln Leu Asn Lys Gly Cys Ser Tyr Ile Thr Asn Gln Asp Ala Asp
                405                 410                 415

Thr Val Thr Ile Asp Asn Thr Val Tyr Gln Leu Ser Lys Val Glu Gly
            420                 425                 430

Glu Gln His Val Ile Lys Gly Arg Pro Val Ser Ser Ser Phe Asp Pro
        435                 440                 445

Ile Lys Phe Pro Glu Asp Gln Phe Asn Val Ala Leu Asp Gln Val Phe
    450                 455                 460

Glu Asn Ile Glu Asn Ser Gln Ala Leu Val Asp Gln Ser Asn Arg Ile
465                 470                 475                 480

Leu Ser Ser Ala Glu Lys Gly Asn Thr Gly Phe Ile Ile Val Ile Ile
                485                 490                 495
```

Leu Ile Ala Val Leu Gly Ser Ser Met Ile Leu Val Ser Ile Phe Ile
500 505 510

Ile Ile Lys Lys Thr Lys Lys Pro Thr Gly Ala Pro Pro Glu Leu Ser
515 520 525

Gly Val Thr Asn Asn Gly Phe Ile Pro His Ser
530 535

<210> 5
<211> 540
<212> PRT
<213> Bovine parainfluenza virus 3

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 5

Met Ile Ile Ile Val Ile Thr Met Ile Leu Ser Leu Thr Pro Ser Ser
1 5 10 15

Leu Cys Gln Ile Asp Ile Thr Lys Leu Gln Ser Val Gly Val Leu Val
20 25 30

Asn Ser Pro Lys Gly Ile Lys Ile Ser Gln Asn Phe Glu Thr Arg Tyr
35 40 45

Leu Ile Leu Ser Leu Ile Pro Lys Ile Glu Asp Ser His Ser Cys Gly
50 55 60

Asn Gln Gln Ile Asp Gln Tyr Lys Lys Leu Leu Asp Arg Leu Ile Ile
65 70 75 80

Pro Leu Tyr Asp Gly Leu Lys Leu Gln Lys Asp Val Ile Val Val Asn
85 90 95

His Glu Ser His Asn Asn Thr Asn Leu Arg Thr Lys Arg Phe Phe Gly
100 105 110

Glu Ile Ile Gly Thr Ile Ala Ile Gly Ile Ala Thr Ser Ala Gln Ile
115 120 125

Thr Ala Ala Val Ala Leu Val Glu Ala Lys Gln Ala Arg Ser Asp Ile
130 135 140

Asp Lys Leu Lys Glu Ala Ile Lys Asp Thr Asn Lys Ala Val Gln Ser
145 150 155 160

Ile Gln Ser Ser Val Gly Asn Leu Ile Val Ala Val Lys Ser Val Gln
165 170 175

Asp Tyr Val Asn Asn Glu Ile Val Pro Ser Ile Thr Arg Leu Gly Cys
180 185 190

Glu Ala Ala Gly Leu Gln Leu Gly Ile Ala Leu Thr Gln His Tyr Ser
195 200 205

Glu Leu Thr Asn Ile Phe Gly Asp Asn Ile Gly Thr Leu Gly Glu Lys
210 215 220

Gly Val Lys Leu Gln Gly Ile Ala Ser Leu Tyr Arg Thr Asn Ile Thr
225 230 235 240

Glu Val Phe Thr Thr Ser Thr Val Asp Gln Tyr Asp Ile Tyr Asp Leu
245 250 255

Leu Phe Thr Glu Ser Ile Lys Met Arg Val Ile Asp Val Asp Leu Ser
260 265 270

Asp Tyr Ser Ile Thr Leu Gln Val Arg Leu Pro Leu Leu Thr Lys Val
275 280 285

Ser Asn Thr Gln Ile Tyr Lys Val Asp Ser Ile Ser Tyr Asn Ile Gln
290 295 300

Gly Lys Glu Trp Tyr Ile Pro Leu Pro His His Ile Met Thr Lys Gly
305 310 315 320

Ala Phe Leu Gly Gly Ala Asp Ile Lys Glu Cys Ile Glu Ser Phe Ser
325 330 335

Asn Tyr Ile Cys Pro Ser Asp Pro Gly Phe Ile Leu Asn His Glu Met
340 345 350

Glu Asn Cys Leu Ser Gly Asn Ile Thr Gln Cys Pro Lys Thr Ile Val
355 360 365

Thr Ser Asp Ile Val Pro Arg Tyr Ala Phe Val Asp Gly Gly Val Ile
370 375 380

Ala Asn Cys Ile Pro Thr Thr Cys Thr Cys Asn Gly Ile Asp Asn Arg
385 390 395 400

Ile Asn Gln Ser Pro Asp Gln Gly Ile Lys Ile Ile Thr Tyr Lys Glu
405 410 415

Cys Gln Ile Val Gly Ile Asn Gly Met Leu Phe Lys Thr Asn Gln Glu
420 425 430

Gly Thr Leu Ala Lys Tyr Thr Phe Asp Asn Ile Lys Leu Asn Asn Ser
435 440 445

Val Ala Leu Asn Pro Ile Asp Ile Ser Leu Glu Leu Asn Lys Ala Lys

```
    450                 455                 460

Ser Asp Leu Glu Glu Ser Lys Arg Trp Ile Glu Lys Ser Asn Gln Lys
465                 470                 475                 480

Leu Asp Ser Ile Gly Ser Trp His Gln Ser Ser Val Thr Ile Ile Ile
                485                 490                 495

Ile Ile Val Met Ile Val Val Leu Leu Ile Ile Asn Ala Ile Ile Ile
            500                 505                 510

Met Ile Met Ile Arg Tyr Leu Arg Asp Arg Asn Arg His Leu Asn Asn
        515                 520                 525

Lys Asp Ser Glu Pro Tyr Val Leu Thr Asn Arg Gln
    530                 535                 540


<210>  6
<211>  538
<212>  PRT
<213>  mumps virus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  6

Met Lys Val Phe Leu Val Thr Cys Leu Gly Phe Ala Val Phe Ser Ser
1               5                   10                  15

Ser Val Cys Val Asn Ile Asn Ile Leu Gln Gln Ile Gly Tyr Ile Lys
            20                  25                  30

Gln Gln Val Arg Gln Leu Ser Tyr Tyr Ser Gln Ser Ser Ser Ser Tyr
        35                  40                  45

Ile Val Val Lys Leu Leu Pro Asn Ile Gln Pro Thr Asp Asn Ser Cys
    50                  55                  60

Glu Phe Lys Ser Val Thr Gln Tyr Asn Lys Thr Leu Ser Asn Leu Leu
65                  70                  75                  80

Leu Pro Ile Ala Glu Asn Ile Asn Asn Ile Ala Ser Pro Ser Ser Gly
                85                  90                  95

Ser Arg Arg His Lys Arg Phe Ala Gly Ile Ala Ile Gly Ile Ala Ala
            100                 105                 110

Leu Gly Val Ala Thr Ala Ala Gln Val Thr Ala Ala Val Ser Leu Val
        115                 120                 125

Gln Ala Gln Thr Asn Ala Arg Ala Ile Ala Ala Met Lys Asn Ser Ile
    130                 135                 140
```

```
Gln Ala Thr Asn Arg Ala Val Phe Glu Val Lys Glu Gly Thr Gln Arg
145                 150                 155                 160

Leu Ala Ile Ala Val Gln Ala Ile Gln Asp His Ile Asn Thr Ile Met
                165                 170                 175

Asn Thr Gln Leu Asn Asn Met Ser Cys Gln Ile Leu Asp Asn Gln Leu
            180                 185                 190

Ala Thr Ser Leu Gly Leu Tyr Leu Thr Glu Leu Thr Thr Val Phe Gln
        195                 200                 205

Pro Gln Leu Ile Asn Pro Ala Leu Ser Pro Ile Ser Ile Gln Ala Leu
    210                 215                 220

Arg Ser Leu Leu Gly Ser Met Thr Pro Ala Val Val Gln Ala Thr Leu
225                 230                 235                 240

Ser Thr Ser Ile Ser Ala Ala Glu Ile Leu Ser Ala Gly Leu Met Glu
                245                 250                 255

Gly Gln Ile Val Ser Val Leu Leu Asp Glu Met Gln Met Ile Val Lys
            260                 265                 270

Ile Asn Ile Pro Thr Ile Val Thr Gln Ser Asn Ala Leu Val Ile Asp
        275                 280                 285

Phe Tyr Ser Ile Ser Ser Phe Ile Asn Asn Gln Glu Ser Ile Ile Gln
    290                 295                 300

Leu Pro Asp Arg Ile Leu Glu Ile Gly Asn Glu Gln Trp Ser Tyr Pro
305                 310                 315                 320

Ala Lys Asn Cys Lys Leu Thr Arg His His Ile Phe Cys Gln Tyr Asn
                325                 330                 335

Glu Ala Glu Arg Leu Ser Leu Glu Ser Lys Leu Cys Leu Ala Gly Asn
            340                 345                 350

Ile Ser Ala Cys Val Phe Ser Pro Ile Ala Gly Ser Tyr Met Arg Arg
        355                 360                 365

Phe Val Ala Leu Asp Gly Thr Ile Val Ala Asn Cys Arg Ser Leu Thr
    370                 375                 380

Cys Leu Cys Lys Ser Pro Ser Tyr Pro Ile Tyr Gln Pro Asp His His
385                 390                 395                 400

Ala Val Thr Thr Ile Asp Leu Thr Ala Cys Gln Thr Leu Ser Leu Asp
                405                 410                 415
```

Gly Leu Asp Phe Ser Ile Val Ser Leu Ser Asn Ile Thr Tyr Ala Glu
420 425 430

Asn Leu Thr Ile Ser Leu Ser Gln Thr Ile Asn Thr Gln Pro Ile Asp
435 440 445

Ile Ser Thr Glu Leu Ser Lys Val Asn Ala Ser Leu Gln Asn Ala Val
450 455 460

Lys Tyr Ile Lys Glu Ser Asn His Gln Leu Gln Ser Val Asn Val Asn
465 470 475 480

Ser Lys Ile Gly Ala Ile Ile Val Ala Ala Leu Val Leu Ser Ile Leu
485 490 495

Ser Ile Ile Ile Ser Leu Leu Phe Cys Cys Trp Ala Tyr Val Ala Thr
500 505 510

Lys Glu Ile Arg Arg Ile Asn Phe Lys Thr Asn His Ile Asn Thr Ile
515 520 525

Ser Ser Ser Val Asp Asp Leu Ile Arg Tyr
530 535

<210> 7
<211> 671
<212> PRT
<213> canine distemper virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 7

Met Asn Pro His Glu Gln Thr Ile Pro Met His Glu Lys Ile Pro Lys
1 5 10 15

Arg Ser Lys Thr Gln Thr His Thr Gln Gln Asp Leu Pro Gln Gln His
20 25 30

Ser Thr Lys Ser Ala Glu Ser Lys Thr Ser Arg Ala Arg His Ser Ile
35 40 45

Thr Ser Ala Gln Arg Ser Thr His Tyr Asp Pro Arg Thr Ala Asp Trp
50 55 60

Pro Asp Tyr Tyr Ile Met Lys Arg Thr Arg Ser Cys Lys Gln Ala Ser
65 70 75 80

Tyr Arg Ser Asp Asn Ile Pro Ala His Gly Asp His Asp Gly Ile Ile
85 90 95

```
His His Thr Pro Glu Ser Val Ser Gln Gly Ala Lys Ser Arg Leu Lys
            100                 105                 110

Met Gly Gln Ser Asn Ala Val Lys Ser Gly Ser Gln Cys Thr Trp Leu
        115                 120                 125

Val Leu Trp Cys Ile Gly Val Ala Ser Leu Phe Leu Cys Ser Lys Ala
    130                 135                 140

Gln Ile His Trp Asn Asn Leu Ser Thr Ile Gly Ile Ile Gly Thr Asp
145                 150                 155                 160

Ser Val His Tyr Lys Ile Met Thr Arg Pro Ser His Gln Tyr Leu Val
                165                 170                 175

Ile Lys Leu Met Pro Asn Val Ser Leu Ile Asp Asn Cys Thr Lys Ala
            180                 185                 190

Glu Leu Asp Glu Tyr Glu Lys Leu Leu Ser Ser Ile Leu Glu Pro Ile
        195                 200                 205

Asn Gln Ala Leu Thr Leu Met Thr Lys Asn Val Lys Pro Leu Gln Ser
    210                 215                 220

Val Gly Ser Gly Arg Arg Gln Arg Arg Phe Ala Gly Val Val Leu Ala
225                 230                 235                 240

Gly Ala Ala Leu Gly Val Ala Thr Ala Ala Gln Ile Thr Ala Gly Ile
                245                 250                 255

Ala Leu His Gln Ser Asn Leu Asn Ala Gln Ala Ile Gln Ser Leu Arg
            260                 265                 270

Thr Ser Leu Glu Gln Ser Asn Lys Ala Ile Glu Glu Ile Arg Glu Ala
        275                 280                 285

Thr Gln Glu Thr Val Ile Ala Val Gln Gly Val Gln Asp Tyr Val Asn
    290                 295                 300

Asn Glu Leu Val Pro Ala Met Gln His Met Ser Cys Glu Leu Val Gly
305                 310                 315                 320

Gln Arg Leu Gly Leu Lys Leu Leu Arg Tyr Tyr Thr Glu Leu Leu Ser
                325                 330                 335

Ile Phe Gly Pro Ser Leu Arg Asp Pro Ile Ser Ala Glu Ile Ser Ile
            340                 345                 350

Gln Ala Leu Ser Tyr Ala Leu Gly Gly Glu Ile His Lys Ile Leu Glu
        355                 360                 365

Lys Leu Gly Tyr Ser Gly Asn Asp Met Ile Ala Ile Leu Glu Ser Arg
```

370 375 380

Gly Ile Lys Thr Lys Ile Thr His Val Asp Leu Pro Gly Lys Phe Ile
385 390 395 400

Ile Leu Ser Val Ser Tyr Pro Thr Leu Ser Glu Val Lys Gly Val Ile
405 410 415

Val His Arg Leu Glu Ala Val Ser Tyr Asn Ile Gly Ser Gln Glu Trp
420 425 430

Tyr Thr Thr Val Pro Arg Tyr Val Ala Thr Asn Gly Tyr Leu Ile Ser
435 440 445

Asn Phe Asp Glu Ser Ser Cys Val Phe Val Ser Glu Ser Ala Ile Cys
450 455 460

Ser Gln Asn Ser Leu Tyr Pro Met Ser Pro Leu Leu Gln Gln Cys Ile
465 470 475 480

Arg Gly Asp Thr Ser Ser Cys Ala Arg Thr Leu Val Ser Gly Thr Met
485 490 495

Gly Asn Lys Phe Ile Leu Ser Lys Gly Asn Ile Val Ala Asn Cys Ala
500 505 510

Ser Ile Leu Cys Lys Cys Tyr Ser Thr Ser Thr Ile Ile Asn Gln Ser
515 520 525

Pro Asp Lys Leu Leu Thr Phe Ile Ala Ser Asp Thr Cys Pro Leu Val
530 535 540

Glu Ile Asp Gly Val Thr Ile Gln Val Gly Ser Arg Gln Tyr Pro Asp
545 550 555 560

Met Val Tyr Glu Ser Lys Val Ala Leu Gly Pro Ala Ile Ser Leu Glu
565 570 575

Arg Leu Asp Val Gly Thr Asn Leu Gly Asn Ala Leu Lys Lys Leu Asp
580 585 590

Asp Ala Lys Val Leu Ile Asp Ser Ser Asn Gln Ile Leu Glu Thr Val
595 600 605

Arg Arg Ser Ser Phe Asn Phe Gly Ser Leu Leu Ser Val Pro Ile Leu
610 615 620

Ser Cys Thr Ala Leu Ala Leu Leu Leu Leu Ile Cys Cys Cys Lys Arg
625 630 635 640

Arg Tyr Gln Gln Thr His Lys Gln Asn Thr Lys Val Asp Pro Thr Phe
645 650 655

```
Lys Pro Asp Leu Thr Gly Thr Ser Arg Ser Tyr Val Arg Ser Leu
            660                 665                 670


<210>  8
<211>  546
<212>  PRT
<213>  Peste-des-petits-ruminants virus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  8

Met Thr Arg Val Ala Ile Leu Thr Phe Leu Phe Leu Phe Pro Asn Ala
1               5                   10                  15


Val Ala Cys Gln Ile His Trp Gly Asn Leu Ser Lys Ile Gly Ile Val
            20                  25                  30


Gly Thr Gly Ser Ala Ser Tyr Lys Val Met Thr Arg Pro Ser His Gln
        35                  40                  45


Thr Leu Val Ile Lys Leu Met Pro Asn Ile Thr Ala Ile Asp Asn Cys
    50                  55                  60


Thr Lys Ser Glu Ile Ala Glu Tyr Lys Arg Leu Leu Ile Thr Val Leu
65                  70                  75                  80


Lys Pro Val Glu Asp Ala Leu Ser Val Ile Thr Lys Asn Val Arg Pro
                85                  90                  95


Ile Gln Thr Leu Thr Pro Gly Arg Arg Thr Arg Arg Phe Ala Gly Ala
            100                 105                 110


Val Leu Ala Gly Val Ala Leu Gly Val Ala Thr Ala Ala Gln Ile Thr
        115                 120                 125


Ala Gly Val Ala Leu His Gln Ser Leu Met Asn Ser Gln Ala Ile Glu
    130                 135                 140


Ser Leu Lys Thr Ser Leu Glu Lys Ser Asn Gln Ala Ile Glu Glu Ile
145                 150                 155                 160


Arg Leu Ala Asn Lys Glu Thr Ile Leu Ala Val Gln Gly Val Gln Asp
                165                 170                 175


Tyr Ile Asn Asn Glu Leu Val Pro Ser Val His Arg Met Ser Cys Glu
            180                 185                 190


Leu Val Gly His Lys Leu Gly Leu Lys Leu Leu Arg Tyr Tyr Thr Glu
        195                 200                 205
```

```
Ile Leu Ser Ile Phe Gly Pro Ser Leu Arg Asp Pro Ile Ala Ala Glu
    210                 215                 220

Ile Ser Ile Gln Ala Leu Ser Tyr Ala Leu Gly Gly Asp Ile Asn Arg
225                 230                 235                 240

Ile Leu Asp Lys Leu Gly Tyr Ser Gly Gly Asp Phe Leu Ala Ile Leu
                245                 250                 255

Glu Ser Lys Gly Ile Lys Ala Arg Val Thr Tyr Val Asp Thr Arg Asp
            260                 265                 270

Tyr Phe Ile Ile Leu Ser Ile Ala Tyr Pro Thr Leu Ser Glu Ile Lys
        275                 280                 285

Gly Val Ile Val His Lys Ile Glu Ala Ile Thr Tyr Asn Ile Gly Ala
    290                 295                 300

Gln Glu Trp Tyr Thr Thr Ile Pro Lys Tyr Val Ala Thr Gln Gly Tyr
305                 310                 315                 320

Leu Ile Ser Asn Phe Asp Glu Thr Ser Cys Val Phe Thr Pro Asp Gly
                325                 330                 335

Thr Val Cys Ser Gln Asn Ala Leu Tyr Pro Met Ser Pro Leu Leu Gln
            340                 345                 350

Glu Cys Phe Gln Gly Ser Thr Lys Ser Cys Ala Arg Thr Leu Val Ser
        355                 360                 365

Gly Thr Ile Ser Asn Arg Phe Ile Leu Ser Lys Gly Asn Leu Ile Ala
    370                 375                 380

Asn Cys Ala Ser Val Leu Cys Lys Cys Tyr Thr Thr Glu Thr Val Ile
385                 390                 395                 400

Ser Gln Asp Pro Asp Lys Leu Leu Thr Val Val Ala Ser Asp Lys Cys
                405                 410                 415

Pro Val Val Glu Val Asp Gly Val Thr Ile Gln Val Gly Ser Arg Glu
            420                 425                 430

Tyr Pro Asp Ser Val Tyr Leu His Lys Ile Asp Leu Gly Pro Ala Ile
        435                 440                 445

Ser Leu Glu Lys Leu Asp Val Gly Thr Asn Leu Gly Asn Ala Val Thr
    450                 455                 460

Arg Leu Glu Asn Ala Lys Glu Leu Leu Asp Ala Ser Asp Gln Ile Leu
465                 470                 475                 480
```

Lys Thr Val Lys Gly Val Pro Phe Gly Gly Asn Met Tyr Ile Ala Leu
485 490 495

Ala Ala Cys Ile Gly Val Ser Leu Gly Leu Val Thr Leu Ile Cys Cys
500 505 510

Cys Lys Gly Arg Cys Lys Asn Lys Glu Val Pro Ile Ser Lys Ile Asn
515 520 525

Pro Gly Leu Lys Pro Asp Leu Thr Gly Thr Ser Lys Ser Tyr Val Arg
530 535 540

Ser Leu
545

<210> 9
<211> 546
<212> PRT
<213> Hendra virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 9

Met Ala Thr Gln Glu Val Arg Leu Lys Cys Leu Leu Cys Gly Ile Ile
1 5 10 15

Val Leu Val Leu Ser Leu Glu Gly Leu Gly Ile Leu His Tyr Glu Lys
20 25 30

Leu Ser Lys Ile Gly Leu Val Lys Gly Ile Thr Arg Lys Tyr Lys Ile
35 40 45

Lys Ser Asn Pro Leu Thr Lys Asp Ile Val Ile Lys Met Ile Pro Asn
50 55 60

Val Ser Asn Val Ser Lys Cys Thr Gly Thr Val Met Glu Asn Tyr Lys
65 70 75 80

Ser Arg Leu Thr Gly Ile Leu Ser Pro Ile Lys Gly Ala Ile Glu Leu
85 90 95

Tyr Asn Asn Asn Thr His Asp Leu Val Gly Asp Val Lys Leu Ala Gly
100 105 110

Val Val Met Ala Gly Ile Ala Ile Gly Ile Ala Thr Ala Ala Gln Ile
115 120 125

Thr Ala Gly Val Ala Leu Tyr Glu Ala Met Lys Asn Ala Asp Asn Ile
130 135 140

Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala Val Val Lys

```
145                 150                 155                 160

Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr Ala Leu Gln
                165                 170                 175

Asp Tyr Ile Asn Thr Asn Leu Val Pro Thr Ile Asp Gln Ile Ser Cys
            180                 185                 190

Lys Gln Thr Glu Leu Ala Leu Asp Leu Ala Leu Ser Lys Tyr Leu Ser
        195                 200                 205

Asp Leu Leu Phe Val Phe Gly Pro Asn Leu Gln Asp Pro Val Ser Asn
    210                 215                 220

Ser Met Thr Ile Gln Ala Ile Ser Gln Ala Phe Gly Gly Asn Tyr Glu
225                 230                 235                 240

Thr Leu Leu Arg Thr Leu Gly Tyr Ala Thr Glu Asp Phe Asp Asp Leu
                245                 250                 255

Leu Glu Ser Asp Ser Ile Ala Gly Gln Ile Val Tyr Val Asp Leu Ser
            260                 265                 270

Ser Tyr Tyr Ile Ile Val Arg Val Tyr Phe Pro Ile Leu Thr Glu Ile
        275                 280                 285

Gln Gln Ala Tyr Val Gln Glu Leu Leu Pro Val Ser Phe Asn Asn Asp
    290                 295                 300

Asn Ser Glu Trp Ile Ser Ile Val Pro Asn Phe Val Leu Ile Arg Asn
305                 310                 315                 320

Thr Leu Ile Ser Asn Ile Glu Val Lys Tyr Cys Leu Ile Thr Lys Lys
                325                 330                 335

Ser Val Ile Cys Asn Gln Asp Tyr Ala Thr Pro Met Thr Ala Ser Val
            340                 345                 350

Arg Glu Cys Leu Thr Gly Ser Thr Asp Lys Cys Pro Arg Glu Leu Val
        355                 360                 365

Val Ser Ser His Val Pro Arg Phe Ala Leu Ser Gly Gly Val Leu Phe
    370                 375                 380

Ala Asn Cys Ile Ser Val Thr Cys Gln Cys Gln Thr Thr Gly Arg Ala
385                 390                 395                 400

Ile Ser Gln Ser Gly Glu Gln Thr Leu Leu Met Ile Asp Asn Thr Thr
                405                 410                 415

Cys Thr Thr Val Val Leu Gly Asn Ile Ile Ile Ser Leu Gly Lys Tyr
            420                 425                 430
```

```
Leu Gly Ser Ile Asn Tyr Asn Ser Glu Ser Ile Ala Val Gly Pro Pro
        435                 440                 445


Val Tyr Thr Asp Lys Val Asp Ile Ser Ser Gln Ile Ser Ser Met Asn
    450                 455                 460


Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala Gln Lys Ile
465                 470                 475                 480


Leu Asp Thr Val Asn Pro Ser Leu Ile Ser Met Leu Ser Met Ile Ile
                485                 490                 495


Leu Tyr Val Leu Ser Ile Ala Ala Leu Cys Ile Gly Leu Ile Thr Phe
            500                 505                 510


Ile Ser Phe Val Ile Val Glu Lys Lys Arg Gly Asn Tyr Ser Arg Leu
        515                 520                 525


Asp Asp Arg Gln Val Arg Pro Val Ser Asn Gly Asp Leu Tyr Tyr Ile
    530                 535                 540


Gly Thr
545


<210>  10
<211>  565
<212>  PRT
<213>  Sendai virus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  10

Met Ala Thr Tyr Ile Gln Arg Val Gln Cys Ile Ser Ala Leu Leu Ser
1               5                   10                  15


Val Val Leu Thr Thr Leu Val Ser Cys Gln Ile Pro Arg Asp Arg Leu
            20                  25                  30


Ser Asn Ile Gly Val Ile Val Asp Glu Gly Lys Ser Leu Lys Ile Ala
        35                  40                  45


Gly Ser His Glu Ser Arg Tyr Ile Val Leu Ser Leu Val Pro Gly Ile
    50                  55                  60


Asp Leu Glu Asn Gly Cys Gly Thr Ala Gln Val Ile Gln Tyr Lys Ser
65                  70                  75                  80


Leu Leu Asn Arg Leu Leu Ile Pro Leu Arg Asp Ala Leu Asp Leu Gln
                85                  90                  95
```

```
Glu Ala Leu Ile Thr Val Thr Asn Asp Thr Met Thr Gly Ala Asp Val
            100                 105                 110

Pro Gln Ser Arg Phe Phe Gly Ala Val Ile Gly Thr Ile Ala Leu Gly
        115                 120                 125

Val Ala Thr Ser Ala Gln Ile Thr Ala Gly Ile Ala Leu Ala Glu Ala
    130                 135                 140

Arg Glu Ala Lys Arg Asp Ile Ala Leu Ile Lys Glu Ser Met Thr Lys
145                 150                 155                 160

Thr His Lys Ser Ile Glu Leu Leu Gln Asn Ala Val Gly Glu Gln Ile
                165                 170                 175

Leu Ala Leu Lys Thr Leu Gln Asp Phe Val Asn Asp Glu Ile Lys Pro
            180                 185                 190

Ala Ile Ser Glu Leu Gly Cys Glu Thr Ala Ala Leu Arg Leu Gly Ile
        195                 200                 205

Lys Leu Thr Gln His Tyr Ser Glu Leu Leu Thr Ala Phe Gly Ser Asn
    210                 215                 220

Phe Gly Thr Ile Gly Glu Lys Ser Leu Thr Leu Gln Ala Leu Ser Ser
225                 230                 235                 240

Leu Tyr Ser Ala Asn Ile Thr Glu Ile Met Thr Thr Ile Arg Thr Gly
                245                 250                 255

Gln Ser Asn Ile Tyr Asp Val Ile Tyr Thr Glu Gln Ile Lys Gly Thr
            260                 265                 270

Val Ile Asp Val Asp Leu Glu Arg Tyr Met Val Thr Leu Ser Val Lys
        275                 280                 285

Ile Pro Ile Leu Ser Glu Val Pro Gly Val Leu Ile His Lys Ala Ser
    290                 295                 300

Ser Ile Ser Tyr Asn Ile Asp Gly Glu Glu Trp Tyr Val Thr Val Pro
305                 310                 315                 320

Ser His Ile Leu Ser Arg Ala Ser Phe Leu Gly Gly Ala Asn Ile Ala
                325                 330                 335

Asp Cys Val Glu Ser Arg Leu Thr Tyr Ile Cys Pro Arg Asp Pro Ala
            340                 345                 350

Gln Leu Ile Pro Asp Ser Gln Gln Lys Cys Ile Leu Gly Asp Thr Thr
        355                 360                 365
```

```
Arg Cys Pro Val Thr Lys Val Val Asp Asn Ile Ile Pro Lys Phe Ala
    370                 375                 380


Phe Val Asn Gly Gly Val Val Ala Asn Cys Ile Ala Ser Thr Cys Thr
385                 390                 395                 400


Cys Gly Thr Gly Arg Arg Pro Ile Ser Gln Asp Arg Ser Lys Gly Val
                405                 410                 415


Val Phe Leu Thr His Asp Asn Cys Gly Leu Ile Gly Val Asn Gly Ile
            420                 425                 430


Glu Leu Tyr Ala Asn Arg Lys Gly His Asp Ala Thr Trp Gly Val Gln
        435                 440                 445


Asn Leu Thr Val Gly Pro Ala Ile Ala Ile Arg Pro Val Asp Ile Ser
    450                 455                 460


Leu Asn Leu Ala Ala Ala Thr Asp Phe Leu Gln Asp Ser Arg Ala Glu
465                 470                 475                 480


Leu Glu Lys Ala Arg Lys Ile Leu Ser Glu Val Gly Arg Trp Tyr Asn
                485                 490                 495


Ser Gly Ala Thr Leu Ile Thr Ile Ile Val Val Met Ile Val Val Leu
            500                 505                 510


Val Val Ile Ile Val Ile Val Ile Val Leu Tyr Arg Leu Arg Arg Ser
        515                 520                 525


Met Leu Met Ser Asn Pro Ala Gly Arg Ile Ser Arg Asp Thr Tyr Thr
    530                 535                 540


Leu Glu Pro Lys Ile Arg His Met Tyr Thr Asn Gly Gly Phe Asp Ala
545                 550                 555                 560


Met Thr Glu Lys Arg
                565


<210>  11
<211>  555
<212>  PRT
<213>  Human parainfluenza virus 1 strain Washington/1964


<220>
<221>  MISC_FEATURE
<223>  F glycoprotein

<400>  11

Met Gln Lys Ser Glu Ile Leu Phe Leu Val Tyr Ser Ser Leu Leu Leu
1               5                   10                  15


Ser Ser Ser Leu Cys Gln Ile Pro Val Glu Lys Leu Ser Asn Val Gly
```

20 25 30

Val Ile Ile Asn Glu Gly Lys Leu Leu Lys Ile Ala Gly Ser Tyr Glu
35 40 45

Ser Arg Tyr Ile Val Leu Ser Leu Val Pro Ser Ile Asp Leu Gln Asp
50 55 60

Gly Cys Gly Thr Thr Gln Ile Ile Gln Tyr Lys Asn Leu Leu Asn Arg
65 70 75 80

Leu Leu Ile Pro Leu Lys Asp Ala Leu Asp Leu Gln Glu Ser Leu Ile
85 90 95

Thr Ile Thr Asn Asp Thr Thr Val Thr Asn Asp Asn Pro Gln Thr Arg
100 105 110

Phe Phe Gly Ala Val Ile Gly Thr Ile Ala Leu Gly Val Ala Thr Ala
115 120 125

Ala Gln Ile Thr Ala Gly Ile Ala Leu Ala Glu Ala Arg Glu Ala Arg
130 135 140

Lys Asp Ile Ala Leu Ile Lys Asp Ser Ile Val Lys Thr His Asn Ser
145 150 155 160

Val Glu Leu Ile Gln Arg Gly Ile Gly Glu Gln Ile Ile Ala Leu Lys
165 170 175

Thr Leu Gln Asp Phe Val Asn Asp Glu Ile Arg Pro Ala Ile Gly Glu
180 185 190

Leu Arg Cys Glu Thr Thr Ala Leu Lys Leu Gly Ile Lys Leu Thr Gln
195 200 205

His Tyr Ser Glu Leu Ala Thr Ala Phe Ser Ser Asn Leu Gly Thr Ile
210 215 220

Gly Glu Lys Ser Leu Thr Leu Gln Ala Leu Ser Ser Leu Tyr Ser Ala
225 230 235 240

Asn Ile Thr Glu Ile Leu Ser Thr Thr Lys Lys Asp Lys Ser Asp Ile
245 250 255

Tyr Asp Ile Ile Tyr Thr Glu Gln Val Lys Gly Thr Val Ile Asp Val
260 265 270

Asp Leu Glu Lys Tyr Met Val Thr Leu Leu Val Lys Ile Pro Ile Leu
275 280 285

Ser Glu Ile Pro Gly Val Leu Ile Tyr Arg Ala Ser Ser Ile Ser Tyr
290 295 300

```
Asn Ile Glu Gly Glu Glu Trp His Val Ala Ile Pro Asn Tyr Ile Ile
305                 310                 315                 320

Asn Lys Ala Ser Ser Leu Gly Gly Ala Asp Val Thr Asn Cys Ile Glu
                325                 330                 335

Ser Lys Leu Ala Tyr Ile Cys Pro Arg Asp Pro Thr Gln Leu Ile Pro
            340                 345                 350

Asp Asn Gln Gln Lys Cys Ile Leu Gly Asp Val Ser Lys Cys Pro Val
        355                 360                 365

Thr Lys Val Ile Asn Asn Leu Val Pro Lys Phe Ala Phe Ile Asn Gly
    370                 375                 380

Gly Val Val Ala Asn Cys Ile Ala Ser Thr Cys Thr Cys Gly Thr Asn
385                 390                 395                 400

Arg Ile Pro Val Asn Gln Asp Arg Ser Arg Gly Val Thr Phe Leu Thr
                405                 410                 415

Tyr Thr Asn Cys Gly Leu Ile Gly Ile Asn Gly Ile Glu Leu Tyr Ala
            420                 425                 430

Asn Lys Arg Gly Arg Asp Thr Thr Trp Gly Asn Gln Ile Ile Lys Val
        435                 440                 445

Gly Pro Ala Val Ser Ile Arg Pro Val Asp Ile Ser Leu Asn Leu Ala
    450                 455                 460

Ser Ala Thr Asn Phe Leu Glu Glu Ser Lys Thr Glu Leu Met Lys Ala
465                 470                 475                 480

Arg Ala Ile Ile Ser Ala Val Gly Gly Trp His Asn Thr Glu Ser Thr
                485                 490                 495

Gln Ile Ile Met Ile Ile Ile Val Cys Ile Leu Ile Ile Ile Ile Cys
            500                 505                 510

Gly Ile Leu Tyr Tyr Leu Tyr Arg Val Arg Arg Leu Leu Val Met Ile
        515                 520                 525

Asn Ser Thr His Asn Ser Pro Val Asn Ala Tyr Thr Leu Glu Ser Arg
    530                 535                 540

Met Arg Asn Pro Tyr Met Gly Asn Asn Ser Asn
545                 550                 555

<210>  12
<211>  543
<212>  PRT
```

<213> Feline morbillivirus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 12

```
Met Gly Lys Ile Arg Val Ile Ile Ile Ser Ser Leu Leu Leu Ser Asn
1               5                   10                  15

Ile Thr Thr Ala Gln Val Gly Trp Asp Asn Leu Thr Ser Ile Gly Val
            20                  25                  30

Ile Ser Thr Lys Gln Tyr Asp Tyr Lys Ile Thr Thr Leu Asn Thr Asn
        35                  40                  45

Gln Leu Met Val Ile Lys Met Val Pro Asn Ile Ser Ser Ile Ile Asn
    50                  55                  60

Cys Thr Lys Pro Glu Leu Met Lys Tyr Arg Glu Leu Val Leu Gly Val
65                  70                  75                  80

Ile Arg Pro Ile Asn Glu Ser Leu Glu Leu Met Asn Ser Tyr Ile Asn
                85                  90                  95

Met Arg Ala Gly Ser Glu Arg Phe Ile Gly Ala Val Ile Ala Gly Val
            100                 105                 110

Ala Leu Gly Val Ala Thr Ala Ala Gln Ile Thr Ser Gly Ile Ala Leu
        115                 120                 125

His Asn Ser Ile Met Asn Lys Arg Gln Ile Gln Glu Leu Arg Lys Ala
    130                 135                 140

Leu Ser Thr Thr Asn Lys Ala Ile Asp Glu Ile Arg Ile Ala Gly Glu
145                 150                 155                 160

Arg Thr Leu Ile Ala Val Gln Gly Val Gln Asp Tyr Ile Asn Asn Ile
                165                 170                 175

Ile Ile Pro Met Gln Asp Lys Leu Gln Cys Asp Ile Leu Ser Ser Gln
            180                 185                 190

Leu Ala Ile Ala Leu Leu Arg Tyr Tyr Thr Asn Ile Leu Thr Val Phe
        195                 200                 205

Gly Pro Ser Ile Arg Asp Pro Val Thr Ser Ile Ile Ser Ile Gln Ala
    210                 215                 220

Leu Ser Gln Ala Phe Asn Gly Asn Leu Gln Ala Leu Leu Asp Gly Leu
225                 230                 235                 240
```

```
Gly Tyr Thr Gly Arg Asp Leu Arg Asp Leu Leu Glu Ser Arg Ser Ile
                245                 250                 255

Thr Gly Gln Ile Ile His Ala Asp Met Thr Asp Leu Phe Leu Val Leu
            260                 265                 270

Arg Ile Asn Tyr Pro Ser Ile Thr Glu Met Gln Gly Val Thr Ile Tyr
        275                 280                 285

Glu Leu Asn Ser Ile Thr Tyr His Ile Gly Pro Glu Glu Trp Tyr Thr
    290                 295                 300

Ile Met Pro Asn Phe Ile Ala Val Gln Gly Phe Leu Thr Ser Asn Phe
305                 310                 315                 320

Asp Glu Arg Lys Cys Ser Ile Thr Lys Ser Ser Ile Leu Cys Gln Gln
                325                 330                 335

Asn Ser Ile Tyr Pro Met Ser Thr Glu Met Gln Arg Cys Ile Lys Gly
            340                 345                 350

Glu Ile Arg Phe Cys Pro Arg Ser Lys Ala Val Gly Thr Leu Val Asn
        355                 360                 365

Arg Phe Ile Leu Thr Lys Gly Asn Leu Met Ala Asn Cys Leu Gly Val
    370                 375                 380

Ile Cys Arg Cys Tyr Ser Ser Gly Gln Ile Ile Thr Gln Asp Pro Ser
385                 390                 395                 400

Lys Leu Ile Thr Ile Ile Ser Gln Glu Glu Cys Lys Glu Val Gly Val
                405                 410                 415

Asp Gly Ile Arg Ile Met Val Gly Pro Arg Lys Leu Pro Asp Val Ile
            420                 425                 430

Phe Asn Ala Arg Leu Glu Val Gly Val Pro Ile Ser Leu Ser Lys Leu
        435                 440                 445

Asp Val Gly Thr Asp Leu Ala Ile Ala Ser Ala Lys Leu Asn Asn Ser
    450                 455                 460

Lys Ala Leu Leu Glu Gln Ser Asp Lys Ile Leu Asp Ser Met Ser Lys
465                 470                 475                 480

Leu Asp Ser Ile Asn Ser Arg Ile Thr Gly Leu Ile Leu Ala Ile Met
                485                 490                 495

Ala Ile Phe Ile Ile Thr Val Thr Ile Ile Trp Ile Ile Tyr Lys Arg
            500                 505                 510

Cys Arg Asn Lys Asp Asn Lys Phe Ser Thr Ser Ile Glu Pro Leu Tyr
```

515 520 525

Ile Pro Pro Ser Tyr Asn Ser Pro His Ser Val Val Lys Ser Ile
530 535 540

<210> 13
<211> 534
<212> PRT
<213> Avian paramyxovirus 2

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 13

Met Ile Ala Ala Leu Phe Ile Ser Leu Phe Ala Thr Cys Gly Ala Leu
1 5 10 15

Asp Asn Ser Val Leu Ala Pro Val Gly Ile Ala Ser Ala Gln Glu Trp
20 25 30

Gln Leu Ala Ala Tyr Thr Asn Thr Leu Ser Gly Thr Ile Ala Val Arg
35 40 45

Phe Val Pro Val Leu Pro Gly Asn Leu Ser Thr Cys Ala Gln Ala Thr
50 55 60

Leu Ala Glu Tyr Asn Lys Thr Val Thr Asn Ile Leu Gly Pro Leu Lys
65 70 75 80

Glu Asn Leu Glu Thr Leu Leu Ser Glu Pro Thr Lys Thr Ala Ala Arg
85 90 95

Phe Val Gly Ala Ile Ile Gly Thr Val Ala Leu Gly Val Ala Thr Ser
100 105 110

Ala Gln Ile Thr Ala Ala Val Ala Leu Asn Gln Ala Gln Glu Asn Ala
115 120 125

Arg Asn Ile Trp Arg Leu Lys Glu Ser Ile Arg Lys Thr Asn Glu Ala
130 135 140

Val Leu Glu Leu Lys Asp Gly Leu Ala Ser Thr Ala Ile Ala Leu Asp
145 150 155 160

Lys Val Gln Lys Phe Ile Asn Glu Asp Ile Ile Pro Gln Ile Lys Glu
165 170 175

Ile Asp Cys Gln Val Val Ala Asn Lys Leu Gly Val Tyr Leu Ser Leu
180 185 190

Tyr Leu Thr Glu Leu Thr Thr Ile Phe Gly Ala Gln Ile Thr Asn Pro
195 200 205

```
Ala Leu Thr Pro Leu Ser Tyr Gln Ala Leu Tyr Asn Leu Cys Gly Gly
    210                 215                 220

Asp Met Gly Lys Leu Thr Glu Leu Ile Gly Val Lys Ala Lys Asp Ile
225                 230                 235                 240

Asn Ser Leu Tyr Glu Ala Asn Leu Ile Thr Gly Gln Val Ile Gly Tyr
                245                 250                 255

Asp Ser Glu Ser Gln Ile Ile Leu Ile Gln Val Ser Tyr Pro Ser Val
            260                 265                 270

Ser Glu Val Thr Gly Val Arg Ala Thr Glu Leu Val Thr Val Ser Val
        275                 280                 285

Thr Thr Pro Lys Gly Glu Gly Arg Ala Ile Ala Pro Lys Tyr Val Ala
    290                 295                 300

Gln Ser Arg Val Val Thr Glu Glu Leu Asp Thr Ser Thr Cys Arg Phe
305                 310                 315                 320

Ser Lys Thr Thr Leu Tyr Cys Arg Ser Ile Ile Thr Arg Pro Leu Pro
                325                 330                 335

Pro Leu Ile Ala Asn Cys Leu Asn Gly Leu Tyr Gln Asp Cys Gln Tyr
            340                 345                 350

Thr Thr Glu Ile Gly Ala Leu Ser Ser Arg Phe Ile Thr Val Asn Gly
        355                 360                 365

Gly Ile Ile Ala Asn Cys Arg Ala Thr Ile Cys Lys Cys Val Asn Pro
    370                 375                 380

Pro Lys Ile Ile Val Gln Ser Asp Ala Ser Ser Leu Thr Val Ile Asp
385                 390                 395                 400

Ser Ala Ile Cys Lys Asp Val Val Leu Asp Asn Val Gln Leu Arg Leu
                405                 410                 415

Glu Gly Lys Leu Ser Ala Gln Tyr Phe Thr Asn Ile Thr Ile Asp Leu
            420                 425                 430

Ser Gln Ile Thr Thr Ser Gly Ser Leu Asp Ile Ser Ser Glu Ile Gly
        435                 440                 445

Ser Ile Asn Asn Thr Val Asn Lys Val Glu Glu Leu Ile Ala Glu Ser
    450                 455                 460

Asn Ala Trp Leu Gln Ala Val Asn Pro His Leu Val Asn Asn Thr Ser
465                 470                 475                 480
```

Ile Ile Val Leu Cys Val Leu Ala Ala Ile Phe Val Val Trp Leu Val
485 490 495

Ala Leu Thr Gly Cys Leu Ala Tyr Tyr Ile Lys Lys Ser Ser Ala Thr
500 505 510

Arg Met Val Gly Ile Gly Ser Ser Pro Ala Gly Asn Pro Tyr Val Ala
515 520 525

Gln Ser Ala Thr Lys Met
530

<210> 14
<211> 555
<212> PRT
<213> avian paramyxovirus 6

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 14

Met Gly Ala Arg Leu Gly Pro Leu Ala Met Ala Pro Gly Arg Tyr Val
1 5 10 15

Ile Ile Phe Asn Leu Ile Leu Leu His Arg Val Val Ser Leu Asp Asn
20 25 30

Ser Arg Leu Leu Gln Gln Gly Ile Met Ser Ala Thr Glu Arg Glu Ile
35 40 45

Lys Val Tyr Thr Asn Ser Ile Thr Gly Ser Ile Ala Val Arg Leu Ile
50 55 60

Pro Asn Leu Pro Gln Glu Val Leu Lys Cys Ser Ala Gly Gln Ile Lys
65 70 75 80

Ser Tyr Asn Asp Thr Leu Asn Arg Ile Phe Thr Pro Ile Lys Ala Asn
85 90 95

Leu Glu Arg Leu Leu Ala Thr Pro Ser Met Leu Glu Asp Asn Gln Asn
100 105 110

Pro Ala Pro Glu Pro Arg Leu Ile Gly Ala Ile Ile Gly Thr Ala Ala
115 120 125

Leu Gly Leu Ala Thr Ala Ala Gln Val Thr Ala Ala Leu Ala Leu Asn
130 135 140

Gln Ala Gln Asp Asn Ala Lys Ala Ile Leu Asn Leu Lys Glu Ser Ile
145 150 155 160

```
Thr Lys Thr Asn Glu Ala Val Leu Glu Leu Lys Asp Ala Thr Gly Gln
                165                 170                 175

Ile Ala Ile Ala Leu Asp Lys Thr Gln Arg Phe Ile Asn Asp Asn Ile
            180                 185                 190

Leu Pro Ala Ile Asn Asn Leu Thr Cys Glu Val Ala Gly Ala Lys Val
        195                 200                 205

Gly Val Glu Leu Ser Leu Tyr Leu Thr Glu Leu Ser Thr Val Phe Gly
    210                 215                 220

Ser Gln Ile Thr Asn Pro Ala Leu Ser Thr Leu Ser Ile Gln Ala Leu
225                 230                 235                 240

Met Ser Leu Cys Gly Asn Asp Phe Asn Tyr Leu Leu Asn Leu Met Gly
                245                 250                 255

Ala Lys His Ser Asp Leu Gly Ala Leu Tyr Glu Ala Asn Leu Ile Asn
            260                 265                 270

Gly Arg Ile Ile Gln Tyr Asp Gln Ala Ser Gln Ile Met Val Ile Gln
        275                 280                 285

Val Ser Val Pro Ser Ile Ser Ser Ile Ser Gly Leu Arg Leu Thr Glu
    290                 295                 300

Leu Phe Thr Leu Ser Ile Glu Thr Pro Val Gly Glu Gly Lys Ala Val
305                 310                 315                 320

Val Pro Gln Phe Val Val Glu Ser Gly Gln Leu Leu Glu Glu Ile Asp
                325                 330                 335

Thr Gln Ala Cys Thr Leu Thr Asp Thr Thr Ala Tyr Cys Thr Ile Val
            340                 345                 350

Arg Thr Lys Pro Leu Pro Glu Leu Val Ala Gln Cys Leu Arg Gly Asp
        355                 360                 365

Glu Ser Arg Cys Gln Tyr Thr Thr Gly Ile Gly Met Leu Glu Ser Arg
    370                 375                 380

Phe Gly Val Phe Asp Gly Leu Val Ile Ala Asn Cys Lys Ala Thr Ile
385                 390                 395                 400

Cys Arg Cys Leu Ala Pro Glu Met Ile Ile Thr Gln Asn Lys Gly Leu
                405                 410                 415

Pro Leu Thr Val Ile Ser Gln Glu Thr Cys Lys Arg Ile Leu Ile Asp
            420                 425                 430

Gly Val Thr Leu Gln Ile Glu Ala Gln Val Ser Gly Ser Tyr Ser Arg
```

435 440 445

Asn Ile Thr Val Gly Asn Ser Gln Ile Ala Pro Ser Gly Pro Leu Asp
450 455 460

Ile Ser Ser Glu Leu Gly Lys Val Asn Gln Ser Leu Ser Asn Val Glu
465 470 475 480

Asp Leu Ile Asp Gln Ser Asn Gln Leu Leu Asn Arg Val Asn Pro Asn
485 490 495

Ile Val Asn Asn Thr Ala Ile Ile Val Thr Ile Val Leu Leu Val Leu
500 505 510

Leu Val Leu Trp Cys Leu Ala Leu Thr Ile Ser Ile Leu Tyr Val Ser
515 520 525

Lys His Ala Val Arg Met Ile Lys Thr Val Pro Asn Pro Tyr Val Met
530 535 540

Gln Ala Lys Ser Pro Gly Ser Ala Thr Gln Phe
545 550 555

<210> 15
<211> 545
<212> PRT
<213> Fer-de-Lance paramyxovirus

<220>
<221> MISC_FEATURE
<223> fusion protein F

<400> 15

Met Thr Arg Ile Thr Ile Leu Gln Ile Ile Leu Thr Leu Thr Leu Pro
1 5 10 15

Val Met Cys Gln Val Ser Phe Asp Asn Leu Glu Gln Val Gly Val Met
20 25 30

Phe Asp Lys Pro Lys Phe Leu Lys Ile Thr Gly Pro Ala Ser Thr Ala
35 40 45

Thr Met Ile Ile Lys Leu Ile Pro Thr Leu Gly Thr Met Glu Ser Cys
50 55 60

Gly Thr Ser Ala Val Asn Glu Tyr Lys Lys Thr Leu Asp Thr Ile Leu
65 70 75 80

Val Pro Leu Arg Asp Thr Ile Asn Lys Leu Ser Thr Asp Ile Thr Val
85 90 95

Val Glu Gly Thr Ser Asn Ile Ser Asn Lys Arg Glu Lys Arg Phe Val
100 105 110

```
Gly Ile Ala Ile Ala Val Gly Ala Val Ala Leu Ala Thr Ser Ala Gln
        115                 120                 125

Ile Thr Ala Gly Ile Ala Leu Ser Asn Thr Ile Lys Asn Ala Glu Ala
    130                 135                 140

Ile Glu Ser Ile Lys Ser Ser Ile Gln Ala Ser Asn Gln Ala Ile Gln
145                 150                 155                 160

Lys Val Ile Asp Ala Gln Gly Arg Thr Val Thr Val Ile Asn Gly Ile
                165                 170                 175

Gln Asp His Ile Asn Ser Val Ile Asn Pro Ala Leu Asn Gln Leu Gly
            180                 185                 190

Cys Asp Val Ala Lys Asn Thr Leu Ala Ile Ser Leu Thr Gln Tyr Phe
        195                 200                 205

Ser Lys Leu Ser Leu Leu Phe Gly Pro Asn Leu Arg Asn Pro Val Glu
    210                 215                 220

Gln Pro Leu Ser Val Gln Ala Ile Ala Gly Leu Met Asp Gly Asp Ile
225                 230                 235                 240

Asn Ala Val Val Ser Gln Leu Gly Tyr Thr Gln Ser Asp Leu Leu Asp
                245                 250                 255

Leu Leu Ser Thr Glu Ser Ile Val Gly Thr Val Thr Ala Ile Asp Met
            260                 265                 270

Val Asn Tyr Met Ile Gln Ile Glu Met Ser Phe Pro Gln Tyr Ile Thr
        275                 280                 285

Ile Pro Asp Thr Lys Val Leu Glu Gly His Lys Ile Thr Phe Asn Asp
    290                 295                 300

Lys Gly Ser Glu Trp Gln Thr Gln Val Pro Ser Thr Ile Ala Val Arg
305                 310                 315                 320

Asp Ile Leu Ile Ala Gly Val Asp Pro Asp Gly Cys Ser Ile Thr Ser
                325                 330                 335

Thr Ser Tyr Ile Cys Lys Asn Asp Pro Thr Tyr Ala Met Ser Glu Val
            340                 345                 350

Leu Thr Asn Cys Phe Arg Gly Asn Thr Gln Glu Cys Pro Arg Ala Arg
        355                 360                 365

Ile Thr Ser Thr Phe Ala Thr Arg Phe Ala Ile Ala Arg Ser Thr Val
    370                 375                 380
```

```
Ile Ala Asn Cys Val Ala Ala Val Cys Leu Cys Gly Asp Pro Gly Ile
385                 390                 395                 400


Pro Val Val Gln Lys Ala Glu Val Thr Leu Thr Ala Met Thr Leu Asp
                405                 410                 415


Gln Cys Ser Leu Ile Thr Val Asp Gly Leu Gln Ile Lys Pro Ser Lys
            420                 425                 430


Ser Ile Ala Asn Val Thr Ala Asn Phe Gly Asn Ile Thr Leu Gly Pro
        435                 440                 445


Val Val Ser Val Gly Asp Leu Asp Leu Ser Ala Glu Leu Thr Lys Val
    450                 455                 460


Gln Ser Asp Leu Lys Glu Ala Gln Asp Lys Leu Asp Glu Ser Asn Ala
465                 470                 475                 480


Ile Leu Gln Gly Ile Asn Asn Lys Ile Leu Thr Ala Pro Thr Ser Ile
                485                 490                 495


Ala Leu Ile Val Val Ser Val Val Val Ile Leu Leu Ile Ile Gly Met
            500                 505                 510


Ile Ser Trp Leu Val Trp Leu Thr Lys Ala Val Arg Arg Ser Asn Thr
        515                 520                 525


Arg Ser Glu Arg Val Thr Pro Ser Ala Tyr Asn Asn Leu Gly Phe Ile
    530                 535                 540


Lys
545


<210>  16
<211>  566
<212>  PRT
<213>  Avian paramyxovirus 4


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  16

Met Arg Leu Ser Arg Thr Ile Leu Thr Leu Ile Leu Gly Thr Leu Thr
1               5                   10                  15


Gly Tyr Leu Met Gly Ala His Ser Thr Asn Val Asn Glu Gly Pro Lys
            20                  25                  30


Ser Glu Gly Ile Arg Gly Asp Leu Ile Pro Gly Ala Gly Ile Phe Val
        35                  40                  45
```

```
Thr Gln Val Arg Gln Leu Gln Ile Tyr Gln Gln Ser Gly Tyr His Asp
    50                  55                  60

Leu Val Ile Arg Leu Leu Pro Leu Leu Pro Ala Glu Leu Asn Asp Cys
65                  70                  75                  80

Gln Arg Glu Val Val Thr Glu Tyr Asn Asn Thr Val Ser Gln Leu Leu
                85                  90                  95

Gln Pro Ile Lys Thr Asn Leu Asp Thr Leu Leu Ala Asp Gly Gly Thr
            100                 105                 110

Arg Asp Ala Asp Ile Gln Pro Arg Phe Ile Gly Ala Ile Ile Ala Thr
        115                 120                 125

Gly Ala Leu Ala Val Ala Thr Val Ala Glu Val Thr Ala Ala Gln Ala
    130                 135                 140

Leu Ser Gln Ser Lys Thr Asn Ala Gln Asn Ile Leu Lys Leu Arg Asp
145                 150                 155                 160

Ser Ile Gln Ala Thr Asn Gln Ala Val Phe Glu Ile Ser Gln Gly Leu
                165                 170                 175

Glu Ala Thr Ala Thr Val Leu Ser Lys Leu Gln Thr Glu Leu Asn Glu
            180                 185                 190

Asn Ile Ile Pro Ser Leu Asn Asn Leu Ser Cys Ala Ala Met Gly Asn
        195                 200                 205

Arg Leu Gly Val Ser Leu Ser Leu Tyr Leu Thr Leu Met Thr Thr Leu
    210                 215                 220

Phe Gly Asp Gln Ile Thr Asn Pro Val Leu Thr Pro Ile Ser Tyr Ser
225                 230                 235                 240

Thr Leu Ser Ala Met Ala Gly Gly His Ile Gly Pro Val Met Ser Lys
                245                 250                 255

Ile Leu Ala Gly Ser Val Thr Ser Gln Leu Gly Ala Glu Gln Leu Ile
            260                 265                 270

Ala Ser Gly Leu Ile Gln Ser Gln Val Val Gly Tyr Asp Ser Gln Tyr
        275                 280                 285

Gln Leu Leu Val Ile Arg Val Asn Leu Val Arg Ile Gln Glu Val Gln
    290                 295                 300

Asn Thr Arg Val Val Ser Leu Arg Thr Leu Ala Val Asn Arg Asp Gly
305                 310                 315                 320

Gly Leu Tyr Arg Ala Gln Val Pro Pro Glu Val Val Glu Arg Ser Gly
```

325 330 335

Ile Ala Glu Arg Phe Tyr Ala Asp Asp Cys Val Leu Thr Thr Thr Asp
340 345 350

Tyr Ile Cys Ser Ser Ile Arg Ser Ser Arg Leu Asn Pro Glu Leu Val
355 360 365

Lys Cys Leu Ser Gly Ala Leu Asp Ser Cys Thr Phe Glu Arg Glu Ser
370 375 380

Ala Leu Leu Ser Thr Pro Phe Phe Val Tyr Asn Lys Ala Val Val Ala
385 390 395 400

Asn Cys Lys Ala Ala Thr Cys Arg Cys Asn Lys Pro Pro Ser Ile Ile
405 410 415

Ala Gln Tyr Ser Ala Ser Ala Leu Val Thr Ile Thr Thr Asp Thr Cys
420 425 430

Ala Asp Leu Glu Ile Glu Gly Tyr Arg Phe Asn Ile Gln Thr Glu Ser
435 440 445

Asn Ser Trp Val Ala Pro Asn Phe Thr Val Ser Thr Ser Gln Ile Val
450 455 460

Ser Val Asp Pro Ile Asp Ile Ser Ser Asp Ile Ala Lys Ile Asn Ser
465 470 475 480

Ser Ile Glu Ala Ala Arg Glu Gln Leu Glu Leu Ser Asn Gln Ile Leu
485 490 495

Ser Arg Ile Asn Pro Arg Ile Val Asn Asp Glu Ser Leu Ile Ala Ile
500 505 510

Ile Val Thr Ile Val Val Leu Ser Leu Leu Val Ile Gly Leu Ile Val
515 520 525

Val Leu Gly Val Met Tyr Lys Asn Leu Lys Lys Val Gln Arg Ala Gln
530 535 540

Ala Ala Met Met Met Gln Gln Met Ser Ser Ser Gln Pro Val Thr Thr
545 550 555 560

Lys Leu Gly Thr Pro Phe
565

<210> 17
<211> 551
<212> PRT
<213> Human parainfluenza virus 2

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 17

Met His His Leu His Pro Met Ile Val Cys Ile Phe Val Met Tyr Thr
1 5 10 15

Gly Ile Val Gly Ser Asp Ala Ile Ala Gly Asp Gln Leu Leu Asn Ile
20 25 30

Gly Val Ile Gln Ser Lys Ile Arg Ser Leu Met Tyr Tyr Thr Asp Gly
35 40 45

Gly Ala Ser Phe Ile Val Val Lys Leu Leu Pro Asn Leu Pro Pro Ser
50 55 60

Asn Gly Thr Cys Asn Ile Thr Ser Leu Asp Ala Tyr Asn Val Thr Leu
65 70 75 80

Phe Lys Leu Leu Thr Pro Leu Ile Glu Asn Leu Ser Lys Ile Ser Thr
85 90 95

Val Thr Asp Thr Lys Thr Arg Gln Lys Arg Phe Ala Gly Val Val Val
100 105 110

Gly Leu Ala Ala Leu Gly Val Ala Thr Ala Ala Gln Ile Thr Ala Ala
115 120 125

Val Ala Ile Val Lys Ala Asn Ala Asn Ala Ala Ala Ile Asn Asn Leu
130 135 140

Ala Ser Ser Ile Gln Ser Thr Asn Lys Ala Val Ser Asp Val Ile Asp
145 150 155 160

Ala Ser Arg Thr Ile Ala Thr Ala Val Gln Ala Ile Gln Asp Arg Ile
165 170 175

Asn Gly Ala Ile Val Asn Gly Ile Thr Ser Ala Ser Cys Arg Ala His
180 185 190

Asp Ala Leu Ile Gly Ser Ile Leu Asn Leu Tyr Leu Thr Glu Leu Thr
195 200 205

Thr Ile Phe His Asn Gln Ile Thr Asn Pro Ala Leu Thr Pro Leu Ser
210 215 220

Ile Gln Ala Leu Arg Ile Leu Leu Gly Ser Thr Leu Pro Ile Val Ile
225 230 235 240

Glu Ser Lys Leu Asn Thr Asn Phe Asn Thr Ala Glu Leu Leu Ser Ser
245 250 255

```
Gly Leu Leu Thr Gly Gln Ile Ile Ser Ile Ser Pro Met Tyr Met Gln
            260                 265                 270

Met Leu Ile Gln Ile Asn Val Pro Thr Phe Ile Met Gln Pro Gly Ala
        275                 280                 285

Lys Val Ile Asp Leu Ile Ala Ile Ser Ala Asn His Lys Leu Gln Glu
    290                 295                 300

Val Val Val Gln Val Pro Asn Arg Ile Leu Glu Tyr Ala Asn Glu Leu
305                 310                 315                 320

Gln Asn Tyr Pro Ala Asn Asp Cys Val Val Thr Pro Asn Ser Val Phe
                325                 330                 335

Cys Arg Tyr Asn Glu Gly Ser Pro Ile Pro Glu Ser Gln Tyr Gln Cys
            340                 345                 350

Leu Arg Gly Asn Leu Asn Ser Cys Thr Phe Thr Pro Ile Ile Gly Asn
        355                 360                 365

Phe Leu Lys Arg Phe Ala Phe Ala Asn Gly Val Leu Tyr Ala Asn Cys
    370                 375                 380

Lys Ser Leu Leu Cys Arg Cys Ala Asp Pro Pro His Val Val Ser Gln
385                 390                 395                 400

Asp Asp Thr Gln Gly Ile Ser Ile Ile Asp Ile Lys Arg Cys Ser Glu
                405                 410                 415

Met Met Leu Asp Thr Phe Ser Phe Arg Ile Thr Ser Thr Phe Asn Ala
            420                 425                 430

Thr Tyr Val Thr Asp Phe Ser Met Ile Asn Ala Asn Ile Val His Leu
        435                 440                 445

Ser Pro Leu Asp Leu Ser Asn Gln Ile Asn Ser Ile Asn Lys Ser Leu
    450                 455                 460

Lys Ser Ala Glu Asp Trp Ile Ala Asp Ser Asn Phe Phe Ala Asn Gln
465                 470                 475                 480

Ala Arg Thr Ala Lys Thr Leu Tyr Ser Leu Ser Ala Ile Ala Leu Ile
                485                 490                 495

Leu Ser Val Ile Thr Leu Val Val Val Gly Leu Leu Ile Ala Tyr Ile
            500                 505                 510

Ile Lys Leu Val Ser Gln Ile His Gln Phe Arg Ser Leu Ala Ala Thr
        515                 520                 525
```

Thr Met Phe His Arg Glu Asn Pro Ala Phe Phe Ser Lys Asn Asn His
530 535 540

Gly Asn Ile Tyr Gly Ile Ser
545 550

<210> 18
<211> 541
<212> PRT
<213> Porcine rubulavirus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 18

Met Pro Gln Gln Gln Val Ala His Thr Cys Val Met Leu Trp Gly Ile
1 5 10 15

Ile Ser Thr Val Ser Gly Ile Asn Thr Glu Ala Leu Ser Gln Tyr Gly
20 25 30

Val Val Val Thr Asn Val Arg Gln Leu Thr Tyr Tyr Thr Gln Ala Gly
35 40 45

Ser Thr Tyr Leu Ala Val Arg Leu Leu Pro Ser Leu Ala Ser Pro Asp
50 55 60

Gln Ser Cys Ala Leu His Ser Ile Ile Asn Tyr Asn Ala Thr Leu Gln
65 70 75 80

Ala Ile Leu Ser Pro Ile Ala Glu Asn Leu Asn Leu Ile Ser Thr Ala
85 90 95

Leu Arg Glu Gln His Arg Lys Lys Arg Phe Ala Gly Val Ala Ile Gly
100 105 110

Leu Thr Ala Leu Gly Val Ala Thr Ala Ala Gln Ala Thr Ala Ala Val
115 120 125

Ala Leu Val Arg Ala Asn Lys Asn Ala Glu Lys Val Glu Gln Leu Ser
130 135 140

Gln Ala Leu Gly Glu Thr Asn Ala Ala Ile Ser Asp Leu Ile Asp Ala
145 150 155 160

Thr Lys Asn Leu Gly Phe Ala Val Gln Ala Ile Gln Asn Gln Ile Asn
165 170 175

Thr Ala Ile Leu Pro Gln Ile His Asn Leu Ser Cys Gln Val Ile Asp
180 185 190

Ala Gln Leu Gly Asn Ile Leu Ser Leu Tyr Leu Thr Glu Leu Thr Thr

195 200 205

Val Phe Gln Pro Gln Leu Thr Asn Pro Ala Leu Ser Pro Leu Thr Ile
210 215 220

Gln Ala Leu Arg Ala Val Leu Gly Thr Thr Leu Pro Ala Leu Leu Ser
225 230 235 240

Glu Lys Leu Lys Ser Asn Ile Pro Leu Gly Asp Leu Met Ser Ser Gly
245 250 255

Leu Leu Lys Gly Gln Leu Val Gly Leu Asn Leu Gln Asn Met Leu Met
260 265 270

Ile Ile Glu Leu Tyr Ile Pro Thr Leu Ser Thr His Ser Thr Ala Lys
275 280 285

Val Leu Asp Leu Val Thr Ile Ser Ser His Val Asn Gly Arg Glu Val
290 295 300

Glu Ile Gln Val Pro Asn Arg Val Leu Glu Leu Gly Ser Glu Val Leu
305 310 315 320

Gly Tyr Gly Gly Ser Glu Cys Ala Leu Thr Met Ser His Ile Leu Cys
325 330 335

Pro Phe Asn Asp Ala Arg Val Leu Ser Thr Asp Met Lys Tyr Cys Leu
340 345 350

Gln Gly Asn Ile Thr His Cys Ile Phe Ser Pro Val Val Gly Ser Phe
355 360 365

Leu Arg Arg Phe Ala Leu Val Asn Gly Val Val Ile Ala Asn Cys Ala
370 375 380

Asp Met Ser Cys Val Cys Phe Asp Pro Gln Glu Ile Ile Tyr Gln Asn
385 390 395 400

Phe Gln Glu Pro Thr Thr Val Ile Asp Ile Lys Lys Cys Gly Lys Val
405 410 415

Gln Leu Asp Thr Leu Thr Phe Thr Ile Ser Thr Phe Ala Asn Arg Thr
420 425 430

Tyr Gly Pro Pro Ala Tyr Val Pro Pro Asp Asn Ile Ile Gln Ser Glu
435 440 445

Pro Leu Asp Ile Ser Gly Asn Leu Ile Ala Val Asn Asn Ser Leu Ser
450 455 460

Ser Ala Leu Asn His Leu Ala Thr Ser Glu Ile Leu Arg Asn Glu Gln
465 470 475 480

```
Ile Trp Thr Ser Ser Leu Gly Ile Ser Thr Ile Val Ala Leu Val Ile
                485                 490                 495


Ile Gly Ile Leu Ile Ile Cys Leu Val Val Thr Trp Ala Ala Leu Trp
            500                 505                 510


Ala Leu Leu Lys Glu Val Arg Gly Leu Asn Ser Ala Val Asn Ser Gln
        515                 520                 525


Leu Ser Ser Tyr Val Met Gly Asp Lys Phe Ile Arg Tyr
    530                 535                 540


<210>  19
<211>  551
<212>  PRT
<213>  Parainfluenza virus 5


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  19

Met Gly Thr Arg Ile Gln Phe Leu Met Val Ser Cys Leu Leu Ala Gly
1               5                   10                  15


Thr Gly Ser Leu Asp Pro Ala Ala Leu Met Gln Ile Gly Val Ile Pro
            20                  25                  30


Thr Asn Val Arg Gln Leu Met Tyr Tyr Thr Glu Ala Ser Ser Ala Phe
        35                  40                  45


Ile Val Val Lys Leu Met Pro Thr Ile Asp Ser Pro Ile Ser Gly Cys
    50                  55                  60


Asn Ile Thr Ser Ile Ser Ser Tyr Asn Ala Thr Met Thr Lys Leu Leu
65                  70                  75                  80


Gln Pro Ile Gly Glu Asn Leu Glu Thr Ile Arg Tyr Gln Leu Ile Pro
                85                  90                  95


Thr Arg Arg Arg Arg Arg Phe Val Gly Val Val Ile Gly Leu Ala Ala
            100                 105                 110


Leu Gly Val Ala Thr Ala Ala Gln Val Thr Ala Ala Val Ala Leu Val
        115                 120                 125


Lys Ala Asn Lys Asn Ala Ala Ala Ile Leu Asn Leu Lys Asn Ala Ile
    130                 135                 140


Gln Lys Thr Asn Ala Ala Val Ala Asp Val Val Gln Ala Thr Gln Ser
145                 150                 155                 160
```

```
Leu Gly Thr Ala Val Gln Ala Val Gln Asp His Ile Asn Ser Val Val
                165                 170                 175

Ser Pro Ala Ile Thr Ala Ala Asn Cys Lys Ala Gln Asp Ala Ile Ile
            180                 185                 190

Gly Ser Ile Leu Asn Leu Tyr Leu Thr Glu Leu Thr Thr Ile Phe His
        195                 200                 205

Asn Gln Ile Thr Asn Pro Ala Leu Ser Pro Ile Thr Ile Gln Ala Leu
    210                 215                 220

Arg Ile Leu Leu Gly Ser Thr Leu Pro Thr Val Val Arg Lys Ser Phe
225                 230                 235                 240

Asn Thr Gln Ile Ser Ala Ala Glu Leu Leu Ser Ser Gly Leu Leu Thr
                245                 250                 255

Gly Gln Ile Val Gly Leu Asp Leu Thr Tyr Met Gln Met Val Ile Lys
            260                 265                 270

Ile Glu Leu Pro Thr Leu Thr Val Gln Pro Ala Thr Gln Ile Ile Asp
        275                 280                 285

Leu Val Thr Ile Ser Ala Phe Ile Asn Asn Arg Glu Val Met Ala Gln
    290                 295                 300

Leu Pro Thr Arg Ile Ile Val Thr Gly Ser Leu Ile Gln Ala Tyr Pro
305                 310                 315                 320

Ala Ser Gln Cys Thr Ile Thr Pro Asn Thr Val Tyr Cys Arg Tyr Asn
                325                 330                 335

Asp Ala Gln Val Leu Ser Asp Asp Thr Met Ala Cys Leu Gln Gly Asn
            340                 345                 350

Leu Thr Arg Cys Thr Phe Ser Pro Val Val Gly Ser Phe Leu Thr Arg
        355                 360                 365

Phe Val Leu Phe Asp Gly Ile Val Tyr Ala Asn Cys Arg Ser Met Leu
    370                 375                 380

Cys Lys Cys Met Gln Pro Ala Ala Val Ile Leu Gln Pro Ser Ser Ser
385                 390                 395                 400

Pro Val Thr Val Ile Asp Met His Lys Cys Val Ser Leu Gln Leu Asp
                405                 410                 415

Asn Leu Arg Phe Thr Ile Thr Gln Leu Ala Asn Ile Thr Tyr Asn Ser
            420                 425                 430
```

Thr Ile Lys Leu Glu Thr Ser Gln Ile Leu Pro Ile Asp Pro Leu Asp
435 440 445

Ile Ser Gln Asn Leu Ala Ala Val Asn Lys Ser Leu Ser Asp Ala Leu
450 455 460

Gln His Leu Ala Gln Ser Asp Thr Tyr Leu Ser Ala Ile Thr Ser Ala
465 470 475 480

Thr Thr Thr Ser Val Leu Ser Ile Ile Ala Ile Cys Leu Gly Ser Leu
485 490 495

Gly Leu Ile Leu Ile Ile Leu Ile Ser Val Val Val Trp Lys Leu Leu
500 505 510

Thr Ile Val Ala Ala Asn Arg Asn Arg Met Glu Asn Phe Val Tyr His
515 520 525

Asn Ser Ala Phe His His Ser Arg Ser Asp Leu Ser Glu Lys Asn Gln
530 535 540

Pro Ala Thr Leu Gly Thr Arg
545 550

<210> 20
<211> 537
<212> PRT
<213> Murine pneumonia virus

<220>
<221> MISC_FEATURE
<223> fusion glycoprotein precursor

<400> 20

Met Ile Pro Gly Arg Ile Phe Leu Val Leu Leu Val Ile Phe Asn Thr
1 5 10 15

Lys Pro Ile His Pro Asn Thr Leu Thr Glu Lys Phe Tyr Glu Ser Thr
20 25 30

Cys Ser Val Glu Thr Ala Gly Tyr Lys Ser Ala Leu Arg Thr Gly Trp
35 40 45

His Met Thr Val Met Ser Ile Lys Leu Ser Gln Ile Asn Ile Glu Ser
50 55 60

Cys Lys Ser Ser Asn Ser Leu Leu Ala His Glu Leu Ala Ile Tyr Ser
65 70 75 80

Ser Ala Val Asp Glu Leu Arg Thr Leu Ser Ser Asn Ala Leu Lys Ser
85 90 95

Lys Arg Lys Lys Arg Phe Leu Gly Leu Ile Leu Gly Leu Gly Ala Ala

100 105 110

Val Thr Ala Gly Val Ala Leu Ala Lys Thr Val Gln Leu Glu Ser Glu
115 120 125

Ile Ala Leu Ile Arg Asp Ala Val Arg Asn Thr Asn Glu Ala Val Val
130 135 140

Ser Leu Thr Asn Gly Met Ser Val Leu Ala Lys Val Val Asp Asp Leu
145 150 155 160

Lys Asn Phe Ile Ser Lys Glu Leu Leu Pro Lys Ile Asn Arg Val Ser
165 170 175

Cys Asp Val His Asp Ile Thr Ala Val Ile Arg Phe Gln Gln Leu Asn
180 185 190

Lys Arg Leu Leu Glu Val Ser Arg Glu Phe Ser Ser Asn Ala Gly Leu
195 200 205

Thr His Thr Val Ser Ser Phe Met Leu Thr Asp Arg Glu Leu Thr Ser
210 215 220

Ile Val Gly Gly Met Ala Val Ser Ala Gly Gln Lys Glu Ile Met Leu
225 230 235 240

Ser Ser Lys Ala Ile Met Arg Arg Asn Gly Leu Ala Ile Leu Ser Ser
245 250 255

Val Asn Ala Asp Thr Leu Val Tyr Val Ile Gln Leu Pro Leu Phe Gly
260 265 270

Val Met Asp Thr Asp Cys Trp Val Ile Arg Ser Ser Ile Asp Cys His
275 280 285

Asn Ile Ala Asp Lys Tyr Ala Cys Leu Ala Arg Ala Asp Asn Gly Trp
290 295 300

Tyr Cys His Asn Ala Gly Ser Leu Ser Tyr Phe Pro Ser Pro Thr Asp
305 310 315 320

Cys Glu Ile His Asn Gly Tyr Ala Phe Cys Asp Thr Leu Lys Ser Leu
325 330 335

Thr Val Pro Val Thr Ser Arg Glu Cys Asn Ser Asn Met Tyr Thr Thr
340 345 350

Asn Tyr Asp Cys Lys Ile Ser Thr Ser Lys Thr Tyr Val Ser Thr Ala
355 360 365

Val Leu Thr Thr Met Gly Cys Leu Val Ser Cys Tyr Gly His Asn Ser
370 375 380

```
Cys Thr Val Ile Asn Asn Asp Lys Gly Ile Ile Arg Thr Leu Pro Asp
385                 390                 395                 400


Gly Cys His Tyr Ile Ser Asn Lys Gly Val Asp Arg Val Gln Val Gly
                405                 410                 415


Asn Thr Val Tyr Tyr Leu Ser Lys Glu Val Gly Lys Ser Ile Val Val
            420                 425                 430


Arg Gly Glu Pro Leu Val Leu Lys Tyr Asp Pro Leu Ser Phe Pro Asp
        435                 440                 445


Asp Lys Phe Asp Val Ala Ile Arg Asp Val Glu His Ser Ile Asn Gln
    450                 455                 460


Thr Arg Thr Phe Leu Lys Ala Ser Asp Gln Leu Leu Asp Leu Ser Glu
465                 470                 475                 480


Asn Arg Glu Asn Lys Asn Leu Asn Lys Ser Tyr Ile Leu Thr Thr Leu
                485                 490                 495


Leu Phe Val Val Met Leu Ile Ile Ile Met Ala Val Ile Gly Phe Ile
            500                 505                 510


Leu Tyr Lys Val Leu Lys Met Ile Arg Asp Asn Lys Leu Lys Ser Lys
        515                 520                 525


Ser Thr Pro Gly Leu Thr Val Leu Ser
    530                 535


<210>  21
<211>  543
<212>  PRT
<213>  Human parainfluenza virus 4a


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  21

Met Gly Val Lys Gly Leu Ser Leu Ile Met Ile Gly Leu Leu Ile Ser
1               5                   10                  15


Pro Ile Thr Asn Leu Asp Ile Thr His Leu Met Asn Leu Gly Thr Val
            20                  25                  30


Pro Thr Ala Ile Arg Ser Leu Val Tyr Tyr Thr Tyr Thr Lys Pro Ser
        35                  40                  45


Tyr Leu Thr Val Asp Leu Ile Pro Asn Leu Lys Asn Leu Asp Gln Asn
    50                  55                  60
```

```
Cys Asn Tyr Ser Ser Leu Asn Tyr Tyr Asn Lys Thr Ala Leu Ser Leu
65                  70                  75                  80

Ile Gln Pro Ile Ala Asp Asn Ile Asn Arg Leu Thr Lys Pro Ile Thr
                85                  90                  95

Ser Ser Glu Ile Gln Ser Arg Phe Phe Gly Ala Val Ile Gly Thr Ile
            100                 105                 110

Ala Leu Gly Val Ala Thr Ala Ala Gln Val Thr Ala Ala Ile Gly Leu
        115                 120                 125

Ala Lys Ala Gln Glu Asn Ala Lys Leu Ile Leu Thr Leu Lys Lys Ala
    130                 135                 140

Ala Thr Glu Thr Asn Glu Ala Val Arg Asp Leu Ala Asn Ser Asn Lys
145                 150                 155                 160

Ile Val Val Lys Met Ile Ser Ala Ile Gln Asn Gln Ile Asn Thr Ile
                165                 170                 175

Ile Gln Pro Ala Ile Asp Gln Ile Asn Cys Gln Ile Lys Asp Leu Gln
            180                 185                 190

Val Ala Asn Ile Leu Asn Leu Tyr Leu Thr Glu Ile Thr Thr Val Phe
        195                 200                 205

His Asn Gln Leu Thr Asn Pro Ala Leu Glu Ser Ile Ser Ile Gln Ala
    210                 215                 220

Leu Lys Ser Leu Leu Gly Pro Thr Leu Pro Glu Val Leu Ser Lys Leu
225                 230                 235                 240

Asp Leu Asn Asn Ile Ser Ala Ala Ser Val Met Ala Ser Gly Leu Ile
                245                 250                 255

Lys Gly Gln Ile Ile Ala Val Asp Ile Pro Thr Met Thr Leu Val Leu
            260                 265                 270

Met Val Gln Ile Pro Ser Ile Ser Pro Leu Arg Gln Ala Lys Ile Ile
        275                 280                 285

Asp Leu Thr Ser Ile Thr Ile His Thr Asn Ser Gln Glu Val Gln Ala
    290                 295                 300

Val Val Pro Ala Arg Phe Leu Glu Ile Gly Ser Glu Ile Leu Gly Phe
305                 310                 315                 320

Asp Gly Ser Val Cys Gln Ile Thr Lys Asp Thr Ile Phe Cys Pro Tyr
                325                 330                 335
```

Asn Asp Ala Tyr Glu Leu Pro Ile Gln Gln Lys Arg Cys Leu Gln Gly
340 345 350

Gln Thr Arg Asp Cys Val Phe Thr Pro Val Ala Gly Thr Phe Pro Arg
355 360 365

Arg Phe Leu Thr Thr Tyr Gly Thr Ile Val Ala Asn Cys Arg Asp Leu
370 375 380

Val Cys Ser Cys Leu Arg Pro Pro Gln Ile Ile Tyr Gln Pro Asp Glu
385 390 395 400

Asn Pro Val Thr Ile Ile Asp Lys Asp Leu Cys Thr Thr Leu Thr Leu
405 410 415

Asp Ser Ile Thr Ile Glu Ile Gln Lys Ser Ile Asn Ser Thr Phe Arg
420 425 430

Arg Glu Val Val Leu Glu Ser Thr Gln Val Arg Ser Leu Thr Pro Leu
435 440 445

Asp Leu Ser Thr Asp Leu Asn Gln Tyr Asn Gln Leu Leu Lys Ser Ala
450 455 460

Glu Asp His Ile Gln Arg Ser Thr Asp Tyr Leu Asn Ser Ile Asn Pro
465 470 475 480

Ser Ile Val Asn Asn Asn Ala Ile Ile Ile Leu Ile Ile Leu Cys Ile
485 490 495

Leu Leu Ile Leu Thr Val Thr Ile Cys Ile Ile Trp Leu Lys Tyr Leu
500 505 510

Thr Lys Glu Val Lys Asn Val Ala Arg Asn Gln Arg Leu Asn Arg Asp
515 520 525

Ala Asp Leu Phe Tyr Lys Ile Pro Ser Gln Ile Pro Val Pro Arg
530 535 540

<210> 22
<211> 538
<212> PRT
<213> Tioman virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 22

Met Arg Ile Ala Leu Thr Ala Val Ile Val Ser Ile His Phe Asp Leu
1 5 10 15

Ala Phe Pro Met Asn Lys Asn Ser Leu Leu Ser Val Gly Leu Val His

20 25 30

Lys Ser Val Lys Asn Leu Tyr Phe Tyr Ser Gln Gly Ser Pro Ser Tyr
35 40 45

Ile Val Val Lys Leu Val Pro Thr Leu Gly Asn Val Pro Gly Asn Cys
50 55 60

Thr Leu Asn Ser Leu Val Arg Tyr Lys Ser Thr Val Ser Ser Leu Leu
65 70 75 80

Ser Pro Leu Ala Glu Asn Leu Glu Tyr Leu Gln Lys Thr Leu Thr Val
85 90 95

Ser Arg Gly Gly Arg Arg Arg Arg Phe Ala Gly Val Ala Ile Gly Leu
100 105 110

Ala Ala Leu Gly Val Ala Ala Ala Ala Gln Ala Thr Ala Ala Val Ala
115 120 125

Leu Val Glu Ala Arg Gln Asn Ala Ala Gln Ile Gln Ser Leu Ser Glu
130 135 140

Ala Ile Gln Asn Thr Asn Leu Ala Val Asn Glu Leu Lys Thr Ala Ile
145 150 155 160

Gly Ala Ser Ala Thr Ala Ile Gln Ala Ile Gln Thr Gln Ile Asn Glu
165 170 175

Val Ile Asn Pro Ala Ile Asn Arg Leu Ser Cys Glu Ile Leu Asp Ala
180 185 190

Gln Leu Ala Ser Met Leu Asn Leu Tyr Leu Ile His Leu Thr Thr Val
195 200 205

Phe Gln Asn Gln Leu Thr Asn Pro Ala Leu Thr Pro Leu Ser Ile Gln
210 215 220

Ser Leu Gln Ser Leu Leu Gln Gly Thr Ser Ser Val Leu Thr Asn Ile
225 230 235 240

Thr Ser Ser Ser Lys Leu Ala Leu Asn Asp Ala Leu Val Thr Gly Leu
245 250 255

Ile Thr Gly Gln Val Val Gly Leu Asn Met Thr Ser Leu Gln Ile Val
260 265 270

Ile Ala Ala Tyr Val Pro Ser Val Ala Lys Leu Ser Asn Ala Val Val
275 280 285

His Asn Phe Ile Arg Ile Thr Thr Ser Val Asn Gly Thr Glu Val Ile
290 295 300

```
Ile Gln Ser Pro Thr Ile Ile Met Glu Gln Asn Glu Val Met Tyr Asp
305                 310                 315                 320

Leu Lys Thr Gly His Cys Thr Glu Ser Asp Leu Asn Ile Tyr Cys Pro
                325                 330                 335

Tyr Val Asp Ala Gln Leu Leu Ser Pro Gly Met Thr Asn Cys Ile Asn
            340                 345                 350

Gly Arg Leu Asn Asp Cys Thr Phe Ser Lys Val Val Gly Ser Phe Pro
        355                 360                 365

Thr Arg Phe Ala Ala Val Glu Gly Ala Ile Leu Ala Asn Cys Lys Tyr
    370                 375                 380

Leu Gln Cys Asn Cys Leu Thr Pro Pro Tyr Ile Ile Thr Pro Leu Asn
385                 390                 395                 400

Gly Glu Met Ile Ser Met Ile Asp Leu Ser Lys Cys Gln Arg Leu Asp
                405                 410                 415

Leu Gly Thr Ile Val Phe Asp Ile Asn Asn Pro Val Asn Val Thr Phe
            420                 425                 430

Asn Gly Asn Tyr Arg Ala Asp Val Gly Gln Met Ile Val Thr Asn Pro
        435                 440                 445

Leu Asp Ile Ser Ala Glu Leu Asn Gln Ile Asn Thr Ser Leu Ser Asn
    450                 455                 460

Ala Gln Gly Phe Leu Ser Lys Ser Asp Ala Trp Leu His Val Ser Gln
465                 470                 475                 480

Trp Val Thr Asn Ser Gly Thr Ile Phe Ile Ile Leu Ile Ile Gly Leu
                485                 490                 495

Ile Val Gly Ile Val Tyr Met Ile Ile Asn Thr Tyr Val Val Val Gln
            500                 505                 510

Ile Ile Lys Glu Ile Asn Arg Met Arg Thr Ser Asp Arg Ala His Leu
        515                 520                 525

Leu Lys Gly Ser Ile Ser Ser Ile Ser Thr
    530                 535


<210>  23
<211>  543
<212>  PRT
<213>  Avian paramyxovirus 8


<220>
```

<221> MISC_FEATURE
<223> fusion protein

<400> 23

```
Met Gly Gln Ile Ser Val Tyr Leu Ile Asn Ser Val Leu Leu Leu Leu
1               5                   10                  15


Val Tyr Pro Val Asn Ser Ile Asp Asn Thr Leu Ile Ala Pro Ile Gly
            20                  25                  30


Val Ala Ser Ala Asn Glu Trp Gln Leu Ala Ala Tyr Thr Thr Ser Leu
        35                  40                  45


Ser Gly Thr Ile Ala Val Arg Phe Leu Pro Val Leu Pro Asp Asn Met
    50                  55                  60


Thr Thr Cys Leu Arg Glu Thr Ile Thr Thr Tyr Asn Asn Thr Val Asn
65                  70                  75                  80


Asn Ile Leu Gly Pro Leu Lys Ser Asn Leu Asp Ala Leu Leu Ser Ser
                85                  90                  95


Glu Thr Tyr Pro Gln Thr Arg Leu Ile Gly Ala Val Ile Gly Ser Ile
            100                 105                 110


Ala Leu Gly Val Ala Thr Ser Ala Gln Ile Thr Ala Ala Val Ala Leu
        115                 120                 125


Lys Gln Ala Gln Asp Asn Ala Arg Asn Ile Leu Ala Leu Lys Glu Ala
    130                 135                 140


Leu Ser Lys Thr Asn Glu Ala Val Lys Glu Leu Ser Ser Gly Leu Gln
145                 150                 155                 160


Gln Thr Ala Ile Ala Leu Gly Lys Ile Gln Ser Phe Val Asn Glu Glu
                165                 170                 175


Ile Leu Pro Ser Ile Asn Gln Leu Ser Cys Glu Val Thr Ala Asn Lys
            180                 185                 190


Leu Gly Val Tyr Leu Ser Leu Tyr Leu Thr Glu Leu Thr Thr Ile Phe
        195                 200                 205


Gly Ala Gln Leu Thr Asn Pro Ala Leu Thr Ser Leu Ser Tyr Gln Ala
    210                 215                 220


Leu Tyr Asn Leu Cys Gly Gly Asn Met Ala Met Leu Thr Gln Lys Ile
225                 230                 235                 240


Gly Ile Lys Gln Gln Asp Val Asn Ser Leu Tyr Glu Ala Gly Leu Ile
                245                 250                 255
```

```
Thr Gly Gln Val Ile Gly Tyr Asp Ser Gln Tyr Gln Leu Leu Val Ile
            260                 265                 270

Gln Val Asn Tyr Pro Ser Ile Ser Glu Val Thr Gly Val Arg Ala Thr
        275                 280                 285

Glu Leu Val Thr Val Ser Val Thr Thr Asp Lys Gly Glu Gly Lys Ala
    290                 295                 300

Ile Val Pro Gln Phe Val Ala Glu Ser Arg Val Thr Ile Glu Glu Leu
305                 310                 315                 320

Asp Val Ala Ser Cys Lys Phe Ser Ser Thr Thr Leu Tyr Cys Arg Gln
                325                 330                 335

Val Asn Thr Arg Ala Leu Pro Pro Leu Val Ala Ser Cys Leu Arg Gly
            340                 345                 350

Asn Tyr Asp Asp Cys Gln Tyr Thr Thr Glu Ile Gly Ala Leu Ser Ser
        355                 360                 365

Arg Tyr Ile Thr Leu Asp Gly Gly Val Leu Val Asn Cys Lys Ser Ile
    370                 375                 380

Val Cys Arg Cys Leu Asn Pro Ser Lys Ile Ile Ser Gln Asn Thr Asn
385                 390                 395                 400

Ala Ala Val Thr Tyr Val Asp Ala Thr Ile Cys Lys Thr Ile Gln Leu
                405                 410                 415

Asp Asp Ile Gln Leu Gln Leu Glu Gly Ser Leu Ser Ser Val Tyr Ala
            420                 425                 430

Arg Asn Ile Ser Ile Glu Ile Ser Gln Val Thr Thr Ser Gly Ser Leu
        435                 440                 445

Asp Ile Ser Ser Glu Ile Gly Asn Ile Asn Asn Thr Val Asn Arg Val
    450                 455                 460

Glu Asp Leu Ile His Gln Ser Glu Glu Trp Leu Ala Lys Val Asn Pro
465                 470                 475                 480

His Ile Val Asn Asn Thr Thr Leu Ile Val Leu Cys Val Leu Ser Ala
                485                 490                 495

Leu Ala Val Ile Trp Leu Ala Val Leu Thr Ala Ile Ile Ile Tyr Leu
            500                 505                 510

Arg Thr Lys Leu Lys Thr Ile Ser Ala Leu Ala Val Thr Asn Thr Ile
        515                 520                 525

Gln Ser Asn Pro Tyr Val Asn Gln Thr Lys Arg Glu Ser Lys Phe
```

530 535 540

<210> 24
<211> 546
<212> PRT
<213> Tailam virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 24

Met Lys Leu Ser Val Val Tyr Thr Thr Leu Leu Val Ser Thr Phe Tyr
1 5 10 15

Ser Asp Leu Ala Arg Ser Gln Leu Ala Leu Ser Glu Leu Thr Lys Ile
20 25 30

Gly Val Ile Pro Gly Arg Ser Tyr Asp Leu Lys Ile Ser Thr Gln Ala
35 40 45

Ser Tyr Gln Tyr Met Val Val Lys Leu Ile Pro Asn Leu Thr Gly Leu
50 55 60

Asn Asn Cys Thr Asn Gly Thr Ile Glu Ala Tyr Lys Lys Met Leu Asn
65 70 75 80

Arg Leu Leu Ser Pro Ile Asp Ala Ala Leu Arg Lys Met Lys Asp Ala
85 90 95

Val Asn Asp Lys Pro Pro Glu Ser Val Gly Asn Val Lys Phe Trp Gly
100 105 110

Ala Val Ile Gly Gly Val Ala Leu Gly Val Ala Thr Ser Ala Gln Ile
115 120 125

Thr Ala Gly Val Ala Leu His Asn Ser Ile Gln Asn Ala Asn Ala Ile
130 135 140

Leu Ala Leu Lys Asp Ser Ile Arg Gln Ser Asn Lys Ala Ile Gln Glu
145 150 155 160

Leu Gln Thr Ala Met Ser Thr Thr Val Val Val Leu Asn Ala Leu Gln
165 170 175

Asp Gln Ile Asn Asn Gln Leu Val Pro Ala Ile Asn Ser Leu Gly Cys
180 185 190

Gln Val Val Ala Asn Thr Leu Gly Leu Lys Leu Asn Gln Tyr Phe Ser
195 200 205

Glu Ile Ser Leu Val Phe Gly Pro Asn Leu Arg Asp Pro Thr Ser Glu
210 215 220

```
Thr Leu Ser Ile Gln Ala Leu Ser Arg Ala Phe Asn Gly Asp Phe Asp
225                 230                 235                 240

Ser Met Leu Ser Lys Leu Lys Tyr Asp Asp Ser Asp Phe Leu Asp Leu
                245                 250                 255

Leu Glu Ser Asp Ser Ile Arg Gly Arg Ile Ile Asp Val Ser Leu Ser
            260                 265                 270

Asp Tyr Leu Ile Thr Ile Gln Ile Glu Tyr Pro Ala Leu Leu Ser Ile
        275                 280                 285

Lys Asp Ala Val Ile Gln Thr Phe Asn Leu Ile Ser Tyr Asn Thr Arg
    290                 295                 300

Gly Thr Glu Trp Ile Ser Ile Phe Pro Lys Gln Leu Leu Val Arg Gly
305                 310                 315                 320

Thr Tyr Ile Ser Asn Ile Asp Ile Ser Gln Cys Val Ile Ala Ala Thr
                325                 330                 335

Ser Ile Ile Cys Lys Ser Asp Thr Ser Thr Pro Ile Ser Ser Ala Thr
            340                 345                 350

Trp Ser Cys Ala Thr Gly Asn Ile Thr Asn Cys Ala Arg Thr Arg Val
        355                 360                 365

Val Asn Ala His Val Pro Arg Phe Ala Leu Tyr Gly Gly Val Val Phe
    370                 375                 380

Ala Asn Cys Ala Pro Val Val Cys Lys Cys Gln Asp Pro Leu Tyr Ser
385                 390                 395                 400

Ile Asn Gln Glu Pro Lys Val Thr Asn Val Met Val Asp Val Asp Ala
                405                 410                 415

Cys Lys Glu Met Tyr Leu Asp Gly Leu Tyr Ile Thr Leu Gly Lys Thr
            420                 425                 430

Gln Ile Ser Arg Ala Met Tyr Ala Glu Asp Val Ser Leu Gly Gly Pro
        435                 440                 445

Ile Ser Val Asp Pro Ile Asp Leu Gly Asn Glu Ile Asn Ser Ile Asn
    450                 455                 460

Ser Ala Ile Asn Arg Ser Glu Glu His Leu Asn His Ala Asn Glu Leu
465                 470                 475                 480

Leu Asp Lys Val Asn Pro Arg Ile Val Asn Val Lys Thr Phe Gly Val
                485                 490                 495
```

Met Ile Gly Leu Leu Val Leu Val Val Leu Trp Cys Val Ile Thr Leu
500 505 510

Val Trp Leu Ile Cys Leu Thr Lys Gln Leu Ala Arg Thr Ala Tyr Ala
515 520 525

Gly Ser Met Gly Ser Arg Ala Ser Thr Val Asn Ser Leu Ser Gly Phe
530 535 540

Val Gly
545

<210> 25
<211> 557
<212> PRT
<213> Porcine parainfluenza virus 1

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 25

Met Gln Val Thr Thr Leu Arg Pro Ala Ile Ile Leu Ser Ile Ala Leu
1 5 10 15

Leu Val Thr Gly Gln Val Pro Arg Asp Lys Leu Ala Asn Leu Gly Ile
20 25 30

Ile Ile Lys Asp Ser Lys Ala Leu Lys Ile Ala Gly Ser Tyr Glu Asn
35 40 45

Arg Tyr Ile Val Leu Ser Leu Val Pro Thr Ile Asp Asn Val Asn Gly
50 55 60

Cys Gly Ser Ile Gln Ile Ala Lys Tyr Lys Glu Met Leu Glu Arg Leu
65 70 75 80

Leu Ile Pro Ile Lys Asp Ala Leu Asp Leu Gln Glu Ser Leu Ile Val
85 90 95

Ile Asp Asn Glu Thr Val Asn Asn Asn Tyr Ser Pro Gln Tyr Arg Phe
100 105 110

Val Gly Ala Ile Ile Gly Thr Ile Ala Leu Gly Val Ala Thr Ala Ala
115 120 125

Gln Val Thr Ala Gly Val Ala Leu Met Glu Ala Arg Glu Ala Lys Arg
130 135 140

Asp Ile Ser Met Leu Lys Glu Ala Ile Glu Lys Thr Gln Asn Ser Ile
145 150 155 160

```
Glu Lys Leu Gln Asn Ser Ala Gly Glu Gln Ile Leu Ala Leu Lys Met
                165                 170                 175

Leu Gln Asp Tyr Val Asn Gly Glu Ile Lys Pro Ala Ile Glu Glu Leu
            180                 185                 190

Gly Cys Glu Thr Ala Ala Leu Lys Leu Gly Ile Ala Leu Thr Gln His
        195                 200                 205

Tyr Thr Glu Leu Thr Asn Ala Phe Gly Ser Asn Leu Gly Ser Ile Gly
    210                 215                 220

Glu Lys Ser Leu Thr Leu Gln Ala Leu Ser Ser Leu Tyr Lys Thr Asn
225                 230                 235                 240

Ile Thr Asn Ile Leu Thr Ala Thr Asn Leu Gly Lys Thr Asp Ile Tyr
                245                 250                 255

Asp Ile Ile Tyr Ala Glu Gln Val Lys Gly Arg Val Ile Asp Val Asp
            260                 265                 270

Leu Lys Arg Tyr Met Val Thr Ile Ser Val Lys Ile Pro Ile Leu Ser
        275                 280                 285

Glu Ile Pro Gly Val Leu Ile Tyr Glu Val Ser Ser Ile Ser Tyr Asn
    290                 295                 300

Ile Asp Gly Ala Glu Trp Tyr Ala Ala Val Pro Asp His Ile Leu Ser
305                 310                 315                 320

Lys Ser Ala Tyr Ile Gly Gly Ala Asp Ile Ser Asp Cys Ile Glu Ser
                325                 330                 335

Arg Leu Thr Tyr Ile Cys Pro Gln Asp Pro Ala Gln Ile Ile Ala Asp
            340                 345                 350

Asn Gln Gln Gln Cys Phe Phe Gly His Leu Asp Lys Cys Pro Ile Thr
        355                 360                 365

Lys Val Ile Asp Asn Leu Val Pro Lys Phe Ala Phe Ile Asn Gly Gly
    370                 375                 380

Val Val Ala Asn Cys Ile Ala Ser Thr Cys Thr Cys Gly Glu Glu Arg
385                 390                 395                 400

Ile Gln Val Ser Gln Asp Arg Asn Lys Gly Val Thr Phe Leu Thr His
                405                 410                 415

Asn Asn Cys Gly Leu Ile Gly Ile Asn Gly Ile Glu Phe His Ala Asn
            420                 425                 430

Lys Lys Gly Ser Asp Ala Thr Trp Asn Val Ser Pro Ile Gly Val Gly
```

435 440 445

Pro Ala Val Ser Leu Arg Pro Val Asp Ile Ser Leu Gln Ile Val Ala
450 455 460

Ala Thr Asn Phe Leu Asn Ser Ser Arg Lys Asp Leu Met Lys Ala Lys
465 470 475 480

Glu Ile Leu Asn Gln Val Gly Asn Leu Lys Asp Leu Thr Thr Ile Thr
485 490 495

Ile Ile Asn Ile Val Ile Ile Ile Ile Leu Leu Ile Cys Val Ile Gly
500 505 510

Leu Gly Ile Leu Tyr His Gln Leu Arg Ser Ala Leu Gly Met Arg Asp
515 520 525

Lys Met Ser Val Leu Asn Asn Ser Ser Tyr Ser Leu Glu Pro Arg Thr
530 535 540

Ala Gln Val Gln Val Ile Lys Pro Thr Ser Phe Met Gly
545 550 555

<210> 26
<211> 538
<212> PRT
<213> Avian metapneumovirus

<220>
<221> MISC_FEATURE
<223> F protein

<400> 26

Met Asp Val Arg Ile Cys Leu Leu Leu Phe Leu Ile Ser Asn Pro Ser
1 5 10 15

Ser Cys Ile Gln Glu Thr Tyr Asn Glu Glu Ser Cys Ser Thr Val Thr
20 25 30

Arg Gly Tyr Lys Ser Val Leu Arg Thr Gly Trp Tyr Thr Asn Val Phe
35 40 45

Asn Leu Glu Ile Gly Asn Val Glu Asn Ile Thr Cys Asn Asp Gly Pro
50 55 60

Ser Leu Ile Asp Thr Glu Leu Val Leu Thr Lys Asn Ala Leu Arg Glu
65 70 75 80

Leu Lys Thr Val Ser Ala Asp Gln Val Ala Lys Glu Ser Arg Leu Ser
85 90 95

Ser Pro Arg Arg Arg Arg Phe Val Leu Gly Ala Ile Ala Leu Gly Val
100 105 110

```
Ala Thr Ala Ala Ala Val Thr Ala Gly Val Ala Leu Ala Lys Thr Ile
        115                 120                 125

Arg Leu Glu Gly Glu Val Lys Ala Ile Lys Asn Ala Leu Arg Asn Thr
    130                 135                 140

Asn Glu Ala Val Ser Thr Leu Gly Asn Gly Val Arg Val Leu Ala Thr
145                 150                 155                 160

Ala Val Asn Asp Leu Lys Glu Phe Ile Ser Lys Lys Leu Thr Pro Ala
                165                 170                 175

Ile Asn Gln Asn Lys Cys Asn Ile Ala Asp Ile Lys Met Ala Ile Ser
            180                 185                 190

Phe Gly Gln Asn Asn Arg Arg Phe Leu Asn Val Val Arg Gln Phe Ser
        195                 200                 205

Asp Ser Ala Gly Ile Thr Ser Ala Val Ser Leu Asp Leu Met Thr Asp
    210                 215                 220

Asp Glu Leu Val Arg Ala Ile Asn Arg Met Pro Thr Ser Ser Gly Gln
225                 230                 235                 240

Ile Ser Leu Met Leu Asn Asn Arg Ala Met Val Arg Arg Lys Gly Phe
                245                 250                 255

Gly Ile Leu Ile Gly Val Tyr Asp Gly Thr Val Val Tyr Met Val Gln
            260                 265                 270

Leu Pro Ile Phe Gly Val Ile Glu Thr Pro Cys Trp Arg Val Val Ala
        275                 280                 285

Ala Pro Leu Cys Arg Lys Arg Arg Gly Asn Tyr Ala Cys Ile Leu Arg
    290                 295                 300

Glu Asp Gln Gly Trp Tyr Cys Thr Asn Ala Gly Ser Thr Ala Tyr Tyr
305                 310                 315                 320

Pro Asn Lys Asp Asp Cys Glu Val Arg Asp Asp Tyr Val Phe Cys Asp
                325                 330                 335

Thr Ala Ala Gly Ile Asn Val Ala Leu Glu Val Asp Gln Cys Asn Tyr
            340                 345                 350

Asn Ile Ser Thr Ser Lys Tyr Pro Cys Lys Val Ser Thr Gly Arg His
        355                 360                 365

Pro Val Ser Met Val Ala Leu Thr Pro Leu Gly Gly Leu Val Ser Cys
    370                 375                 380
```

```
Tyr Glu Ser Val Ser Cys Ser Ile Gly Ser Asn Lys Val Gly Ile Ile
385                 390                 395                 400


Lys Gln Leu Gly Lys Gly Cys Thr His Ile Pro Asn Asn Glu Ala Asp
                405                 410                 415


Thr Ile Thr Ile Asp Asn Thr Val Tyr Gln Leu Ser Lys Val Val Gly
            420                 425                 430


Glu Gln Arg Thr Ile Lys Gly Ala Pro Val Val Asn Asn Phe Asn Pro
        435                 440                 445


Ile Leu Phe Pro Val Asp Gln Phe Asn Val Ala Leu Asp Gln Val Phe
    450                 455                 460


Glu Ser Ile Asp Arg Ser Gln Asp Leu Ile Asp Lys Ser Asn Asp Leu
465                 470                 475                 480


Leu Gly Ala Asp Ala Lys Ser Lys Ala Gly Ile Ala Ile Ala Ile Val
                485                 490                 495


Val Leu Val Ile Leu Gly Ile Phe Phe Leu Leu Ala Val Ile Tyr Tyr
            500                 505                 510


Cys Ser Arg Val Arg Lys Thr Lys Pro Lys His Asp Tyr Pro Ala Thr
        515                 520                 525


Thr Gly His Ser Ser Met Ala Tyr Val Ser
    530                 535


<210>  27
<211>  545
<212>  PRT
<213>  Avian paramyxovirus 13 goose/Kazakhstan/5751/2013


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  27

Met Ala Arg Phe Ser Trp Glu Ile Phe Arg Leu Ser Thr Ile Leu Leu
1               5                   10                  15


Ile Ala Gln Thr Cys Gln Gly Ser Ile Asp Gly Arg Leu Thr Leu Ala
            20                  25                  30


Ala Gly Ile Val Pro Val Gly Asp Arg Pro Ile Ser Ile Tyr Thr Ser
        35                  40                  45


Ser Gln Thr Gly Ile Ile Val Val Lys Leu Ile Pro Asn Leu Pro Asp
    50                  55                  60
```

```
Asn Lys Lys Asp Cys Ala Lys Gln Ser Leu Gln Ser Tyr Asn Glu Thr
65                  70                  75                  80

Leu Ser Arg Ile Leu Thr Pro Leu Ala Thr Ala Met Ser Ala Ile Arg
                85                  90                  95

Gly Asn Ser Thr Thr Gln Val Arg Glu Asn Arg Leu Val Gly Ala Ile
            100                 105                 110

Ile Gly Ser Val Ala Leu Gly Val Ala Thr Ala Ala Gln Ile Thr Ala
        115                 120                 125

Ala Thr Ala Leu Ile Gln Ala Asn Gln Asn Ala Ala Asn Ile Ala Arg
    130                 135                 140

Leu Ala Asn Ser Ile Ala Lys Thr Asn Glu Ala Val Thr Asp Leu Thr
145                 150                 155                 160

Glu Gly Leu Gly Thr Leu Ala Ile Gly Val Gly Lys Leu Gln Asp Tyr
                165                 170                 175

Val Asn Glu Gln Phe Asn Asn Thr Ala Val Ala Ile Asp Cys Leu Thr
            180                 185                 190

Leu Glu Ser Arg Leu Gly Ile Gln Leu Ser Leu Tyr Leu Thr Glu Leu
        195                 200                 205

Met Gly Val Phe Gly Asn Gln Leu Thr Ser Pro Ala Leu Thr Pro Ile
    210                 215                 220

Thr Ile Gln Ala Leu Tyr Asn Leu Ala Gly Gly Asn Leu Asn Ala Leu
225                 230                 235                 240

Leu Ser Arg Leu Gly Ala Ser Glu Thr Gln Leu Gly Ser Leu Ile Asn
                245                 250                 255

Ser Gly Leu Ile Lys Gly Met Pro Ile Met Tyr Asp Asp Ala Asn Lys
            260                 265                 270

Leu Leu Ala Val Gln Val Glu Leu Pro Ser Ile Gly Lys Leu Asn Gly
        275                 280                 285

Ala Arg Ser Thr Leu Leu Glu Thr Leu Ala Val Asp Thr Thr Arg Gly
    290                 295                 300

Pro Ser Ser Pro Ile Ile Pro Ser Ala Val Ile Glu Ile Gly Gly Ala
305                 310                 315                 320

Met Glu Glu Leu Asp Leu Ser Pro Cys Ile Thr Thr Asp Leu Asp Met
                325                 330                 335

Phe Cys Thr Lys Ile Ile Ser Tyr Pro Leu Ser Gln Ser Thr Leu Ser
```

340 345 350

Cys Leu Asn Gly Asn Leu Ser Asp Cys Val Phe Ser Arg Ser Glu Gly
355 360 365

Val Leu Ser Thr Pro Tyr Met Thr Ile Lys Gly Lys Ile Val Ala Asn
370 375 380

Cys Lys Gln Val Ile Cys Arg Cys Met Asp Pro Pro Gln Ile Leu Ser
385 390 395 400

Gln Asn Tyr Gly Glu Ala Leu Leu Leu Ile Asp Glu Asn Thr Cys Arg
405 410 415

Ser Leu Glu Leu Ser Gly Val Ile Leu Lys Leu Ala Gly Thr Tyr Glu
420 425 430

Ser Glu Tyr Thr Arg Asn Leu Thr Val Asp Pro Ser Gln Val Ile Ile
435 440 445

Thr Gly Pro Leu Asp Ile Ser Ala Glu Leu Ser Lys Val Asn Gln Ser
450 455 460

Ile Asp Ser Ala Lys Glu Asn Ile Ala Glu Ser Asn Lys Phe Leu Ser
465 470 475 480

Gln Val Asn Val Lys Leu Leu Ser Ser Ser Ala Met Ile Thr Tyr Ile
485 490 495

Val Ala Thr Val Val Cys Leu Ile Ile Ala Ile Thr Gly Cys Val Ile
500 505 510

Gly Ile Tyr Thr Leu Thr Lys Leu Lys Ser Gln Gln Lys Thr Leu Leu
515 520 525

Trp Leu Gly Asn Asn Ala Glu Met His Gly Ser Arg Ser Lys Thr Ser
530 535 540

Phe
545

<210> 28
<211> 554
<212> PRT
<213> Menangle virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 28

Met Met Pro Arg Val Leu Gly Met Ile Val Leu Tyr Leu Thr His Ser
1 5 10 15

```
Gln Ile Leu Cys Ile Asn Arg Asn Thr Leu Tyr Gln Ile Gly Leu Ile
            20                  25                  30

His Arg Ser Val Lys Lys Val Asn Phe Tyr Ser Gln Gly Ser Pro Ser
        35                  40                  45

Tyr Ile Val Val Lys Leu Val Pro Thr Leu Ala Ala Ile Pro Pro Asn
    50                  55                  60

Cys Ser Ile Lys Ser Leu Gln Arg Tyr Lys Glu Thr Val Thr Ser Leu
65                  70                  75                  80

Val Gln Pro Ile Ser Asp Asn Leu Gly Tyr Leu Gln Asp Lys Leu Val
                85                  90                  95

Thr Gly Gln Ser Arg Arg Arg Arg Arg Phe Ala Gly Val Ala Ile Gly
            100                 105                 110

Leu Ala Ala Leu Gly Val Ala Ala Ala Ala Gln Ala Thr Ala Ala Val
        115                 120                 125

Ala Leu Val Glu Thr Arg Glu Asn Ala Gly Lys Ile Gln Ala Leu Ser
    130                 135                 140

Glu Ser Ile Gln Asn Thr Asn Gln Ala Val His Ser Leu Lys Thr Ala
145                 150                 155                 160

Leu Gly Phe Ser Ala Thr Ala Ile Gln Ala Ile Gln Asn Gln Val Asn
                165                 170                 175

Glu Val Ile Asn Pro Ala Ile Asn Lys Leu Ser Cys Glu Val Leu Asp
            180                 185                 190

Ser Gln Leu Ala Ser Met Leu Asn Leu Tyr Leu Ile His Leu Thr Thr
        195                 200                 205

Val Phe Gln Thr Gln Leu Thr Asn Pro Ala Leu Thr Pro Leu Ser Ile
    210                 215                 220

Gln Ala Leu Thr Ser Val Leu Gln Gly Thr Ser Gly Val Leu Met Asn
225                 230                 235                 240

Ser Thr Asn Ser Thr Leu Thr Gln Pro Ile Asp Leu Leu Ala Thr Gly
                245                 250                 255

Leu Ile Thr Gly Gln Ile Ile Ser Val Asn Met Thr Ser Leu Gln Leu
            260                 265                 270

Ile Ile Ala Thr Phe Met Pro Ser Ile Ala Glu Leu Pro Asn Ala Val
        275                 280                 285
```

```
Leu His Ser Phe Phe Arg Ile Thr Thr Ser Val Asn Leu Thr Glu Val
    290                 295                 300

Met Ile Gln Ser Pro Glu Phe Ile Met Glu Gln Asn Gly Val Phe Tyr
305                 310                 315                 320

Asp Phe Asn Thr Ala His Cys Gln Leu Gly Asp Asn Asn Val Tyr Cys
                325                 330                 335

Pro Tyr Ile Asp Ala Ala Arg Leu Ser Ser Met Met Thr Asn Cys Ile
            340                 345                 350

Asn Gly Asn Leu Gly Glu Cys Val Phe Ser Arg Val Ile Gly Ser Phe
        355                 360                 365

Pro Ser Arg Phe Val Ser Leu Asn Gly Ala Ile Leu Ala Asn Cys Lys
    370                 375                 380

Phe Met Arg Cys Asn Cys Leu Ser Pro Glu Lys Ile Ile Thr Pro Leu
385                 390                 395                 400

Asp Gly Glu Met Ile Ser Leu Ile Asp Leu Arg Val Cys Gln Lys Leu
                405                 410                 415

Thr Leu Gly Thr Ile Thr Phe Glu Ile Ser Gln Pro Val Asn Val Ser
            420                 425                 430

Phe Gln Gly Gly Phe Val Ala Asn Ala Gly Gln Ile Ile Val Thr Asn
        435                 440                 445

Pro Phe Asp Ile Ser Ala Glu Leu Gly Gln Ile Asn Asn Ser Leu Asn
    450                 455                 460

Asp Ala Gln Gly Phe Leu Asp Gln Ser Asn Asn Trp Leu Lys Val Ser
465                 470                 475                 480

Gly Trp Ile Asn Asn Ser Gly Ser Leu Phe Ile Ala Gly Ile Val Val
                485                 490                 495

Ile Gly Leu Ile Val Leu Cys Ile Val Ile Ile Ile Tyr Ile Asn Val
            500                 505                 510

Gln Ile Ile Arg Glu Val Asn Arg Leu Arg Ser Phe Ile Tyr Arg Asp
        515                 520                 525

Tyr Val Leu Asp His Asp Lys Ala Pro Tyr Ser Pro Glu Ser Ser Ser
    530                 535                 540

Pro His Arg Lys Ser Leu Lys Thr Val Ser
545                 550
```

<210> 29
<211> 544
<212> PRT
<213> Avian paramyxovirus 5

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 29

Met Leu Gln Leu Pro Leu Thr Ile Leu Leu Ser Ile Leu Ser Ala His
1 5 10 15

Gln Ser Leu Cys Leu Asp Asn Ser Lys Leu Ile His Ala Gly Ile Met
20 25 30

Ser Thr Thr Glu Arg Glu Val Asn Val Tyr Ala Gln Ser Ile Thr Gly
35 40 45

Ser Ile Val Val Arg Leu Ile Pro Asn Ile Pro Ser Asn His Lys Ser
50 55 60

Cys Ala Thr Ser Gln Ile Lys Leu Tyr Asn Asp Thr Leu Thr Arg Leu
65 70 75 80

Leu Thr Pro Ile Lys Ala Asn Leu Glu Gly Leu Ile Ser Ala Val Ser
85 90 95

Gln Asp Gln Ser Gln Asn Ser Gly Lys Arg Lys Lys Arg Phe Val Gly
100 105 110

Ala Val Ile Gly Ala Ala Ala Leu Gly Leu Ala Thr Ala Ala Gln Val
115 120 125

Thr Ala Thr Val Ala Leu Asn Gln Ala Gln Glu Asn Ala Arg Asn Ile
130 135 140

Leu Arg Leu Lys Asn Ser Ile Gln Lys Thr Asn Glu Ala Val Met Glu
145 150 155 160

Leu Lys Asp Ala Val Gly Gln Thr Ala Val Ala Ile Asp Lys Thr Gln
165 170 175

Ala Phe Ile Asn Asn Gln Ile Leu Pro Ala Ile Ser Asn Leu Ser Cys
180 185 190

Glu Val Leu Gly Asn Lys Ile Gly Val Gln Leu Ser Leu Tyr Leu Thr
195 200 205

Glu Leu Thr Thr Val Phe Gly Asn Gln Leu Thr Asn Pro Ala Leu Thr
210 215 220

Thr Leu Ser Leu Gln Ala Leu Tyr Asn Leu Cys Gly Asp Asp Phe Asn

225 230 235 240

Tyr Leu Ile Asn Leu Leu Asn Ala Lys Asn Arg Asn Leu Ala Ser Leu
245 250 255

Tyr Glu Ala Asn Leu Ile Gln Gly Arg Ile Thr Gln Tyr Asp Ser Met
260 265 270

Asn Gln Leu Leu Ile Ile Gln Val Gln Ile Pro Ser Ile Ser Thr Val
275 280 285

Ser Gly Met Arg Val Thr Glu Leu Phe Thr Leu Ser Val Asp Thr Pro
290 295 300

Ile Gly Glu Gly Lys Ala Leu Val Pro Lys Tyr Val Leu Ser Ser Gly
305 310 315 320

Arg Ile Met Glu Glu Val Asp Leu Ser Ser Cys Ala Ile Thr Ser Thr
325 330 335

Ser Val Phe Cys Ser Ser Ile Ile Ser Arg Pro Leu Pro Leu Glu Thr
340 345 350

Ile Asn Cys Leu Asn Gly Asn Val Thr Gln Cys Gln Phe Thr Ala Asn
355 360 365

Thr Gly Thr Leu Glu Ser Arg Tyr Ala Val Ile Gly Gly Leu Val Ile
370 375 380

Ala Asn Cys Lys Ala Ile Val Cys Arg Cys Leu Asn Pro Pro Gly Val
385 390 395 400

Ile Ala Gln Asn Leu Gly Leu Pro Ile Thr Ile Ile Ser Ser Asn Thr
405 410 415

Cys Gln Arg Ile Asn Leu Glu Gln Ile Thr Leu Ser Leu Gly Asn Ser
420 425 430

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

Ala Pro Ser Asn Pro Leu Asp Ile Ser Val Glu Leu Asn Arg Val Asn
450 455 460

Thr Ser Leu Ser Lys Val Glu Ser Leu Ile Lys Glu Ser Asn Ser Ile
465 470 475 480

Leu Asp Ser Val Asn Pro Gln Ile Leu Asn Val Lys Thr Val Ile Ile
485 490 495

Leu Ala Val Ile Ile Gly Leu Ile Val Val Trp Cys Phe Ile Leu Thr
500 505 510

```
Cys Leu Ile Val Arg Gly Phe Met Leu Leu Val Lys Gln Gln Lys Phe
        515                 520                 525


Lys Gly Leu Ser Val Gln Asn Asn Pro Tyr Val Ser Asn Asn Ser His
    530                 535                 540


<210>  30
<211>  557
<212>  PRT
<213>  Cedar virus


<220>
<221>  MISC_FEATURE
<223>  fusion glycoprotein

<400>  30

Met Ser Asn Lys Arg Thr Thr Val Leu Ile Ile Ile Ser Tyr Thr Leu
1               5                   10                  15


Phe Tyr Leu Asn Asn Ala Ala Ile Val Gly Phe Asp Phe Asp Lys Leu
            20                  25                  30


Asn Lys Ile Gly Val Val Gln Gly Arg Val Leu Asn Tyr Lys Ile Lys
        35                  40                  45


Gly Asp Pro Met Thr Lys Asp Leu Val Leu Lys Phe Ile Pro Asn Ile
    50                  55                  60


Val Asn Ile Thr Glu Cys Val Arg Glu Pro Leu Ser Arg Tyr Asn Glu
65                  70                  75                  80


Thr Val Arg Arg Leu Leu Leu Pro Ile His Asn Met Leu Gly Leu Tyr
                85                  90                  95


Leu Asn Asn Thr Asn Ala Lys Met Thr Gly Leu Met Ile Ala Gly Val
            100                 105                 110


Ile Met Gly Gly Ile Ala Ile Gly Ile Ala Thr Ala Ala Gln Ile Thr
        115                 120                 125


Ala Gly Phe Ala Leu Tyr Glu Ala Lys Lys Asn Thr Glu Asn Ile Gln
    130                 135                 140


Lys Leu Thr Asp Ser Ile Met Lys Thr Gln Asp Ser Ile Asp Lys Leu
145                 150                 155                 160


Thr Asp Ser Val Gly Thr Ser Ile Leu Ile Leu Asn Lys Leu Gln Thr
                165                 170                 175


Tyr Ile Asn Asn Gln Leu Val Pro Asn Leu Glu Leu Leu Ser Cys Arg
            180                 185                 190
```

```
Gln Asn Lys Ile Glu Phe Asp Leu Met Leu Thr Lys Tyr Leu Val Asp
        195                 200                 205

Leu Met Thr Val Ile Gly Pro Asn Ile Asn Asn Pro Val Asn Lys Asp
    210                 215                 220

Met Thr Ile Gln Ser Leu Ser Leu Leu Phe Asp Gly Asn Tyr Asp Ile
225                 230                 235                 240

Met Met Ser Glu Leu Gly Tyr Thr Pro Gln Asp Phe Leu Asp Leu Ile
                245                 250                 255

Glu Ser Lys Ser Ile Thr Gly Gln Ile Ile Tyr Val Asp Met Glu Asn
            260                 265                 270

Leu Tyr Val Val Ile Arg Thr Tyr Leu Pro Thr Leu Ile Glu Val Pro
        275                 280                 285

Asp Ala Gln Ile Tyr Glu Phe Asn Lys Ile Thr Met Ser Ser Asn Gly
    290                 295                 300

Gly Glu Tyr Leu Ser Thr Ile Pro Asn Phe Ile Leu Ile Arg Gly Asn
305                 310                 315                 320

Tyr Met Ser Asn Ile Asp Val Ala Thr Cys Tyr Met Thr Lys Ala Ser
                325                 330                 335

Val Ile Cys Asn Gln Asp Tyr Ser Leu Pro Met Ser Gln Asn Leu Arg
            340                 345                 350

Ser Cys Tyr Gln Gly Glu Thr Glu Tyr Cys Pro Val Glu Ala Val Ile
        355                 360                 365

Ala Ser His Ser Pro Arg Phe Ala Leu Thr Asn Gly Val Ile Phe Ala
    370                 375                 380

Asn Cys Ile Asn Thr Ile Cys Arg Cys Gln Asp Asn Gly Lys Thr Ile
385                 390                 395                 400

Thr Gln Asn Ile Asn Gln Phe Val Ser Met Ile Asp Asn Ser Thr Cys
                405                 410                 415

Asn Asp Val Met Val Asp Lys Phe Thr Ile Lys Val Gly Lys Tyr Met
            420                 425                 430

Gly Arg Lys Asp Ile Asn Asn Ile Asn Ile Gln Ile Gly Pro Gln Ile
        435                 440                 445

Ile Ile Asp Lys Val Asp Leu Ser Asn Glu Ile Asn Lys Met Asn Gln
    450                 455                 460
```

Ser Leu Lys Asp Ser Ile Phe Tyr Leu Arg Glu Ala Lys Arg Ile Leu
465 470 475 480

Asp Ser Val Asn Ile Ser Leu Ile Ser Pro Ser Val Gln Leu Phe Leu
485 490 495

Ile Ile Ile Ser Val Leu Ser Phe Ile Ile Leu Leu Ile Ile Ile Val
500 505 510

Tyr Leu Tyr Cys Lys Ser Lys His Ser Tyr Lys Tyr Asn Lys Phe Ile
515 520 525

Asp Asp Pro Asp Tyr Tyr Asn Asp Tyr Lys Arg Glu Arg Ile Asn Gly
530 535 540

Lys Ala Ser Lys Ser Asn Asn Ile Tyr Tyr Val Gly Asp
545 550 555

<210> 31
<211> 546
<212> PRT
<213> Rinderpest morbillivirus

<220>
<221> MISC_FEATURE
<223> F protein

<400> 31

Met Gly Ile Leu Phe Ala Ala Leu Leu Ala Met Thr Asn Pro His Leu
1 5 10 15

Ala Thr Gly Gln Ile His Trp Gly Asn Leu Ser Lys Ile Gly Val Val
20 25 30

Gly Thr Gly Ser Ala Ser Tyr Lys Val Met Thr Gln Ser Ser His Gln
35 40 45

Ser Leu Val Ile Lys Leu Met Pro Asn Val Thr Ala Ile Asp Asn Cys
50 55 60

Thr Lys Thr Glu Ile Met Glu Tyr Lys Arg Leu Leu Gly Thr Val Leu
65 70 75 80

Lys Pro Ile Arg Glu Ala Leu Asn Ala Ile Thr Lys Asn Ile Lys Pro
85 90 95

Ile Gln Ser Ser Thr Thr Ser Arg Arg His Lys Arg Phe Ala Gly Val
100 105 110

Val Leu Ala Gly Ala Ala Leu Gly Val Ala Thr Ala Ala Gln Ile Thr
115 120 125

Ala Gly Ile Ala Leu His Gln Ser Met Met Asn Ser Gln Ala Ile Glu

130 135 140

Ser Leu Lys Ala Ser Leu Glu Thr Thr Asn Gln Ala Ile Glu Glu Ile
145 150 155 160

Arg Gln Ala Gly Gln Glu Met Val Leu Ala Val Gln Gly Val Gln Asp
165 170 175

Tyr Ile Asn Asn Glu Leu Val Pro Ala Met Gly Gln Leu Ser Cys Glu
180 185 190

Ile Val Gly Gln Lys Leu Gly Leu Lys Leu Leu Arg Tyr Tyr Thr Glu
195 200 205

Ile Leu Ser Leu Phe Gly Pro Ser Leu Arg Asp Pro Val Ser Ala Glu
210 215 220

Leu Ser Ile Gln Ala Leu Ser Tyr Ala Leu Gly Gly Asp Ile Asn Lys
225 230 235 240

Ile Leu Glu Lys Leu Gly Tyr Ser Gly Ser Asp Leu Leu Ala Ile Leu
245 250 255

Glu Ser Lys Gly Ile Lys Ala Lys Ile Thr Tyr Val Asp Ile Glu Ser
260 265 270

Tyr Phe Ile Val Leu Ser Ile Ala Tyr Pro Ser Leu Ser Glu Ile Lys
275 280 285

Gly Val Ile Val His Arg Leu Glu Ser Val Ser Tyr Asn Ile Gly Ser
290 295 300

Gln Glu Trp Tyr Thr Thr Val Pro Arg Tyr Val Ala Thr Gln Gly Tyr
305 310 315 320

Leu Ile Ser Asn Phe Asp Asp Thr Pro Cys Ala Phe Thr Pro Glu Gly
325 330 335

Thr Ile Cys Ser Gln Asn Ala Leu Tyr Pro Met Ser Pro Leu Leu Gln
340 345 350

Glu Cys Phe Arg Gly Ser Thr Arg Ser Cys Ala Arg Thr Leu Val Ser
355 360 365

Gly Ser Ile Gly Asn Arg Phe Ile Leu Ser Lys Gly Asn Leu Ile Ala
370 375 380

Asn Cys Ala Ser Ile Leu Cys Lys Cys Tyr Thr Thr Gly Ser Ile Ile
385 390 395 400

Ser Gln Asp Pro Asp Lys Ile Leu Thr Tyr Ile Ala Ala Asp Gln Cys
405 410 415

```
Pro Val Val Glu Val Gly Gly Val Thr Ile Gln Val Gly Ser Arg Glu
            420                 425                 430


Tyr Ser Asp Ala Val Tyr Leu His Glu Ile Asp Leu Gly Pro Pro Ile
        435                 440                 445


Ser Leu Glu Lys Leu Asp Val Gly Thr Asn Leu Trp Asn Ala Val Thr
    450                 455                 460


Lys Leu Glu Lys Ala Lys Asp Leu Leu Asp Ser Ser Asp Leu Ile Leu
465                 470                 475                 480


Glu Asn Ile Lys Gly Val Ser Val Thr Asn Thr Gly Tyr Ile Leu Val
                485                 490                 495


Gly Val Gly Leu Ile Ala Val Val Gly Ile Leu Ile Ile Thr Cys Cys
            500                 505                 510


Cys Lys Lys Arg Arg Ser Asp Asn Lys Val Ser Thr Met Val Leu Asn
        515                 520                 525


Pro Gly Leu Arg Pro Asp Leu Thr Gly Thr Ser Lys Ser Tyr Val Arg
    530                 535                 540


Ser Leu
545


<210>  32
<211>  536
<212>  PRT
<213>  Avian paramyxovirus 10


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  32

Met Thr Arg Thr Arg Leu Leu Phe Leu Leu Thr Cys Tyr Ile Pro Gly
1               5                   10                  15


Ala Val Ser Leu Asp Asn Ser Ile Leu Ala Pro Ala Gly Ile Ile Ser
            20                  25                  30


Ala Ser Glu Arg Gln Ile Ala Ile Tyr Thr Gln Thr Leu Gln Gly Thr
        35                  40                  45


Ile Ala Leu Arg Phe Ile Pro Val Leu Pro Gln Asn Leu Ser Ser Cys
    50                  55                  60


Ala Lys Asp Thr Leu Glu Ser Tyr Asn Ser Thr Val Ser Asn Leu Leu
65                  70                  75                  80
```

```
Leu Pro Ile Ala Glu Asn Leu Asn Ala Leu Leu Lys Asp Ala Asp Lys
                85                  90                  95

Pro Ser Gln Arg Ile Ile Gly Ala Ile Ile Gly Ser Val Ala Leu Gly
            100                 105                 110

Val Ala Thr Thr Ala Gln Val Thr Ala Ala Leu Ala Met Thr Gln Ala
        115                 120                 125

Gln Gln Asn Ala Arg Asn Ile Trp Lys Leu Lys Glu Ser Ile Lys Asn
    130                 135                 140

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

Ile Ala Leu Asp Lys Val Gln Ser Phe Ile Asn Ser Glu Ile Leu Pro
                165                 170                 175

Gln Ile Asn Gln Leu Gly Cys Glu Val Ala Ala Asn Lys Leu Gly Ile
            180                 185                 190

Phe Leu Ser Leu Tyr Leu Thr Glu Ile Thr Thr Val Phe Lys Asn Gln
        195                 200                 205

Ile Thr Asn Pro Ala Leu Ser Thr Leu Ser Tyr Gln Ala Leu Tyr Asn
    210                 215                 220

Leu Cys Gly Gly Asn Met Ala Ala Leu Thr Lys Gln Ile Gly Ile Lys
225                 230                 235                 240

Asp Thr Glu Ile Asn Ser Leu Tyr Glu Ala Glu Leu Ile Thr Gly Gln
                245                 250                 255

Val Ile Gly Tyr Asp Ser Ala Asp Gln Ile Leu Leu Ile Gln Val Ser
            260                 265                 270

Tyr Pro Ser Val Ser Arg Val Gln Gly Val Arg Ala Val Glu Leu Leu
        275                 280                 285

Thr Val Ser Val Ala Thr Pro Lys Gly Glu Gly Lys Ala Ile Ala Pro
    290                 295                 300

Ser Phe Ile Ala Gln Ser Asn Ile Ile Ala Glu Glu Leu Asp Thr Gln
305                 310                 315                 320

Pro Cys Lys Phe Ser Lys Thr Thr Leu Tyr Cys Arg Gln Val Asn Thr
                325                 330                 335

Arg Thr Leu Pro Val Arg Val Ala Asn Cys Leu Lys Gly Lys Tyr Asn
            340                 345                 350
```

```
Asp Cys Gln Tyr Thr Thr Glu Ile Gly Ala Leu Ala Ser Arg Tyr Val
        355                 360                 365


Thr Ile Thr Asn Gly Val Val Ala Asn Cys Arg Ser Ile Ile Cys Arg
    370                 375                 380


Cys Leu Asp Pro Glu Gly Ile Val Ala Gln Asn Ser Asp Ala Ala Ile
385                 390                 395                 400


Thr Val Ile Asp Arg Ser Thr Cys Lys Leu Ile Gln Leu Gly Asp Ile
                405                 410                 415


Thr Leu Arg Leu Glu Gly Lys Leu Ser Ser Ser Tyr Ser Lys Asn Ile
            420                 425                 430


Thr Ile Asp Ile Ser Gln Val Thr Thr Ser Gly Ser Leu Asp Ile Ser
        435                 440                 445


Ser Glu Leu Gly Ser Ile Asn Asn Thr Ile Thr Lys Val Glu Asp Leu
    450                 455                 460


Ile Ser Lys Ser Asn Asp Trp Leu Ser Lys Val Asn Pro Thr Leu Ile
465                 470                 475                 480


Ser Asn Asp Thr Ile Ile Ala Leu Cys Val Ile Ala Gly Ile Val Val
                485                 490                 495


Ile Trp Leu Val Ile Ile Thr Ile Leu Ser Tyr Tyr Ile Leu Ile Lys
            500                 505                 510


Leu Lys Asn Val Ala Leu Leu Ser Thr Met Pro Lys Lys Asp Leu Asn
        515                 520                 525


Pro Tyr Val Asn Asn Thr Lys Phe
    530                 535


<210>  33
<211>  552
<212>  PRT
<213>  Dolphin morbillivirus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  33

Met Ala Ala Ser Asn Gly Gly Val Met Tyr Gln Ser Phe Leu Thr Ile
1               5                   10                  15


Ile Ile Leu Val Ile Met Thr Glu Gly Gln Ile His Trp Gly Asn Leu
            20                  25                  30


Ser Lys Ile Gly Ile Val Gly Thr Gly Ser Ala Ser Tyr Lys Val Met
```

35 40 45

Thr Arg Pro Asn His Gln Tyr Leu Val Ile Lys Leu Met Pro Asn Val
50 55 60

Thr Met Ile Asp Asn Cys Thr Arg Thr Glu Val Thr Glu Tyr Arg Lys
65 70 75 80

Leu Leu Lys Thr Val Leu Glu Pro Val Lys Asn Ala Leu Thr Val Ile
85 90 95

Thr Lys Asn Ile Lys Pro Ile Gln Ser Leu Thr Thr Ser Arg Arg Ser
100 105 110

Lys Arg Phe Ala Gly Val Val Leu Ala Gly Val Ala Leu Gly Val Ala
115 120 125

Thr Ala Ala Gln Ile Thr Ala Gly Val Ala Leu His Gln Ser Ile Met
130 135 140

Asn Ser Gln Ser Ile Asp Asn Leu Arg Thr Ser Leu Glu Lys Ser Asn
145 150 155 160

Gln Ala Ile Glu Glu Ile Arg Gln Ala Ser Gln Glu Thr Val Leu Ala
165 170 175

Val Gln Gly Val Gln Asp Phe Ile Asn Asn Glu Leu Ile Pro Ser Met
180 185 190

His Gln Leu Ser Cys Glu Met Leu Gly Gln Lys Leu Gly Leu Lys Leu
195 200 205

Leu Arg Tyr Tyr Thr Glu Ile Leu Ser Ile Phe Gly Pro Ser Leu Arg
210 215 220

Asp Pro Val Ser Ala Glu Ile Ser Ile Gln Ala Leu Ser Tyr Ala Leu
225 230 235 240

Gly Gly Asp Ile Asn Lys Ile Leu Glu Lys Leu Gly Tyr Ser Gly Ala
245 250 255

Asp Leu Leu Ala Ile Leu Glu Ser Arg Gly Ile Lys Ala Lys Val Thr
260 265 270

His Val Asp Leu Glu Gly Tyr Phe Ile Val Leu Ser Ile Ala Tyr Pro
275 280 285

Thr Leu Ser Glu Val Lys Gly Val Ile Val His Lys Leu Glu Ala Val
290 295 300

Ser Tyr Asn Leu Gly Ser Gln Glu Trp Tyr Thr Thr Leu Pro Lys Tyr
305 310 315 320

```
Val Ala Thr Asn Gly Tyr Leu Ile Ser Asn Phe Asp Glu Ser Ser Cys
                325                 330                 335

Ala Phe Met Ser Glu Val Thr Ile Cys Ser Gln Asn Ala Leu Tyr Pro
            340                 345                 350

Met Ser Pro Leu Leu Gln Gln Cys Leu Arg Gly Ser Thr Ala Ser Cys
        355                 360                 365

Ala Arg Ser Leu Val Ser Gly Thr Ile Gly Asn Arg Phe Ile Leu Ser
    370                 375                 380

Lys Gly Asn Leu Ile Ala Asn Cys Ala Ser Val Leu Cys Lys Cys Tyr
385                 390                 395                 400

Ser Thr Gly Thr Ile Ile Ser Gln Asp Pro Asp Lys Leu Leu Thr Phe
                405                 410                 415

Val Ala Ala Asp Lys Cys Pro Leu Val Glu Val Asp Gly Ile Thr Ile
            420                 425                 430

Gln Val Gly Ser Arg Glu Tyr Pro Asp Ser Val Tyr Val Ser Arg Ile
        435                 440                 445

Asp Leu Gly Pro Ala Ile Ser Leu Glu Lys Leu Asp Val Gly Thr Asn
    450                 455                 460

Leu Gly Ser Ala Leu Thr Lys Leu Asp Asn Ala Lys Asp Leu Leu Asp
465                 470                 475                 480

Ser Ser Asn Gln Ile Leu Glu Asn Val Arg Arg Ser Ser Phe Gly Gly
                485                 490                 495

Ala Met Tyr Ile Gly Ile Leu Val Cys Ala Gly Ala Leu Val Ile Leu
            500                 505                 510

Cys Val Leu Val Tyr Cys Cys Arg Arg His Cys Arg Lys Arg Val Gln
        515                 520                 525

Thr Pro Pro Lys Ala Thr Pro Gly Leu Lys Pro Asp Leu Thr Gly Thr
    530                 535                 540

Thr Lys Ser Tyr Val Arg Ser Leu
545                 550


<210>  34
<211>  552
<212>  PRT
<213>  Mossman virus


<220>
```

<221> MISC_FEATURE
<223> fusion protein

<400> 34

```
Met Ser Asn Tyr Phe Pro Ala Arg Val Ile Ile Ile Val Ser Leu Ile
1               5                   10                  15


Thr Ala Val Ser Cys Gln Ile Ser Phe Gln Asn Leu Ser Thr Ile Gly
            20                  25                  30


Val Phe Lys Phe Lys Glu Tyr Asp Tyr Arg Val Ser Gly Asp Tyr Asn
        35                  40                  45


Glu Gln Phe Leu Ala Ile Lys Met Val Pro Asn Val Thr Gly Val Glu
    50                  55                  60


Asn Cys Thr Ala Ser Leu Ile Asp Glu Tyr Arg His Val Ile Tyr Asn
65                  70                  75                  80


Leu Leu Gln Pro Ile Asn Thr Thr Leu Thr Ala Ser Thr Ser Asn Val
                85                  90                  95


Asp Pro Tyr Ala Gly Asn Lys Lys Phe Phe Gly Ala Val Ile Ala Gly
            100                 105                 110


Val Ala Leu Gly Val Ala Thr Ala Ala Gln Val Thr Ala Gly Val Ala
        115                 120                 125


Leu Tyr Glu Ala Arg Gln Asn Ala Ala Ala Ile Ala Glu Ile Lys Glu
    130                 135                 140


Ser Leu His Tyr Thr His Lys Ala Ile Glu Ser Leu Gln Ile Ser Gln
145                 150                 155                 160


Lys Gln Thr Val Val Ala Ile Gln Gly Ile Gln Asp Gln Ile Asn Thr
                165                 170                 175


Asn Ile Ile Pro Gln Ile Asn Ala Leu Thr Cys Glu Ile Ala Asn Gln
            180                 185                 190


Arg Leu Arg Leu Met Leu Leu Gln Tyr Tyr Thr Glu Met Leu Ser Ser
        195                 200                 205


Phe Gly Pro Ile Ile Gln Asp Pro Leu Ser Gly His Ile Thr Val Gln
    210                 215                 220


Ala Leu Ser Gln Ala Ala Gly Gly Asn Ile Thr Gly Leu Met Arg Glu
225                 230                 235                 240


Leu Gly Tyr Ser Ser Lys Asp Leu Arg Tyr Ile Leu Ser Val Asn Gly
                245                 250                 255
```

```
Ile Ser Ala Asn Ile Ile Asp Ala Asp Pro Glu Ile Gly Ser Ile Ile
            260                 265                 270

Leu Arg Ile Arg Tyr Pro Ser Met Ile Lys Ile Pro Asp Val Ala Val
        275                 280                 285

Met Glu Leu Ser Tyr Leu Ala Tyr His Ala Ala Gly Gly Asp Trp Leu
    290                 295                 300

Thr Val Gly Pro Arg Phe Ile Leu Lys Arg Gly Tyr Ser Leu Ser Asn
305                 310                 315                 320

Leu Asp Ile Thr Ser Cys Thr Ile Gly Glu Asp Phe Leu Leu Cys Ser
                325                 330                 335

Lys Asp Val Ser Ser Pro Met Ser Leu Ala Thr Gln Ser Cys Leu Arg
            340                 345                 350

Gly Asp Thr Gln Met Cys Ser Arg Thr Ala Val Gln Asp Arg Glu Ala
        355                 360                 365

Pro Arg Phe Leu Leu Leu Gln Gly Asn Leu Ile Val Asn Cys Met Ser
    370                 375                 380

Val Asn Cys Lys Cys Glu Asp Pro Glu Glu Thr Ile Thr Gln Asp Pro
385                 390                 395                 400

Ala Tyr Pro Leu Met Val Leu Gly Ser Asp Thr Cys Lys Ile His Tyr
                405                 410                 415

Ile Asp Gly Ile Arg Ile Lys Leu Gly Lys Val Gln Leu Pro Pro Ile
            420                 425                 430

Thr Val Leu Asn Thr Leu Ser Leu Gly Pro Ile Val Val Leu Asn Pro
        435                 440                 445

Ile Asp Val Ser Asn Gln Leu Ser Leu Val Glu Thr Thr Val Lys Glu
    450                 455                 460

Ser Glu Asp His Leu Lys Asn Ala Ile Gly Ala Leu Arg Ser Gln Ser
465                 470                 475                 480

Arg Val Gly Gly Val Gly Ile Val Ala Ile Val Gly Leu Ile Ile Ala
                485                 490                 495

Thr Val Ser Leu Val Val Leu Val Ile Ser Gly Cys Cys Leu Val Lys
            500                 505                 510

Tyr Phe Ser Arg Thr Ala Thr Leu Glu Ser Ser Leu Thr Thr Ile Glu
        515                 520                 525

His Gly Pro Thr Leu Ala Pro Lys Ser Gly Pro Ile Ile Pro Thr Tyr
```

530 535 540

Ile Asn Pro Val Tyr Arg His Asp
545 550

<210> 35
<211> 544
<212> PRT
<213> J-virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 35

Met Lys Pro Val Ala Leu Ile Tyr Leu Thr Ile Leu Ala Phe Thr Val
1 5 10 15

Lys Val Arg Ser Gln Leu Ala Leu Ser Asp Leu Thr Lys Ile Gly Ile
20 25 30

Ile Pro Ala Lys Ser Tyr Glu Leu Lys Ile Ser Thr Gln Ala Ala Gln
35 40 45

Gln Leu Met Val Ile Lys Leu Ile Pro Asn Val Asn Gly Leu Thr Asn
50 55 60

Cys Thr Ile Pro Val Met Asp Ser Tyr Lys Lys Met Leu Asp Arg Ile
65 70 75 80

Leu Lys Pro Ile Asp Asp Ala Leu Asn His Val Lys Asn Ala Ile Gln
85 90 95

Asp Lys Gln Gly Asp Gly Val Pro Gly Val Arg Phe Trp Gly Ala Ile
100 105 110

Ile Gly Gly Val Ala Leu Gly Val Ala Thr Ser Ala Gln Ile Thr Ala
115 120 125

Gly Val Ala Leu His Asn Ser Ile Gln Asn Ala Asn Ala Ile Leu Gln
130 135 140

Leu Lys Glu Ser Ile Arg Asn Ser Asn Lys Ala Ile Glu Glu Leu Gln
145 150 155 160

Ala Gly Leu Gln Ser Thr Val Leu Val Ile Asn Ala Leu Gln Asp Gln
165 170 175

Ile Asn Ser Gln Leu Val Pro Ala Ile Asn Thr Leu Gly Cys Ser Val
180 185 190

Ile Ala Asn Thr Leu Gly Leu Arg Leu Asn Gln Tyr Phe Ser Glu Ile
195 200 205

```
Ser Leu Val Phe Gly Pro Asn Leu Arg Asp Pro Thr Ser Gln Thr Leu
    210                 215                 220

Ser Ile Gln Ala Ile Ala Lys Ala Phe Asn Gly Asp Phe Asp Ser Met
225                 230                 235                 240

Met Lys Lys Met His Tyr Thr Asp Ser Asp Phe Leu Asp Leu Leu Glu
                245                 250                 255

Ser Asp Ser Ile Arg Gly Arg Ile Ile Ser Val Ser Leu Glu Asp Tyr
            260                 265                 270

Leu Ile Ile Ile Gln Ile Asp Tyr Pro Gly Leu Thr Thr Ile Pro Asn
        275                 280                 285

Ser Val Val Gln Thr Phe Asn Leu Ile Thr Tyr Asn Tyr Lys Gly Thr
    290                 295                 300

Glu Trp Glu Ser Ile Phe Pro Arg Glu Leu Leu Ile Arg Gly Ser Tyr
305                 310                 315                 320

Ile Ser Asn Ile Asp Ile Ser Gln Cys Val Gly Thr Ser Lys Ser Met
                325                 330                 335

Ile Cys Lys Ser Asp Thr Ser Thr Thr Ile Ser Pro Ala Thr Trp Ala
            340                 345                 350

Cys Ala Thr Gly Asn Leu Thr Ser Cys Ala Arg Thr Arg Val Val Asn
        355                 360                 365

Ser His Ser Thr Arg Phe Ala Leu Ser Gly Gly Val Leu Phe Ala Asn
    370                 375                 380

Cys Ala Pro Ile Ala Cys Arg Cys Gln Asp Pro Gln Tyr Ser Ile Asn
385                 390                 395                 400

Gln Glu Pro Lys Thr Thr Asn Val Met Val Thr Ser Glu Asp Cys Lys
                405                 410                 415

Glu Leu Tyr Ile Asp Gly Phe Tyr Leu Thr Leu Gly Lys Lys Met Leu
            420                 425                 430

Asp Arg Ala Met Tyr Ala Glu Asp Val Ala Leu Gly Gly Ser Val Ser
        435                 440                 445

Val Asp Pro Ile Asp Ile Gly Asn Glu Leu Asn Ser Ile Asn Glu Ser
    450                 455                 460

Ile Asn Lys Ser His Glu Tyr Leu Asp Lys Ala Asn Glu Leu Leu Glu
465                 470                 475                 480
```

```
Gln Val Asn Pro Asn Ile Val Asn Val Ser Ser Phe Ser Phe Ile Leu
                485                 490                 495


Val Ile Ser Ile Leu Leu Ile Ile Trp Phe Ile Val Thr Leu Val Trp
            500                 505                 510


Leu Ile Tyr Leu Thr Lys His Met Asn Phe Ile Val Gly Lys Val Ala
        515                 520                 525


Met Gly Ser Arg Ser Ser Thr Val Asn Ser Leu Ser Gly Phe Val Gly
    530                 535                 540


<210>  36
<211>  537
<212>  PRT
<213>  Mapuera virus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  36

Met Arg Ser Ser Leu Phe Leu Val Leu Thr Leu Leu Val Pro Phe Ala
1               5                   10                  15


His Ser Ile Asp Ser Ile Thr Leu Glu Gln Tyr Gly Thr Val Ile Thr
            20                  25                  30


Ser Val Arg Ser Leu Ala Tyr Phe Leu Glu Thr Asn Pro Thr Tyr Ile
        35                  40                  45


Ser Val Arg Leu Met Pro Ala Ile Gln Thr Asp Ser Ser His Cys Ser
    50                  55                  60


Tyr His Ser Ile Glu Asn Tyr Asn Leu Thr Leu Thr Lys Leu Leu Leu
65                  70                  75                  80


Pro Leu Gln Glu Asn Leu His Gln Ile Thr Asp Ser Leu Ser Ser Arg
                85                  90                  95


Arg Arg Lys Lys Arg Phe Ala Gly Val Ala Val Gly Leu Ala Ala Leu
            100                 105                 110


Gly Val Ala Thr Ala Ala Gln Val Thr Ala Ala Ile Ala Val Val Lys
        115                 120                 125


Ala Lys Glu Asn Ser Ala Lys Ile Ala Gln Leu Thr Ser Ala Ile Ser
    130                 135                 140


Glu Thr Asn Arg Ala Val Gln Asp Leu Ile Glu Gly Ser Lys Gln Leu
145                 150                 155                 160
```

```
Ala Val Ala Val Gln Ala Ile Gln Asp Gln Ile Asn Asn Val Ile Gln
                165                 170                 175

Pro Gln Leu Thr Asn Leu Ser Cys Gln Val Ala Asp Ala Gln Val Gly
            180                 185                 190

Thr Ile Leu Asn Met Tyr Leu Thr Glu Leu Thr Thr Val Phe His Pro
        195                 200                 205

Gln Ile Thr Asn Ser Ala Leu Thr Pro Ile Thr Ile Gln Ala Leu Arg
    210                 215                 220

Ser Leu Leu Gly Ser Thr Leu Pro Gln Val Val Thr Ser Thr Ile Lys
225                 230                 235                 240

Thr Asp Val Pro Leu Gln Asp Leu Leu Thr Ser Gly Leu Leu Lys Gly
                245                 250                 255

Gln Ile Val Tyr Leu Asp Leu Gln Ser Met Ile Met Val Val Ser Val
            260                 265                 270

Ser Val Pro Thr Ile Ala Leu His Ser Met Ala Lys Val Tyr Thr Leu
        275                 280                 285

Lys Ala Ile Ser Ala His Val Asn Asn Ala Glu Val Gln Met Gln Val
    290                 295                 300

Pro Ser Arg Val Met Glu Leu Gly Ser Glu Ile Met Gly Tyr Asp Ile
305                 310                 315                 320

Asp Gln Cys Glu Glu Thr Ser Arg Tyr Leu Phe Cys Pro Tyr Asn Gly
                325                 330                 335

Gly Ser Ile Leu Ser Ala Thr Met Lys Met Cys Leu Asn Gly Asn Ile
            340                 345                 350

Ser Gln Cys Val Phe Thr Pro Ile Tyr Gly Ser Phe Leu Gln Arg Phe
        355                 360                 365

Val Leu Val Asp Gly Val Ile Val Ala Asn Cys Arg Asp Met Thr Cys
    370                 375                 380

Ala Cys Lys Ser Pro Ser Lys Ile Ile Thr Gln Pro Asp Ser Leu Pro
385                 390                 395                 400

Val Thr Ile Ile Asp Ser Thr Ser Cys Ser Asn Leu Val Leu Asp Thr
                405                 410                 415

Leu Glu Leu Pro Ile Ile Ser Ile Asn Asn Ala Thr Tyr Arg Pro Val
            420                 425                 430

Gln Tyr Val Gly Pro Asn Gln Ile Ile Phe Ser Gln Pro Leu Asp Leu
```

435 440 445

Leu Ser Gln Leu Gly Lys Ile Asn Ser Ser Leu Ser Asp Ala Ile Glu
450 455 460

His Leu Ala Lys Ser Asp Glu Ile Leu Glu Gln Ile Gln Trp Asp Ser
465 470 475 480

Pro Gln Gly Tyr Thr Leu Ile Ala Leu Thr Ser Val Leu Ala Phe Val
485 490 495

Val Val Ala Ile Val Gly Leu Leu Ile Ser Thr Arg Tyr Leu Ile Phe
500 505 510

Glu Ile Arg Arg Ile Asn Thr Thr Leu Thr Gln Gln Leu Ser Ser Tyr
515 520 525

Val Leu Ser Asn Lys Ile Ile Gln Tyr
530 535

<210> 37
<211> 566
<212> PRT
<213> Nariva virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 37

Met Ala Glu Gln Glu Lys Thr Pro Leu Arg Tyr Lys Ile Leu Leu Ile
1 5 10 15

Ile Ile Val Ile Asn His Tyr Asn Ile Thr Asn Val Phe Gly Gln Ile
20 25 30

His Leu Ala Asn Leu Ser Ser Ile Gly Val Phe Val Thr Lys Thr Leu
35 40 45

Asp Tyr Arg Thr Thr Ser Asp Pro Thr Glu Gln Leu Leu Val Ile Asn
50 55 60

Met Leu Pro Asn Ile Ser Asn Ile Gln Asp Cys Ala Gln Gly Val Val
65 70 75 80

Asn Glu Tyr Lys His Leu Ile Ser Ser Leu Leu Thr Pro Ile Asn Asp
85 90 95

Thr Leu Asp Leu Ile Thr Ser Asn Ile Asn Pro Tyr Ser Gly Arg Asn
100 105 110

Lys Leu Phe Gly Glu Ile Ile Ala Gly Ala Ala Leu Thr Val Ala Thr
115 120 125

```
Ser Ala Gln Ile Thr Ala Gly Val Ala Leu Tyr Glu Ala Arg Gln Asn
    130                 135                 140

Ala Lys Asp Ile Ala Ala Ile Lys Glu Ser Leu Gly Tyr Ala Tyr Lys
145                 150                 155                 160

Ala Ile Asp Lys Leu Thr Thr Ala Thr Arg Glu Ile Thr Val Val Ile
                165                 170                 175

Asn Glu Leu Gln Asp Gln Ile Asn Asn Arg Leu Ile Pro Arg Ile Asn
            180                 185                 190

Asp Leu Ala Cys Glu Val Trp Ala Thr Arg Leu Gln Ala Met Leu Leu
        195                 200                 205

Gln Tyr Tyr Ala Glu Ile Phe Ser Val Ile Gly Pro Asn Leu Gln Asp
    210                 215                 220

Pro Leu Ser Gly Lys Ile Ser Ile Gln Ala Leu Ala Arg Ala Ala Gly
225                 230                 235                 240

Gly Asn Ile Lys Leu Met Val Asp Glu Leu Asn Tyr Ser Gly Gln Asp
                245                 250                 255

Leu Ser Arg Leu Val Lys Val Gly Ala Ile Lys Gly Gln Ile Ile Asp
            260                 265                 270

Ala Asp Pro Ser Leu Gly Val Val Ile Ile Lys Met Arg Tyr Pro Asn
        275                 280                 285

Ile Ile Lys Ile Pro Asn Val Ala Ile Ser Glu Leu Ser Tyr Val Ser
    290                 295                 300

Tyr Ser Ser Asp Gly Gln Asp Trp Ile Thr Thr Gly Pro Asn Tyr Ile
305                 310                 315                 320

Val Thr Arg Gly Tyr Ser Ile Ala Asn Ile Gln Thr Ser Ser Cys Ser
                325                 330                 335

Val Gly Asp Asp Phe Val Leu Cys Asp Arg Asp Met Thr Tyr Pro Met
            340                 345                 350

Ser Gln Val Thr Gln Asp Cys Leu Arg Gly Asn Ile Ala Leu Cys Ser
        355                 360                 365

Arg Met Val Val Arg Asp Arg Glu Ala Pro Arg Tyr Leu Ile Leu Gln
    370                 375                 380

Gly Asn Met Val Ala Asn Cys Met Ser Ile Thr Cys Arg Cys Glu Glu
385                 390                 395                 400
```

```
Pro Glu Ser Glu Ile Tyr Gln Ser Pro Asp Gln Pro Leu Thr Leu Leu
                405                 410                 415

Thr Arg Asp Thr Cys Asp Thr His Val Val Asp Gly Ile Arg Ile Arg
            420                 425                 430

Leu Gly Val Arg Lys Leu Pro Thr Ile Ser Val Ile Asn Asn Ile Thr
        435                 440                 445

Leu Gly Pro Ile Ile Thr Thr Asp Pro Ile Asp Val Ser Asn Gln Leu
    450                 455                 460

Asn Ala Val Val Ser Thr Ile Asp Gln Ser Ala Glu Leu Leu His Gln
465                 470                 475                 480

Ala Gln Arg Val Leu Ser Glu Arg Ala Arg Gly Ala Arg Asp His Ile
                485                 490                 495

Leu Ala Thr Ala Ala Ile Val Ile Cys Val Val Leu Ala Val Leu Ile
            500                 505                 510

Leu Val Leu Leu Ile Gly Leu Val Tyr Leu Tyr Arg Thr Gln Asn Glu
        515                 520                 525

Ile Leu Val Lys Thr Thr Met Leu Glu Gln Val Pro Thr Phe Ala Pro
    530                 535                 540

Lys Ser Phe Pro Met Glu Ser Gln Ile Tyr Ser Gly Lys Thr Asn Lys
545                 550                 555                 560

Gly Tyr Asp Pro Ala Glu
                565
```

<210> 38
<211> 662
<212> PRT
<213> Bat paramyxovirus Eid_hel/GH-M74a/GHA/2009

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 38

```
Met Lys Lys Lys Thr Asp Asn Pro Thr Ile Ser Lys Arg Gly His Asn
1               5                   10                  15

His Ser Arg Gly Ile Lys Ser Arg Ala Leu Leu Arg Glu Thr Asp Asn
            20                  25                  30

Tyr Ser Asn Gly Leu Ile Val Glu Asn Leu Val Arg Asn Cys His His
        35                  40                  45
```

```
Pro Ser Lys Asn Asn Leu Asn Tyr Thr Lys Thr Gln Lys Arg Asp Ser
    50                  55                  60

Thr Ile Pro Tyr Arg Val Glu Glu Arg Lys Gly His Tyr Pro Lys Ile
65                  70                  75                  80

Lys His Leu Ile Asp Lys Ser Tyr Lys His Ile Lys Arg Gly Lys Arg
                85                  90                  95

Arg Asn Gly His Asn Gly Asn Ile Ile Thr Ile Ile Leu Leu Leu Ile
            100                 105                 110

Leu Ile Leu Lys Thr Gln Met Ser Glu Gly Ala Ile His Tyr Glu Thr
        115                 120                 125

Leu Ser Lys Ile Gly Leu Ile Lys Gly Ile Thr Arg Glu Tyr Lys Val
    130                 135                 140

Lys Gly Thr Pro Ser Ser Lys Asp Ile Val Ile Lys Leu Ile Pro Asn
145                 150                 155                 160

Val Thr Gly Leu Asn Lys Cys Thr Asn Ile Ser Met Glu Asn Tyr Lys
                165                 170                 175

Glu Gln Leu Asp Lys Ile Leu Ile Pro Ile Asn Asn Ile Ile Glu Leu
            180                 185                 190

Tyr Ala Asn Ser Thr Lys Ser Ala Pro Gly Asn Ala Arg Phe Ala Gly
        195                 200                 205

Val Ile Ile Ala Gly Val Ala Leu Gly Val Ala Ala Ala Ala Gln Ile
    210                 215                 220

Thr Ala Gly Ile Ala Leu His Glu Ala Arg Gln Asn Ala Glu Arg Ile
225                 230                 235                 240

Asn Leu Leu Lys Asp Ser Ile Ser Ala Thr Asn Asn Ala Val Ala Glu
                245                 250                 255

Leu Gln Glu Ala Thr Gly Gly Ile Val Asn Val Ile Thr Gly Met Gln
            260                 265                 270

Asp Tyr Ile Asn Thr Asn Leu Val Pro Gln Ile Asp Lys Leu Gln Cys
        275                 280                 285

Ser Gln Ile Lys Thr Ala Leu Asp Ile Ser Leu Ser Gln Tyr Tyr Ser
    290                 295                 300

Glu Ile Leu Thr Val Phe Gly Pro Asn Leu Gln Asn Pro Val Thr Thr
305                 310                 315                 320

Ser Met Ser Ile Gln Ala Ile Ser Gln Ser Phe Gly Gly Asn Ile Asp
```

325 330 335

Leu Leu Leu Asn Leu Leu Gly Tyr Thr Ala Asn Asp Leu Leu Asp Leu
340 345 350

Leu Glu Ser Lys Ser Ile Thr Gly Gln Ile Thr Tyr Ile Asn Leu Glu
355 360 365

His Tyr Phe Met Val Ile Arg Val Tyr Tyr Pro Ile Met Thr Thr Ile
370 375 380

Ser Asn Ala Tyr Val Gln Glu Leu Ile Lys Ile Ser Phe Asn Val Asp
385 390 395 400

Gly Ser Glu Trp Val Ser Leu Val Pro Ser Tyr Ile Leu Ile Arg Asn
405 410 415

Ser Tyr Leu Ser Asn Ile Asp Ile Ser Glu Cys Leu Ile Thr Lys Asn
420 425 430

Ser Val Ile Cys Arg His Asp Phe Ala Met Pro Met Ser Tyr Thr Leu
435 440 445

Lys Glu Cys Leu Thr Gly Asp Thr Glu Lys Cys Pro Arg Glu Ala Val
450 455 460

Val Thr Ser Tyr Val Pro Arg Phe Ala Ile Ser Gly Gly Val Ile Tyr
465 470 475 480

Ala Asn Cys Leu Ser Thr Thr Cys Gln Cys Tyr Gln Thr Gly Lys Val
485 490 495

Ile Ala Gln Asp Gly Ser Gln Thr Leu Met Met Ile Asp Asn Gln Thr
500 505 510

Cys Ser Ile Val Arg Ile Glu Glu Ile Leu Ile Ser Thr Gly Lys Tyr
515 520 525

Leu Gly Ser Gln Glu Tyr Asn Thr Met His Val Ser Val Gly Asn Pro
530 535 540

Val Phe Thr Asp Lys Leu Asp Ile Thr Ser Gln Ile Ser Asn Ile Asn
545 550 555 560

Gln Ser Ile Glu Gln Ser Lys Phe Tyr Leu Asp Lys Ser Lys Ala Ile
565 570 575

Leu Asp Lys Ile Asn Leu Asn Leu Ile Gly Ser Val Pro Ile Ser Ile
580 585 590

Leu Phe Ile Ile Ala Ile Leu Ser Leu Ile Leu Ser Ile Ile Thr Phe
595 600 605

```
Val Ile Val Met Ile Ile Val Arg Arg Tyr Asn Lys Tyr Thr Pro Leu
    610                 615                 620

Ile Asn Ser Asp Pro Ser Ser Arg Arg Ser Thr Ile Gln Asp Val Tyr
625                 630                 635                 640

Ile Ile Pro Asn Pro Gly Glu His Ser Ile Arg Ser Ala Ala Arg Ser
                645                 650                 655

Ile Asp Arg Asp Arg Asp
            660
```

<210> 39
<211> 539
<212> PRT
<213> Avian paramyxovirus 7

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 39

```
Met Arg Val Arg Pro Leu Ile Ile Ile Leu Val Leu Leu Val Leu Leu
1               5                   10                  15

Trp Leu Asn Ile Leu Pro Val Ile Gly Leu Asp Asn Ser Lys Ile Ala
            20                  25                  30

Gln Ala Gly Ile Ile Ser Ala Gln Glu Tyr Ala Val Asn Val Tyr Ser
        35                  40                  45

Gln Ser Asn Glu Ala Tyr Ile Ala Leu Arg Thr Val Pro Tyr Ile Pro
    50                  55                  60

Pro His Asn Leu Ser Cys Phe Gln Asp Leu Ile Asn Thr Tyr Asn Thr
65                  70                  75                  80

Thr Ile Gln Asn Ile Phe Ser Pro Ile Gln Asp Gln Ile Thr Ser Ile
                85                  90                  95

Thr Ser Ala Ser Thr Leu Pro Ser Ser Arg Phe Ala Gly Leu Val Val
            100                 105                 110

Gly Ala Ile Ala Leu Gly Val Ala Thr Ser Ala Gln Ile Thr Ala Ala
        115                 120                 125

Val Ala Leu Thr Lys Ala Gln Gln Asn Ala Gln Glu Ile Ile Arg Leu
    130                 135                 140

Arg Asp Ser Ile Gln Asn Thr Ile Asn Ala Val Asn Asp Ile Thr Val
145                 150                 155                 160
```

```
Gly Leu Ser Ser Ile Gly Val Ala Leu Ser Lys Val Gln Asn Tyr Leu
                165                 170                 175

Asn Asp Val Ile Asn Pro Ala Leu Gln Asn Leu Ser Cys Gln Val Ser
            180                 185                 190

Ala Leu Asn Leu Gly Ile Gln Leu Asn Leu Tyr Leu Thr Glu Ile Thr
        195                 200                 205

Thr Ile Phe Gly Pro Gln Ile Thr Asn Pro Ser Leu Thr Pro Leu Ser
    210                 215                 220

Ile Gln Ala Leu Tyr Thr Leu Ala Gly Asp Asn Leu Met Gln Phe Leu
225                 230                 235                 240

Thr Arg Tyr Gly Tyr Gly Glu Thr Ser Val Ser Ser Ile Leu Glu Ser
                245                 250                 255

Gly Leu Ile Ser Ala Gln Ile Val Ser Phe Asp Lys Gln Thr Gly Ile
            260                 265                 270

Ala Ile Leu Tyr Val Thr Leu Pro Ser Ile Ala Thr Leu Ser Gly Ser
        275                 280                 285

Arg Val Thr Lys Leu Met Ser Val Ser Val Gln Thr Gly Val Gly Glu
    290                 295                 300

Gly Ser Ala Ile Val Pro Ser Tyr Val Ile Gln Gln Gly Thr Val Ile
305                 310                 315                 320

Glu Glu Phe Ile Pro Asp Ser Cys Ile Phe Thr Arg Ser Asp Val Tyr
                325                 330                 335

Cys Thr Gln Leu Tyr Ser Lys Leu Leu Pro Asp Ser Ile Leu Gln Cys
            340                 345                 350

Leu Gln Gly Ser Met Ala Asp Cys Gln Phe Thr Arg Ser Leu Gly Ser
        355                 360                 365

Phe Ala Asn Arg Phe Met Thr Val Ala Gly Gly Val Ile Ala Asn Cys
    370                 375                 380

Gln Thr Val Leu Cys Arg Cys Tyr Asn Pro Val Met Ile Ile Pro Gln
385                 390                 395                 400

Asn Asn Gly Ile Ala Val Thr Leu Ile Asp Gly Ser Leu Cys Lys Glu
                405                 410                 415

Leu Glu Leu Glu Gly Ile Arg Leu Thr Met Ala Asp Pro Val Phe Ala
            420                 425                 430
```

Ser Tyr Ser Arg Asp Leu Ile Ile Asn Gly Asn Gln Phe Ala Pro Ser
435 440 445

Asp Ala Leu Asp Ile Ser Ser Glu Leu Gly Gln Leu Asn Asn Ser Ile
450 455 460

Ser Ser Ala Thr Asp Asn Leu Gln Lys Ala Gln Glu Ser Leu Asn Lys
465 470 475 480

Ser Ile Ile Pro Ala Ala Thr Ser Ser Trp Leu Ile Ile Leu Leu Phe
485 490 495

Val Leu Val Ser Ile Ser Leu Val Ile Gly Cys Ile Ser Ile Tyr Phe
500 505 510

Ile Tyr Lys His Ser Thr Thr Asn Arg Ser Arg Asn Leu Ser Ser Asp
515 520 525

Ile Ile Ser Asn Pro Tyr Ile Gln Lys Ala Asn
530 535

<210> 40
<211> 539
<212> PRT
<213> Tuhoko virus 2

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 40

Met Ala Pro Cys Val Leu Phe Leu Ser Ser Leu Leu Leu Ile Ser Thr
1 5 10 15

Ile Ser Pro Ser His Gly Ile Asn Gln Pro Ala Leu Arg Arg Ile Gly
20 25 30

Ala Ile Val Ser Ser Val Lys Gln Leu Lys Phe Tyr Ser Lys Thr Lys
35 40 45

Pro Asn Tyr Ile Ile Val Lys Leu Leu Pro Thr Ile Asn Leu Ser Lys
50 55 60

Ser Asn Cys Asn Leu Thr Ser Ile Asn Arg Tyr Lys Glu Ser Val Ile
65 70 75 80

Glu Ile Ile Lys Pro Leu Ala Asp Asn Ile Asp Asn Leu Asn Gln Lys
85 90 95

Leu Leu Pro Lys Asn Arg Arg Lys Arg Met Ala Gly Val Ala Ile Gly
100 105 110

Leu Ala Ala Leu Gly Val Ala Ala Ala Ala Gln Ala Thr Ala Ala Val

```
        115                 120                 125

Ala Leu Val Glu Ala Arg Lys Asn Thr Gln Met Ile Gln Ser Leu Ala
    130                 135                 140

Asp Ser Ile Gln Asp Thr Asn Ala Ala Val Gln Ala Val Asn Ile Gly
145                 150                 155                 160

Leu Gln Asn Ser Ala Val Ala Ile Gln Ala Ile Gln Asn Gln Ile Asn
                165                 170                 175

Asn Val Ile Asn Pro Ala Leu Asp Arg Leu Asn Cys Glu Val Leu Asp
            180                 185                 190

Ala Gln Ile Ala Ser Ile Leu Asn Leu Tyr Leu Ile Lys Ser Val Thr
        195                 200                 205

Ile Phe Gln Asn Gln Leu Thr Asn Pro Ala Leu Gln Gln Leu Ser Ile
    210                 215                 220

Gln Met Leu Ser Ile Val Met Gln Asp Thr Ala Lys Ile Leu Gly Asn
225                 230                 235                 240

Phe Thr Ile Gly Asp Lys Phe Asp Gln His Asp Leu Leu Gly Ser Gly
                245                 250                 255

Leu Ile Thr Gly Gln Val Val Gly Val Asn Leu Thr Asn Leu Gln Leu
            260                 265                 270

Ile Ile Ala Ala Phe Ile Pro Ser Ile Ala Pro Leu Pro Gln Ala Tyr
        275                 280                 285

Ile Ile Asp Leu Ile Ser Ile Thr Ile Ser Val Asn Asp Thr Glu Ala
    290                 295                 300

Val Ile Gln Ile Pro Glu Arg Ile Met Glu His Gly Ser Ser Ile Tyr
305                 310                 315                 320

Gln Phe Gly Gly Lys Gln Cys Val Tyr Gly Gln Phe Ser Ala Tyr Cys
                325                 330                 335

Pro Phe Ser Asp Ala Val Leu Met Thr Gln Asp Leu Gln Leu Cys Met
            340                 345                 350

Lys Gly Asn Ile Glu His Cys Ile Phe Ser Ser Val Leu Gly Ser Phe
        355                 360                 365

Pro Asn Arg Phe Ala Ser Val Asp Gly Val Phe Tyr Ala Asn Cys Lys
    370                 375                 380

Tyr Met Ser Cys Ala Cys Ser Asp Pro Leu Gln Val Ile His Gln Asp
385                 390                 395                 400
```

```
Asp Ser Val Asn Leu Met Val Ile Asp Ser Ser Val Cys Arg Ser Leu
                405                 410                 415


Thr Leu Gly His Val Thr Phe Pro Ile Ile Ala Phe Ser Asn Val Ser
            420                 425                 430


Tyr Gln Met Lys Thr Asn Ile Ser Ile Glu Gln Met Ile Val Thr Ser
        435                 440                 445


Pro Leu Asp Leu Ser Thr Glu Leu Lys Gln Ile Asn Asn Ser Val Asn
    450                 455                 460


Ile Ala Asn Thr Phe Leu Asp Ser Ser Asn Arg Ala Leu Lys Thr Ser
465                 470                 475                 480


Ile Phe Gly Thr Ser Ser Gln Ile Ile Leu Ile Val Leu Leu Ile Phe
                485                 490                 495


Thr Cys Leu Leu Ile Leu Tyr Val Ile Phe Leu Thr Tyr Ile Ile Lys
            500                 505                 510


Ile Leu Ile Lys Glu Val Lys Arg Leu Arg Asp Gly Asn Ser Arg Thr
        515                 520                 525


Gly Ser Lys Leu Ser Phe Ile Asn Pro Asp Val
    530                 535


<210>  41
<211>  537
<212>  PRT
<213>  Tuhoko viirus 3


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  41

Met Leu Trp Leu Thr Ile Leu Ile Ala Leu Val Gly Asn His Glu Ser
1               5                   10                  15


Thr Cys Met Asn Ile Asn Phe Leu Gln Ser Leu Gly Gln Ile Asn Ser
            20                  25                  30


Gln Lys Arg Phe Leu Asn Phe Tyr Thr Gln Gln Pro Pro Ser Tyr Met
        35                  40                  45


Val Ile Arg Leu Val Pro Thr Leu Gln Leu Ser Ala Asn Asn Cys Thr
    50                  55                  60


Leu Gly Ser Ile Val Arg Tyr Arg Asn Ala Ile Lys Glu Leu Ile Gln
65                  70                  75                  80
```

```
Pro Met Asp Glu Asn Leu Arg Trp Leu Ser Ser Asn Leu Ile Pro Gln
                85                  90                  95

Arg Arg Gly Lys Arg Phe Ala Gly Val Ala Val Gly Leu Ala Ala Leu
            100                 105                 110

Gly Val Ala Val Ala Ala Gln Ala Thr Ala Ala Val Ala Leu Val Glu
        115                 120                 125

Ala Arg Ala Asn Ala Glu Lys Ile Ala Ser Met Ser Gln Ser Ile Gln
    130                 135                 140

Glu Thr Asn Lys Ala Val Thr Ser Leu Ser Gln Ala Val Ser Ala Ser
145                 150                 155                 160

Gly Ile Ala Ile Gln Ala Ile Gln Asn Glu Ile Asn Asn Val Ile His
                165                 170                 175

Pro Ile Leu Asn Gln Val Gln Cys Asp Val Leu Asp Ala Arg Val Gly
            180                 185                 190

Asn Ile Leu Asn Leu Tyr Leu Ile Lys Val Thr Thr Ile Phe Gln Asn
        195                 200                 205

Gln Leu Thr Asn Pro Ala Leu Gln Arg Leu Ser Thr Gln Ala Leu Ser
    210                 215                 220

Met Leu Met Gln Ser Thr Ser Ser Tyr Leu Arg Asn Leu Ser Ser Ser
225                 230                 235                 240

Glu Ser Ala Ile Asn Ala Asp Leu Ser Met Thr Asn Leu Ile Glu Ala
                245                 250                 255

Gln Ile Val Gly Ile Asn Met Thr Asn Leu Gln Leu Val Leu Ala Val
            260                 265                 270

Phe Ile Pro Ser Ile Ala Arg Leu Asn Gly Ala Leu Leu Tyr Asp Phe
        275                 280                 285

Ile Ser Ile Thr Ile Ser Ser Asn Gln Thr Glu Val Met Leu Gln Ile
    290                 295                 300

Pro His Arg Val Leu Glu Ile Gly Asn Ser Leu Tyr Thr Phe Glu Gly
305                 310                 315                 320

Thr Gln Cys Glu Met Thr Lys Leu Asn Ala Tyr Cys Leu Tyr Ser Asp
                325                 330                 335

Ala Ile Pro Val Thr Glu Ser Leu Arg Asp Cys Met Asn Gly Leu Phe
            340                 345                 350
```

Ser Gln Cys Gly Phe Val Arg Ile Ile Gly Ser Phe Ala Asn Arg Phe
355 360 365

Ala Ser Val Asn Gly Val Ile Tyr Ala Asn Cys Lys His Leu Thr Cys
370 375 380

Ser Cys Leu Gln Pro Asp Glu Ile Ile Thr Gln Asp Thr Asn Val Pro
385 390 395 400

Leu Thr Ile Ile Asp Thr Lys Arg Cys Thr Lys Ile Ser Leu Gly His
405 410 415

Leu Thr Phe Thr Ile Arg Glu Tyr Ala Asn Val Thr Tyr Ser Leu Arg
420 425 430

Thr Glu Ile Ala Asn Ser Gln Ile Thr Val Val Ser Pro Leu Asp Leu
435 440 445

Ser Ser Gln Leu Thr Thr Ile Asn Asn Ser Leu Ala Asp Ala Thr Asn
450 455 460

His Ile Met Asn Ser Asp Arg Ile Leu Asp Arg Leu Asn Ser Gly Leu
465 470 475 480

Tyr Ser Lys Trp Val Ile Ile Phe Leu Ile Cys Ala Ser Ile Val Ser
485 490 495

Leu Ile Gly Leu Val Phe Leu Gly Phe Leu Ile Arg Gly Leu Ile Leu
500 505 510

Glu Leu Arg Ser Lys His Arg Ser Asn Leu Asn Lys Ala Ser Thr Tyr
515 520 525

Ser Ile Asp Ser Ser Ile Gly Leu Thr
530 535

<210> 42
<211> 545
<212> PRT
<213> Mojiang virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 42

Met Ala Leu Asn Lys Asn Met Phe Ser Ser Leu Phe Leu Gly Tyr Leu
1 5 10 15

Leu Val Tyr Ala Thr Thr Val Gln Ser Ser Ile His Tyr Asp Ser Leu
20 25 30

Ser Lys Val Gly Val Ile Lys Gly Leu Thr Tyr Asn Tyr Lys Ile Lys

35 40 45

Gly Ser Pro Ser Thr Lys Leu Met Val Val Lys Leu Ile Pro Asn Ile
50 55 60

Asp Ser Val Lys Asn Cys Thr Gln Lys Gln Tyr Asp Glu Tyr Lys Asn
65 70 75 80

Leu Val Arg Lys Ala Leu Glu Pro Val Lys Met Ala Ile Asp Thr Met
85 90 95

Leu Asn Asn Val Lys Ser Gly Asn Asn Lys Tyr Arg Phe Ala Gly Ala
100 105 110

Ile Met Ala Gly Val Ala Leu Gly Val Ala Thr Ala Ala Thr Val Thr
115 120 125

Ala Gly Ile Ala Leu His Arg Ser Asn Glu Asn Ala Gln Ala Ile Ala
130 135 140

Asn Met Lys Ser Ala Ile Gln Asn Thr Asn Glu Ala Val Lys Gln Leu
145 150 155 160

Gln Leu Ala Asn Lys Gln Thr Leu Ala Val Ile Asp Thr Ile Arg Gly
165 170 175

Glu Ile Asn Asn Asn Ile Ile Pro Val Ile Asn Gln Leu Ser Cys Asp
180 185 190

Thr Ile Gly Leu Ser Val Gly Ile Arg Leu Thr Gln Tyr Tyr Ser Glu
195 200 205

Ile Ile Thr Ala Phe Gly Pro Ala Leu Gln Asn Pro Val Asn Thr Arg
210 215 220

Ile Thr Ile Gln Ala Ile Ser Ser Val Phe Asn Gly Asn Phe Asp Glu
225 230 235 240

Leu Leu Lys Ile Met Gly Tyr Thr Ser Gly Asp Leu Tyr Glu Ile Leu
245 250 255

His Ser Glu Leu Ile Arg Gly Asn Ile Ile Asp Val Asp Val Asp Ala
260 265 270

Gly Tyr Ile Ala Leu Glu Ile Glu Phe Pro Asn Leu Thr Leu Val Pro
275 280 285

Asn Ala Val Val Gln Glu Leu Met Pro Ile Ser Tyr Asn Ile Asp Gly
290 295 300

Asp Glu Trp Val Thr Leu Val Pro Arg Phe Val Leu Thr Arg Thr Thr
305 310 315 320

Leu Leu Ser Asn Ile Asp Thr Ser Arg Cys Thr Ile Thr Asp Ser Ser
325 330 335

Val Ile Cys Asp Asn Asp Tyr Ala Leu Pro Met Ser His Glu Leu Ile
340 345 350

Gly Cys Leu Gln Gly Asp Thr Ser Lys Cys Ala Arg Glu Lys Val Val
355 360 365

Ser Ser Tyr Val Pro Lys Phe Ala Leu Ser Asp Gly Leu Val Tyr Ala
370 375 380

Asn Cys Leu Asn Thr Ile Cys Arg Cys Met Asp Thr Asp Thr Pro Ile
385 390 395 400

Ser Gln Ser Leu Gly Ala Thr Val Ser Leu Leu Asp Asn Lys Arg Cys
405 410 415

Ser Val Tyr Gln Val Gly Asp Val Leu Ile Ser Val Gly Ser Tyr Leu
420 425 430

Gly Asp Gly Glu Tyr Asn Ala Asp Asn Val Glu Leu Gly Pro Pro Ile
435 440 445

Val Ile Asp Lys Ile Asp Ile Gly Asn Gln Leu Ala Gly Ile Asn Gln
450 455 460

Thr Leu Gln Glu Ala Glu Asp Tyr Ile Glu Lys Ser Glu Glu Phe Leu
465 470 475 480

Lys Gly Val Asn Pro Ser Ile Ile Thr Leu Gly Ser Met Val Val Leu
485 490 495

Tyr Ile Phe Met Ile Leu Ile Ala Ile Val Ser Val Ile Ala Leu Val
500 505 510

Leu Ser Ile Lys Leu Thr Val Lys Gly Asn Val Val Arg Gln Gln Phe
515 520 525

Thr Tyr Thr Gln His Val Pro Ser Met Glu Asn Ile Asn Tyr Val Ser
530 535 540

His
545

<210> 43
<211> 546
<212> PRT
<213> Avian paramyxovirus 12

<220>

<221> MISC_FEATURE
<223> fusion protein

<400> 43

```
Met Ala Ile Pro Val Pro Ser Ser Thr Ala Leu Met Ile Phe Asn Ile
1               5                   10                  15


Leu Val Ser Leu Ala Pro Ala Ser Ala Leu Asp Gly Arg Leu Leu Leu
            20                  25                  30


Gly Ala Gly Ile Val Pro Thr Gly Asp Arg Gln Val Asn Val Tyr Thr
        35                  40                  45


Ser Ser Gln Thr Gly Ile Ile Ala Leu Lys Leu Leu Pro Asn Leu Pro
    50                  55                  60


Lys Asp Lys Glu Asn Cys Ala Glu Val Ser Ile Arg Ser Tyr Asn Glu
65                  70                  75                  80


Thr Leu Thr Arg Ile Leu Thr Pro Leu Ala Gln Ser Met Ala Ala Ile
                85                  90                  95


Arg Gly Asn Ser Thr Val Ser Thr Arg Gly Arg Glu Pro Arg Leu Val
            100                 105                 110


Gly Ala Ile Ile Gly Gly Val Ala Leu Gly Val Ala Thr Ala Ala Gln
        115                 120                 125


Ile Thr Ala Ala Thr Ala Leu Ile Gln Ala Asn Gln Asn Ala Glu Asn
    130                 135                 140


Ile Ala Arg Leu Ala Lys Gly Leu Ala Ala Thr Asn Glu Ala Val Thr
145                 150                 155                 160


Asp Leu Thr Lys Gly Val Gly Ser Leu Ala Ile Gly Val Gly Lys Leu
                165                 170                 175


Gln Asp Tyr Val Asn Glu Gln Phe Asn Arg Thr Gly Glu Ala Ile Glu
            180                 185                 190


Cys Leu Thr Ile Glu Ser Arg Val Gly Val Gln Leu Ser Leu Tyr Leu
        195                 200                 205


Thr Glu Val Ile Gly Val Phe Gly Asp Gln Ile Thr Ser Pro Ala Leu
    210                 215                 220


Ser Asp Ile Ser Ile Gln Ala Leu Tyr Asn Leu Ala Gly Gly Asn Leu
225                 230                 235                 240


Asn Val Leu Leu Gln Lys Met Gly Ile Glu Gly Thr Gln Leu Gly Ser
                245                 250                 255
```

```
Leu Ile Asn Ser Gly Leu Ile Lys Gly Arg Pro Ile Met Tyr Asp Asp
            260                 265                 270

Gly Asn Lys Ile Leu Gly Ile Gln Val Thr Leu Pro Ser Val Gly Arg
        275                 280                 285

Ile Asn Gly Ala Arg Ala Thr Leu Leu Glu Ala Ile Ala Val Ala Thr
    290                 295                 300

Pro Lys Gly Asn Ala Ser Pro Leu Ile Pro Arg Ala Val Ile Ser Val
305                 310                 315                 320

Gly Ser Leu Val Glu Glu Leu Asp Met Thr Pro Cys Val Leu Thr Pro
                325                 330                 335

Thr Asp Ile Phe Cys Thr Arg Ile Leu Ser Tyr Pro Leu Ser Asp Ser
            340                 345                 350

Leu Thr Thr Cys Leu Lys Gly Asn Leu Ser Ser Cys Val Phe Ser Arg
        355                 360                 365

Thr Glu Gly Ala Leu Ser Thr Pro Tyr Val Ser Val His Gly Lys Ile
    370                 375                 380

Val Ala Asn Cys Lys Ser Val Val Cys Arg Cys Val Glu Pro Gln Gln
385                 390                 395                 400

Ile Ile Ser Gln Asn Tyr Gly Glu Ala Leu Ser Leu Ile Asp Glu Ser
                405                 410                 415

Leu Cys Arg Ile Leu Glu Leu Asn Gly Val Ile Leu Lys Met Asp Gly
            420                 425                 430

Gln Phe Thr Ser Glu Tyr Thr Lys Asn Ile Thr Ile Asp Pro Val Gln
        435                 440                 445

Val Ile Ile Ser Gly Pro Ile Asp Ile Ser Ser Glu Leu Ser Gln Val
    450                 455                 460

Asn Gln Ser Leu Asp Ser Ala Leu Glu Asn Ile Lys Glu Ser Asn Ser
465                 470                 475                 480

Tyr Leu Ser Lys Val Asn Val Lys Leu Ile Ser Ser Ser Ala Met Ile
                485                 490                 495

Thr Tyr Ile Val Ile Thr Val Ile Cys Leu Ile Leu Thr Phe Val Ala
            500                 505                 510

Leu Val Leu Gly Ile Tyr Ser Tyr Thr Lys Ile Arg Ser Gln Gln Lys
        515                 520                 525

Thr Leu Ile Trp Met Gly Asn Asn Ile Ala Arg Ser Lys Glu Gly Asn
```

530 535 540

Arg Phe
545

<210> 44
<211> 538
<212> PRT
<213> Avian paramyxovirus 3

<220>
<221> MISC_FEATURE
<223> fusion protein

<220>
<221> misc_feature
<222> (287)..(287)
<223> Xaa can be any naturally occurring amino acid

<400> 44

Met Ala Ser Pro Met Val Pro Leu Leu Ile Ile Thr Val Val Pro Ala
1 5 10 15

Leu Ile Ser Ser Gln Ser Ala Asn Ile Asp Lys Leu Ile Gln Ala Gly
20 25 30

Ile Ile Met Gly Ser Gly Lys Glu Leu His Ile Tyr Gln Glu Ser Gly
35 40 45

Ser Leu Asp Leu Tyr Leu Arg Leu Leu Pro Val Ile Pro Ser Asn Leu
50 55 60

Ser His Cys Gln Ser Glu Val Ile Thr Gln Tyr Asn Ser Thr Val Thr
65 70 75 80

Arg Leu Leu Ser Pro Ile Ala Lys Asn Leu Asn His Leu Leu Gln Pro
85 90 95

Arg Pro Ser Gly Arg Leu Phe Gly Ala Val Ile Gly Ser Ile Ala Leu
100 105 110

Gly Val Ala Thr Ser Ala Gln Ile Ser Ala Ala Ile Ala Leu Val Arg
115 120 125

Ala Gln Gln Asn Ala Asn Asp Ile Leu Ala Leu Lys Ala Ala Ile Gln
130 135 140

Ser Ser Asn Glu Ala Ile Lys Gln Leu Thr Tyr Gly Gln Glu Lys Gln
145 150 155 160

Leu Leu Ala Ile Ser Lys Ile Gln Lys Ala Val Asn Glu Gln Val Ile
165 170 175

Pro Ala Leu Thr Ala Leu Asp Cys Ala Val Leu Gly Asn Lys Leu Ala

180 185 190

Ala Gln Leu Asn Leu Tyr Leu Ile Glu Met Thr Thr Ile Phe Gly Asp
195 200 205

Gln Ile Asn Asn Pro Val Leu Thr Pro Ile Pro Leu Ser Tyr Leu Leu
210 215 220

Arg Leu Thr Gly Ser Glu Leu Asn Asp Val Leu Leu Gln Gln Thr Arg
225 230 235 240

Ser Ser Leu Ser Leu Ile His Leu Val Ser Lys Gly Leu Leu Ser Gly
245 250 255

Gln Ile Ile Gly Tyr Asp Pro Ser Val Gln Gly Ile Ile Ile Arg Ile
260 265 270

Gly Leu Ile Arg Thr Gln Arg Ile Asp Arg Ser Leu Val Phe Xaa Pro
275 280 285

Tyr Val Leu Pro Ile Thr Ile Ser Ser Asn Ile Ala Thr Pro Ile Ile
290 295 300

Pro Asp Cys Val Val Lys Lys Gly Val Ile Ile Glu Gly Met Leu Lys
305 310 315 320

Ser Asn Cys Ile Glu Leu Glu Arg Asp Ile Ile Cys Lys Thr Ile Asn
325 330 335

Thr Tyr Gln Ile Thr Lys Glu Thr Arg Ala Cys Leu Gln Gly Asn Ile
340 345 350

Thr Met Cys Lys Tyr Gln Gln Ser Arg Thr Gln Leu Ser Thr Pro Phe
355 360 365

Ile Thr Tyr Asn Gly Val Val Ile Ala Asn Cys Asp Leu Val Ser Cys
370 375 380

Arg Cys Ile Arg Pro Pro Met Ile Ile Thr Gln Val Lys Gly Tyr Pro
385 390 395 400

Leu Thr Ile Ile Asn Arg Asn Leu Cys Thr Glu Leu Ser Val Asp Asn
405 410 415

Leu Ile Leu Asn Ile Glu Thr Asn His Asn Phe Ser Leu Asn Pro Thr
420 425 430

Ile Ile Asp Ser Gln Ser Arg Leu Ile Ala Thr Ser Pro Leu Glu Ile
435 440 445

Asp Ala Leu Ile Gln Asp Ala Gln His His Ala Ala Ala Ala Leu Leu
450 455 460

```
Lys Val Glu Glu Ser Asn Ala His Leu Leu Arg Val Thr Gly Leu Gly
465                 470                 475                 480


Ser Ser Ser Trp His Ile Ile Leu Ile Leu Thr Leu Leu Val Cys Thr
                485                 490                 495


Ile Ala Trp Leu Ile Gly Leu Ser Ile Tyr Val Cys Arg Ile Lys Asn
            500                 505                 510


Asp Asp Ser Thr Asp Lys Glu Pro Thr Thr Gln Ser Ser Asn Arg Gly
        515                 520                 525


Ile Gly Val Gly Ser Ile Gln Tyr Met Thr
    530                 535


<210>  45
<211>  552
<212>  PRT
<213>  Salem virus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  45

Met Asn Pro Leu Asn Gln Thr Leu Ile Ala Lys Val Leu Gly Phe Leu
1               5                   10                  15


Leu Leu Ser Ser Ser Phe Thr Val Gly Gln Ile Gly Phe Glu Asn Leu
            20                  25                  30


Thr Arg Ile Gly Val His Gln Val Lys Gln Tyr Gly Tyr Lys Leu Ala
        35                  40                  45


His Tyr Asn Ser His Gln Leu Leu Leu Ile Arg Met Ile Pro Thr Val
    50                  55                  60


Asn Gly Thr His Asn Cys Thr His Gln Val Ile Thr Arg Tyr Arg Glu
65                  70                  75                  80


Met Val Arg Glu Ile Ile Thr Pro Ile Lys Gly Ala Leu Asp Ile Met
                85                  90                  95


Lys Lys Ala Val Ser Pro Asp Leu Val Gly Ala Arg Ile Phe Gly Ala
            100                 105                 110


Ile Val Ala Gly Ala Ala Leu Gly Ile Ala Thr Ser Ala Gln Ile Thr
        115                 120                 125


Ala Gly Val Ala Leu His Arg Thr Lys Leu Asn Gly Gln Glu Ile Ser
    130                 135                 140
```

```
Lys Leu Lys Glu Ala Val Ser Leu Thr Asn Glu Ala Val Glu Gln Leu
145                 150                 155                 160

Gln Tyr Ser Gln Gly Lys Ser Ile Leu Ala Ile Gln Gly Ile Gln Asp
                165                 170                 175

Phe Ile Asn Phe Asn Val Val Pro Leu Leu Glu Glu His Thr Cys Gly
            180                 185                 190

Ile Ala Lys Leu His Leu Glu Met Ala Leu Met Glu Tyr Phe Gln Lys
        195                 200                 205

Leu Ile Leu Val Phe Gly Pro Asn Leu Arg Asp Pro Ile Gly Ser Thr
    210                 215                 220

Ile Gly Ile Gln Ala Leu Ala Thr Leu Phe Gln Asn Asn Met Phe Glu
225                 230                 235                 240

Val Ser Leu Arg Leu Gly Tyr Ala Gly Asp Asp Leu Glu Asp Val Leu
                245                 250                 255

Gln Ser Asn Ser Ile Arg Ala Asn Ile Ile Glu Ala Glu Pro Asp Ser
            260                 265                 270

Gly Phe Ile Val Leu Ala Ile Arg Tyr Pro Thr Leu Thr Leu Val Glu
        275                 280                 285

Asp Gln Val Ile Thr Glu Leu Ala His Ile Thr Phe Asn Asp Gly Pro
    290                 295                 300

Gln Glu Trp Val Ala Thr Ile Pro Gln Phe Val Thr Tyr Arg Gly Leu
305                 310                 315                 320

Val Leu Ala Asn Ile Asp Val Ser Thr Cys Thr Phe Thr Glu Arg Asn
                325                 330                 335

Val Ile Cys Ala Arg Asp Gln Thr Tyr Pro Met Ile Ile Asp Leu Gln
            340                 345                 350

Leu Cys Met Arg Gly Asn Ile Ala Lys Cys Gly Arg Thr Arg Val Thr
        355                 360                 365

Gly Ser Thr Ala Ser Arg Phe Leu Leu Lys Asp Gly Asn Met Tyr Ala
    370                 375                 380

Asn Cys Ile Ala Thr Met Cys Arg Cys Met Ser Ser Ser Ser Ile Ile
385                 390                 395                 400

Asn Gln Glu Pro Ser His Leu Thr Thr Leu Ile Val Lys Glu Thr Cys
                405                 410                 415
```

Ser Glu Val Met Ile Asp Thr Ile Arg Ile Thr Leu Gly Glu Arg Lys
420 425 430

His Pro Pro Ile Asp Tyr Gln Thr Thr Ile Thr Leu Gly Gln Pro Ile
435 440 445

Ala Leu Ala Pro Leu Asp Val Gly Thr Glu Leu Ala Asn Ala Val Ser
450 455 460

Tyr Leu Asn Lys Ser Lys Val Leu Leu Glu His Ser Asn Glu Val Leu
465 470 475 480

Ser Ser Val Ser Thr Ala His Thr Ser Leu Thr Ala Thr Ile Val Leu
485 490 495

Gly Ile Val Val Gly Gly Leu Ala Ile Leu Ile Val Val Met Phe Leu
500 505 510

Phe Leu Glu Ala Gln Val Ile Lys Val Gln Arg Ala Met Met Leu Cys
515 520 525

Pro Ile Thr Asn His Gly Tyr Leu Pro Asn Glu Asp Leu Leu Thr Arg
530 535 540

Gly His Ser Ile Pro Thr Ile Gly
545 550

<210> 46
<211> 551
<212> PRT
<213> Avian paramyxovirus 9

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 46

Met Gly Tyr Phe His Leu Leu Leu Ile Leu Thr Ala Ile Ala Ile Ser
1 5 10 15

Ala His Leu Cys Tyr Thr Thr Thr Leu Asp Gly Arg Lys Leu Leu Gly
20 25 30

Ala Gly Ile Val Ile Thr Glu Glu Lys Gln Val Arg Val Tyr Thr Ala
35 40 45

Ala Gln Ser Gly Thr Ile Val Leu Arg Ser Phe Arg Val Val Ser Leu
50 55 60

Asp Arg Tyr Ser Cys Met Glu Ser Thr Ile Glu Ser Tyr Asn Lys Thr
65 70 75 80

Val Tyr Asn Ile Leu Ala Pro Leu Gly Asp Ala Ile Arg Arg Ile Gln

85 90 95

Ala Ser Gly Val Ser Val Glu Arg Ile Arg Glu Gly Arg Ile Phe Gly
100 105 110

Ala Ile Leu Gly Gly Val Ala Leu Gly Val Ala Thr Ala Ala Gln Ile
115 120 125

Thr Ala Ala Ile Ala Leu Ile Gln Ala Asn Glu Asn Ala Lys Asn Ile
130 135 140

Leu Arg Ile Lys Asp Ser Ile Thr Lys Thr Asn Glu Ala Val Arg Asp
145 150 155 160

Val Thr Asn Gly Val Ser Gln Leu Thr Ile Ala Val Gly Lys Leu Gln
165 170 175

Asp Phe Val Asn Lys Glu Phe Asn Lys Thr Thr Glu Ala Ile Asn Cys
180 185 190

Val Gln Ala Ala Gln Gln Leu Gly Val Glu Leu Ser Leu Tyr Leu Thr
195 200 205

Glu Ile Thr Thr Val Phe Gly Pro Gln Ile Thr Ser Pro Ala Leu Ser
210 215 220

Lys Leu Thr Ile Gln Ala Leu Tyr Asn Leu Ala Gly Val Ser Leu Asp
225 230 235 240

Val Leu Leu Gly Arg Leu Gly Ala Asp Asn Ser Gln Leu Ser Ser Leu
245 250 255

Val Ser Ser Gly Leu Ile Thr Gly Gln Pro Ile Leu Tyr Asp Ser Glu
260 265 270

Ser Gln Ile Leu Ala Leu Gln Val Ser Leu Pro Ser Ile Ser Asp Leu
275 280 285

Arg Gly Val Arg Ala Thr Tyr Leu Asp Thr Leu Ala Val Asn Thr Ala
290 295 300

Ala Gly Leu Ala Ser Ala Met Ile Pro Lys Val Val Ile Gln Ser Asn
305 310 315 320

Asn Ile Val Glu Glu Leu Asp Thr Thr Ala Cys Ile Ala Ala Glu Ala
325 330 335

Asp Leu Tyr Cys Thr Arg Ile Thr Thr Phe Pro Ile Ala Ser Ala Val
340 345 350

Ser Ala Cys Ile Leu Gly Asp Val Ser Gln Cys Leu Tyr Ser Lys Thr
355 360 365

```
Asn Gly Val Leu Thr Thr Pro Tyr Val Ala Val Lys Gly Lys Ile Val
    370                 375                 380

Ala Asn Cys Lys His Val Thr Cys Arg Cys Val Asp Pro Thr Ser Ile
385                 390                 395                 400

Ile Ser Gln Asn Tyr Gly Glu Ala Ala Thr Leu Ile Asp Asp Gln Leu
                405                 410                 415

Cys Lys Val Ile Asn Leu Asp Gly Val Ser Ile Gln Leu Ser Gly Thr
            420                 425                 430

Phe Glu Ser Thr Tyr Val Arg Asn Val Ser Ile Ser Ala Asn Lys Val
        435                 440                 445

Ile Val Ser Ser Ser Ile Asp Ile Ser Asn Glu Leu Glu Asn Val Asn
    450                 455                 460

Ser Ser Leu Ser Ser Ala Leu Glu Lys Leu Asp Glu Ser Asp Ala Ala
465                 470                 475                 480

Leu Ser Lys Val Asn Val His Leu Thr Ser Thr Ser Ala Met Ala Thr
                485                 490                 495

Tyr Ile Val Leu Thr Val Ile Ala Leu Ile Leu Gly Phe Val Gly Leu
            500                 505                 510

Gly Leu Gly Cys Phe Ala Met Ile Lys Val Lys Ser Gln Ala Lys Thr
        515                 520                 525

Leu Leu Trp Leu Gly Ala His Ala Asp Arg Ser Tyr Ile Leu Gln Ser
    530                 535                 540

Lys Pro Ala Gln Ser Ser Thr
545                 550


<210>  47
<211>  533
<212>  PRT
<213>  Achimota virus 1


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  47

Met Trp Ile Met Ile Ile Leu Ser Leu Phe Gln Ile Ile Pro Gly Val
1               5                   10                  15

Thr Pro Ile Asn Ser Lys Val Leu Thr Gln Leu Gly Val Ile Thr Lys
            20                  25                  30
```

```
His Thr Arg Gln Leu Lys Phe Tyr Ser His Ser Thr Pro Ser Tyr Leu
        35                  40                  45

Val Val Lys Leu Val Pro Thr Ile Asn Thr Glu Ser Thr Val Cys Asn
    50                  55                  60

Phe Thr Ser Leu Ser Arg Tyr Lys Asp Ser Val Arg Glu Leu Ile Thr
65                  70                  75                  80

Pro Leu Ala Lys Asn Ile Asp Asn Leu Asn Ser Ile Leu Thr Ile Pro
                85                  90                  95

Lys Arg Arg Lys Arg Met Ala Gly Val Val Ile Gly Leu Ala Ala Leu
            100                 105                 110

Gly Val Ala Ala Ala Ala Gln Ala Thr Ala Ala Val Ala Leu Ile Glu
        115                 120                 125

Ala Lys Lys Asn Thr Glu Gln Ile Gln Ala Leu Ser Glu Ser Ile Gln
    130                 135                 140

Asn Thr Asn Lys Ala Val Ser Ser Ile Glu Lys Gly Leu Ser Ser Ala
145                 150                 155                 160

Ala Ile Ala Val Gln Ala Ile Gln Asn Gln Ile Asn Asn Val Ile Asn
                165                 170                 175

Pro Ala Leu Thr Ala Leu Asp Cys Gly Val Thr Asp Ala Gln Leu Gly
            180                 185                 190

Asn Ile Leu Asn Leu Tyr Leu Ile Lys Thr Leu Thr Val Phe Gln Lys
        195                 200                 205

Gln Ile Thr Asn Pro Ala Leu Gln Pro Leu Ser Ile Gln Ala Leu Asn
    210                 215                 220

Ile Ile Met Gln Glu Thr Ser Ser Val Leu Arg Asn Phe Thr Lys Thr
225                 230                 235                 240

Asp Glu Ile Glu His Thr Asp Leu Leu Thr Ser Gly Leu Ile Thr Gly
                245                 250                 255

Gln Val Val Gly Val Asn Leu Thr Asn Leu Gln Leu Ile Ile Ala Ala
            260                 265                 270

Phe Ile Pro Ser Ile Ala Pro Leu Asn Gln Ala Tyr Ile Leu Asp Phe
        275                 280                 285

Ile Arg Ile Thr Val Asn Ile Asn Asn Ser Glu Ser Met Ile Gln Ile
    290                 295                 300
```

Pro Glu Arg Ile Met Glu His Gly Ile Ser Leu Tyr Gln Phe Gly Gly
305 310 315 320

Asp Gln Cys Thr Phe Ser Asp Trp Ser Ala Tyr Cys Pro Tyr Ser Asp
325 330 335

Ala Thr Leu Met Ala Pro Gly Leu Gln Asn Cys Phe Arg Gly Gln Ala
340 345 350

Ala Asp Cys Val Phe Ser Thr Val Met Gly Ser Phe Pro Asn Arg Phe
355 360 365

Val Ser Val Gln Gly Val Phe Tyr Val Asn Cys Lys Phe Ile Arg Cys
370 375 380

Ala Cys Thr Gln Pro Gln Arg Leu Ile Thr Gln Asp Asp Ser Leu Ser
385 390 395 400

Leu Thr Gln Ile Asp Ala Lys Thr Cys Arg Met Leu Thr Leu Gly Phe
405 410 415

Val Gln Phe Ser Ile Asn Glu Tyr Ala Asn Val Thr Tyr Ser Phe Lys
420 425 430

Asn Asn Val Thr Ala Gly Gln Leu Ile Met Thr Asn Pro Ile Asp Leu
435 440 445

Ser Thr Glu Ile Lys Gln Met Asn Asp Ser Val Asp Glu Ala Ala Arg
450 455 460

Tyr Ile Glu Lys Ser Asn Ala Ala Leu Asn Lys Leu Met Tyr Gly Gly
465 470 475 480

Arg Ser Asp Ile Val Thr Thr Val Leu Leu Val Gly Phe Ile Leu Leu
485 490 495

Val Val Tyr Val Ile Phe Val Thr Tyr Ile Leu Lys Ile Leu Met Lys
500 505 510

Glu Val Ala Arg Leu Arg Asn Ser Asn His Pro Asp Leu Ile Lys Pro
515 520 525

Tyr Asn Tyr Pro Met
530

<210> 48
<211> 531
<212> PRT
<213> Achimota virus 2

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 48

```
Met Leu Asn Ser Phe Tyr Gln Ile Ile Cys Leu Ala Val Cys Leu Thr
1               5                   10                  15

Thr Tyr Thr Val Ile Ser Ile Asp Gln His Asn Leu Leu Lys Ala Gly
            20                  25                  30

Val Ile Val Lys Ser Ile Lys Gly Leu Asn Phe Tyr Ser Arg Gly Gln
        35                  40                  45

Ala Asn Tyr Ile Ile Val Lys Leu Ile Pro Asn Val Asn Val Thr Asp
    50                  55                  60

Thr Asp Cys Asp Ile Gly Ser Ile Lys Arg Tyr Asn Glu Thr Val Tyr
65                  70                  75                  80

Ser Leu Ile Lys Pro Leu Ala Asp Asn Ile Asp Tyr Leu Arg Thr Gln
                85                  90                  95

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

Leu Thr Ala Leu Gly Val Ala Thr Ala Ala Gln Val Thr Ala Ala Val
        115                 120                 125

Ala Leu Val Lys Ala Gln Glu Asn Ala Arg Lys Leu Asp Ala Leu Ala
    130                 135                 140

Asp Ser Ile Gln Ala Thr Asn Glu Ala Val Gln Asp Leu Ser Thr Gly
145                 150                 155                 160

Leu Gln Ala Gly Ala Ile Ala Ile Gln Ala Ile Gln Ser Glu Ile Asn
                165                 170                 175

His Val Ile Asn Pro Ala Leu Glu Arg Leu Ser Cys Glu Ile Ile Asp
            180                 185                 190

Thr Arg Val Ala Ser Ile Leu Asn Leu Tyr Leu Ile Arg Leu Thr Thr
        195                 200                 205

Val Phe His Arg Gln Leu Val Asn Pro Ala Leu Thr Pro Leu Ser Ile
    210                 215                 220

Gln Ala Leu Asn His Leu Leu Gln Gly Glu Thr Glu Gly Leu Val Lys
225                 230                 235                 240

Asn Glu Ser Lys Met Thr Asp Ser Lys Ile Asp Leu Leu Met Ser Gly
                245                 250                 255

Leu Ile Thr Gly Gln Val Val Gly Val Asn Ile Lys His Met Gln Leu
            260                 265                 270
```

```
Met Ile Ala Val Phe Val Pro Thr Thr Ala Gln Leu Pro Asn Ala Tyr
        275                 280                 285

Val Ile Asn Leu Leu Thr Ile Thr Ala Asn Ile Asn Asn Ser Glu Val
    290                 295                 300

Leu Val Gln Leu Pro Asn Gln Ile Leu Glu Arg Ser Gly Ile Ile Tyr
305                 310                 315                 320

Gln Phe Arg Gly Lys Asp Cys Val Ser Ser Pro Asn His Met Tyr Cys
                325                 330                 335

Pro Tyr Ser Asp Ala Ser Ile Leu Ser Pro Glu Leu Gln Leu Cys Leu
            340                 345                 350

Gln Gly Arg Leu Glu Met Cys Leu Phe Thr Gln Val Val Gly Ser Phe
        355                 360                 365

Pro Thr Arg Phe Ala Ser Asp Lys Gly Ile Val Tyr Ala Asn Cys Arg
    370                 375                 380

His Leu Gln Cys Ala Cys Ser Glu Pro Glu Gly Ile Ile Tyr Gln Asp
385                 390                 395                 400

Asp Thr Ser Ala Ile Thr Gln Ile Asp Ala Ser Lys Cys Ser Thr Leu
                405                 410                 415

Lys Leu Asp Met Leu Thr Phe Lys Leu Ser Thr Tyr Ala Asn Lys Thr
            420                 425                 430

Phe Asp Ala Ser Phe Ser Val Gly Lys Asp Gln Met Leu Val Thr Asn
        435                 440                 445

Leu Leu Asp Leu Ser Ala Glu Leu Lys Thr Met Asn Ala Ser Val Ala
    450                 455                 460

His Ala Asn Lys Leu Ile Asp Lys Ser Asn Leu Leu Ile Gln Ser Asn
465                 470                 475                 480

Ala Leu Ile Gly His Ser Asn Thr Ile Phe Ile Val Val Ile Val Ile
                485                 490                 495

Leu Ala Val Met Val Leu Tyr Leu Ile Ile Val Thr Tyr Ile Ile Lys
            500                 505                 510

Val Ile Met Val Glu Val Ser Arg Leu Lys Arg Met Asn Ile Tyr Ser
        515                 520                 525

Ile Asp Lys
    530
```

<210> 49
<211> 532
<212> PRT
<213> Tuhoko virus 1

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 49

Met Val Thr Ile Ile Lys Pro Leu Ile Leu Leu Val Thr Val Ile Leu
1 5 10 15

Gln Ile Ser Gly His Ile Asp Thr Thr Ala Leu Thr Ser Ile Gly Ala
20 25 30

Val Ile Ala Ser Ser Lys Glu Ile Met Tyr Tyr Ala Gln Ser Thr Pro
35 40 45

Asn Tyr Ile Val Ile Lys Leu Ile Pro Asn Leu Pro Asn Ile Pro Ser
50 55 60

Gln Cys Asn Phe Ser Ser Ile Ala Tyr Tyr Asn Lys Thr Leu Leu Asp
65 70 75 80

Leu Phe Thr Pro Ile Ser Asp Asn Ile Asn Met Leu His Gln Arg Leu
85 90 95

Ser Asn Thr Gly Arg Asn Arg Arg Phe Ala Gly Val Ala Ile Gly Leu
100 105 110

Ala Ala Leu Gly Val Ala Thr Ala Ala Gln Val Thr Ala Ala Phe Ala
115 120 125

Leu Val Glu Ala Lys Ser Asn Thr Ala Lys Ile Ala Gln Ile Gly Gln
130 135 140

Ala Ile Gln Asn Thr Asn Ala Ala Ile Asn Ser Leu Asn Ala Gly Ile
145 150 155 160

Gly Gly Ala Val Thr Ala Ile Gln Ala Ile Gln Thr Gln Ile Asn Gly
165 170 175

Ile Ile Thr Asp Gln Ile Asn Ala Ala Thr Cys Thr Ala Leu Asp Ala
180 185 190

Gln Ile Gly Thr Leu Leu Asn Met Tyr Leu Leu Gln Leu Thr Thr Thr
195 200 205

Phe Gln Pro Gln Ile Gln Asn Pro Ala Leu Gln Pro Leu Ser Ile Gln
210 215 220

```
Ala Leu His Arg Ile Met Gln Gly Thr Ser Ile Val Leu Ser Asn Leu
225                 230                 235                 240

Thr Asp Ser Ser Lys Tyr Gly Leu Asn Asp Ala Leu Ser Ala Gly Leu
                245                 250                 255

Ile Thr Gly Gln Ile Val Ser Val Asp Leu Arg Leu Met Gln Ile Thr
            260                 265                 270

Ile Ala Ala Asn Val Pro Thr Leu Ser Arg Leu Glu Asn Ala Ile Ala
        275                 280                 285

His Asp Ile Met Arg Ile Thr Thr Asn Val Asn Asn Thr Glu Val Ile
    290                 295                 300

Val Gln Leu Pro Glu Thr Ile Met Glu His Ala Gly Arg Leu Tyr Gln
305                 310                 315                 320

Phe Asn Lys Asp His Cys Leu Ser Ser Thr Gln Arg Phe Phe Cys Pro
                325                 330                 335

Tyr Ser Asp Ala Lys Leu Leu Thr Ser Lys Ile Ser Ser Cys Leu Ser
            340                 345                 350

Gly Ile Arg Gly Asp Cys Ile Phe Ser Pro Val Val Gly Asn Phe Ala
        355                 360                 365

Thr Arg Phe Ile Ser Val Lys Gly Val Ile Ile Ala Asn Cys Lys Phe
    370                 375                 380

Ile Arg Cys Thr Cys Leu Gln Pro Glu Gly Ile Ile Ser Gln Leu Asp
385                 390                 395                 400

Asp His Thr Leu Thr Val Ile Asp Leu Lys Leu Cys Asn Lys Leu Asp
                405                 410                 415

Leu Gly Leu Ile Gln Phe Asp Leu Gln Val Leu Ser Asn Ile Ser Tyr
            420                 425                 430

Glu Met Thr Leu Asn Thr Ser Gln Asn Gln Leu Ile Leu Thr Asp Pro
        435                 440                 445

Leu Asp Leu Ser Ser Glu Leu Gln Thr Met Asn Gln Ser Ile Asn Asn
    450                 455                 460

Ala Ala Asn Phe Ile Glu Lys Ser Asn Ser Leu Leu Asn Ser Ser Thr
465                 470                 475                 480

Tyr Glu Phe Asn Arg Ser Val Ala Leu Leu Val Ala Leu Ile Leu Leu
                485                 490                 495

Ser Leu Thr Ile Leu Tyr Val Ile Val Leu Thr Cys Val Val Lys Leu
```

500 505 510

Leu Val His Glu Val Ser Lys Asn Arg Arg His Ile Gln Asp Leu Glu
515 520 525

Ser His His Lys
530

<210> 50
<211> 631
<212> PRT
<213> Phpcine distemper virus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 50

Met Thr Arg Val Lys Lys Leu Pro Val Pro Thr Asn Pro Pro Met His
1 5 10 15

His Ser Leu Asp Ser Pro Phe Leu Asn Pro Glu His Ala Thr Gly Lys
20 25 30

Ile Ser Ile Thr Asp Asp Thr Ser Ser Gln Leu Thr Asn Phe Leu Tyr
35 40 45

His Lys Tyr His Lys Thr Thr Ile Asn His Leu Ser Arg Thr Ile Ser
50 55 60

Gly Thr Asp Pro Pro Ser Ala Lys Leu Asn Lys Phe Gly Ser Pro Ile
65 70 75 80

Leu Ser Thr Tyr Gln Ile Arg Ser Ala Leu Trp Trp Ile Ala Met Val
85 90 95

Ile Leu Val His Cys Val Met Gly Gln Ile His Trp Thr Asn Leu Ser
100 105 110

Thr Ile Gly Ile Ile Gly Thr Asp Ser Ser His Tyr Lys Ile Met Thr
115 120 125

Arg Ser Ser His Gln Tyr Leu Val Leu Lys Leu Met Pro Asn Val Ser
130 135 140

Ile Ile Asp Asn Cys Thr Lys Ala Glu Leu Asp Glu Tyr Glu Lys Leu
145 150 155 160

Leu Asn Ser Val Leu Glu Pro Ile Asn Gln Ala Leu Thr Leu Met Thr
165 170 175

Lys Asn Val Lys Ser Leu Gln Ser Leu Gly Ser Gly Arg Arg Gln Arg
180 185 190

```
Arg Phe Ala Gly Val Val Ile Ala Gly Ala Ala Leu Gly Val Ala Thr
        195                 200                 205

Ala Ala Gln Ile Thr Ala Gly Val Ala Leu Tyr Gln Ser Asn Leu Asn
    210                 215                 220

Ala Gln Ala Ile Gln Ser Leu Arg Ala Ser Leu Glu Gln Ser Asn Lys
225                 230                 235                 240

Ala Ile Asp Glu Val Arg Gln Ala Ser Gln Asn Ile Ile Ile Ala Val
                245                 250                 255

Gln Gly Val Gln Asp Tyr Val Asn Asn Glu Ile Val Pro Ala Leu Gln
            260                 265                 270

His Met Ser Cys Glu Leu Ile Gly Gln Arg Leu Gly Leu Lys Leu Leu
        275                 280                 285

Arg Tyr Tyr Thr Glu Leu Leu Ser Val Phe Gly Pro Ser Leu Arg Asp
    290                 295                 300

Pro Val Ser Ala Glu Ile Ser Ile Gln Ala Leu Ser Tyr Ala Leu Gly
305                 310                 315                 320

Gly Glu Ile His Lys Ile Leu Glu Lys Leu Gly Tyr Ser Gly Asn Asp
                325                 330                 335

Met Val Ala Ile Leu Glu Thr Lys Gly Ile Arg Ala Lys Ile Thr His
            340                 345                 350

Val Asp Leu Ser Gly Lys Phe Ile Val Leu Ser Ile Ser Tyr Pro Thr
        355                 360                 365

Leu Ser Glu Val Lys Gly Val Val Val His Arg Leu Glu Ala Val Ser
    370                 375                 380

Tyr Asn Ile Gly Ser Gln Glu Trp Tyr Thr Thr Val Pro Arg Tyr Val
385                 390                 395                 400

Ala Thr Asn Gly Tyr Leu Ile Ser Asn Phe Asp Glu Ser Ser Cys Val
                405                 410                 415

Phe Val Ser Glu Ser Ala Ile Cys Ser Gln Asn Ser Leu Tyr Pro Met
            420                 425                 430

Ser Pro Ile Leu Gln Gln Cys Leu Arg Gly Glu Thr Ala Ser Cys Ala
        435                 440                 445

Arg Thr Leu Val Ser Gly Thr Leu Gly Asn Lys Phe Ile Leu Ser Lys
    450                 455                 460
```

```
Gly Asn Ile Ile Ala Asn Cys Ala Ser Ile Leu Cys Lys Cys His Ser
465                 470                 475                 480


Thr Ser Lys Ile Ile Asn Gln Ser Pro Asp Lys Leu Leu Thr Phe Ile
                485                 490                 495


Ala Ser Asp Thr Cys Ser Leu Val Glu Ile Asp Gly Val Thr Ile Gln
            500                 505                 510


Val Gly Ser Arg Gln Tyr Pro Asp Val Val Tyr Ala Ser Lys Val Ile
        515                 520                 525


Leu Gly Pro Ala Ile Ser Leu Glu Arg Leu Asp Val Gly Thr Asn Leu
    530                 535                 540


Gly Ser Ala Leu Lys Lys Leu Asn Asp Ala Lys Val Leu Ile Glu Ser
545                 550                 555                 560


Ser Asp Gln Ile Leu Asp Thr Val Lys Asn Ser Tyr Leu Ser Leu Gly
                565                 570                 575


Thr Leu Ile Ala Leu Pro Val Ser Ile Gly Leu Gly Leu Ile Leu Leu
            580                 585                 590


Leu Leu Ile Cys Cys Cys Lys Lys Arg Tyr Gln His Leu Phe Ser Gln
        595                 600                 605


Ser Thr Lys Val Ala Pro Val Phe Lys Pro Asp Leu Thr Gly Thr Ser
    610                 615                 620


Lys Ser Tyr Val Arg Ser Leu
625                 630


<210>  51
<211>  541
<212>  PRT
<213>  Caprine parainfluenza virus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  51

Met Ile Lys Lys Ile Ile Cys Ile Phe Ser Met Pro Ile Leu Leu Ser
1               5                   10                  15


Phe Cys Gln Val Asp Ile Ile Lys Leu Gln Arg Val Gly Ile Leu Val
            20                  25                  30


Ser Lys Pro Lys Ser Ile Lys Ile Ser Gln Asn Phe Glu Thr Arg Tyr
        35                  40                  45
```

Leu Val Leu Asn Leu Ile Pro Asn Ile Glu Asn Ala Gln Ser Cys Gly
50 55 60

Asp Gln Gln Ile Lys Gln Tyr Lys Lys Leu Leu Asp Arg Leu Ile Ile
65 70 75 80

Pro Leu Tyr Asp Gly Leu Arg Leu Gln Gln Asp Ile Ile Val Val Asp
85 90 95

Asn Asn Leu Lys Asn Asn Thr Asn His Arg Ala Lys Arg Phe Phe Gly
100 105 110

Glu Ile Ile Gly Thr Ile Ala Leu Gly Val Ala Thr Ser Ala Gln Ile
115 120 125

Thr Ala Ala Val Ala Leu Val Glu Ala Lys Gln Ala Arg Ser Asp Ile
130 135 140

Glu Arg Val Lys Asn Ala Val Arg Asp Thr Asn Lys Ala Val Gln Ser
145 150 155 160

Ile Gln Gly Ser Val Gly Asn Leu Ile Val Ala Val Lys Ser Val Gln
165 170 175

Asp Tyr Val Asn Asn Glu Ile Val Pro Ser Ile Lys Arg Leu Gly Cys
180 185 190

Glu Ala Ala Gly Leu Gln Leu Gly Ile Ala Leu Thr Gln His Tyr Ser
195 200 205

Glu Leu Thr Asn Ile Phe Gly Asp Asn Ile Gly Thr Leu Lys Glu Lys
210 215 220

Gly Ile Lys Leu Gln Gly Ile Ala Ser Leu Tyr His Thr Asn Ile Thr
225 230 235 240

Glu Ile Phe Thr Thr Ser Thr Val Asp Gln Tyr Asp Ile Tyr Asp Leu
245 250 255

Leu Phe Thr Glu Ser Ile Lys Met Arg Val Ile Asp Val Asp Leu Asn
260 265 270

Asp Tyr Ser Ile Thr Leu Gln Val Arg Leu Pro Leu Leu Thr Lys Ile
275 280 285

Ser Asp Ala Gln Ile Tyr Asn Val Asp Ser Val Ser Tyr Asn Ile Gly
290 295 300

Gly Thr Glu Trp Tyr Ile Pro Leu Pro Arg Asn Ile Met Thr Lys Gly
305 310 315 320

Ala Phe Leu Gly Gly Ala Asn Leu Gln Asp Cys Ile Glu Ser Phe Ser

325 330 335

Asp Tyr Ile Cys Pro Ser Asp Pro Gly Phe Ile Leu Asn Arg Asp Ile
340 345 350

Glu Asn Cys Leu Ser Gly Asn Ile Thr Gln Cys Pro Lys Thr Leu Val
355 360 365

Ile Ser Asp Ile Val Pro Arg Tyr Ala Phe Val Asp Gly Gly Val Ile
370 375 380

Ala Asn Cys Leu Ser Thr Thr Cys Thr Cys Asn Gly Ile Asp Asn Arg
385 390 395 400

Ile Asn Gln Ala Pro Asp Gln Gly Ile Lys Ile Ile Thr Tyr Lys Asp
405 410 415

Cys Gln Thr Ile Gly Ile Asn Gly Met Leu Phe Lys Thr Asn Gln Glu
420 425 430

Gly Thr Leu Ala Ala Tyr Thr Pro Val Asp Ile Thr Leu Asn Asn Ser
435 440 445

Val Asn Leu Asp Pro Ile Asp Leu Ser Ile Glu Leu Asn Arg Ala Arg
450 455 460

Ser Asp Leu Ala Glu Ser Lys Glu Trp Ile Lys Arg Ser Glu Ala Lys
465 470 475 480

Leu Asp Ser Val Gly Ser Trp Tyr Gln Ser Ser Thr Thr Glu Ile Ile
485 490 495

Gln Ile Val Met Ile Ile Val Leu Phe Ile Ile Asn Ile Ile Val Leu
500 505 510

Ile Val Leu Ile Lys Tyr Ser Arg Ser Gln Asn Gln Ser Met Asn Asn
515 520 525

His Met Asn Glu Pro Tyr Ile Leu Thr Asn Lys Val Gln
530 535 540

<210> 52
<211> 553
<212> PRT
<213> Tupaia paramyxovirus

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 52

Met Ala Ser Leu Leu Lys Thr Ile Cys Tyr Ile Tyr Leu Ile Thr Tyr
1 5 10 15

```
Ala Lys Leu Glu Pro Thr Pro Lys Ser Gln Leu Asp Leu Asp Ser Leu
            20                  25                  30

Ala Ser Ile Gly Val Val Asp Ala Gly Lys Tyr Asn Tyr Lys Leu Met
        35                  40                  45

Thr Thr Gly Ser Glu Lys Leu Met Val Ile Lys Leu Val Pro Asn Ile
    50                  55                  60

Thr Tyr Ala Thr Asn Cys Asn Leu Thr Ala His Thr Ala Tyr Thr Lys
65                  70                  75                  80

Met Ile Glu Arg Leu Leu Thr Pro Ile Asn Gln Ser Leu Tyr Glu Met
                85                  90                  95

Arg Ser Val Ile Thr Glu Arg Asp Gly Gly Thr Ile Phe Trp Gly Ala
            100                 105                 110

Ile Ile Ala Gly Ala Ala Leu Gly Val Ala Thr Ala Ala Ala Ile Thr
        115                 120                 125

Ala Gly Val Ala Leu His Arg Ala Glu Gln Asn Ala Arg Asn Ile Ala
    130                 135                 140

Ala Leu Lys Asp Ala Leu Arg Asn Ser Asn Glu Ala Ile Gln His Leu
145                 150                 155                 160

Lys Asp Ala Gln Gly His Thr Val Leu Ala Ile Gln Gly Leu Gln Glu
                165                 170                 175

Gln Ile Asn Asn Asn Ile Ile Pro Lys Leu Lys Glu Ser His Cys Leu
            180                 185                 190

Gly Val Asn Asn Gln Leu Gly Leu Leu Leu Asn Gln Tyr Tyr Ser Glu
        195                 200                 205

Ile Leu Thr Val Phe Gly Pro Asn Leu Gln Asn Pro Val Ser Ala Ser
    210                 215                 220

Leu Thr Ile Gln Ala Ile Ala Lys Ala Phe Asn Gly Asp Phe Asn Ser
225                 230                 235                 240

Leu Met Thr Asn Leu Asn Tyr Asp Pro Thr Asp Leu Leu Asp Ile Leu
                245                 250                 255

Glu Ser Asn Ser Ile Asn Gly Arg Ile Ile Asp Val Asn Leu Asn Glu
            260                 265                 270

Lys Tyr Ile Ala Leu Ser Ile Glu Ile Pro Asn Phe Ile Thr Leu Thr
        275                 280                 285
```

```
Asp Ala Lys Ile Gln Thr Phe Asn Arg Ile Thr Tyr Gly Tyr Gly Ser
    290                 295                 300

Asn Glu Trp Leu Thr Leu Ile Pro Asp Asn Ile Leu Glu Tyr Gly Asn
305                 310                 315                 320

Leu Ile Ser Asn Val Asp Leu Thr Ser Cys Val Lys Thr Lys Ser Ser
                325                 330                 335

Tyr Ile Cys Asn Gln Asp Thr Ser Tyr Pro Ile Ser Ser Glu Leu Thr
            340                 345                 350

Arg Cys Leu Arg Gly Asp Thr Ser Ser Cys Pro Arg Thr Pro Val Val
        355                 360                 365

Asn Ser Arg Ala Pro Thr Phe Ala Leu Ser Gly Gly His Ile Tyr Ala
    370                 375                 380

Asn Cys Ala Lys Ala Ala Cys Arg Cys Glu Lys Pro Pro Met Ala Ile
385                 390                 395                 400

Val Gln Pro Ala Thr Ser Thr Leu Thr Phe Leu Thr Glu Lys Glu Cys
                405                 410                 415

Gln Glu Val Val Ile Asp Gln Ile Asn Ile Gln Leu Ala Pro Asn Arg
            420                 425                 430

Leu Asn Lys Thr Ile Ile Thr Asp Gly Ile Asp Leu Gly Pro Glu Val
        435                 440                 445

Ile Ile Asn Pro Ile Asp Val Ser Ala Glu Leu Gly Asn Ile Glu Leu
    450                 455                 460

Glu Met Asp Lys Thr Gln Lys Ala Leu Asp Arg Ser Asn Lys Ile Leu
465                 470                 475                 480

Asp Ser Met Ile Thr Glu Val Thr Pro Asp Lys Leu Leu Ile Ala Met
                485                 490                 495

Ile Val Val Phe Gly Ile Leu Leu Leu Trp Leu Phe Gly Val Ser Tyr
            500                 505                 510

Tyr Ala Phe Lys Ile Trp Ser Lys Leu His Phe Leu Asp Ser Tyr Val
        515                 520                 525

Tyr Ser Leu Arg Asn Pro Ser His His Arg Ser Asn Gly His Gln Asn
    530                 535                 540

His Ser Phe Ser Thr Asp Ile Ser Gly
545                 550
```

<210> 53
<211> 543
<212> PRT
<213> Avian paramyxovirus 3

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 53

Met Gln Pro Gly Ser Ala Leu His Leu Pro His Leu Tyr Ile Ile Ile
1 5 10 15

Ala Leu Val Ser Asp Gly Thr Leu Gly Gln Thr Ala Lys Ile Asp Arg
20 25 30

Leu Ile Gln Ala Gly Ile Val Leu Gly Ser Gly Lys Glu Leu His Ile
35 40 45

Ser Gln Asp Ser Gly Thr Leu Asp Leu Phe Val Arg Leu Leu Pro Val
50 55 60

Leu Pro Ser Asn Leu Ser His Cys Gln Leu Glu Ala Ile Thr Gln Tyr
65 70 75 80

Asn Lys Thr Val Thr Arg Leu Leu Ala Pro Ile Gly Lys Asn Leu Glu
85 90 95

Gln Val Leu Gln Ala Arg Pro Arg Gly Arg Leu Phe Gly Pro Ile Ile
100 105 110

Gly Ser Ile Ala Leu Gly Val Ala Thr Ser Ala Gln Ile Thr Ala Ala
115 120 125

Ile Ala Leu Val Arg Ala Gln Gln Asn Ala Asn Asp Ile Leu Ala Leu
130 135 140

Lys Asn Ala Leu Gln Ser Ser Asn Glu Ala Ile Arg Gln Leu Thr Tyr
145 150 155 160

Gly Gln Asp Lys Gln Leu Leu Ala Ile Ser Lys Ile Gln Lys Ala Val
165 170 175

Asn Glu Gln Ile Leu Pro Ala Leu Asp Gln Leu Asp Cys Ala Val Leu
180 185 190

Gly Thr Lys Leu Ala Val Gln Leu Asn Leu Tyr Leu Ile Glu Met Thr
195 200 205

Thr Ile Phe Gly Glu Gln Ile Asn Asn Pro Val Leu Ala Thr Ile Pro
210 215 220

Leu Ser Tyr Ile Leu Arg Leu Thr Gly Ala Glu Leu Asn Asn Val Leu

225 230 235 240

Met Lys Gln Ala Arg Ser Ser Leu Ser Leu Val Gln Leu Val Ser Lys
245 250 255

Gly Leu Leu Ser Gly Gln Val Ile Gly Tyr Asp Pro Ser Val Gln Gly
260 265 270

Leu Ile Ile Arg Val Asn Leu Met Arg Thr Gln Lys Ile Asp Arg Ala
275 280 285

Leu Val Tyr Gln Pro Tyr Val Leu Pro Ile Thr Leu Asn Ser Asn Ile
290 295 300

Val Thr Pro Ile Ala Pro Glu Cys Val Ile Gln Lys Gly Thr Ile Ile
305 310 315 320

Glu Gly Met Ser Arg Lys Asp Cys Thr Glu Leu Glu Gln Asp Ile Ile
325 330 335

Cys Arg Thr Val Thr Thr Tyr Thr Leu Ala Arg Asp Thr Arg Leu Cys
340 345 350

Leu Gln Gly Asn Ile Ser Ser Cys Arg Tyr Gln Gln Ser Gly Thr Gln
355 360 365

Leu His Thr Pro Phe Ile Thr Tyr Asn Gly Ala Val Ile Ala Asn Cys
370 375 380

Asp Leu Val Ser Cys Arg Cys Leu Arg Pro Pro Met Ile Ile Thr Gln
385 390 395 400

Val Lys Gly Tyr Pro Leu Thr Ile Ile Thr Arg Ser Val Cys Gln Glu
405 410 415

Leu Ser Val Asp Asn Leu Val Leu Asn Ile Glu Thr His His Asn Phe
420 425 430

Ser Leu Asn Pro Thr Ile Ile Asp Pro Leu Thr Arg Val Ile Ala Thr
435 440 445

Thr Pro Leu Glu Ile Asp Ser Leu Ile Gln Glu Ala Gln Asp His Ala
450 455 460

Asn Ala Ala Leu Ala Lys Val Glu Glu Ser Asp Lys Tyr Leu Arg Ala
465 470 475 480

Val Thr Gly Gly Asn Tyr Ser Asn Trp Tyr Ile Val Leu Val Ile Val
485 490 495

Leu Leu Phe Gly Asn Leu Gly Trp Ser Leu Leu Leu Thr Val Leu Leu
500 505 510

```
Cys Arg Ser Arg Lys Gln Gln Arg Arg Tyr Gln Gln Asp Asp Ser Val
        515                 520                 525


Gly Ser Glu Arg Gly Val Gly Val Gly Thr Ile Gln Tyr Met Ser
    530                 535                 540


<210>  54
<211>  541
<212>  PRT
<213>  Avian paramyxovirus 14


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  54

Met Glu Lys Gly Thr Val Leu Phe Leu Ala Ala Leu Thr Leu Tyr Asn
1               5                   10                  15


Val Lys Ala Leu Asp Asn Thr Lys Leu Leu Gly Ala Gly Ile Ala Ser
            20                  25                  30


Gly Lys Glu His Glu Leu Lys Ile Tyr Gln Ser Ser Val Asn Gly Tyr
        35                  40                  45


Ile Ala Val Lys Leu Ile Pro Phe Leu Pro Ser Thr Lys Arg Glu Cys
    50                  55                  60


Tyr Asn Glu Gln Leu Lys Asn Tyr Asn Ala Thr Ile Asn Arg Leu Met
65                  70                  75                  80


Gly Pro Ile Asn Asp Asn Ile Lys Leu Val Leu Ser Gly Val Lys Thr
                85                  90                  95


Arg Thr Arg Glu Gly Lys Leu Ile Gly Ala Ile Ile Gly Thr Ala Ala
            100                 105                 110


Leu Gly Leu Ala Thr Ala Ala Gln Val Thr Ala Ala Ile Ala Leu Glu
        115                 120                 125


Gln Ala Gln Asp Asn Ala Arg Ala Ile Leu Thr Leu Lys Glu Ser Ile
    130                 135                 140


Arg Asn Thr Asn Asn Ala Val Ser Glu Leu Lys Thr Gly Leu Ser Glu
145                 150                 155                 160


Val Ser Ile Ala Leu Ser Lys Thr Gln Asp Tyr Ile Asn Thr Gln Ile
                165                 170                 175


Met Pro Ala Leu Ser Asn Leu Ser Cys Glu Ile Val Gly Leu Lys Ile
            180                 185                 190
```

```
Gly Ile Gln Leu Ser Gln Tyr Leu Thr Glu Val Thr Ala Val Phe Gly
        195                 200                 205

Asn Gln Ile Thr Asn Pro Ala Leu Gln Pro Leu Ser Met Gln Ala Leu
    210                 215                 220

Tyr Gln Leu Cys Gly Gly Asp Phe Ser Leu Leu Leu Asp Lys Ile Gly
225                 230                 235                 240

Ala Asp Arg Asn Glu Leu Glu Ser Leu Tyr Glu Ala Asn Leu Val Thr
                245                 250                 255

Gly Arg Ile Val Gln Tyr Asp Thr Ala Asp Gln Leu Val Ile Ile Gln
            260                 265                 270

Val Ser Ile Pro Ser Val Ser Thr Leu Ser Gly Tyr Arg Val Thr Glu
        275                 280                 285

Leu Gln Ser Ile Ser Val Asp Met Asp His Gly Glu Gly Lys Ala Val
    290                 295                 300

Ile Pro Arg Tyr Ile Val Thr Ser Gly Arg Val Ile Glu Glu Met Asp
305                 310                 315                 320

Ile Ser Pro Cys Val Leu Thr Ala Thr Ala Val Tyr Cys Asn Arg Leu
                325                 330                 335

Leu Thr Thr Ser Leu Pro Glu Ser Val Leu Lys Cys Leu Asp Gly Asp
            340                 345                 350

His Ser Ser Cys Thr Tyr Thr Ser Asn Ser Gly Val Leu Glu Thr Arg
        355                 360                 365

Tyr Ile Ala Phe Asp Gly Met Leu Ile Ala Asn Cys Arg Ser Ile Val
    370                 375                 380

Cys Lys Cys Leu Asp Pro Pro Tyr Ile Ile Pro Gln Asn Lys Gly Lys
385                 390                 395                 400

Pro Leu Thr Ile Ile Ser Lys Glu Val Cys Lys Lys Val Thr Leu Asp
                405                 410                 415

Gly Ile Thr Leu Leu Ile Asp Ala Glu Phe Thr Gly Glu Tyr Gly Leu
            420                 425                 430

Asn Ile Thr Ile Gly Pro Asp Gln Phe Ala Pro Ser Gly Ala Leu Asp
        435                 440                 445

Ile Ser Thr Glu Leu Gly Lys Leu Asn Asn Ser Ile Asn Lys Ala Glu
    450                 455                 460
```

Asp Tyr Ile Asp Lys Ser Asn Glu Leu Leu Asn Arg Val Asn Val Asp
465 470 475 480

Ile Val Asn Asp Thr Ala Val Ile Val Leu Cys Val Met Ser Ala Leu
485 490 495

Val Val Val Trp Cys Ile Gly Leu Thr Val Gly Leu Ile Tyr Val Ser
500 505 510

Lys Asn Thr Leu Arg Ala Val Ala Ile Lys Gly Thr Ser Ile Glu Asn
515 520 525

Pro Tyr Val Ser Ser Gly Lys His Ala Lys Asn Ser Ser
530 535 540

<210> 55
<211> 551
<212> PRT
<213> Avian paramyxovirus UPO216

<220>
<221> MISC_FEATURE
<223> fusion protein

<400> 55

Met Ile Phe Thr Met Tyr His Val Thr Val Leu Leu Leu Leu Ser Leu
1 5 10 15

Leu Thr Leu Pro Leu Gly Ile Gln Leu Ala Arg Ala Ser Ile Asp Gly
20 25 30

Arg Gln Leu Ala Ala Ala Gly Ile Val Val Thr Gly Glu Lys Ala Ile
35 40 45

Asn Leu Tyr Thr Ser Ser Gln Thr Gly Thr Ile Val Val Lys Leu Leu
50 55 60

Pro Asn Val Pro Gln Gly Arg Glu Ala Cys Met Arg Asp Pro Leu Thr
65 70 75 80

Ser Tyr Asn Lys Thr Leu Thr Ser Leu Leu Ser Pro Leu Gly Glu Ala
85 90 95

Ile Arg Arg Ile His Glu Ser Thr Thr Glu Thr Ala Gly Leu Val Gln
100 105 110

Ala Arg Leu Val Gly Ala Ile Ile Gly Ser Val Ala Leu Gly Val Ala
115 120 125

Thr Ser Ala Gln Ile Thr Ala Ala Ala Ala Leu Ile Gln Ala Asn Lys
130 135 140

Asn Ala Glu Asn Ile Leu Lys Leu Lys Gln Ser Ile Ala Ala Thr Asn

145 150 155 160

Glu Ala Val His Glu Val Thr Asp Gly Leu Ser Gln Leu Ala Val Ala
165 170 175

Val Gly Lys Met Gln Asp Phe Ile Asn Thr Gln Phe Asn Asn Thr Ala
180 185 190

Gln Glu Ile Asp Cys Ile Arg Ile Ser Gln Gln Leu Gly Val Glu Leu
195 200 205

Asn Leu Tyr Leu Thr Glu Leu Thr Thr Val Phe Gly Pro Gln Ile Thr
210 215 220

Ser Pro Ala Leu Ser Pro Leu Ser Ile Gln Ala Leu Tyr Asn Leu Ala
225 230 235 240

Gly Gly Asn Leu Asp Val Leu Leu Ser Lys Ile Gly Val Gly Asn Asn
245 250 255

Gln Leu Ser Ala Leu Ile Ser Ser Gly Leu Ile Ser Gly Ser Pro Ile
260 265 270

Leu Tyr Asp Ser Gln Thr Gln Leu Leu Gly Ile Gln Val Thr Leu Pro
275 280 285

Ser Val Ser Ser Leu Asn Asn Met Arg Ala Ile Phe Leu Glu Thr Leu
290 295 300

Ser Val Ser Thr Asp Lys Gly Phe Ala Ala Ala Leu Ile Pro Lys Val
305 310 315 320

Val Thr Thr Val Gly Thr Val Thr Glu Glu Leu Asp Thr Ser Tyr Cys
325 330 335

Ile Glu Thr Asp Ile Asp Leu Phe Cys Thr Arg Ile Val Thr Phe Pro
340 345 350

Met Ser Pro Gly Ile Tyr Ala Cys Leu Asn Gly Asn Thr Ser Glu Cys
355 360 365

Met Tyr Ser Lys Thr Gln Gly Ala Leu Thr Thr Pro Tyr Met Ser Val
370 375 380

Lys Gly Ser Ile Val Ala Asn Cys Lys Met Thr Thr Cys Arg Cys Ala
385 390 395 400

Asp Pro Ala Ser Ile Ile Ser Gln Asn Tyr Gly Glu Ala Val Ser Leu
405 410 415

Ile Asp Ser Ser Val Cys Arg Val Ile Thr Leu Asp Gly Val Thr Leu
420 425 430

```
Arg Leu Ser Gly Ser Phe Asp Ser Thr Tyr Gln Lys Asn Ile Thr Ile
        435                 440                 445


Arg Asp Ser Gln Val Ile Ile Thr Gly Ser Leu Asp Ile Ser Thr Glu
    450                 455                 460


Leu Gly Asn Val Asn Asn Ser Ile Asn Asn Ala Leu Asp Lys Ile Glu
465                 470                 475                 480


Glu Ser Asn Gln Ile Leu Glu Ser Val Asn Val Ser Leu Thr Ser Thr
                485                 490                 495


Asn Ala Leu Ile Val Tyr Ile Ile Cys Thr Ala Leu Ala Leu Ile Cys
            500                 505                 510


Gly Ile Thr Gly Leu Ile Leu Ser Cys Tyr Ile Met Tyr Lys Met Arg
        515                 520                 525


Ser Gln Gln Lys Thr Leu Met Trp Leu Gly Asn Asn Thr Leu Asp Gln
    530                 535                 540


Met Arg Ala Gln Thr Lys Met
545                 550


<210>  56
<211>  552
<212>  PRT
<213>  Atlantic salmon paramyxovirus


<220>
<221>  MISC_FEATURE
<223>  fusion protein

<400>  56

Met Asp Gly Pro Lys Phe Arg Phe Val Leu Leu Ile Leu Leu Thr Ala
1               5                   10                  15


Pro Ala Arg Gly Gln Val Asp Tyr Asp Lys Leu Leu Lys Val Gly Ile
            20                  25                  30


Phe Glu Lys Gly Thr Ala Asn Leu Lys Ile Ser Val Ser Ser Gln Gln
        35                  40                  45


Arg Tyr Met Val Ile Lys Met Met Pro Asn Leu Gly Pro Met Asn Gln
    50                  55                  60


Cys Gly Ile Lys Glu Val Asn Leu Tyr Lys Glu Ser Ile Leu Arg Leu
65                  70                  75                  80


Ile Thr Pro Ile Ser Thr Thr Leu Asn Tyr Ile Lys Ser Glu Ile Gln
                85                  90                  95
```

```
Val Glu Arg Glu Val Ala Leu Gln Pro Asn Gly Thr Ile Val Arg Phe
            100                 105                 110

Phe Gly Leu Ile Val Ala Ala Gly Ala Leu Thr Leu Ala Thr Ser Ala
        115                 120                 125

Gln Ile Thr Ala Gly Ile Ala Leu His Asn Ser Leu Glu Asn Ala Lys
    130                 135                 140

Ala Ile Lys Gly Leu Thr Asp Ala Ile Lys Glu Ser Asn Leu Ala Ile
145                 150                 155                 160

Gln Lys Ile Gln Asp Ala Thr Ala Gly Thr Val Ile Ala Leu Asn Ala
                165                 170                 175

Leu Gln Asp Gln Val Asn Thr Asn Ile Ile Pro Ala Ile Asn Thr Leu
            180                 185                 190

Gly Cys Thr Ala Ala Gly Asn Thr Leu Gly Ile Ala Leu Thr Arg Tyr
        195                 200                 205

Tyr Ser Glu Leu Ile Met Ile Phe Gly Pro Ser Leu Gly Asn Pro Val
    210                 215                 220

Glu Ala Pro Leu Thr Ile Gln Ala Leu Ala Gly Ala Phe Asn Gly Asp
225                 230                 235                 240

Leu His Gly Met Ile Arg Glu Tyr Gly Tyr Thr Pro Ser Asp Ile Glu
                245                 250                 255

Asp Ile Leu Arg Thr Asn Ser Val Thr Gly Arg Val Ile Asp Val Asp
            260                 265                 270

Leu Val Gly Met Asn Ile Val Leu Glu Ile Asn Leu Pro Thr Leu Tyr
        275                 280                 285

Thr Leu Arg Asp Thr Lys Ile Val Asn Leu Gly Lys Ile Thr Tyr Asn
    290                 295                 300

Val Asp Gly Ser Glu Trp Gln Thr Leu Val Pro Glu Trp Leu Ala Ile
305                 310                 315                 320

Arg Asn Thr Leu Met Gly Gly Val Asp Leu Ser Arg Cys Val Val Ser
                325                 330                 335

Ser Arg Asp Leu Ile Cys Lys Gln Asp Pro Val Phe Ser Leu Asp Thr
            340                 345                 350

Ser Ile Ile Ser Cys Leu Asn Gly Asn Thr Glu Ser Cys Pro Arg Asn
        355                 360                 365
```

Arg Val Val Asn Ser Val Ala Pro Arg Tyr Ala Val Ile Arg Gly Asn
370 375 380

Ile Leu Ala Asn Cys Ile Ser Thr Thr Cys Leu Cys Gly Asp Pro Gly
385 390 395 400

Val Pro Ile Ile Gln Lys Gly Asp Asn Thr Leu Thr Ala Met Ser Ile
405 410 415

Asn Asp Cys Lys Leu Val Gly Val Asp Gly Tyr Val Phe Arg Pro Gly
420 425 430

Pro Lys Ala Val Asn Val Thr Phe Asn Leu Pro His Leu Asn Leu Gly
435 440 445

Pro Glu Val Asn Val Asn Pro Val Asp Ile Ser Gly Ala Leu Gly Lys
450 455 460

Val Glu Gln Asp Leu Ala Ser Ser Arg Asp His Leu Ala Lys Ser Glu
465 470 475 480

Lys Ile Leu Ser Gly Ile Asn Pro Asn Ile Ile Asn Thr Glu Met Val
485 490 495

Leu Val Ala Val Ile Leu Ser Leu Val Cys Ala Met Val Val Ile Gly
500 505 510

Ile Val Cys Trp Leu Ser Ile Leu Thr Lys Trp Val Arg Ser Cys Arg
515 520 525

Ala Asp Cys Arg Arg Pro Asn Lys Gly Pro Asp Leu Gly Pro Ile Met
530 535 540

Ser Ser Gln Asp Asn Leu Ser Phe
545 550

<210> 57
<211> 546
<212> PRT
<213> Nipah virus

<220>
<221> MISC_FEATURE
<223> fusion glycoprotein F0

<400> 57

Met Val Val Ile Leu Asp Lys Arg Cys Tyr Cys Asn Leu Leu Ile Leu
1 5 10 15

Ile Leu Met Ile Ser Glu Cys Ser Val Gly Ile Leu His Tyr Glu Lys
20 25 30

Leu Ser Lys Ile Gly Leu Val Lys Gly Val Thr Arg Lys Tyr Lys Ile

```
        35                  40                  45

Lys Ser Asn Pro Leu Thr Lys Asp Ile Val Ile Lys Met Ile Pro Asn
    50                  55                  60

Val Ser Asn Met Ser Gln Cys Thr Gly Ser Val Met Glu Asn Tyr Lys
65                  70                  75                  80

Thr Arg Leu Asn Gly Ile Leu Thr Pro Ile Lys Gly Ala Leu Glu Ile
                85                  90                  95

Tyr Lys Asn Asn Thr His Asp Leu Val Gly Asp Val Arg Leu Ala Gly
            100                 105                 110

Val Ile Met Ala Gly Val Ala Ile Gly Ile Ala Thr Ala Ala Gln Ile
        115                 120                 125

Thr Ala Gly Val Ala Leu Tyr Glu Ala Met Lys Asn Ala Asp Asn Ile
    130                 135                 140

Asn Lys Leu Lys Ser Ser Ile Glu Ser Thr Asn Glu Ala Val Val Lys
145                 150                 155                 160

Leu Gln Glu Thr Ala Glu Lys Thr Val Tyr Val Leu Thr Ala Leu Gln
                165                 170                 175

Asp Tyr Ile Asn Thr Asn Leu Val Pro Thr Ile Asp Lys Ile Ser Cys
            180                 185                 190

Lys Gln Thr Glu Leu Ser Leu Asp Leu Ala Leu Ser Lys Tyr Leu Ser
        195                 200                 205

Asp Leu Leu Phe Val Phe Gly Pro Asn Leu Gln Asp Pro Val Ser Asn
    210                 215                 220

Ser Met Thr Ile Gln Ala Ile Ser Gln Ala Phe Gly Gly Asn Tyr Glu
225                 230                 235                 240

Thr Leu Leu Arg Thr Leu Gly Tyr Ala Thr Glu Asp Phe Asp Asp Leu
                245                 250                 255

Leu Glu Ser Asp Ser Ile Thr Gly Gln Ile Ile Tyr Val Asp Leu Ser
            260                 265                 270

Ser Tyr Tyr Ile Ile Val Arg Val Tyr Phe Pro Ile Leu Thr Glu Ile
        275                 280                 285

Gln Gln Ala Tyr Ile Gln Glu Leu Leu Pro Val Ser Phe Asn Asn Asp
    290                 295                 300

Asn Ser Glu Trp Ile Ser Ile Val Pro Asn Phe Ile Leu Val Arg Asn
305                 310                 315                 320
```

```
Thr Leu Ile Ser Asn Ile Glu Ile Gly Phe Cys Leu Ile Thr Lys Arg
                325                 330                 335

Ser Val Ile Cys Asn Gln Asp Tyr Ala Thr Pro Met Thr Asn Asn Met
            340                 345                 350

Arg Glu Cys Leu Thr Gly Ser Thr Glu Lys Cys Pro Arg Glu Leu Val
        355                 360                 365

Val Ser Ser His Val Pro Arg Phe Ala Leu Ser Asn Gly Val Leu Phe
    370                 375                 380

Ala Asn Cys Ile Ser Val Thr Cys Gln Cys Gln Thr Thr Gly Arg Ala
385                 390                 395                 400

Ile Ser Gln Ser Gly Glu Gln Thr Leu Leu Met Ile Asp Asn Thr Thr
                405                 410                 415

Cys Pro Thr Ala Val Leu Gly Asn Val Ile Ile Ser Leu Gly Lys Tyr
            420                 425                 430

Leu Gly Ser Val Asn Tyr Asn Ser Glu Gly Ile Ala Ile Gly Pro Pro
        435                 440                 445

Val Phe Thr Asp Lys Val Asp Ile Ser Ser Gln Ile Ser Ser Met Asn
    450                 455                 460

Gln Ser Leu Gln Gln Ser Lys Asp Tyr Ile Lys Glu Ala Gln Arg Leu
465                 470                 475                 480

Leu Asp Thr Val Asn Pro Ser Leu Ile Ser Met Leu Ser Met Ile Ile
                485                 490                 495

Leu Tyr Val Leu Ser Ile Ala Ser Leu Cys Ile Gly Leu Ile Thr Phe
            500                 505                 510

Ile Ser Phe Ile Ile Val Glu Lys Lys Arg Asn Thr Tyr Ser Arg Leu
        515                 520                 525

Glu Asp Arg Arg Val Arg Pro Thr Ser Ser Gly Asp Leu Tyr Tyr Ile
    530                 535                 540

Gly Thr
545
```

<210> 58
<211> 321
<212> PRT
<213> Human respiratory syncytial virus

<220>

<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 58

```
Met Ser Lys Thr Lys Asp Gln Arg Thr Ala Lys Thr Leu Glu Arg Thr
1               5                   10                  15


Trp Asp Thr Leu Asn His Leu Leu Phe Ile Ser Ser Cys Leu Tyr Lys
            20                  25                  30


Leu Asn Leu Lys Ser Ile Ala Gln Ile Thr Leu Ser Ile Leu Ala Met
        35                  40                  45


Ile Ile Ser Thr Ser Leu Ile Ile Ala Ala Ile Ile Phe Ile Ala Ser
    50                  55                  60


Ala Asn His Lys Val Thr Leu Thr Thr Ala Ile Ile Gln Asp Ala Thr
65                  70                  75                  80


Asn Gln Ile Lys Asn Thr Thr Pro Thr Tyr Leu Thr Gln Asn Pro Gln
                85                  90                  95


Leu Gly Ile Ser Phe Ser Asn Leu Ser Gly Thr Thr Leu Gln Ser Thr
            100                 105                 110


Thr Ile Leu Ala Ser Thr Thr Pro Ser Ala Glu Ser Thr Pro Gln Ser
        115                 120                 125


Thr Thr Val Lys Ile Ile Asn Thr Thr Thr Thr Gln Ile Leu Pro Ser
    130                 135                 140


Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Gln Asn Lys Pro Asn
145                 150                 155                 160


Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
                165                 170                 175


Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys
            180                 185                 190


Lys Pro Gly Lys Lys Thr Thr Thr Lys Pro Thr Lys Lys Pro Thr Leu
        195                 200                 205


Lys Thr Thr Lys Lys Asp Pro Lys Pro Gln Thr Thr Lys Pro Lys Glu
    210                 215                 220


Ala Leu Thr Thr Lys Pro Thr Gly Lys Pro Thr Ile Asn Thr Thr Lys
225                 230                 235                 240


Thr Asn Ile Arg Thr Thr Leu Leu Thr Ser Asn Thr Lys Gly Asn Pro
                245                 250                 255
```

Glu His Thr Ser Gln Glu Glu Thr Leu His Ser Thr Thr Ser Glu Gly
260 265 270

Tyr Leu Ser Pro Ser Gln Val Tyr Thr Thr Ser Gly Gln Glu Glu Thr
275 280 285

Leu His Ser Thr Thr Ser Glu Gly Tyr Leu Ser Pro Ser Gln Val Tyr
290 295 300

Thr Thr Ser Glu Tyr Leu Ser Gln Ser Leu Ser Ser Ser Asn Thr Thr
305 310 315 320

Lys

<210> 59
<211> 616
<212> PRT
<213> Newcastle disease virus

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 59

Met Glu Arg Gly Val Ser Gln Val Ala Leu Glu Asn Asp Glu Arg Glu
1 5 10 15

Ala Lys Asn Thr Trp Arg Leu Val Phe Arg Val Thr Val Leu Phe Leu
20 25 30

Thr Ile Val Thr Leu Ala Ile Ser Ala Ala Ala Leu Ala Phe Ser Met
35 40 45

Asn Ala Ser Thr Pro Gln Asp Leu Glu Gly Ile Pro Val Ala Ile Ser
50 55 60

Lys Val Glu Asp Lys Ile Thr Ser Ala Leu Gly Ala Ser Gln Asp Val
65 70 75 80

Met Asp Arg Ile Tyr Lys Gln Val Ala Leu Glu Ser Pro Leu Ala Leu
85 90 95

Leu Asn Thr Glu Ser Thr Ile Met Asn Ala Leu Thr Ser Leu Ser Tyr
100 105 110

Gln Ile Asn Gly Ala Ala Asn Ala Ser Gly Cys Gly Ala Pro Val Pro
115 120 125

Asp Pro Asp Tyr Ile Gly Gly Ile Gly Lys Glu Leu Ile Val Asp Asp
130 135 140

Thr Ser Asp Val Thr Ser Phe Tyr Pro Ser Ala Phe Gln Glu His Leu

145 150 155 160

Asn Phe Ile Pro Ala Pro Thr Thr Gly Ser Gly Cys Thr Arg Ile Pro
165 170 175

Ser Phe Asp Met Ser Ala Thr His Tyr Cys Tyr Thr His Asn Val Ile
180 185 190

Leu Ser Gly Cys Arg Asp His Ser His Ser His Gln Tyr Leu Ala Leu
195 200 205

Gly Val Leu Arg Thr Ser Ala Thr Gly Arg Val Phe Phe Ser Thr Leu
210 215 220

Arg Ser Ile Asn Leu Asp Asp Thr Gln Asn Arg Lys Ser Cys Ser Val
225 230 235 240

Ser Ala Thr Pro Leu Gly Cys Asp Met Leu Cys Ser Lys Val Thr Glu
245 250 255

Thr Glu Glu Glu Asp Tyr Gln Ser Thr Asp Pro Thr Leu Met Val His
260 265 270

Gly Arg Leu Gly Phe Asp Gly Gln Tyr His Glu Arg Asp Leu Asp Val
275 280 285

His Thr Leu Phe Gly Asp Trp Val Ala Asn Tyr Pro Gly Val Gly Gly
290 295 300

Gly Ser Phe Ile Asn Asn Arg Val Trp Phe Pro Val Tyr Gly Gly Leu
305 310 315 320

Lys Pro Gly Ser Pro Thr Asp Lys Arg Gln Glu Gly Gln Tyr Ala Ile
325 330 335

Tyr Lys Arg Tyr Asn Asp Thr Cys Pro Asp Asp Gln Glu Tyr Gln Val
340 345 350

Arg Met Ala Lys Ser Ala Tyr Lys Pro Asn Arg Phe Gly Gly Lys Arg
355 360 365

Val Gln Gln Ala Ile Leu Ser Ile Gly Val Ser Thr Thr Leu Ala Asp
370 375 380

Asp Pro Val Leu Thr Val Thr Ser Asn Thr Ile Thr Leu Met Gly Ala
385 390 395 400

Glu Gly Arg Val Met Thr Val Gly Thr Ser His Tyr Leu Tyr Gln Arg
405 410 415

Gly Ser Ser Tyr Tyr Ser Pro Ala Ile Leu Tyr Pro Leu Thr Ile Ala
420 425 430

```
Asn Lys Thr Ala Thr Leu Gln Asp Pro Tyr Lys Phe Asn Ala Phe Thr
        435                 440                 445


Arg Pro Gly Ser Val Pro Cys Gln Ala Ser Ala Arg Cys Pro Asn Ser
    450                 455                 460


Cys Val Thr Gly Val Tyr Thr Asp Pro Tyr Pro Ile Val Phe His Lys
465                 470                 475                 480


Asn His Thr Leu Arg Gly Val Phe Gly Thr Met Leu Asp Asp Glu Gln
                485                 490                 495


Ala Arg Leu Asn Pro Val Ser Ala Val Phe Asp Ser Ile Ala Arg Ser
            500                 505                 510


Arg Val Thr Arg Val Ser Ser Ser Ser Thr Lys Ala Ala Tyr Thr Thr
        515                 520                 525


Ser Thr Cys Phe Lys Val Val Lys Thr Gly Lys Val Tyr Cys Leu Ser
    530                 535                 540


Ile Ala Glu Ile Ser Asn Thr Leu Phe Gly Glu Phe Arg Ile Val Pro
545                 550                 555                 560


Leu Leu Val Glu Ile Leu Arg Asp Glu Gly Arg Ser Glu Ala Arg Ser
                565                 570                 575


Ala Leu Thr Thr Gln Gly His Pro Gly Trp Asn Asp Glu Val Val Asp
            580                 585                 590


Pro Ile Phe Cys Ala Val Thr Asn Gln Thr Asp His Arg Gln Lys Leu
        595                 600                 605


Glu Glu Tyr Ala Gln Ser Trp Pro
    610                 615


<210>  60
<211>  316
<212>  PRT
<213>  Human respiratory syncytial virus


<220>
<221>  MISC_FEATURE
<223>  receptor binding glycoprotein

<400>  60

Met Ser Lys Asn Lys Asn Gln Arg Thr Ala Arg Thr Leu Glu Lys Thr
1               5                   10                  15


Trp Asp Thr Leu Asn His Leu Ile Val Ile Ser Ser Cys Leu Tyr Lys
            20                  25                  30
```

```
Leu Asn Leu Lys Ser Ile Ala Gln Ile Ala Leu Ser Val Leu Ala Met
        35                  40                  45

Ile Ile Ser Thr Ser Leu Ile Ile Ala Ala Ile Ile Phe Ile Ile Ser
    50                  55                  60

Ala Asn His Lys Val Thr Leu Thr Thr Val Thr Val Gln Thr Ile Lys
65                  70                  75                  80

Asn His Thr Glu Lys Asn Ile Thr Thr Tyr Leu Thr Gln Val Ser Pro
                85                  90                  95

Glu Arg Val Ser Pro Ser Lys Gln Pro Thr Thr Thr Pro Pro Ile His
            100                 105                 110

Thr Asn Ser Ala Thr Ile Ser Pro Asn Thr Lys Ser Glu Thr His His
        115                 120                 125

Thr Thr Ala Gln Thr Lys Gly Arg Thr Thr Thr Pro Thr Gln Asn Asn
    130                 135                 140

Lys Pro Ser Thr Lys Pro Arg Pro Lys Asn Pro Pro Lys Lys Pro Lys
145                 150                 155                 160

Asp Asp Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
                165                 170                 175

Gly Asn Asn Gln Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Asn Asn
            180                 185                 190

Lys Pro Lys Lys Lys Pro Thr Thr Lys Pro Thr Asn Lys Pro Pro Thr
        195                 200                 205

Lys Thr Thr Asn Lys Arg Asp Pro Lys Thr Pro Ala Lys Thr Leu Lys
    210                 215                 220

Lys Glu Thr Thr Ile Asn Pro Thr Thr Lys Lys Pro Thr Pro Lys Thr
225                 230                 235                 240

Thr Glu Arg Asp Thr Ser Thr Pro Gln Ser Thr Val Leu Asp Thr Thr
                245                 250                 255

Thr Ser Lys His Thr Glu Arg Asp Thr Ser Thr Pro Gln Ser Thr Val
            260                 265                 270

Leu Asp Thr Thr Thr Ser Lys His Thr Ile Gln Gln Gln Ser Leu His
        275                 280                 285

Ser Ile Thr Pro Glu Asn Thr Pro Asn Ser Thr Gln Thr Pro Thr Ala
    290                 295                 300
```

Ser Glu Pro Ser Thr Ser Asn Ser Thr Gln Lys Leu
305 310 315

<210> 61
<211> 621
<212> PRT
<213> Measels virus

<220>
<221> MISC_FEATURE
<223> Hemagglutinin

<400> 61

Met Ser Pro His Arg Asp Arg Ile Asn Ala Phe Tyr Arg Asp Asn Pro
1 5 10 15

His Pro Lys Gly Ser Arg Ile Val Ile Asn Arg Glu His Leu Met Ile
20 25 30

Asp Arg Pro Tyr Val Leu Leu Ala Val Leu Phe Val Met Phe Leu Ser
35 40 45

Leu Ile Gly Leu Leu Ala Ile Ala Gly Ile Arg Leu His Arg Ala Ala
50 55 60

Ile Tyr Thr Ala Glu Ile His Lys Ser Leu Ser Thr Asn Leu Asp Val
65 70 75 80

Thr Asn Ser Ile Glu His Gln Val Lys Asp Val Leu Thr Pro Leu Phe
85 90 95

Lys Ile Ile Gly Asp Glu Val Gly Leu Arg Thr Pro Gln Arg Phe Thr
100 105 110

Asp Leu Val Lys Phe Ile Ser Asp Lys Ile Lys Phe Leu Asn Pro Asp
115 120 125

Arg Glu Tyr Asp Phe Arg Asp Leu Thr Trp Cys Ile Asn Pro Pro Glu
130 135 140

Arg Ile Lys Leu Asp Tyr Asp Gln Tyr Cys Ala Asp Val Ala Ala Glu
145 150 155 160

Glu Leu Met Asn Ala Leu Val Asn Ser Thr Leu Leu Glu Ala Arg Ala
165 170 175

Thr Asn Gln Phe Leu Ala Val Ser Lys Gly Asn Cys Ser Gly Pro Thr
180 185 190

Thr Ile Arg Gly Gln Phe Ser Asn Met Ser Leu Ser Leu Leu Asp Leu
195 200 205

Tyr Leu Ser Arg Gly Tyr Asn Val Ser Ser Ile Val Thr Met Thr Ser

210 215 220

Gln Gly Met Tyr Gly Gly Thr Tyr Leu Val Gly Lys Pro Asn Leu Ser
225 230 235 240

Ser Lys Gly Ser Glu Leu Ser Gln Leu Ser Met His Arg Val Phe Glu
245 250 255

Val Gly Val Ile Arg Asn Pro Gly Leu Gly Ala Pro Val Phe His Met
260 265 270

Thr Asn Tyr Phe Glu Gln Pro Val Ser Asn Asp Phe Ser Asn Cys Met
275 280 285

Val Ala Leu Gly Glu Leu Arg Phe Ala Ala Leu Cys His Arg Glu Asp
290 295 300

Ser Val Thr Val Pro Tyr Gln Gly Ser Gly Lys Gly Val Ser Phe Gln
305 310 315 320

Leu Val Lys Leu Gly Val Trp Lys Ser Pro Thr Asp Met Gln Ser Trp
325 330 335

Val Pro Leu Ser Thr Asp Asp Pro Val Ile Asp Arg Leu Tyr Leu Ser
340 345 350

Ser His Arg Gly Val Ile Ala Asp Asn Gln Ala Lys Trp Ala Val Pro
355 360 365

Thr Thr Arg Thr Asp Asp Lys Leu Arg Met Glu Thr Cys Phe Gln Gln
370 375 380

Ala Cys Lys Gly Lys Asn Gln Ala Leu Cys Glu Asn Pro Glu Trp Ala
385 390 395 400

Pro Leu Lys Asp Asn Arg Ile Pro Ser Tyr Gly Val Leu Ser Val Asn
405 410 415

Leu Ser Leu Thr Val Glu Leu Lys Ile Lys Ile Ala Ser Gly Phe Gly
420 425 430

Pro Leu Ile Thr His Gly Ser Gly Met Asp Leu Tyr Lys Thr Asn His
435 440 445

Asp Asn Val Tyr Trp Leu Thr Ile Pro Pro Met Lys Asn Leu Ala Leu
450 455 460

Gly Val Ile Asn Thr Leu Glu Trp Ile Pro Arg Phe Lys Val Ser Pro
465 470 475 480

Asn Leu Phe Thr Val Pro Ile Lys Glu Ala Gly Glu Asp Cys His Ala
485 490 495

```
Pro Thr Tyr Leu Pro Ala Glu Val Asp Gly Asp Val Lys Leu Ser Ser
            500                 505                 510


Asn Leu Val Ile Leu Pro Gly Gln Asp Leu Gln Tyr Val Leu Ala Thr
        515                 520                 525


Tyr Asp Thr Ser Arg Val Glu His Ala Val Val Tyr Tyr Val Tyr Ser
    530                 535                 540


Pro Ser Arg Ser Phe Ser Tyr Phe Tyr Pro Phe Arg Leu Pro Ile Lys
545                 550                 555                 560


Gly Val Pro Ile Glu Leu Gln Val Glu Cys Phe Thr Trp Asp Gln Lys
                565                 570                 575


Leu Trp Cys Arg His Phe Cys Val Leu Ala Asp Ser Glu Ser Gly Gly
            580                 585                 590


His Ile Thr His Ser Gly Met Val Gly Met Gly Val Ser Cys Thr Val
        595                 600                 605


Thr Arg Glu Asp Gly Thr Asn Arg Arg Gln Gly Cys Gln
    610                 615                 620


<210>  62
<211>  582
<212>  PRT
<213>  Mumps virus


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin neuraminidase

<400>  62

Met Glu Pro Ser Lys Leu Phe Thr Met Ser Asp Asn Ala Thr Phe Ala
1               5                   10                  15


Pro Gly Pro Val Ile Asn Ala Ala Asp Lys Lys Thr Phe Arg Thr Cys
            20                  25                  30


Phe Arg Ile Leu Val Leu Ser Val Gln Ala Val Thr Leu Ile Leu Val
        35                  40                  45


Ile Val Thr Leu Gly Glu Leu Val Arg Met Ile Asn Asp Gln Gly Leu
    50                  55                  60


Ser Asn Gln Leu Ser Ser Ile Ala Asp Lys Ile Arg Glu Ser Ala Thr
65                  70                  75                  80


Met Ile Ala Ser Ala Val Gly Val Met Asn Gln Val Ile His Gly Val
                85                  90                  95
```

```
Thr Val Ser Leu Pro Leu Gln Ile Glu Gly Asn Gln Asn Gln Leu Leu
            100                 105                 110

Ser Thr Leu Ala Thr Ile Cys Thr Gly Lys Lys Gln Val Ser Asn Cys
        115                 120                 125

Ser Thr Asn Ile Pro Leu Val Asn Asp Leu Arg Phe Ile Asn Gly Ile
    130                 135                 140

Asn Lys Phe Ile Ile Glu Asp Tyr Ala Thr His Asp Phe Ser Ile Gly
145                 150                 155                 160

His Pro Leu Asn Met Pro Ser Phe Ile Pro Thr Ala Thr Ser Pro Asn
                165                 170                 175

Gly Cys Thr Arg Ile Pro Ser Phe Ser Leu Gly Lys Thr His Trp Cys
            180                 185                 190

Tyr Thr His Asn Val Ile Asn Ala Asn Cys Lys Asp His Thr Ser Ser
        195                 200                 205

Asn Gln Tyr Ile Ser Met Gly Ile Leu Val Gln Thr Ala Ser Gly Tyr
    210                 215                 220

Pro Met Phe Lys Thr Leu Lys Ile Gln Tyr Leu Ser Asp Gly Leu Asn
225                 230                 235                 240

Arg Lys Ser Cys Ser Ile Ala Thr Val Pro Asp Gly Cys Ala Met Tyr
                245                 250                 255

Cys Tyr Val Ser Thr Gln Leu Glu Thr Asp Asp Tyr Ala Gly Ser Ser
            260                 265                 270

Pro Pro Thr Gln Lys Leu Thr Leu Leu Phe Tyr Asn Asp Thr Val Thr
        275                 280                 285

Glu Arg Thr Ile Ser Pro Thr Gly Leu Glu Gly Asn Trp Ala Thr Leu
    290                 295                 300

Val Pro Gly Val Gly Ser Gly Ile Tyr Phe Glu Asn Lys Leu Ile Phe
305                 310                 315                 320

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

Ala Arg Glu Phe Phe Arg Pro Val Asn Pro Tyr Asn Pro Cys Ser Gly
            340                 345                 350

Pro Gln Gln Asp Leu Asp Gln Arg Ala Leu Arg Ser Tyr Phe Pro Ser
        355                 360                 365
```

```
Tyr Phe Ser Asn Arg Arg Val Gln Ser Ala Phe Leu Val Cys Ala Trp
    370                 375                 380


Asn Gln Ile Leu Val Thr Asn Cys Glu Leu Val Val Pro Ser Asn Asn
385                 390                 395                 400


Gln Thr Leu Met Gly Ala Glu Gly Arg Val Leu Leu Ile Asn Asn Arg
                405                 410                 415


Leu Leu Tyr Tyr Gln Arg Ser Thr Ser Trp Trp Pro Tyr Glu Leu Leu
            420                 425                 430


Tyr Glu Ile Ser Phe Thr Phe Thr Asn Ser Gly Gln Ser Ser Val Asn
        435                 440                 445


Met Ser Trp Ile Pro Ile Tyr Ser Phe Thr Arg Pro Gly Ser Gly Asn
    450                 455                 460


Cys Ser Gly Glu Asn Val Cys Pro Thr Ala Cys Val Ser Gly Val Tyr
465                 470                 475                 480


Leu Asp Pro Trp Pro Leu Thr Pro Tyr Ser His Gln Ser Gly Ile Asn
                485                 490                 495


Arg Asn Phe Tyr Phe Thr Gly Ala Leu Leu Asn Ser Ser Thr Thr Arg
            500                 505                 510


Val Asn Pro Thr Leu Tyr Val Ser Ala Leu Asn Asn Leu Lys Val Leu
        515                 520                 525


Ala Pro Tyr Gly Asn Gln Gly Leu Phe Ala Ser Tyr Thr Thr Thr Thr
    530                 535                 540


Cys Phe Gln Asp Thr Gly Asp Ala Ser Val Tyr Cys Val Tyr Ile Met
545                 550                 555                 560


Glu Leu Ala Ser Asn Ile Val Gly Glu Phe Gln Ile Leu Pro Val Leu
                565                 570                 575


Thr Arg Leu Thr Ile Thr
            580


<210>  63
<211>  572
<212>  PRT
<213>  Human respirovirus 3


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin neuraminidase

<400>  63

Met Glu Tyr Trp Lys His Thr Asn His Gly Lys Asp Ala Gly Asn Glu
```

```
1               5                   10                  15

Leu Glu Thr Ala Thr Ala Thr His Gly Asn Arg Leu Thr Asn Lys Ile
            20                  25                  30

Thr Tyr Ile Leu Trp Thr Ile Thr Leu Val Leu Leu Ser Ile Val Phe
        35                  40                  45

Ile Ile Val Leu Ile Asn Ser Ile Lys Ser Glu Lys Ala His Glu Ser
    50                  55                  60

Leu Leu Gln Asp Ile Asn Asn Glu Phe Met Glu Val Thr Glu Lys Ile
65                  70                  75                  80

Gln Val Ala Ser Asp Asn Thr Asn Asp Leu Ile Gln Ser Gly Val Asn
                85                  90                  95

Thr Arg Leu Leu Thr Ile Gln Ser His Val Gln Asn Tyr Ile Pro Ile
            100                 105                 110

Ser Leu Thr Gln Gln Ile Ser Asp Leu Arg Lys Phe Ile Ser Glu Ile
        115                 120                 125

Thr Ile Arg Asn Asp Asn Gln Glu Val Pro Pro Gln Arg Ile Thr His
    130                 135                 140

Asp Val Gly Ile Lys Pro Leu Asn Pro Asp Asp Phe Trp Arg Cys Thr
145                 150                 155                 160

Ser Gly Leu Pro Ser Leu Met Arg Thr Pro Lys Ile Arg Leu Met Pro
                165                 170                 175

Gly Pro Gly Leu Leu Ala Met Pro Thr Thr Val Asp Gly Cys Val Arg
            180                 185                 190

Thr Pro Ser Leu Val Ile Asn Asp Leu Ile Tyr Ala Tyr Thr Ser Asn
        195                 200                 205

Leu Ile Thr Arg Gly Cys Gln Asp Ile Gly Lys Ser Tyr Gln Val Leu
    210                 215                 220

Gln Ile Gly Ile Ile Thr Val Asn Ser Asp Leu Val Pro Asp Leu Asn
225                 230                 235                 240

Pro Arg Ile Ser His Thr Phe Asn Ile Asn Asp Asn Arg Lys Ser Cys
                245                 250                 255

Ser Leu Ala Leu Leu Asn Thr Asp Val Tyr Gln Leu Cys Ser Thr Pro
            260                 265                 270

Lys Val Asp Glu Arg Ser Asp Tyr Ala Ser Ser Gly Ile Glu Asp Ile
        275                 280                 285
```

```
Val Leu Asp Ile Val Asn Tyr Asp Gly Ser Ile Ser Thr Thr Arg Phe
    290                 295                 300

Lys Asn Asn Asn Ile Ser Phe Asp Gln Pro Tyr Ala Ala Leu Tyr Pro
305                 310                 315                 320

Ser Val Gly Pro Gly Ile Tyr Tyr Lys Gly Lys Ile Ile Phe Leu Gly
                325                 330                 335

Tyr Gly Gly Leu Glu His Pro Ile Asn Glu Asn Ala Ile Cys Asn Thr
            340                 345                 350

Thr Gly Cys Pro Gly Lys Thr Gln Arg Asp Cys Asn Gln Ala Ser His
        355                 360                 365

Ser Pro Trp Phe Ser Asp Arg Arg Met Val Asn Ser Ile Ile Val Val
    370                 375                 380

Asp Lys Gly Leu Asn Ser Val Pro Lys Leu Lys Val Trp Thr Ile Ser
385                 390                 395                 400

Met Arg Gln Asn Tyr Trp Gly Ser Glu Gly Arg Leu Leu Leu Leu Gly
                405                 410                 415

Asn Lys Ile Tyr Ile Tyr Thr Arg Ser Thr Ser Trp His Ser Lys Leu
            420                 425                 430

Gln Leu Gly Ile Ile Asp Ile Thr Asp Tyr Ser Asp Ile Arg Ile Lys
        435                 440                 445

Trp Thr Trp His Asn Val Leu Ser Arg Pro Gly Asn Asn Glu Cys Pro
    450                 455                 460

Trp Gly His Ser Cys Pro Asp Gly Cys Ile Thr Gly Val Tyr Thr Asp
465                 470                 475                 480

Ala Tyr Pro Leu Asn Pro Thr Gly Ser Ile Val Ser Ser Val Ile Leu
                485                 490                 495

Asp Ser Gln Lys Ser Arg Val Asn Pro Val Ile Thr Tyr Ser Thr Ala
            500                 505                 510

Thr Glu Arg Val Asn Glu Leu Ala Ile Arg Asn Lys Thr Leu Ser Ala
        515                 520                 525

Gly Tyr Thr Thr Thr Ser Cys Ile Thr His Tyr Asn Lys Gly Tyr Cys
    530                 535                 540

Phe His Ile Val Glu Ile Asn His Lys Ser Leu Asn Thr Phe Gln Pro
545                 550                 555                 560
```

```
Met Leu Phe Lys Thr Glu Ile Pro Lys Ser Cys Ser
                565                 570


<210>  64
<211>  220
<212>  PRT
<213>  Human metapneumovirus


<220>
<221>  MISC_FEATURE
<223>  attachment glycoprotein G

<400>  64

Met Glu Val Lys Val Glu Asn Ile Arg Ala Ile Asp Met Leu Lys Ala
1               5                   10                  15


Arg Val Lys Asn Arg Val Ala Arg Ser Lys Cys Phe Lys Asn Ala Ser
            20                  25                  30


Leu Ile Leu Ile Gly Ile Thr Thr Leu Ser Ile Ala Leu Asn Ile Tyr
        35                  40                  45


Leu Ile Ile Asn Tyr Thr Ile Gln Lys Thr Thr Ser Glu Ser Glu His
    50                  55                  60


His Thr Ser Ser Pro Pro Thr Glu Ser Asn Lys Glu Thr Ser Thr Ile
65                  70                  75                  80


Pro Ile Asp Asn Pro Asp Ile Thr Pro Asn Ser Gln His Pro Thr Gln
                85                  90                  95


Gln Ser Thr Glu Ser Leu Thr Leu Tyr Pro Ala Ser Ser Met Ser Pro
            100                 105                 110


Ser Glu Thr Glu Pro Ala Ser Thr Pro Gly Ile Thr Asn Arg Leu Ser
        115                 120                 125


Leu Ala Asp Arg Ser Thr Thr Gln Pro Ser Glu Ser Arg Thr Lys Thr
    130                 135                 140


Asn Ser Thr Val His Lys Lys Asn Lys Lys Asn Ile Ser Ser Thr Ile
145                 150                 155                 160


Ser Arg Thr Gln Ser Pro Pro Arg Thr Thr Ala Lys Ala Val Ser Arg
                165                 170                 175


Thr Thr Ala Leu Arg Met Ser Ser Thr Gly Glu Arg Pro Thr Thr Thr
            180                 185                 190


Ser Val Gln Ser Asp Ser Ser Thr Thr Ala Gln Asn His Glu Glu Thr
        195                 200                 205
```

Gly Pro Ala Asn Pro Gln Ala Ser Val Ser Thr Met
210 215 220

<210> 65
<211> 607
<212> PRT
<213> Canine distemper virus

<220>
<221> MISC_FEATURE
<223> hemagglutinin

<400> 65

Met Leu Ser Tyr Gln Asp Lys Val Gly Ala Phe Tyr Lys Asp Asn Ala
1 5 10 15

Arg Ala Asn Ser Ser Lys Leu Ser Leu Val Thr Glu Glu Gln Gly Gly
20 25 30

Arg Arg Pro Pro Tyr Leu Leu Phe Val Leu Leu Ile Leu Leu Val Gly
35 40 45

Ile Leu Ala Leu Leu Ala Ile Ala Gly Val Arg Phe Arg Gln Val Ser
50 55 60

Thr Ser Asn Val Glu Phe Gly Arg Leu Leu Lys Asp Asp Leu Glu Lys
65 70 75 80

Ser Glu Ala Val His His Gln Val Met Asp Val Leu Thr Pro Leu Phe
85 90 95

Lys Ile Ile Gly Asp Glu Ile Gly Leu Arg Leu Pro Gln Lys Leu Asn
100 105 110

Glu Ile Lys Gln Phe Ile Leu Gln Lys Thr Asn Phe Phe Asn Pro Asn
115 120 125

Arg Glu Phe Asp Phe Arg Asp Leu His Trp Cys Ile Asn Pro Pro Ser
130 135 140

Lys Ile Lys Val Asn Phe Thr Asn Tyr Cys Asp Ala Ile Gly Val Arg
145 150 155 160

Lys Ser Ile Ala Ser Ala Ala Asn Pro Ile Leu Leu Ser Ala Leu Ser
165 170 175

Gly Gly Arg Gly Asp Ile Phe Pro Pro Tyr Arg Cys Ser Gly Ala Thr
180 185 190

Thr Ser Val Gly Arg Val Phe Pro Leu Ser Val Ser Leu Ser Met Ser
195 200 205

Leu Ile Ser Lys Thr Ser Glu Ile Ile Ser Met Leu Thr Ala Ile Ser

210 215 220

Asp Gly Val Tyr Gly Lys Thr Tyr Leu Leu Val Pro Asp Tyr Ile Glu
225 230 235 240

Arg Glu Phe Asp Thr Gln Lys Ile Arg Val Phe Glu Ile Gly Phe Ile
245 250 255

Lys Arg Trp Leu Asn Asp Met Pro Leu Leu Gln Thr Thr Asn Tyr Met
260 265 270

Val Leu Pro Glu Asn Ser Lys Ala Lys Val Cys Thr Ile Ala Val Gly
275 280 285

Glu Leu Thr Leu Ala Ser Leu Cys Val Asp Glu Ser Thr Val Leu Leu
290 295 300

Tyr His Asp Ser Asn Gly Ser Gln Asp Ser Ile Leu Val Val Thr Leu
305 310 315 320

Gly Ile Phe Gly Ala Thr Pro Met Asn Gln Val Glu Glu Val Ile Pro
325 330 335

Val Ala His Pro Ser Val Glu Arg Ile His Ile Thr Asn His Arg Gly
340 345 350

Phe Ile Lys Asp Ser Val Ala Thr Trp Met Val Pro Ala Leu Val Ser
355 360 365

Glu Gln Gln Glu Gly Gln Lys Asn Cys Leu Glu Ser Ala Cys Gln Arg
370 375 380

Lys Ser Tyr Pro Met Cys Asn Gln Thr Ser Trp Glu Pro Phe Gly Gly
385 390 395 400

Val Gln Leu Pro Ser Tyr Gly Arg Leu Thr Leu Pro Leu Asp Ala Ser
405 410 415

Ile Asp Leu Gln Leu Asn Ile Ser Phe Thr Tyr Gly Pro Val Ile Leu
420 425 430

Asn Gly Asp Gly Met Asp Tyr Tyr Glu Asn Pro Leu Leu Asp Ser Gly
435 440 445

Trp Leu Thr Ile Pro Pro Lys Asn Gly Thr Ile Leu Gly Leu Ile Asn
450 455 460

Lys Ala Ser Arg Gly Asp Gln Phe Thr Val Thr Pro His Val Leu Thr
465 470 475 480

Phe Ala Pro Arg Glu Ser Ser Gly Asn Cys Tyr Leu Pro Ile Gln Thr
485 490 495

Ser Gln Ile Met Asp Lys Asp Val Leu Thr Glu Ser Asn Leu Val Val
500 505 510

Leu Pro Thr Gln Asn Phe Arg Tyr Val Val Ala Thr Tyr Asp Ile Ser
515 520 525

Arg Glu Asn His Ala Ile Val Tyr Tyr Val Tyr Asp Pro Ile Arg Thr
530 535 540

Ile Ser Tyr Thr Tyr Pro Phe Arg Leu Thr Thr Lys Gly Arg Pro Asp
545 550 555 560

Phe Leu Arg Ile Glu Cys Phe Val Trp Asp Asp Asp Leu Trp Cys His
565 570 575

Gln Phe Tyr Arg Phe Glu Ser Asp Ile Thr Asn Ser Thr Thr Ser Val
580 585 590

Glu Asp Leu Val Arg Ile Arg Phe Ser Cys Asn Arg Ser Lys Pro
595 600 605

<210> 66
<211> 609
<212> PRT
<213> Peste-des-petits-ruminants virus

<220>
<221> MISC_FEATURE
<223> hemagglutinin

<400> 66

Met Ser Ala Gln Arg Glu Arg Ile Asn Ala Phe Tyr Lys Asp Asn Pro
1 5 10 15

His Asn Lys Asn His Arg Val Ile Leu Asp Arg Glu Arg Leu Val Ile
20 25 30

Glu Arg Pro Tyr Ile Leu Leu Gly Val Leu Leu Val Met Phe Leu Ser
35 40 45

Leu Ile Gly Leu Leu Ala Ile Ala Gly Ile Arg Leu His Arg Ala Thr
50 55 60

Val Gly Thr Ser Glu Ile Gln Ser Arg Leu Asn Thr Asn Ile Glu Leu
65 70 75 80

Thr Glu Ser Ile Asp His Gln Thr Lys Asp Val Leu Thr Pro Leu Phe
85 90 95

Lys Ile Ile Gly Asp Glu Val Gly Ile Arg Ile Pro Gln Lys Phe Ser
100 105 110

```
Asp Leu Val Lys Phe Ile Ser Asp Lys Ile Lys Phe Leu Asn Pro Asp
        115                 120                 125

Arg Glu Tyr Asp Phe Arg Asp Leu Arg Trp Cys Met Asn Pro Pro Glu
    130                 135                 140

Arg Val Lys Ile Asn Phe Asp Gln Phe Cys Glu Tyr Lys Ala Ala Val
145                 150                 155                 160

Lys Ser Ile Glu His Ile Phe Glu Ser Pro Leu Asn Lys Ser Lys Lys
                165                 170                 175

Leu Gln Ser Leu Thr Leu Gly Pro Gly Thr Gly Cys Leu Gly Arg Thr
            180                 185                 190

Val Thr Arg Ala His Phe Ser Glu Leu Thr Leu Thr Leu Met Asp Leu
        195                 200                 205

Asp Leu Glu Met Lys His Asn Val Ser Ser Val Phe Thr Val Val Glu
    210                 215                 220

Glu Gly Leu Phe Gly Arg Thr Tyr Thr Val Trp Arg Ser Asp Ala Arg
225                 230                 235                 240

Asp Pro Ser Thr Asp Leu Gly Ile Gly His Phe Leu Arg Val Phe Glu
                245                 250                 255

Ile Gly Leu Val Arg Asp Leu Gly Leu Gly Pro Pro Val Phe His Met
            260                 265                 270

Thr Asn Tyr Leu Thr Val Asn Met Ser Asp Asp Tyr Arg Arg Cys Leu
        275                 280                 285

Leu Ala Val Gly Glu Leu Lys Leu Thr Ala Leu Cys Ser Ser Ser Glu
    290                 295                 300

Thr Val Thr Leu Gly Glu Arg Gly Val Pro Lys Arg Glu Pro Leu Val
305                 310                 315                 320

Val Val Ile Leu Asn Leu Ala Gly Pro Thr Leu Gly Gly Glu Leu Tyr
                325                 330                 335

Ser Val Leu Pro Thr Ser Asp Leu Met Val Glu Lys Leu Tyr Leu Ser
            340                 345                 350

Ser His Arg Gly Ile Ile Lys Asp Asp Glu Ala Asn Trp Val Val Pro
        355                 360                 365

Ser Thr Asp Val Arg Asp Leu Gln Asn Lys Gly Glu Cys Leu Val Glu
    370                 375                 380
```

Ala Cys Lys Thr Arg Pro Pro Ser Phe Cys Asn Gly Thr Gly Ser Gly
385 390 395 400

Pro Trp Ser Glu Gly Arg Ile Pro Ala Tyr Gly Val Ile Arg Val Ser
405 410 415

Leu Asp Leu Ala Ser Asp Pro Gly Val Val Ile Thr Ser Val Phe Gly
420 425 430

Pro Leu Ile Pro His Leu Ser Gly Met Asp Leu Tyr Asn Asn Pro Phe
435 440 445

Ser Arg Ala Val Trp Leu Ala Val Pro Pro Tyr Glu Gln Ser Phe Leu
450 455 460

Gly Met Ile Asn Thr Ile Gly Phe Pro Asn Arg Ala Glu Val Met Pro
465 470 475 480

His Ile Leu Thr Thr Glu Ile Arg Gly Pro Arg Gly Arg Cys His Val
485 490 495

Pro Ile Glu Leu Ser Arg Arg Val Asp Asp Asp Ile Lys Ile Gly Ser
500 505 510

Asn Met Val Ile Leu Pro Thr Ile Asp Leu Arg Tyr Ile Thr Ala Thr
515 520 525

Tyr Asp Val Ser Arg Ser Glu His Ala Ile Val Tyr Tyr Ile Tyr Asp
530 535 540

Thr Gly Arg Ser Ser Ser Tyr Phe Tyr Pro Val Arg Leu Asn Phe Lys
545 550 555 560

Gly Asn Pro Leu Ser Leu Arg Ile Glu Cys Phe Pro Trp Arg His Lys
565 570 575

Val Trp Cys Tyr His Asp Cys Leu Ile Tyr Asn Thr Ile Thr Asp Glu
580 585 590

Glu Val His Thr Arg Gly Leu Thr Gly Ile Glu Val Thr Cys Asn Pro
595 600 605

Val

<210> 67
<211> 575
<212> PRT
<213> Sendai virus

<220>
<221> MISC_FEATURE
<223> hemagglutinin neuraminidase protein

<400> 67

```
Met Asp Gly Asp Arg Ser Lys Arg Asp Ser Tyr Trp Ser Thr Ser Pro
1               5                   10                  15


Gly Gly Ser Thr Thr Lys Leu Val Ser Asp Ser Glu Arg Ser Gly Lys
            20                  25                  30


Val Asp Thr Trp Leu Leu Ile Leu Ala Phe Thr Gln Trp Ala Leu Ser
        35                  40                  45


Ile Ala Thr Val Ile Ile Cys Ile Val Ile Ala Ala Arg Gln Gly Tyr
    50                  55                  60


Ser Met Glu Arg Tyr Ser Met Thr Val Glu Ala Leu Asn Thr Ser Asn
65                  70                  75                  80


Lys Glu Val Lys Glu Ser Leu Thr Ser Leu Ile Arg Gln Glu Val Ile
                85                  90                  95


Thr Arg Ala Ala Asn Ile Gln Ser Ser Val Gln Thr Gly Ile Pro Val
            100                 105                 110


Leu Leu Asn Lys Asn Ser Arg Asp Val Ile Arg Leu Ile Glu Lys Ser
        115                 120                 125


Cys Asn Arg Gln Glu Leu Thr Gln Leu Cys Asp Ser Thr Ile Ala Val
    130                 135                 140


His His Ala Glu Gly Ile Ala Pro Leu Glu Pro His Ser Phe Trp Arg
145                 150                 155                 160


Cys Pro Ala Gly Glu Pro Tyr Leu Ser Ser Asp Pro Glu Val Ser Leu
                165                 170                 175


Leu Pro Gly Pro Ser Leu Leu Ser Gly Ser Thr Thr Ile Ser Gly Cys
            180                 185                 190


Val Arg Leu Pro Ser Leu Ser Ile Gly Glu Ala Ile Tyr Ala Tyr Ser
        195                 200                 205


Ser Asn Leu Ile Thr Gln Gly Cys Ala Asp Ile Gly Lys Ser Tyr Gln
    210                 215                 220


Val Leu Gln Leu Gly Tyr Ile Ser Leu Asn Ser Asp Met Phe Pro Asp
225                 230                 235                 240


Leu Asn Pro Val Val Ser His Thr Tyr Asp Ile Asn Asp Asn Arg Lys
                245                 250                 255


Ser Cys Ser Val Val Ala Thr Gly Thr Arg Gly Tyr Gln Leu Cys Ser
            260                 265                 270
```

```
Met Pro Ile Val Asp Glu Arg Thr Asp Tyr Ser Ser Asp Gly Ile Glu
        275                 280                 285

Asp Leu Val Leu Asp Ile Leu Asp Leu Lys Gly Arg Thr Lys Ser His
    290                 295                 300

Arg Tyr Ser Asn Ser Glu Ile Asp Leu Asp His Pro Phe Ser Ala Leu
305                 310                 315                 320

Tyr Pro Ser Val Gly Ser Gly Ile Ala Thr Glu Gly Ser Leu Ile Phe
                325                 330                 335

Leu Gly Tyr Gly Gly Leu Thr Thr Pro Leu Gln Gly Asp Thr Lys Cys
            340                 345                 350

Arg Ile Gln Gly Cys Gln Gln Val Ser Gln Asp Thr Cys Asn Glu Ala
        355                 360                 365

Leu Lys Ile Thr Trp Leu Gly Gly Lys Gln Val Val Ser Val Leu Ile
    370                 375                 380

Gln Val Asn Asp Tyr Leu Ser Glu Arg Pro Arg Ile Arg Val Thr Thr
385                 390                 395                 400

Ile Pro Ile Thr Gln Asn Tyr Leu Gly Ala Glu Gly Arg Leu Leu Lys
                405                 410                 415

Leu Gly Asp Gln Val Tyr Ile Tyr Thr Arg Ser Ser Gly Trp His Ser
            420                 425                 430

Gln Leu Gln Ile Gly Val Leu Asp Val Ser His Pro Leu Thr Ile Ser
        435                 440                 445

Trp Thr Pro His Glu Ala Leu Ser Arg Pro Gly Asn Glu Asp Cys Asn
    450                 455                 460

Trp Tyr Asn Thr Cys Pro Lys Glu Cys Ile Ser Gly Val Tyr Thr Asp
465                 470                 475                 480

Ala Tyr Pro Leu Ser Pro Asp Ala Ala Asn Val Ala Thr Val Thr Leu
                485                 490                 495

Tyr Ala Asn Thr Ser Arg Val Asn Pro Thr Ile Met Tyr Ser Asn Thr
            500                 505                 510

Thr Asn Ile Ile Asn Met Leu Arg Ile Lys Asp Val Gln Leu Glu Ala
        515                 520                 525

Ala Tyr Thr Thr Thr Ser Cys Ile Thr His Phe Gly Lys Gly Tyr Cys
    530                 535                 540
```

```
Phe His Ile Ile Glu Ile Asn Gln Lys Ser Leu Asn Thr Leu Gln Pro
545                 550                 555                 560


Met Leu Phe Lys Thr Ser Ile Pro Lys Leu Cys Lys Ala Glu Ser
                565                 570                 575


<210>  68
<211>  575
<212>  PRT
<213>  Human parainfluenza virus 1 strain Washington/1964


<220>
<221>  MISC_FEATURE
<223>  HN glycoprotein

<400>  68

Met Ala Glu Lys Gly Lys Thr Asn Ser Ser Tyr Trp Ser Thr Thr Arg
1               5                   10                  15


Asn Asp Asn Ser Thr Val Asn Thr His Ile Asn Thr Pro Ala Gly Arg
            20                  25                  30


Thr His Ile Trp Leu Leu Ile Ala Thr Thr Met His Thr Val Leu Ser
        35                  40                  45


Phe Ile Ile Met Ile Leu Cys Ile Asp Leu Ile Ile Lys Gln Asp Thr
    50                  55                  60


Cys Met Lys Thr Asn Ile Met Thr Val Ser Ser Met Asn Glu Ser Ala
65                  70                  75                  80


Lys Ile Ile Lys Glu Thr Ile Thr Glu Leu Ile Arg Gln Glu Val Ile
                85                  90                  95


Ser Arg Thr Ile Asn Ile Gln Ser Ser Val Gln Ser Gly Ile Pro Ile
            100                 105                 110


Leu Leu Asn Lys Gln Ser Arg Asp Leu Thr Gln Leu Ile Glu Lys Ser
        115                 120                 125


Cys Asn Arg Gln Glu Leu Ala Gln Ile Cys Glu Asn Thr Ile Ala Ile
    130                 135                 140


His His Ala Asp Gly Ile Ser Pro Leu Asp Pro His Asp Phe Trp Arg
145                 150                 155                 160


Cys Pro Val Gly Glu Pro Leu Leu Ser Asn Asn Pro Asn Ile Ser Leu
                165                 170                 175


Leu Pro Gly Pro Ser Leu Leu Ser Gly Ser Thr Thr Ile Ser Gly Cys
            180                 185                 190
```

```
Val Arg Leu Pro Ser Leu Ser Ile Gly Asp Ala Ile Tyr Ala Tyr Ser
        195                 200                 205

Ser Asn Leu Ile Thr Gln Gly Cys Ala Asp Ile Gly Lys Ser Tyr Gln
    210                 215                 220

Val Leu Gln Leu Gly Tyr Ile Ser Leu Asn Ser Asp Met Tyr Pro Asp
225                 230                 235                 240

Leu Asn Pro Val Ile Ser His Thr Tyr Asp Ile Asn Asp Asn Arg Lys
                245                 250                 255

Ser Cys Ser Val Ile Ala Ala Gly Thr Arg Gly Tyr Gln Leu Cys Ser
            260                 265                 270

Leu Pro Thr Val Asn Glu Thr Thr Asp Tyr Ser Ser Glu Gly Ile Glu
        275                 280                 285

Asp Leu Val Phe Asp Ile Leu Asp Leu Lys Gly Lys Thr Lys Ser His
    290                 295                 300

Arg Tyr Lys Asn Glu Asp Ile Thr Phe Asp His Pro Phe Ser Ala Met
305                 310                 315                 320

Tyr Pro Ser Val Gly Ser Gly Ile Lys Ile Glu Asn Thr Leu Ile Phe
                325                 330                 335

Leu Gly Tyr Gly Gly Leu Thr Thr Pro Leu Gln Gly Asp Thr Lys Cys
            340                 345                 350

Val Ile Asn Arg Cys Thr Asn Val Asn Gln Ser Val Cys Asn Asp Ala
        355                 360                 365

Leu Lys Ile Thr Trp Leu Lys Lys Arg Gln Val Val Asn Val Leu Ile
    370                 375                 380

Arg Ile Asn Asn Tyr Leu Ser Asp Arg Pro Lys Ile Val Val Glu Thr
385                 390                 395                 400

Ile Pro Ile Thr Gln Asn Tyr Leu Gly Ala Glu Gly Arg Leu Leu Lys
                405                 410                 415

Leu Gly Lys Lys Ile Tyr Ile Tyr Thr Arg Ser Ser Gly Trp His Ser
            420                 425                 430

Asn Leu Gln Ile Gly Ser Leu Asp Ile Asn Asn Pro Met Thr Ile Lys
        435                 440                 445

Trp Ala Pro His Glu Val Leu Ser Arg Pro Gly Asn Gln Asp Cys Asn
    450                 455                 460

Trp Tyr Asn Arg Cys Pro Arg Glu Cys Ile Ser Gly Val Tyr Thr Asp
```

465 470 475 480

Ala Tyr Pro Leu Ser Pro Asp Ala Val Asn Val Ala Thr Thr Thr Leu
485 490 495

Tyr Ala Asn Thr Ser Arg Val Asn Pro Thr Ile Met Tyr Ser Asn Thr
500 505 510

Ser Glu Ile Ile Asn Met Leu Arg Leu Lys Asn Val Gln Leu Glu Ala
515 520 525

Ala Tyr Thr Thr Thr Ser Cys Ile Thr His Phe Gly Lys Gly Tyr Cys
530 535 540

Phe His Ile Val Glu Ile Asn Gln Ala Ser Leu Asn Thr Leu Gln Pro
545 550 555 560

Met Leu Phe Lys Thr Ser Ile Pro Lys Ile Cys Lys Ile Thr Ser
565 570 575

<210> 69
<211> 240
<212> PRT
<213> Human metapneumovirus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein G

<400> 69

Met Glu Val Arg Val Glu Asn Ile Arg Ala Ile Asp Met Phe Lys Ala
1 5 10 15

Lys Met Lys Asn Arg Ile Arg Ser Ser Lys Cys Tyr Arg Asn Ala Thr
20 25 30

Leu Ile Leu Ile Gly Leu Thr Ala Leu Ser Met Ala Leu Asn Ile Phe
35 40 45

Leu Ile Ile Asp Tyr Ala Thr Leu Lys Asn Met Thr Lys Val Glu His
50 55 60

Cys Val Asn Met Pro Pro Val Glu Pro Ser Lys Lys Ser Pro Met Thr
65 70 75 80

Ser Ala Ala Asp Leu Asn Thr Lys Leu Asn Pro Gln Gln Ala Thr Gln
85 90 95

Leu Thr Thr Glu Asp Ser Thr Ser Leu Ala Ala Thr Ser Glu Asn His
100 105 110

Leu His Thr Glu Thr Thr Pro Thr Ser Asp Ala Thr Ile Ser Gln Gln
115 120 125

```
Ala Thr Asp Glu His Thr Thr Leu Leu Arg Pro Ile Asn Arg Gln Thr
    130                 135                 140

Thr Gln Thr Thr Thr Glu Lys Lys Pro Thr Gly Ala Thr Thr Lys Lys
145                 150                 155                 160

Asp Lys Glu Lys Glu Thr Thr Thr Arg Thr Thr Ser Thr Ala Ala Thr
                165                 170                 175

Gln Thr Leu Asn Thr Thr Asn Gln Thr Ser Asn Gly Arg Glu Ala Thr
            180                 185                 190

Thr Thr Ser Ala Arg Ser Arg Asn Gly Ala Thr Thr Gln Asn Ser Asp
        195                 200                 205

Gln Thr Ile Gln Ala Ala Asp Pro Ser Ser Lys Pro Tyr His Thr Gln
    210                 215                 220

Thr Asn Thr Thr Thr Ala His Asn Thr Asp Thr Ser Ser Leu Ser Ser
225                 230                 235                 240

<210>  70
<211>  604
<212>  PRT
<213>  Hendra virus

<220>
<221>  MISC_FEATURE
<223>  glycoprotein

<400>  70

Met Met Ala Asp Ser Lys Leu Val Ser Leu Asn Asn Asn Leu Ser Gly
1               5                   10                  15

Lys Ile Lys Asp Gln Gly Lys Val Ile Lys Asn Tyr Tyr Gly Thr Met
            20                  25                  30

Asp Ile Lys Lys Ile Asn Asp Gly Leu Leu Asp Ser Lys Ile Leu Gly
        35                  40                  45

Ala Phe Asn Thr Val Ile Ala Leu Leu Gly Ser Ile Ile Ile Ile Val
    50                  55                  60

Met Asn Ile Met Ile Ile Gln Asn Tyr Thr Arg Thr Thr Asp Asn Gln
65                  70                  75                  80

Ala Leu Ile Lys Glu Ser Leu Gln Ser Val Gln Gln Gln Ile Lys Ala
                85                  90                  95

Leu Thr Asp Lys Ile Gly Thr Glu Ile Gly Pro Lys Val Ser Leu Ile
            100                 105                 110
```

```
Asp Thr Ser Ser Thr Ile Thr Ile Pro Ala Asn Ile Gly Leu Leu Gly
        115                 120                 125

Ser Lys Ile Ser Gln Ser Thr Ser Ser Ile Asn Glu Asn Val Asn Asp
    130                 135                 140

Lys Cys Lys Phe Thr Leu Pro Pro Leu Lys Ile His Glu Cys Asn Ile
145                 150                 155                 160

Ser Cys Pro Asn Pro Leu Pro Phe Arg Glu Tyr Arg Pro Ile Ser Gln
                165                 170                 175

Gly Val Ser Asp Leu Val Gly Leu Pro Asn Gln Ile Cys Leu Gln Lys
            180                 185                 190

Thr Thr Ser Thr Ile Leu Lys Pro Arg Leu Ile Ser Tyr Thr Leu Pro
        195                 200                 205

Ile Asn Thr Arg Glu Gly Val Cys Ile Thr Asp Pro Leu Leu Ala Val
    210                 215                 220

Asp Asn Gly Phe Phe Ala Tyr Ser His Leu Glu Lys Ile Gly Ser Cys
225                 230                 235                 240

Thr Arg Gly Ile Ala Lys Gln Arg Ile Ile Gly Val Gly Glu Val Leu
                245                 250                 255

Asp Arg Gly Asp Lys Val Pro Ser Met Phe Met Thr Asn Val Trp Thr
            260                 265                 270

Pro Pro Asn Pro Ser Thr Ile His His Cys Ser Ser Thr Tyr His Glu
        275                 280                 285

Asp Phe Tyr Tyr Thr Leu Cys Ala Val Ser His Val Gly Asp Pro Ile
    290                 295                 300

Leu Asn Ser Thr Ser Trp Thr Glu Ser Leu Ser Leu Ile Arg Leu Ala
305                 310                 315                 320

Val Arg Pro Lys Ser Asp Ser Gly Asp Tyr Asn Gln Lys Tyr Ile Ala
                325                 330                 335

Ile Thr Lys Val Glu Arg Gly Lys Tyr Asp Lys Val Met Pro Tyr Gly
            340                 345                 350

Pro Ser Gly Ile Lys Gln Gly Asp Thr Leu Tyr Phe Pro Ala Val Gly
        355                 360                 365

Phe Leu Pro Arg Thr Glu Phe Gln Tyr Asn Asp Ser Asn Cys Pro Ile
    370                 375                 380
```

Ile His Cys Lys Tyr Ser Lys Ala Glu Asn Cys Arg Leu Ser Met Gly
385 390 395 400

Val Asn Ser Lys Ser His Tyr Ile Leu Arg Ser Gly Leu Leu Lys Tyr
405 410 415

Asn Leu Ser Leu Gly Gly Asp Ile Ile Leu Gln Phe Ile Glu Ile Ala
420 425 430

Asp Asn Arg Leu Thr Ile Gly Ser Pro Ser Lys Ile Tyr Asn Ser Leu
435 440 445

Gly Gln Pro Val Phe Tyr Gln Ala Ser Tyr Ser Trp Asp Thr Met Ile
450 455 460

Lys Leu Gly Asp Val Asp Thr Val Asp Pro Leu Arg Val Gln Trp Arg
465 470 475 480

Asn Asn Ser Val Ile Ser Arg Pro Gly Gln Ser Gln Cys Pro Arg Phe
485 490 495

Asn Val Cys Pro Glu Val Cys Trp Glu Gly Thr Tyr Asn Asp Ala Phe
500 505 510

Leu Ile Asp Arg Leu Asn Trp Val Ser Ala Gly Val Tyr Leu Asn Ser
515 520 525

Asn Gln Thr Ala Glu Asn Pro Val Phe Ala Val Phe Lys Asp Asn Glu
530 535 540

Ile Leu Tyr Gln Val Pro Leu Ala Glu Asp Asp Thr Asn Ala Gln Lys
545 550 555 560

Thr Ile Thr Asp Cys Phe Leu Leu Glu Asn Val Ile Trp Cys Ile Ser
565 570 575

Leu Val Glu Ile Tyr Asp Thr Gly Asp Ser Val Ile Arg Pro Lys Leu
580 585 590

Phe Ala Val Lys Ile Pro Ala Gln Cys Ser Glu Ser
595 600

<210> 71
<211> 602
<212> PRT
<213> Nipah virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 71

Met Pro Ala Glu Asn Lys Lys Val Arg Phe Glu Asn Thr Thr Ser Asp

```
1               5                   10                  15

Lys Gly Lys Ile Pro Ser Lys Val Ile Lys Ser Tyr Tyr Gly Thr Met
            20                  25                  30

Asp Ile Lys Lys Ile Asn Glu Gly Leu Leu Asp Ser Lys Ile Leu Ser
        35                  40                  45

Ala Phe Asn Thr Val Ile Ala Leu Leu Gly Ser Ile Val Ile Ile Val
    50                  55                  60

Met Asn Ile Met Ile Ile Gln Asn Tyr Thr Arg Ser Thr Asp Asn Gln
65                  70                  75                  80

Ala Val Ile Lys Asp Ala Leu Gln Gly Ile Gln Gln Gln Ile Lys Gly
                85                  90                  95

Leu Ala Asp Lys Ile Gly Thr Glu Ile Gly Pro Lys Val Ser Leu Ile
            100                 105                 110

Asp Thr Ser Ser Thr Ile Thr Ile Pro Ala Asn Ile Gly Leu Leu Gly
        115                 120                 125

Ser Lys Ile Ser Gln Ser Thr Ala Ser Ile Asn Glu Asn Val Asn Glu
    130                 135                 140

Lys Cys Lys Phe Thr Leu Pro Pro Leu Lys Ile His Glu Cys Asn Ile
145                 150                 155                 160

Ser Cys Pro Asn Pro Leu Pro Phe Arg Glu Tyr Arg Pro Gln Thr Glu
                165                 170                 175

Gly Val Ser Asn Leu Val Gly Leu Pro Asn Asn Ile Cys Leu Gln Lys
            180                 185                 190

Thr Ser Asn Gln Ile Leu Lys Pro Lys Leu Ile Ser Tyr Thr Leu Pro
        195                 200                 205

Val Val Gly Gln Ser Gly Thr Cys Ile Thr Asp Pro Leu Leu Ala Met
    210                 215                 220

Asp Glu Gly Tyr Phe Ala Tyr Ser His Leu Glu Arg Ile Gly Ser Cys
225                 230                 235                 240

Ser Arg Gly Val Ser Lys Gln Arg Ile Ile Gly Val Gly Glu Val Leu
                245                 250                 255

Asp Arg Gly Asp Glu Val Pro Ser Leu Phe Met Thr Asn Val Trp Thr
            260                 265                 270

Pro Pro Asn Pro Asn Thr Val Tyr His Cys Ser Ala Val Tyr Asn Asn
        275                 280                 285
```

```
Glu Phe Tyr Tyr Val Leu Cys Ala Val Ser Thr Val Gly Asp Pro Ile
    290                 295                 300

Leu Asn Ser Thr Tyr Trp Ser Gly Ser Leu Met Met Thr Arg Leu Ala
305                 310                 315                 320

Val Lys Pro Lys Ser Asn Gly Gly Gly Tyr Asn Gln His Gln Leu Ala
                325                 330                 335

Leu Arg Ser Ile Glu Lys Gly Arg Tyr Asp Lys Val Met Pro Tyr Gly
            340                 345                 350

Pro Ser Gly Ile Lys Gln Gly Asp Thr Leu Tyr Phe Pro Ala Val Gly
        355                 360                 365

Phe Leu Val Arg Thr Glu Phe Lys Tyr Asn Asp Ser Asn Cys Pro Ile
    370                 375                 380

Thr Lys Cys Gln Tyr Ser Lys Pro Glu Asn Cys Arg Leu Ser Met Gly
385                 390                 395                 400

Ile Arg Pro Asn Ser His Tyr Ile Leu Arg Ser Gly Leu Leu Lys Tyr
                405                 410                 415

Asn Leu Ser Asp Gly Glu Asn Pro Lys Val Val Phe Ile Glu Ile Ser
            420                 425                 430

Asp Gln Arg Leu Ser Ile Gly Ser Pro Ser Lys Ile Tyr Asp Ser Leu
        435                 440                 445

Gly Gln Pro Val Phe Tyr Gln Ala Ser Phe Ser Trp Asp Thr Met Ile
    450                 455                 460

Lys Phe Gly Asp Val Leu Thr Val Asn Pro Leu Val Val Asn Trp Arg
465                 470                 475                 480

Asn Asn Thr Val Ile Ser Arg Pro Gly Gln Ser Gln Cys Pro Arg Phe
                485                 490                 495

Asn Thr Cys Pro Glu Ile Cys Trp Glu Gly Val Tyr Asn Asp Ala Phe
            500                 505                 510

Leu Ile Asp Arg Ile Asn Trp Ile Ser Ala Gly Val Phe Leu Asp Ser
        515                 520                 525

Asn Gln Thr Ala Glu Asn Pro Val Phe Thr Val Phe Lys Asp Asn Glu
    530                 535                 540

Ile Leu Tyr Arg Ala Gln Leu Ala Ser Glu Asp Thr Asn Ala Gln Lys
545                 550                 555                 560
```

```
Thr Ile Thr Asn Cys Phe Leu Leu Lys Asn Lys Ile Trp Cys Ile Ser
                565                 570                 575


Leu Val Glu Ile Tyr Asp Thr Gly Asp Asn Val Ile Arg Pro Lys Leu
            580                 585                 590


Phe Ala Val Lys Ile Pro Glu Gln Cys Thr
        595                 600


<210>  72
<211>  572
<212>  PRT
<213>  Swine parainfluenza virus 3


<220>
<221>  MISC_FEATURE
<223>  hrmagglutinin neuraminidase HN

<400>  72

Met Glu Tyr Trp Lys His Thr Asn Ser Thr Lys Asp Thr Asn Asn Glu
1               5                   10                  15


Leu Gly Thr Thr Arg Asp Arg His Ser Ser Lys Ala Thr Asn Ile Ile
            20                  25                  30


Met Tyr Ile Phe Trp Thr Thr Thr Ser Thr Ile Leu Ser Val Ile Phe
        35                  40                  45


Ile Met Ile Leu Ile Asn Leu Ile Gln Glu Asn Asn His Asn Lys Leu
    50                  55                  60


Met Leu Gln Glu Ile Lys Lys Glu Phe Ala Val Ile Asp Thr Lys Ile
65                  70                  75                  80


Gln Lys Thr Ser Asp Asp Ile Ser Thr Ser Ile Gln Ser Gly Ile Asn
                85                  90                  95


Thr Arg Leu Leu Thr Ile Gln Ser His Val Gln Asn Tyr Ile Pro Leu
            100                 105                 110


Ser Leu Thr Gln Gln Met Ser Asp Leu Arg Lys Phe Ile Asn Asp Leu
        115                 120                 125


Thr Thr Lys Arg Glu His Gln Glu Val Pro Ile Gln Arg Met Thr His
    130                 135                 140


Asp Ser Gly Ile Glu Pro Leu Asn Pro Asp Lys Phe Trp Arg Cys Thr
145                 150                 155                 160


Ser Gly Asn Pro Ser Leu Thr Ser Ser Pro Lys Ile Arg Leu Ile Pro
                165                 170                 175
```

```
Gly Pro Gly Leu Leu Ala Thr Ser Thr Thr Val Asn Gly Cys Ile Arg
            180                 185                 190

Ile Pro Ser Leu Ala Ile Asn Asn Leu Ile Tyr Ala Tyr Thr Ser Asn
        195                 200                 205

Leu Ile Thr Gln Gly Cys Gln Asp Ile Gly Lys Ser Tyr Gln Val Leu
    210                 215                 220

Gln Ile Gly Ile Ile Thr Ile Asn Ser Asp Leu Val Pro Asp Leu Asn
225                 230                 235                 240

Pro Arg Val Thr His Thr Phe Asn Ile Asp Asp Asn Arg Lys Ser Cys
                245                 250                 255

Ser Leu Ala Leu Leu Asn Thr Asp Val Tyr Gln Leu Cys Ser Thr Pro
            260                 265                 270

Lys Val Asp Glu Arg Ser Asp Tyr Ala Ser Thr Gly Ile Glu Asp Ile
        275                 280                 285

Val Leu Asp Ile Val Thr Ser Asn Gly Leu Ile Ile Thr Thr Arg Phe
    290                 295                 300

Thr Asn Asn Asn Ile Thr Phe Asp Lys Pro Tyr Ala Ala Leu Tyr Pro
305                 310                 315                 320

Ser Val Gly Pro Gly Ile Tyr Tyr Lys Asp Lys Val Ile Phe Leu Gly
                325                 330                 335

Tyr Gly Gly Leu Glu His Glu Glu Asn Gly Asp Val Ile Cys Asn Thr
            340                 345                 350

Thr Gly Cys Pro Gly Lys Thr Gln Arg Asp Cys Asn Gln Ala Ser Tyr
        355                 360                 365

Ser Pro Trp Phe Ser Asn Arg Arg Met Val Asn Ser Ile Ile Val Val
    370                 375                 380

Asp Lys Ser Ile Asp Thr Thr Phe Ser Leu Arg Val Trp Thr Ile Pro
385                 390                 395                 400

Met Arg Gln Asn Tyr Trp Gly Ser Glu Gly Arg Leu Leu Leu Leu Gly
                405                 410                 415

Asp Arg Ile Tyr Ile Tyr Thr Arg Ser Thr Ser Trp His Ser Lys Leu
            420                 425                 430

Gln Leu Gly Val Ile Asp Ile Ser Asp Tyr Asn Asn Ile Arg Ile Asn
        435                 440                 445

Trp Thr Trp His Asn Val Leu Ser Arg Pro Gly Asn Asp Glu Cys Pro
```

```
     450                 455                 460

 Trp Gly His Ser Cys Pro Asp Gly Cys Ile Thr Gly Val Tyr Thr Asp
 465                 470                 475                 480

 Ala Tyr Pro Leu Asn Pro Ser Gly Ser Val Val Ser Ser Val Ile Leu
                 485                 490                 495

 Asp Ser Gln Lys Ser Arg Glu Asn Pro Ile Ile Thr Tyr Ser Thr Ala
             500                 505                 510

 Thr Asn Arg Val Asn Glu Leu Ala Ile Tyr Asn Arg Thr Leu Pro Ala
         515                 520                 525

 Ala Tyr Thr Thr Thr Asn Cys Ile Thr His Tyr Asp Lys Gly Tyr Cys
     530                 535                 540

 Phe His Ile Val Glu Ile Asn His Arg Ser Leu Asn Thr Phe Gln Pro
 545                 550                 555                 560

 Met Leu Phe Lys Thr Glu Val Pro Lys Asn Cys Ser
                 565                 570


 <210>  73
 <211>  231
 <212>  PRT
 <213>  human metapneumovirus


 <220>
 <221>  MISC_FEATURE
 <223>  attachment glycoprotein G

 <400>  73

 Met Glu Val Arg Val Glu Asn Ile Arg Ala Ile Asp Met Phe Lys Ala
 1               5                   10                  15

 Lys Ile Lys Asn Arg Ile Arg Ser Ser Arg Cys Tyr Arg Asn Ala Thr
             20                  25                  30

 Leu Ile Leu Ile Gly Leu Thr Ala Leu Ser Met Ala Leu Asn Ile Phe
         35                  40                  45

 Leu Ile Ile Asp His Ala Thr Leu Arg Asn Met Ile Lys Thr Glu Asn
     50                  55                  60

 Cys Ala Asn Met Pro Ser Ala Glu Pro Ser Lys Lys Thr Pro Met Thr
 65                  70                  75                  80

 Ser Thr Ala Gly Pro Ser Thr Lys Pro Asn Pro Gln Gln Ala Thr Gln
                 85                  90                  95

 Trp Thr Thr Glu Asn Ser Thr Ser Pro Ala Ala Thr Leu Glu Gly His
             100                 105                 110
```

```
Pro Tyr Thr Gly Thr Thr Gln Thr Pro Asp Thr Thr Ala Pro Gln Gln
        115                 120                 125


Thr Thr Asp Lys His Thr Ala Leu Pro Lys Ser Thr Asn Glu Gln Ile
    130                 135                 140


Thr Gln Thr Thr Thr Glu Lys Lys Thr Thr Arg Ala Thr Thr Gln Lys
145                 150                 155                 160


Arg Glu Lys Arg Lys Glu Asn Thr Asn Gln Thr Thr Ser Thr Ala Ala
                165                 170                 175


Thr Gln Thr Thr Asn Thr Thr Asn Gln Thr Arg Asn Ala Ser Glu Thr
            180                 185                 190


Ile Thr Thr Ser Asp Gly Pro Arg Ile Asp Thr Thr Thr Gln Ser Ser
        195                 200                 205


Glu Gln Thr Ala Arg Ala Thr Glu Pro Gly Ser Ser Pro Tyr His Ala
    210                 215                 220


Arg Arg Gly Ala Gly Pro Arg
225                 230


<210>  74
<211>  595
<212>  PRT
<213>  Feline morbillivirus


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin protein

<400>  74

Met Lys Asn Ile Asn Ile Lys Tyr Tyr Lys Asp Ser Asn Arg Tyr Leu
1               5                   10                  15


Gly Lys Ile Leu Asp Glu His Lys Ile Val Asn Ser Gln Leu Tyr Ser
            20                  25                  30


Leu Ser Ile Lys Val Ile Thr Ile Ile Ala Ile Ile Val Ser Leu Ile
        35                  40                  45


Ala Thr Ile Met Thr Ile Ile Asn Ala Thr Ser Gly Arg Thr Thr Leu
    50                  55                  60


Asn Ser Asn Thr Asp Ile Leu Leu Asn Gln Arg Asp Glu Ile His Ser
65                  70                  75                  80


Ile His Glu Met Ile Phe Asp Arg Val Tyr Pro Leu Ile Thr Ala Met
                85                  90                  95
```

```
Ser Thr Glu Leu Gly Leu His Ile Pro Thr Leu Leu Asp Glu Leu Thr
            100                 105                 110

Lys Ala Ile Asp Gln Lys Ile Lys Ile Met Asn Pro Pro Val Asp Thr
        115                 120                 125

Val Thr Ser Asp Leu Ser Trp Cys Ile Lys Pro Pro Asn Gly Ile Ile
    130                 135                 140

Ile Asp Pro Lys Gly Tyr Cys Glu Ser Met Glu Leu Ser Lys Thr Tyr
145                 150                 155                 160

Lys Leu Leu Leu Asp Gln Leu Asp Val Ser Arg Lys Lys Ser Leu Thr
                165                 170                 175

Ile Asn Arg Lys Asn Ile Asn Gln Cys Gln Leu Val Asp Asp Ser Glu
            180                 185                 190

Ile Ile Phe Ala Thr Val Asn Ile Gln Ser Thr Pro Arg Phe Leu Asn
        195                 200                 205

Phe Gly His Thr Val Ser Asn Gln Arg Ile Thr Phe Gly Gln Gly Thr
    210                 215                 220

Tyr Ser Ser Thr Tyr Ile Leu Thr Ile Gln Glu Asp Gly Ile Thr Asp
225                 230                 235                 240

Val Gln Tyr Arg Val Phe Glu Ile Gly Tyr Ile Ser Asp Gln Phe Gly
                245                 250                 255

Val Phe Pro Ser Leu Ile Val Ser Arg Val Leu Pro Ile Arg Met Val
            260                 265                 270

Leu Gly Met Glu Ser Cys Thr Leu Thr Ser Asp Arg Gln Gly Gly Tyr
        275                 280                 285

Phe Leu Cys Met Asn Thr Leu Thr Arg Ser Ile Tyr Asp Tyr Val Asn
    290                 295                 300

Ile Arg Asp Leu Lys Ser Leu Tyr Ile Thr Leu Pro His Tyr Gly Lys
305                 310                 315                 320

Val Asn Tyr Thr Tyr Phe Asn Phe Gly Lys Ile Arg Ser Pro His Glu
                325                 330                 335

Ile Asp Lys Leu Trp Leu Thr Ser Asp Arg Gly Gln Ile Ile Ser Gly
            340                 345                 350

Tyr Phe Ala Ala Phe Val Thr Ile Thr Ile Arg Asn Tyr Asn Asn Tyr
        355                 360                 365
```

```
Pro Tyr Lys Cys Leu Asn Asn Pro Cys Phe Asp Asn Ser Glu Asn Tyr
    370                 375                 380


Cys Arg Gly Trp Tyr Lys Asn Ile Thr Gly Thr Asp Asp Val Pro Ile
385                 390                 395                 400


Leu Ala Tyr Leu Leu Val Glu Met Tyr Asp Glu Glu Gly Pro Leu Ile
                405                 410                 415


Thr Leu Val Ala Ile Pro Pro Tyr Asn Tyr Thr Ala Pro Ser His Asn
            420                 425                 430


Ser Leu Tyr Tyr Asp Asp Lys Ile Asn Lys Leu Ile Met Thr Thr Ser
        435                 440                 445


His Ile Gly Tyr Ile Gln Ile Asn Glu Val His Glu Val Ile Val Gly
    450                 455                 460


Asp Asn Leu Lys Ala Ile Leu Leu Asn Arg Leu Ser Asp Glu His Pro
465                 470                 475                 480


Asn Leu Thr Ala Cys Arg Leu Asn Gln Gly Ile Lys Glu Gln Tyr Lys
                485                 490                 495


Ser Asp Gly Met Ile Ile Ser Asn Ser Ala Leu Ile Asp Ile Gln Glu
            500                 505                 510


Arg Met Tyr Ile Thr Val Lys Ala Ile Pro Pro Val Gly Asn Tyr Asn
        515                 520                 525


Phe Thr Val Glu Leu His Ser Arg Ser Asn Thr Ser Tyr Ile Leu Leu
    530                 535                 540


Pro Lys Gln Phe Asn Ala Lys Tyr Asp Lys Leu His Leu Glu Cys Phe
545                 550                 555                 560


Asn Trp Asp Lys Ser Trp Trp Cys Ala Leu Ile Pro Gln Phe Ser Leu
                565                 570                 575


Ser Trp Asn Glu Ser Leu Ser Val Asp Thr Ala Ile Phe Asn Leu Ile
            580                 585                 590


Asn Cys Lys
        595
```

<210> 75
<211> 613
<212> PRT
<213> Avian paramyxovirus 6

<400> 75

```
Met Ala Ser Pro Ser Glu Leu Asn Arg Ser Gln Ala Thr Leu Tyr Glu
1               5                   10                  15
```

```
Gly Asp Pro Asn Ser Lys Arg Thr Trp Arg Thr Val Tyr Arg Ala Ser
            20                  25                  30

Thr Leu Ile Leu Asp Leu Ala Ile Leu Cys Val Ser Ile Val Ala Ile
        35                  40                  45

Val Arg Met Ser Thr Leu Thr Pro Ser Asp Val Thr Asp Ser Ile Ser
    50                  55                  60

Ser Ser Ile Thr Ser Leu Ser Asp Thr Tyr Gln Ser Val Trp Ser Asp
65                  70                  75                  80

Thr His Gln Lys Val Asn Ser Ile Phe Lys Glu Val Gly Ile Ser Ile
                85                  90                  95

Pro Val Thr Leu Asp Lys Met Gln Val Glu Met Gly Thr Ala Val Asn
            100                 105                 110

Ile Ile Thr Asp Ala Val Arg Gln Leu Gln Gly Val Asn Gly Ser Ala
        115                 120                 125

Gly Phe Ser Ile Thr Asn Ser Pro Glu Tyr Ser Gly Gly Ile Asp Ala
    130                 135                 140

Leu Ile Tyr Pro Gln Lys Ser Leu Asn Gly Lys Ser Leu Ala Ile Ser
145                 150                 155                 160

Asp Leu Leu Glu His Pro Ser Phe Ile Pro Ala Pro Thr Thr Ser His
                165                 170                 175

Gly Cys Thr Arg Ile Pro Thr Phe His Leu Gly Tyr Arg His Trp Cys
            180                 185                 190

Tyr Ser His Asn Thr Ile Glu Ser Gly Cys His Asp Ala Gly Glu Ser
        195                 200                 205

Ile Met Tyr Leu Ser Met Gly Ala Val Gly Val Gly His Gln Gly Lys
    210                 215                 220

Pro Val Phe Thr Thr Ser Ala Ala Val Ile Leu Asp Asp Gly Lys Asn
225                 230                 235                 240

Arg Lys Ser Cys Ser Val Val Ala Asn Pro Asn Gly Cys Asp Val Leu
                245                 250                 255

Cys Ser Leu Val Lys Gln Thr Glu Asp Gln Asp Tyr Ala Asp Pro Thr
            260                 265                 270

Pro Thr Pro Met Ile His Gly Arg Leu His Phe Asn Gly Thr Tyr Thr
        275                 280                 285
```

```
Glu Ser Met Leu Asp Gln Ser Leu Phe Thr Gly His Trp Val Ala Gln
    290                 295                 300

Tyr Pro Ala Val Gly Ser Gly Ser Val Ser His Gly Arg Leu Phe Phe
305                 310                 315                 320

Pro Leu Tyr Gly Gly Ile Ser Lys Ser Ser Ser Leu Phe Pro Lys Leu
                325                 330                 335

Arg Ala His Ala Tyr Phe Thr His Asn Glu Glu Leu Glu Cys Lys Asn
            340                 345                 350

Leu Thr Ser Lys Gln Arg Glu Asp Leu Phe Asn Ala Tyr Met Pro Gly
        355                 360                 365

Lys Ile Ala Gly Ser Leu Trp Ala Gln Gly Ile Val Ile Cys Asn Leu
    370                 375                 380

Thr Thr Leu Ala Asp Cys Lys Ile Ala Val Ala Asn Thr Ser Thr Met
385                 390                 395                 400

Met Met Ala Ala Glu Gly Arg Leu Gln Leu Val Gln Asp Lys Val Val
                405                 410                 415

Leu Tyr Gln Arg Ser Ser Ser Trp Trp Pro Val Leu Ile Tyr Tyr Asp
            420                 425                 430

Ile Leu Val Ser Glu Leu Val Asn Ala Arg His Leu Asp Ile Val Asn
        435                 440                 445

Trp Val Pro Tyr Pro Gln Ser Lys Phe Pro Arg Pro Thr Trp Thr Lys
    450                 455                 460

Gly Leu Cys Glu Lys Pro Ser Ile Cys Pro Ala Val Cys Val Thr Gly
465                 470                 475                 480

Val Tyr Gln Asp Val Trp Val Val Ser Val Gly Asp Phe Ser Asn Glu
                485                 490                 495

Thr Val Val Ile Gly Gly Tyr Leu Glu Ala Ala Ser Glu Arg Lys Asp
            500                 505                 510

Pro Trp Ile Ala Ala Ala Asn Gln Tyr Asn Trp Leu Thr Arg Arg Gln
        515                 520                 525

Leu Phe Thr Ala Gln Thr Glu Ala Ala Tyr Ser Ser Thr Thr Cys Phe
    530                 535                 540

Arg Asn Thr His Gln Asp Lys Val Phe Cys Leu Thr Ile Met Glu Val
545                 550                 555                 560
```

```
Thr Asp Asn Leu Leu Gly Asp Trp Arg Ile Ala Pro Leu Leu Tyr Glu
                565                 570                 575


Val Thr Val Val Asp Arg Gln Gln Ser Ser Arg Lys Ala Val Ala Met
            580                 585                 590


Ser Glu Ala His Arg Thr Arg Phe Lys Tyr Tyr Ser Pro Glu Asn Lys
        595                 600                 605


Phe Thr Pro Gln His
    610


<210>  76
<211>  564
<212>  PRT
<213>  Fer-de-Lance paramyxovirus


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin neuraminidase protein

<400>  76

Met Asp Pro Lys Ser Tyr Tyr Cys Asn Glu Asp Leu Arg Ser Asp Gly
1               5                   10                  15


Gly Glu Lys Ser Pro Gly Gly Asp Leu Tyr Lys Gly Ile Ile Leu Val
            20                  25                  30


Ser Thr Val Ile Ser Leu Ile Ile Ala Ile Ile Ser Leu Ala Phe Ile
        35                  40                  45


Ile Asp Asn Lys Ile Asn Ile Gln Ser Leu Asp Pro Leu Arg Gly Leu
    50                  55                  60


Glu Asp Ser Tyr Leu Val Pro Ile Lys Asp Lys Ser Glu Ser Ile Ser
65                  70                  75                  80


Gln Asp Ile Gln Glu Gly Ile Phe Pro Arg Leu Asn Leu Ile Thr Ala
                85                  90                  95


Ala Thr Thr Thr Thr Ile Pro Arg Ser Ile Ala Ile Gln Thr Lys Asp
            100                 105                 110


Leu Ser Asp Leu Ile Met Asn Arg Cys Tyr Pro Ser Val Val Asn Asn
        115                 120                 125


Asp Thr Ser Cys Asp Val Leu Ala Gly Ala Ile His Ser Asn Leu Phe
    130                 135                 140


Ser Gln Leu Asp Pro Ser Thr Tyr Trp Thr Cys Ser Ser Gly Thr Pro
145                 150                 155                 160


Thr Met Asn Gln Thr Val Lys Leu Leu Pro Asp Asn Ser Gln Ile Pro
```

```
                165                 170                 175

Gly Ser Thr Tyr Ser Thr Gly Cys Val Arg Ile Pro Thr Phe Ser Leu
            180                 185                 190

Gly Ser Met Ile Tyr Ser Tyr Ser His Asn Val Ile Tyr Glu Gly Cys
        195                 200                 205

Asn Asp His Ser Lys Ser Ser Gln Tyr Trp Gln Leu Gly Tyr Ile Ser
    210                 215                 220

Thr Ser Lys Thr Gly Glu Pro Leu Gln Gln Val Ser Arg Thr Leu Thr
225                 230                 235                 240

Leu Asn Asn Gly Leu Asn Arg Lys Ser Cys Ser Thr Val Ala Gln Gly
                245                 250                 255

Arg Gly Ala Tyr Leu Leu Cys Thr Asn Val Val Glu Asp Glu Arg Thr
            260                 265                 270

Asp Tyr Ser Thr Glu Gly Ile Gln Asp Leu Thr Leu Asp Tyr Ile Asp
        275                 280                 285

Ile Phe Gly Ala Glu Arg Ser Tyr Arg Tyr Thr Asn Asn Glu Val Asp
    290                 295                 300

Leu Asp Arg Pro Tyr Ala Ala Leu Tyr Pro Ser Val Gly Ser Gly Thr
305                 310                 315                 320

Val Tyr Asn Asp Arg Ile Leu Phe Leu Gly Tyr Gly Gly Leu Met Thr
                325                 330                 335

Pro Tyr Gly Asp Gln Ala Met Cys Gln Ala Pro Glu Cys Thr Ser Ala
            340                 345                 350

Thr Gln Glu Gly Cys Asn Ser Asn Gln Leu Ile Gly Tyr Phe Ser Gly
        355                 360                 365

Arg Gln Ile Val Asn Cys Ile Ile Glu Ile Ile Thr Val Gly Thr Glu
    370                 375                 380

Lys Pro Ile Ile Arg Val Arg Thr Ile Pro Asn Ser Gln Val Trp Leu
385                 390                 395                 400

Gly Ala Glu Gly Arg Ile Gln Thr Leu Gly Gly Val Leu Tyr Leu Tyr
                405                 410                 415

Ile Arg Ser Ser Gly Trp His Ala Leu Ala Gln Thr Gly Ile Ile Leu
            420                 425                 430

Thr Leu Asp Pro Ile Arg Ile Ser Trp Ile Val Asn Thr Gly Tyr Ser
        435                 440                 445
```

```
Arg Pro Gly Asn Gly Pro Cys Ser Ala Ser Ser Arg Cys Pro Ala Gln
    450                 455                 460


Cys Ile Thr Gly Val Tyr Thr Asp Ile Phe Pro Leu Ser Gln Asn Tyr
465                 470                 475                 480


Gly Tyr Leu Ala Thr Val Thr Leu Leu Ser Gly Val Asp Arg Val Asn
                485                 490                 495


Pro Val Ile Ser Tyr Gly Thr Ser Thr Gly Arg Val Ala Asp Ser Gln
            500                 505                 510


Leu Thr Ser Ser Ser Gln Val Ala Ala Tyr Thr Thr Thr Thr Cys Phe
        515                 520                 525


Thr Phe Asn Gln Lys Gly Tyr Cys Tyr His Ile Ile Glu Leu Ser Pro
    530                 535                 540


Ala Thr Leu Gly Ile Phe Gln Pro Val Leu Val Val Thr Glu Ile Pro
545                 550                 555                 560


Lys Ile Cys Ser



<210>  77
<211>  569
<212>  PRT
<213>  Avian paramyxovirus 4


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin neruaminidase

<400>  77

Met Gln Gly Asn Met Glu Gly Ser Arg Asp Asn Leu Thr Val Asp Asp
1               5                   10                  15


Glu Leu Lys Thr Thr Trp Arg Leu Ala Tyr Arg Val Val Ser Leu Leu
            20                  25                  30


Leu Met Val Ser Ala Leu Ile Ile Ser Ile Val Ile Leu Thr Arg Asp
        35                  40                  45


Asn Ser Gln Ser Ile Ile Thr Ala Ile Asn Gln Ser Ser Asp Ala Asp
    50                  55                  60


Ser Lys Trp Gln Thr Gly Ile Glu Gly Lys Ile Thr Ser Ile Met Thr
65                  70                  75                  80


Asp Thr Leu Asp Thr Arg Asn Ala Ala Leu Leu His Ile Pro Leu Gln
                85                  90                  95
```

```
Leu Asn Thr Leu Glu Ala Asn Leu Leu Ser Ala Leu Gly Gly Asn Thr
            100                 105                 110

Gly Ile Gly Pro Gly Asp Leu Glu His Cys Arg Tyr Pro Val His Asp
        115                 120                 125

Thr Ala Tyr Leu His Gly Val Asn Arg Leu Leu Ile Asn Gln Thr Ala
    130                 135                 140

Asp Tyr Thr Ala Glu Gly Pro Leu Asp His Val Asn Phe Ile Pro Ala
145                 150                 155                 160

Pro Val Thr Thr Thr Gly Cys Thr Arg Ile Pro Ser Phe Ser Val Ser
                165                 170                 175

Ser Ser Ile Trp Cys Tyr Thr His Asn Val Ile Glu Thr Gly Cys Asn
            180                 185                 190

Asp His Ser Gly Ser Asn Gln Tyr Ile Ser Met Gly Val Ile Lys Arg
        195                 200                 205

Ala Gly Asn Gly Leu Pro Tyr Phe Ser Thr Val Val Ser Lys Tyr Leu
    210                 215                 220

Thr Asp Gly Leu Asn Arg Lys Ser Cys Ser Val Ala Ala Gly Ser Gly
225                 230                 235                 240

His Cys Tyr Leu Leu Cys Ser Leu Val Ser Glu Pro Glu Pro Asp Asp
                245                 250                 255

Tyr Val Ser Pro Asp Pro Thr Pro Met Arg Leu Gly Val Leu Thr Trp
            260                 265                 270

Asp Gly Ser Tyr Thr Glu Gln Ala Val Pro Glu Arg Ile Phe Lys Asn
        275                 280                 285

Ile Trp Ser Ala Asn Tyr Pro Gly Val Gly Ser Gly Ala Ile Val Gly
    290                 295                 300

Asn Lys Val Leu Phe Pro Phe Tyr Gly Gly Val Arg Asn Gly Ser Thr
305                 310                 315                 320

Pro Glu Val Met Asn Arg Gly Arg Tyr Tyr Tyr Ile Gln Asp Pro Asn
                325                 330                 335

Asp Tyr Cys Pro Asp Pro Leu Gln Asp Gln Ile Leu Arg Ala Glu Gln
            340                 345                 350

Ser Tyr Tyr Pro Thr Arg Phe Gly Arg Arg Met Val Met Gln Gly Val
        355                 360                 365
```

Leu Ala Cys Pro Val Ser Asn Asn Ser Thr Ile Ala Ser Gln Cys Gln
370 375 380

Ser Tyr Tyr Phe Asn Asn Ser Leu Gly Phe Ile Gly Ala Glu Ser Arg
385 390 395 400

Ile Tyr Tyr Leu Asn Gly Asn Ile Tyr Leu Tyr Gln Arg Ser Ser Ser
405 410 415

Trp Trp Pro His Pro Gln Ile Tyr Leu Leu Asp Ser Arg Ile Ala Ser
420 425 430

Pro Gly Thr Gln Asn Ile Asp Ser Gly Val Asn Leu Lys Met Leu Asn
435 440 445

Val Thr Val Ile Thr Arg Pro Ser Ser Gly Phe Cys Asn Ser Gln Ser
450 455 460

Arg Cys Pro Asn Asp Cys Leu Phe Gly Val Tyr Ser Asp Ile Trp Pro
465 470 475 480

Leu Ser Leu Thr Ser Asp Ser Ile Phe Ala Phe Thr Met Tyr Leu Gln
485 490 495

Gly Lys Thr Thr Arg Ile Asp Pro Ala Trp Ala Leu Phe Ser Asn His
500 505 510

Ala Ile Gly His Glu Ala Arg Leu Phe Asn Lys Glu Val Ser Ala Ala
515 520 525

Tyr Ser Thr Thr Thr Cys Phe Ser Asp Thr Ile Gln Asn Gln Val Tyr
530 535 540

Cys Leu Ser Ile Leu Glu Val Arg Ser Glu Leu Leu Gly Ala Phe Lys
545 550 555 560

Ile Val Pro Phe Leu Tyr Arg Val Leu
565

<210> 78
<211> 571
<212> PRT
<213> Human parainfluenza virus 2

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 78

Met Glu Asp Tyr Ser Asn Leu Ser Leu Lys Ser Ile Pro Lys Arg Thr
1 5 10 15

Cys Arg Ile Ile Phe Arg Thr Ala Thr Ile Leu Gly Ile Cys Thr Leu

20 25 30

Ile Val Leu Cys Ser Ser Ile Leu His Glu Ile Ile His Leu Asp Val
35 40 45

Ser Ser Gly Leu Met Asp Ser Asp Asp Ser Gln Gln Gly Ile Ile Gln
50 55 60

Pro Ile Ile Glu Ser Leu Lys Ser Leu Ile Ala Leu Ala Asn Gln Ile
65 70 75 80

Leu Tyr Asn Val Ala Ile Ile Ile Pro Leu Lys Ile Asp Ser Ile Glu
85 90 95

Thr Val Ile Phe Ser Ala Leu Lys Asp Met His Thr Gly Ser Met Ser
100 105 110

Asn Thr Asn Cys Thr Pro Gly Asn Leu Leu Leu His Asp Ala Ala Tyr
115 120 125

Ile Asn Gly Ile Asn Lys Phe Leu Val Leu Lys Ser Tyr Asn Gly Thr
130 135 140

Pro Lys Tyr Gly Pro Leu Leu Asn Ile Pro Ser Phe Ile Pro Ser Ala
145 150 155 160

Thr Ser Pro Asn Gly Cys Thr Arg Ile Pro Ser Phe Ser Leu Ile Lys
165 170 175

Thr His Trp Cys Tyr Thr His Asn Val Met Leu Gly Asp Cys Leu Asp
180 185 190

Phe Thr Thr Ser Asn Gln Tyr Leu Ala Met Gly Ile Ile Gln Gln Ser
195 200 205

Ala Ala Ala Phe Pro Ile Phe Arg Thr Met Lys Thr Ile Tyr Leu Ser
210 215 220

Asp Gly Ile Asn Arg Lys Ser Cys Ser Val Thr Ala Ile Pro Gly Gly
225 230 235 240

Cys Val Leu Tyr Cys Tyr Val Ala Thr Arg Ser Glu Lys Glu Asp Tyr
245 250 255

Ala Thr Thr Asp Leu Ala Glu Leu Arg Leu Ala Phe Tyr Tyr Tyr Asn
260 265 270

Asp Thr Phe Ile Glu Arg Val Ile Ser Leu Pro Asn Thr Thr Gly Gln
275 280 285

Trp Ala Thr Ile Asn Pro Ala Val Gly Ser Gly Ile Tyr His Leu Gly
290 295 300

```
Phe Ile Leu Phe Pro Val Tyr Gly Gly Leu Ile Ser Gly Thr Pro Ser
305                 310                 315                 320

Tyr Asn Lys Gln Ser Ser Arg Tyr Phe Ile Pro Lys His Pro Asn Ile
                325                 330                 335

Thr Cys Ala Gly Asn Ser Ser Glu Gln Ala Ala Ala Ala Arg Ser Ser
            340                 345                 350

Tyr Val Ile Arg Tyr His Ser Asn Arg Leu Ile Gln Ser Ala Val Leu
        355                 360                 365

Ile Cys Pro Leu Ser Asp Met His Thr Ala Arg Cys Asn Leu Val Met
    370                 375                 380

Phe Asn Asn Ser Gln Val Met Met Gly Ala Glu Gly Arg Leu Tyr Val
385                 390                 395                 400

Ile Asp Asn Asn Leu Tyr Tyr Tyr Gln Arg Ser Ser Ser Trp Trp Ser
                405                 410                 415

Ala Ser Leu Phe Tyr Arg Ile Asn Thr Asp Phe Ser Lys Gly Ile Pro
            420                 425                 430

Pro Ile Ile Glu Ala Gln Trp Val Pro Ser Tyr Gln Val Pro Arg Pro
        435                 440                 445

Gly Val Met Pro Cys Asn Ala Thr Ser Phe Cys Pro Ala Asn Cys Ile
    450                 455                 460

Thr Gly Val Tyr Ala Asp Val Trp Pro Leu Asn Asp Pro Glu Pro Thr
465                 470                 475                 480

Ser Gln Asn Ala Leu Asn Pro Asn Tyr Arg Phe Ala Gly Ala Phe Leu
                485                 490                 495

Arg Asn Glu Ser Asn Arg Thr Asn Pro Thr Phe Tyr Thr Ala Ser Ala
            500                 505                 510

Ser Ala Leu Leu Asn Thr Thr Gly Phe Asn Asn Thr Asn His Lys Ala
        515                 520                 525

Ala Tyr Thr Ser Ser Thr Cys Phe Lys Asn Thr Gly Thr Gln Lys Ile
    530                 535                 540

Tyr Cys Leu Ile Ile Ile Glu Met Gly Ser Ser Leu Leu Gly Glu Phe
545                 550                 555                 560

Gln Ile Ile Pro Phe Leu Arg Glu Leu Ile Pro
                565                 570
```

<210> 79
<211> 565
<212> PRT
<213> Parainfluenza virus 5

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 79

Met Val Ala Glu Asp Ala Pro Val Arg Ala Thr Cys Arg Val Leu Phe
1 5 10 15

Arg Thr Thr Thr Leu Ile Phe Leu Cys Thr Leu Leu Ala Leu Ser Ile
20 25 30

Ser Ile Leu Tyr Glu Ser Leu Ile Thr Gln Lys Gln Ile Met Ser Gln
35 40 45

Ala Gly Ser Thr Gly Ser Asn Ser Gly Leu Gly Ser Ile Thr Asp Leu
50 55 60

Leu Asn Asn Ile Leu Ser Val Ala Asn Gln Ile Ile Tyr Asn Ser Ala
65 70 75 80

Val Ala Leu Pro Leu Gln Leu Asp Thr Leu Glu Ser Thr Leu Leu Thr
85 90 95

Ala Ile Lys Ser Leu Gln Thr Ser Asp Lys Leu Glu Gln Asn Cys Ser
100 105 110

Trp Ser Ala Ala Leu Ile Asn Asp Asn Arg Tyr Ile Asn Gly Ile Asn
115 120 125

Gln Phe Tyr Phe Ser Ile Ala Glu Gly Arg Asn Leu Thr Leu Gly Pro
130 135 140

Leu Leu Asn Met Pro Ser Phe Ile Pro Thr Ala Thr Thr Pro Glu Gly
145 150 155 160

Cys Thr Arg Ile Pro Ser Phe Ser Leu Thr Lys Thr His Trp Cys Tyr
165 170 175

Thr His Asn Val Ile Leu Asn Gly Cys Gln Asp His Val Ser Ser Asn
180 185 190

Gln Phe Val Ser Met Gly Ile Ile Glu Pro Thr Ser Ala Gly Phe Pro
195 200 205

Phe Phe Arg Thr Leu Lys Thr Leu Tyr Leu Ser Asp Gly Val Asn Arg
210 215 220

```
Lys Ser Cys Ser Ile Ser Thr Val Pro Gly Gly Cys Met Met Tyr Cys
225                 230                 235                 240

Phe Val Ser Thr Gln Pro Glu Arg Asp Asp Tyr Phe Ser Ala Ala Pro
                245                 250                 255

Pro Glu Gln Arg Ile Ile Ile Met Tyr Tyr Asn Asp Thr Ile Val Glu
            260                 265                 270

Arg Ile Ile Asn Pro Pro Gly Val Leu Asp Val Trp Ala Thr Leu Asn
        275                 280                 285

Pro Gly Thr Gly Ser Gly Val Tyr Tyr Leu Gly Trp Val Leu Phe Pro
    290                 295                 300

Ile Tyr Gly Gly Val Ile Lys Gly Thr Ser Leu Trp Asn Asn Gln Ala
305                 310                 315                 320

Asn Lys Tyr Phe Ile Pro Gln Met Val Ala Ala Leu Cys Ser Gln Asn
                325                 330                 335

Gln Ala Thr Gln Val Gln Asn Ala Lys Ser Ser Tyr Tyr Ser Ser Trp
            340                 345                 350

Phe Gly Asn Arg Met Ile Gln Ser Gly Ile Leu Ala Cys Pro Leu Arg
        355                 360                 365

Gln Asp Leu Thr Asn Glu Cys Leu Val Leu Pro Phe Ser Asn Asp Gln
    370                 375                 380

Val Leu Met Gly Ala Glu Gly Arg Leu Tyr Met Tyr Gly Asp Ser Val
385                 390                 395                 400

Tyr Tyr Tyr Gln Arg Ser Asn Ser Trp Trp Pro Met Thr Met Leu Tyr
                405                 410                 415

Lys Val Thr Ile Thr Phe Thr Asn Gly Gln Pro Ser Ala Ile Ser Ala
            420                 425                 430

Gln Asn Val Pro Thr Gln Gln Val Pro Arg Pro Gly Thr Gly Asp Cys
        435                 440                 445

Ser Ala Thr Asn Arg Cys Pro Gly Phe Cys Leu Thr Gly Val Tyr Ala
    450                 455                 460

Asp Ala Trp Leu Leu Thr Asn Pro Ser Ser Thr Ser Thr Phe Gly Ser
465                 470                 475                 480

Glu Ala Thr Phe Thr Gly Ser Tyr Leu Asn Thr Ala Thr Gln Arg Ile
                485                 490                 495

Asn Pro Thr Met Tyr Ile Ala Asn Asn Thr Gln Ile Ile Ser Ser Gln
```

500 505 510

Gln Phe Gly Ser Ser Gly Gln Glu Ala Ala Tyr Gly His Thr Thr Cys
515 520 525

Phe Arg Asp Thr Gly Ser Val Met Val Tyr Cys Ile Tyr Ile Ile Glu
530 535 540

Leu Ser Ser Ser Leu Leu Gly Gln Phe Gln Ile Val Pro Phe Ile Arg
545 550 555 560

Gln Val Thr Leu Ser
565

<210> 80
<211> 576
<212> PRT
<213> Porcine rubulavirus

<220>
<221> MISC_FEATURE
<223> attachment protein

<400> 80

Met Ser Gln Leu Gly Thr Asp Gln Ile Met His Leu Ala Gln Pro Ala
1 5 10 15

Ile Ala Arg Arg Thr Trp Arg Leu Cys Phe Arg Ile Phe Ala Leu Phe
20 25 30

Ile Leu Ile Ala Ile Val Ile Thr Gln Ile Phe Met Leu Thr Phe Asp
35 40 45

His Thr Leu Leu Thr Thr Thr Gln Phe Leu Thr Ser Ile Gly Asn Leu
50 55 60

Gln Ser Thr Ile Thr Ser Trp Thr Pro Asp Val Gln Ala Met Leu Ser
65 70 75 80

Ile Ser Asn Gln Leu Ile Tyr Thr Thr Ser Ile Thr Leu Pro Leu Lys
85 90 95

Ile Ser Thr Thr Glu Met Ser Ile Leu Thr Ala Ile Arg Asp His Cys
100 105 110

His Cys Pro Asp Cys Ser Ser Ala Cys Pro Thr Arg Gln Met Leu Leu
115 120 125

Asn Asp Pro Arg Tyr Met Ser Gly Val Asn Gln Phe Ile Gly Ala Pro
130 135 140

Thr Glu Ser Ile Asn Ile Thr Phe Gly Pro Leu Phe Gly Ile Pro Ser
145 150 155 160

```
Phe Ile Pro Thr Ser Thr Thr Thr Gln Gly Cys Thr Arg Ile Pro Ser
                165                 170                 175

Phe Ala Leu Gly Pro Ser His Trp Cys Tyr Thr His Asn Phe Ile Thr
            180                 185                 190

Ala Gly Cys Ala Asp Gly Gly His Ser Asn Gln Tyr Leu Ala Met Gly
        195                 200                 205

Thr Ile Gln Ser Ala Ser Asp Gly Ser Pro Leu Leu Ile Thr Ala Arg
    210                 215                 220

Ser Tyr Tyr Leu Ser Asp Gly Val Asn Arg Lys Ser Cys Ser Ile Ala
225                 230                 235                 240

Val Val Pro Gly Gly Cys Ala Met Tyr Cys Tyr Val Ala Thr Arg Ser
                245                 250                 255

Glu Thr Asp Tyr Tyr Ala Gly Asn Ser Pro Pro Gln Gln Leu Leu Thr
            260                 265                 270

Leu Val Phe Ser Asn Asp Thr Ile Ile Glu Arg Thr Ile His Pro Thr
        275                 280                 285

Gly Leu Ala Asn Gly Trp Val Met Leu Val Pro Gly Val Gly Ser Gly
    290                 295                 300

Thr Leu Tyr Asn Glu Tyr Leu Leu Phe Pro Ala Tyr Gly Gly Met Gln
305                 310                 315                 320

Gln Ile Leu Ala Asn Gln Ser Gly Glu Ile Asn Gln Phe Phe Thr Pro
                325                 330                 335

Tyr Asn Ala Thr Val Arg Cys Ala Met Ala Gln Pro Gln Phe Ser Gln
            340                 345                 350

Arg Ala Ala Ala Ser Tyr Tyr Pro Arg Tyr Phe Ser Asn Arg Trp Ile
        355                 360                 365

Arg Ser Ala Ile Val Ala Cys Pro Tyr Arg Ala Ile Tyr Gln Thr Gln
    370                 375                 380

Cys Thr Leu Ile Pro Leu Pro Asn Arg Met Val Met Met Gly Ser Glu
385                 390                 395                 400

Gly Arg Ile Phe Thr Leu Gly Asp Arg Leu Phe Tyr Tyr Gln Arg Ser
                405                 410                 415

Ser Ser Trp Trp Pro Tyr Pro Leu Leu Tyr Gln Val Gly Leu Asn Phe
            420                 425                 430
```

```
Leu Thr Thr Pro Pro Ser Val Ser Ser Met Thr Gln Val Pro Leu Glu
        435                 440                 445


His Leu Ala Arg Pro Gly Lys Gly Gly Cys Pro Gly Asn Ser His Cys
    450                 455                 460


Pro Ala Thr Cys Val Thr Gly Val Tyr Ala Asp Val Trp Pro Leu Thr
465                 470                 475                 480


Asp Pro Arg Ser Gly Val Gly Gly Thr Ser Leu Val Ala Ala Gly Gly
                485                 490                 495


Leu Asp Ser Thr Ser Glu Arg Met Ala Pro Val Asn Tyr Leu Ala Ile
            500                 505                 510


Gly Glu Ser Leu Leu Ser Lys Thr Tyr Leu Leu Ser Lys Thr Gln Pro
        515                 520                 525


Ala Ala Tyr Thr Thr Thr Thr Cys Phe Arg Asp Thr Asp Thr Gly Lys
    530                 535                 540


Ile Tyr Cys Ile Thr Ile Ala Glu Leu Gly Lys Val Leu Leu Gly Glu
545                 550                 555                 560


Phe Gln Ile Val Pro Phe Leu Arg Glu Ile Lys Ile Gln Ser Arg Tyr
                565                 570                 575


<210>  81
<211>  580
<212>  PRT
<213>  Avian paramyxovirus 2


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase

<400>  81

Met Asp Phe Pro Ser Arg Glu Asn Leu Ala Ala Gly Asp Ile Ser Gly
1               5                   10                  15


Arg Lys Thr Trp Arg Leu Leu Phe Arg Ile Leu Thr Leu Ser Ile Gly
            20                  25                  30


Val Val Cys Leu Ala Ile Asn Ile Ala Thr Ile Ala Lys Leu Asp His
        35                  40                  45


Leu Asp Asn Met Ala Ser Asn Thr Trp Thr Thr Thr Glu Ala Asp Arg
    50                  55                  60


Val Ile Ser Ser Ile Thr Thr Pro Leu Lys Val Pro Val Asn Gln Ile
65                  70                  75                  80
```

```
Asn Asp Met Phe Arg Ile Val Ala Leu Asp Leu Pro Leu Gln Met Thr
                85                  90                  95

Ser Leu Gln Lys Glu Ile Thr Ser Gln Val Gly Phe Leu Ala Glu Ser
            100                 105                 110

Ile Asn Asn Val Leu Ser Lys Asn Gly Ser Ala Gly Leu Val Leu Val
        115                 120                 125

Asn Asp Pro Glu Tyr Ala Gly Gly Ile Ala Val Ser Leu Tyr Gln Gly
    130                 135                 140

Asp Ala Ser Ala Gly Leu Asn Phe Gln Pro Ile Ser Leu Ile Glu His
145                 150                 155                 160

Pro Ser Phe Val Pro Gly Pro Thr Thr Ala Lys Gly Cys Ile Arg Ile
                165                 170                 175

Pro Thr Phe His Met Gly Pro Ser His Trp Cys Tyr Ser His Asn Ile
            180                 185                 190

Ile Ala Ser Gly Cys Gln Asp Ala Ser His Ser Ser Met Tyr Ile Ser
        195                 200                 205

Leu Gly Val Leu Lys Ala Ser Gln Thr Gly Ser Pro Ile Phe Leu Thr
    210                 215                 220

Thr Ala Ser His Leu Val Asp Asp Asn Ile Asn Arg Lys Ser Cys Ser
225                 230                 235                 240

Ile Val Ala Ser Lys Tyr Gly Cys Asp Ile Leu Cys Ser Ile Val Ile
                245                 250                 255

Glu Thr Glu Asn Glu Asp Tyr Arg Ser Asp Pro Ala Thr Ser Met Ile
            260                 265                 270

Ile Gly Arg Leu Phe Phe Asn Gly Ser Tyr Thr Glu Ser Lys Ile Asn
        275                 280                 285

Thr Gly Ser Ile Phe Ser Leu Phe Ser Ala Asn Tyr Pro Ala Val Gly
    290                 295                 300

Ser Gly Ile Val Val Gly Asp Glu Ala Ala Phe Pro Ile Tyr Gly Gly
305                 310                 315                 320

Val Lys Gln Asn Thr Trp Leu Phe Asn Gln Leu Lys Asp Phe Gly Tyr
                325                 330                 335

Phe Thr His Asn Asp Val Tyr Lys Cys Asn Arg Thr Asp Ile Gln Gln
            340                 345                 350

Thr Ile Leu Asp Ala Tyr Arg Pro Pro Lys Ile Ser Gly Arg Leu Trp
```

```
        355                 360                 365

Val Gln Gly Ile Leu Leu Cys Pro Val Ser Leu Arg Pro Asp Pro Gly
    370                 375                 380

Cys Arg Leu Lys Val Phe Asn Thr Ser Asn Val Met Met Gly Ala Glu
385                 390                 395                 400

Ala Arg Leu Ile Gln Val Gly Ser Thr Val Tyr Leu Tyr Gln Arg Ser
                405                 410                 415

Ser Ser Trp Trp Val Val Gly Leu Thr Tyr Lys Leu Asp Val Ser Glu
            420                 425                 430

Ile Thr Ser Gln Thr Gly Asn Thr Leu Asn His Val Asp Pro Ile Ala
        435                 440                 445

His Thr Lys Phe Pro Arg Pro Ser Phe Arg Arg Asp Ala Cys Ala Arg
    450                 455                 460

Pro Asn Ile Cys Pro Ala Val Cys Val Ser Gly Val Tyr Gln Asp Ile
465                 470                 475                 480

Trp Pro Ile Ser Thr Ala Thr Asn Asn Ser Asn Ile Val Trp Val Gly
                485                 490                 495

Gln Tyr Leu Glu Ala Phe Tyr Ser Arg Lys Asp Pro Arg Ile Gly Ile
            500                 505                 510

Ala Thr Gln Tyr Glu Trp Lys Val Thr Asn Gln Leu Phe Asn Ser Asn
        515                 520                 525

Thr Glu Gly Gly Tyr Ser Thr Thr Thr Cys Phe Arg Asn Thr Lys Arg
    530                 535                 540

Asp Lys Ala Tyr Cys Val Val Ile Ser Glu Tyr Ala Asp Gly Val Phe
545                 550                 555                 560

Gly Ser Tyr Arg Ile Val Pro Gln Leu Ile Glu Ile Arg Thr Thr Thr
                565                 570                 575

Gly Lys Ser Glu
            580
```

<210> 82
<211> 236
<212> PRT
<213> Human metapneumovirus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein G

<400> 82

```
Met Glu Val Lys Val Glu Asn Ile Arg Thr Ile Asp Met Leu Lys Ala
1               5                   10                  15

Arg Val Lys Asn Arg Val Ala Arg Ser Lys Cys Phe Lys Asn Ala Ser
            20                  25                  30

Leu Ile Leu Ile Gly Ile Thr Thr Leu Ser Ile Ala Leu Asn Ile Tyr
        35                  40                  45

Leu Ile Ile Asn Tyr Thr Met Gln Glu Asn Thr Ser Glu Ser Glu His
    50                  55                  60

His Thr Ser Ser Ser Pro Met Glu Ser Ser Arg Glu Thr Pro Thr Val
65                  70                  75                  80

Pro Ile Asp Asn Ser Asp Thr Asn Pro Ser Ser Gln Tyr Pro Thr Gln
                85                  90                  95

Gln Ser Thr Glu Gly Ser Thr Leu Tyr Phe Ala Ala Ser Ala Ser Ser
            100                 105                 110

Pro Glu Thr Glu Pro Thr Ser Thr Pro Asp Thr Thr Ser Arg Pro Pro
        115                 120                 125

Phe Val Asp Thr His Thr Thr Pro Pro Ser Ala Ser Arg Thr Lys Thr
    130                 135                 140

Ser Pro Ala Val His Thr Lys Asn Asn Pro Arg Ile Ser Ser Arg Thr
145                 150                 155                 160

His Ser Pro Pro Trp Ala Met Thr Arg Thr Val Arg Arg Thr Thr Thr
                165                 170                 175

Leu Arg Thr Ser Ser Ile Arg Lys Arg Ser Ser Thr Ala Ser Val Gln
            180                 185                 190

Pro Asp Ser Ser Ala Thr Thr His Lys His Glu Glu Ala Ser Pro Val
        195                 200                 205

Ser Pro Gln Thr Ser Ala Ser Thr Thr Arg Pro Gln Arg Lys Ser Met
    210                 215                 220

Glu Ala Ser Thr Ser Thr Thr Tyr Asn Gln Thr Ser
225                 230                 235
```

<210> 83
<211> 414
<212> PRT
<213> Canine pneumovirus

<220>

<221> MISC_FEATURE
<223> attachment protein

<400> 83

```
Met Arg Pro Ala Glu Gln Leu Ile Gln Glu Asn Tyr Lys Leu Thr Ser
1               5                   10                  15


Leu Ser Met Gly Arg Asn Phe Glu Val Ser Gly Ser Thr Thr Asn Leu
            20                  25                  30


Asn Phe Glu Arg Thr Gln Tyr Pro Asp Thr Phe Arg Ala Val Val Lys
        35                  40                  45


Val Asn Gln Met Cys Lys Leu Ile Ala Gly Val Leu Thr Ser Ala Ala
    50                  55                  60


Val Ala Val Cys Val Gly Val Ile Met Tyr Ser Val Phe Thr Ser Asn
65                  70                  75                  80


His Lys Ala Asn Ser Met Gln Asn Ala Thr Ile Arg Asn Ser Thr Ser
                85                  90                  95


Ala Pro Pro Gln Pro Thr Ala Gly Pro Pro Thr Thr Glu Gln Gly Thr
            100                 105                 110


Thr Pro Lys Phe Thr Lys Pro Pro Thr Lys Thr Thr Thr His His Glu
        115                 120                 125


Ile Thr Glu Pro Ala Lys Met Val Thr Pro Ser Glu Asp Pro Tyr Gln
    130                 135                 140


Cys Ser Ser Asn Gly Tyr Leu Asp Arg Pro Asp Leu Pro Glu Asp Phe
145                 150                 155                 160


Lys Leu Val Leu Asp Val Ile Cys Lys Pro Pro Gly Pro Glu His His
                165                 170                 175


Ser Thr Asn Cys Tyr Glu Lys Arg Glu Ile Asn Leu Gly Ser Val Cys
            180                 185                 190


Pro Asp Leu Val Thr Met Lys Ala Asn Met Gly Leu Asn Asn Gly Gly
        195                 200                 205


Gly Glu Glu Ala Ala Pro Tyr Ile Glu Val Ile Thr Leu Ser Thr Tyr
    210                 215                 220


Ser Asn Lys Arg Ala Met Cys Val His Asn Gly Cys Asp Gln Gly Phe
225                 230                 235                 240


Cys Phe Phe Leu Ser Gly Leu Ser Thr Asp Gln Lys Arg Ala Val Leu
                245                 250                 255
```

Glu Leu Gly Gly Gln Gln Ala Ile Met Glu Leu His Tyr Asp Ser Tyr
260 265 270

Trp Lys His Tyr Trp Ser Asn Ser Asn Cys Val Val Pro Arg Thr Asn
275 280 285

Cys Asn Leu Thr Asp Gln Thr Val Ile Leu Phe Pro Ser Phe Asn Asn
290 295 300

Lys Asn Gln Ser Gln Cys Thr Thr Cys Ala Asp Ser Ala Gly Leu Asp
305 310 315 320

Asn Lys Phe Tyr Leu Thr Cys Asp Gly Leu Ser Arg Asn Leu Pro Leu
325 330 335

Val Gly Leu Pro Ser Leu Ser Pro Gln Ala His Lys Ala Ala Leu Lys
340 345 350

Gln Ser Thr Gly Thr Thr Thr Ala Pro Thr Pro Glu Thr Arg Asn Pro
355 360 365

Thr Pro Ala Pro Arg Arg Ser Lys Pro Leu Ser Arg Lys Lys Arg Ala
370 375 380

Leu Cys Gly Val Asp Ser Ser Arg Glu Pro Lys Pro Thr Met Pro Tyr
385 390 395 400

Trp Cys Pro Met Leu Gln Leu Phe Pro Arg Arg Ser Asn Ser
405 410

<210> 84
<211> 228
<212> PRT
<213> Respiratory syncytial virus type A

<220>
<221> MISC_FEATURE
<223> truncated attachment glycoprotein

<400> 84

Met Ser Lys Thr Lys Asp Gln Arg Ala Ala Lys Thr Leu Glu Lys Thr
1 5 10 15

Trp Asp Thr Leu Asn His Leu Leu Phe Ile Ser Ser Cys Leu Tyr Lys
20 25 30

Ser Asn Leu Lys Ser Ile Ala Gln Ile Thr Leu Ser Ile Leu Ala Met
35 40 45

Thr Ile Pro Thr Ser Leu Ile Ile Val Ala Thr Thr Phe Ile Ala Ser
50 55 60

Ala Asn Asn Lys Val Thr Pro Thr Thr Ala Ile Ile Gln Asp Ala Thr

65 70 75 80

Ser Gln Ile Lys Asn Thr Thr Pro Thr His Leu Thr Gln Asn Pro Gln
85 90 95

Pro Gly Ile Ser Phe Phe Asn Leu Ser Gly Thr Ile Ser Gln Thr Thr
100 105 110

Ala Ile Leu Ala Pro Thr Thr Pro Ser Val Glu Pro Ile Leu Gln Ser
115 120 125

Thr Thr Val Lys Thr Lys Asn Thr Thr Thr Thr Gln Ile Gln Pro Ser
130 135 140

Lys Leu Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Asn Lys Pro Asn
145 150 155 160

Asp Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys
165 170 175

Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Ser Lys
180 185 190

Lys Pro Gly Lys Lys Thr Thr Thr Lys Pro Thr Lys Lys Gln Thr Ile
195 200 205

Lys Thr Thr Lys Lys Asp Leu Lys Pro Gln Thr Thr Lys Pro Lys Glu
210 215 220

Ala Pro Thr Thr
225

<210> 85
<211> 577
<212> PRT
<213> Avian paramyxovirus 8

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 85

Met Ser Asn Ile Ala Ser Ser Leu Glu Asn Ile Val Glu Gln Asp Ser
1 5 10 15

Arg Lys Thr Thr Trp Arg Ala Ile Phe Arg Trp Ser Val Leu Leu Ile
20 25 30

Thr Thr Gly Cys Leu Ala Leu Ser Ile Val Ser Ile Val Gln Ile Gly
35 40 45

Asn Leu Lys Ile Pro Ser Val Gly Asp Leu Ala Asp Glu Val Val Thr
50 55 60

```
Pro Leu Lys Thr Thr Leu Ser Asp Thr Leu Arg Asn Pro Ile Asn Gln
65                  70                  75                  80

Ile Asn Asp Ile Phe Arg Ile Val Ala Leu Asp Ile Pro Leu Gln Val
                85                  90                  95

Thr Ser Ile Gln Lys Asp Leu Ala Ser Gln Phe Ser Met Leu Ile Asp
            100                 105                 110

Ser Leu Asn Ala Ile Lys Leu Gly Asn Gly Thr Asn Leu Ile Ile Pro
        115                 120                 125

Thr Ser Asp Lys Glu Tyr Ala Gly Gly Ile Gly Asn Pro Val Phe Thr
    130                 135                 140

Val Asp Ala Gly Gly Ser Ile Gly Phe Lys Gln Phe Ser Leu Ile Glu
145                 150                 155                 160

His Pro Ser Phe Ile Ala Gly Pro Thr Thr Thr Arg Gly Cys Thr Arg
                165                 170                 175

Ile Pro Thr Phe His Met Ser Glu Ser His Trp Cys Tyr Ser His Asn
            180                 185                 190

Ile Ile Ala Ala Gly Cys Gln Asp Ala Ser Ala Ser Ser Met Tyr Ile
        195                 200                 205

Ser Met Gly Val Leu His Val Ser Ser Ser Gly Thr Pro Ile Phe Leu
    210                 215                 220

Thr Thr Ala Ser Glu Leu Ile Asp Asp Gly Val Asn Arg Lys Ser Cys
225                 230                 235                 240

Ser Ile Val Ala Thr Gln Phe Gly Cys Asp Ile Leu Cys Ser Ile Val
                245                 250                 255

Ile Glu Lys Glu Gly Asp Asp Tyr Trp Ser Asp Thr Pro Thr Pro Met
            260                 265                 270

Arg His Gly Arg Phe Ser Phe Asn Gly Ser Phe Val Glu Thr Glu Leu
        275                 280                 285

Pro Val Ser Ser Met Phe Ser Ser Phe Ser Ala Asn Tyr Pro Ala Val
    290                 295                 300

Gly Ser Gly Glu Ile Val Lys Asp Arg Ile Leu Phe Pro Ile Tyr Gly
305                 310                 315                 320

Gly Ile Lys Gln Thr Ser Pro Glu Phe Thr Glu Leu Val Lys Tyr Gly
                325                 330                 335
```

Leu Phe Val Ser Thr Pro Thr Thr Val Cys Gln Ser Ser Trp Thr Tyr
340 345 350

Asp Gln Val Lys Ala Ala Tyr Arg Pro Asp Tyr Ile Ser Gly Arg Phe
355 360 365

Trp Ala Gln Val Ile Leu Ser Cys Ala Leu Asp Ala Val Asp Leu Ser
370 375 380

Ser Cys Ile Val Lys Ile Met Asn Ser Ser Thr Val Met Met Ala Ala
385 390 395 400

Glu Gly Arg Ile Ile Lys Ile Gly Ile Asp Tyr Phe Tyr Tyr Gln Arg
405 410 415

Ser Ser Ser Trp Trp Pro Leu Ala Phe Val Thr Lys Leu Asp Pro Gln
420 425 430

Glu Leu Ala Asp Thr Asn Ser Ile Trp Leu Thr Asn Ser Ile Pro Ile
435 440 445

Pro Gln Ser Lys Phe Pro Arg Pro Ser Tyr Ser Glu Asn Tyr Cys Thr
450 455 460

Lys Pro Ala Val Cys Pro Ala Thr Cys Val Thr Gly Val Tyr Ser Asp
465 470 475 480

Ile Trp Pro Leu Thr Ser Ser Ser Ser Leu Pro Ser Ile Ile Trp Ile
485 490 495

Gly Gln Tyr Leu Asp Ala Pro Val Gly Arg Thr Tyr Pro Arg Phe Gly
500 505 510

Ile Ala Asn Gln Ser His Trp Tyr Leu Gln Glu Asp Ile Leu Pro Thr
515 520 525

Ser Thr Ala Ser Ala Tyr Ser Thr Thr Thr Cys Phe Lys Asn Thr Ala
530 535 540

Arg Asn Arg Val Phe Cys Val Thr Ile Ala Glu Phe Ala Asp Gly Leu
545 550 555 560

Phe Gly Glu Tyr Arg Ile Thr Pro Gln Leu Tyr Glu Leu Val Arg Asn
565 570 575

Asn

<210> 86
<211> 576
<212> PRT
<213> Porcine parainfluenza virus 1

<220>
<221> MISC_FEATURE
<223> hemagglutinin protein

<400> 86

```
Met Glu Glu Thr Lys Val Lys Thr Ser Glu Tyr Trp Ala Arg Ser Pro
1               5                   10                  15


Gln Ile His Ala Thr Asn His Pro Asn Val Gln Asn Arg Glu Lys Ile
            20                  25                  30


Lys Glu Ile Leu Thr Ile Leu Ile Ser Phe Ile Ser Ser Leu Ser Leu
        35                  40                  45


Val Leu Val Ile Ala Val Leu Ile Met Gln Ser Leu His Asn Gly Thr
    50                  55                  60


Ile Leu Arg Cys Lys Asp Val Gly Leu Glu Ser Ile Asn Lys Ser Thr
65                  70                  75                  80


Tyr Ser Ile Ser Asn Ala Ile Leu Asp Val Ile Lys Gln Glu Leu Ile
                85                  90                  95


Thr Arg Ile Ile Asn Thr Gln Ser Ser Val Gln Val Ala Leu Pro Ile
            100                 105                 110


Leu Ile Asn Lys Lys Ile Gln Asp Leu Ser Leu Ile Ile Glu Lys Ser
        115                 120                 125


Ser Lys Val His Gln Asn Ser Pro Thr Cys Ser Gly Val Ala Ala Leu
    130                 135                 140


Thr His Val Glu Gly Ile Lys Pro Leu Asp Pro Asp Asp Tyr Trp Arg
145                 150                 155                 160


Cys Pro Ser Gly Glu Pro Tyr Leu Glu Asp Glu Leu Thr Leu Ser Leu
                165                 170                 175


Ile Pro Gly Pro Ser Met Leu Ala Gly Thr Ser Thr Ile Asp Gly Cys
            180                 185                 190


Val Arg Leu Pro Ser Leu Ala Ile Gly Lys Ser Leu Tyr Ala Tyr Ser
        195                 200                 205


Ser Asn Leu Ile Thr Lys Gly Cys Gln Asp Ile Gly Lys Ser Tyr Gln
    210                 215                 220


Val Leu Gln Leu Gly Ile Ile Thr Leu Asn Ser Asp Leu His Pro Asp
225                 230                 235                 240


Leu Asn Pro Ile Ile Ser His Thr Tyr Asp Ile Asn Asp Asn Arg Lys
```

```
                245                 250                 255

Ser Cys Ser Val Ala Val Ser Glu Thr Lys Gly Tyr Gln Leu Cys Ser
            260                 265                 270

Met Pro Arg Val Asn Glu Lys Thr Asp Tyr Thr Ser Asp Gly Ile Glu
        275                 280                 285

Asp Ile Val Phe Asp Val Leu Asp Leu Lys Gly Ser Ser Arg Ser Phe
    290                 295                 300

Lys Phe Ser Asn Asn Asp Ile Asn Phe Asp His Pro Phe Ser Ala Leu
305                 310                 315                 320

Tyr Pro Ser Val Gly Ser Gly Ile Ile Trp Lys Asn Glu Leu Tyr Phe
                325                 330                 335

Leu Gly Tyr Gly Ala Leu Thr Thr Ala Leu Gln Gly Asn Thr Lys Cys
            340                 345                 350

Asn Leu Met Gly Cys Pro Gly Ala Thr Gln Asp Asn Cys Asn Lys Phe
        355                 360                 365

Ile Ser Ser Ser Trp Leu Tyr Ser Lys Gln Met Val Asn Val Leu Ile
    370                 375                 380

Gln Val Lys Gly Tyr Leu Ser Ser Lys Pro Ser Ile Ile Val Arg Thr
385                 390                 395                 400

Ile Pro Ile Thr Glu Asn Tyr Val Gly Ala Glu Gly Lys Leu Val Gly
                405                 410                 415

Thr Arg Glu Arg Ile Tyr Ile Tyr Thr Arg Ser Thr Gly Trp His Thr
            420                 425                 430

Asn Leu Gln Ile Gly Val Leu Asn Ile Asn His Pro Ile Thr Ile Thr
        435                 440                 445

Trp Thr Asp His Arg Val Leu Ser Arg Pro Gly Arg Ser Pro Cys Ala
    450                 455                 460

Trp Asn Asn Lys Cys Pro Arg Asn Cys Thr Thr Gly Val Tyr Thr Asp
465                 470                 475                 480

Ala Tyr Pro Ile Ser Pro Asp Ala Asn Tyr Val Ala Thr Val Thr Leu
                485                 490                 495

Leu Ser Asn Ser Thr Arg Asn Asn Pro Thr Ile Met Tyr Ser Ser Ser
            500                 505                 510

Asp Arg Val Tyr Asn Met Leu Arg Leu Arg Asn Thr Glu Leu Glu Ala
        515                 520                 525
```

```
Ala Tyr Thr Thr Thr Ser Cys Ile Val His Phe Asp Arg Gly Tyr Cys
    530                 535                 540


Phe His Ile Ile Glu Ile Asn Gln Lys Glu Leu Asn Thr Leu Gln Pro
545                 550                 555                 560


Met Leu Phe Lys Thr Ala Ile Pro Lys Ala Cys Arg Ile Ser Asn Leu
                565                 570                 575


<210>  87
<211>  579
<212>  PRT
<213>  Human parainfluenza virus 4b


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase protein

<400>  87

Met Gln Asp Ser Arg Gly Asn Thr Gln Ile Phe Ser Gln Ala Asn Ser
1               5                   10                  15


Met Val Lys Arg Thr Trp Arg Leu Leu Phe Arg Ile Val Thr Leu Ile
            20                  25                  30


Leu Leu Ile Ser Ile Phe Val Leu Ser Leu Ile Ile Val Leu Gln Ser
        35                  40                  45


Thr Pro Gly Asn Leu Gln Ser Asp Val Asp Ile Ile Arg Lys Glu Leu
    50                  55                  60


Asp Glu Leu Met Glu Asn Phe Glu Thr Thr Ser Lys Ser Leu Leu Ser
65                  70                  75                  80


Val Ala Asn Gln Ile Thr Tyr Asp Val Ser Val Leu Thr Pro Ile Arg
                85                  90                  95


Gln Glu Ala Thr Glu Thr Asn Ile Ile Ala Lys Ile Lys Asp His Cys
            100                 105                 110


Lys Asp Arg Val Val Lys Gly Glu Ser Thr Cys Thr Leu Gly His Lys
        115                 120                 125


Pro Leu His Asp Val Ser Phe Leu Asn Gly Phe Asn Lys Phe Tyr Phe
    130                 135                 140


Thr Tyr Arg Asp Asn Val Gln Ile Arg Leu Asn Pro Leu Leu Asp Tyr
145                 150                 155                 160


Pro Asn Phe Ile Pro Thr Ala Thr Thr Pro His Gly Cys Ile Arg Ile
                165                 170                 175
```

```
Pro Ser Phe Ser Leu Ser Gln Thr His Trp Cys Tyr Thr His Asn Thr
            180                 185                 190

Ile Leu Arg Gly Cys Glu Asp Thr Ala Ser Ser Lys Gln Tyr Val Ser
        195                 200                 205

Leu Gly Thr Leu Gln Thr Leu Glu Asn Gly Asp Pro Tyr Phe Lys Val
    210                 215                 220

Glu Tyr Ser His Tyr Leu Asn Asp Arg Lys Asn Arg Lys Ser Cys Ser
225                 230                 235                 240

Val Val Ala Val Leu Asp Gly Cys Leu Leu Tyr Cys Val Ile Met Thr
                245                 250                 255

Lys Asn Glu Thr Glu Asn Phe Lys Asp Pro Gln Leu Ala Thr Gln Leu
            260                 265                 270

Leu Thr Tyr Ile Ser Tyr Asn Gly Thr Ile Lys Glu Arg Ile Ile Asn
        275                 280                 285

Pro Pro Gly Ser Ser Arg Asp Trp Val His Ile Ser Pro Gly Val Gly
    290                 295                 300

Ser Gly Ile Leu Tyr Ser Asn Tyr Ile Ile Phe Pro Leu Tyr Gly Gly
305                 310                 315                 320

Leu Met Glu Asn Ser Met Ile Tyr Asn Asn Gln Ser Gly Lys Tyr Phe
                325                 330                 335

Phe Pro Asn Ser Thr Lys Leu Pro Cys Ser Asn Lys Thr Ser Glu Lys
            340                 345                 350

Ile Thr Gly Ala Lys Asp Ser Tyr Thr Ile Thr Tyr Phe Ser Lys Arg
        355                 360                 365

Leu Ile Gln Ser Ala Phe Leu Ile Cys Asp Leu Arg Gln Phe Leu Ser
    370                 375                 380

Glu Asp Cys Glu Ile Leu Ile Pro Ser Asn Asp His Met Leu Val Gly
385                 390                 395                 400

Ala Glu Gly Arg Leu Tyr Asn Ile Glu Asn Asn Ile Phe Tyr Tyr Gln
                405                 410                 415

Arg Gly Ser Ser Trp Trp Pro Tyr Pro Ser Leu Tyr Arg Ile Lys Leu
            420                 425                 430

Asn Ser Asn Lys Lys Tyr Pro Arg Ile Ile Glu Ile Lys Phe Thr Lys
        435                 440                 445
```

Ile Glu Ile Ala Pro Arg Pro Gly Asn Lys Asp Cys Pro Gly Asn Lys
450 455 460

Ala Cys Pro Lys Glu Cys Ile Thr Gly Val Tyr Gln Asp Ile Trp Pro
465 470 475 480

Leu Ser Tyr Pro Asn Thr Ala Phe Pro His Lys Lys Arg Ala Tyr Tyr
485 490 495

Thr Gly Phe Tyr Leu Asn Asn Ser Leu Ala Arg Arg Asn Pro Thr Phe
500 505 510

Tyr Thr Ala Asp Asn Leu Asp Tyr His Gln Gln Glu Arg Leu Gly Lys
515 520 525

Phe Asn Leu Thr Ala Gly Tyr Ser Thr Thr Thr Cys Phe Lys Gln Thr
530 535 540

Thr Thr Ala Arg Leu Tyr Cys Leu Tyr Ile Leu Glu Val Gly Asp Ser
545 550 555 560

Val Ile Gly Asp Phe Gln Ile Phe Pro Phe Leu Arg Ser Ile Asp Gln
565 570 575

Ala Ile Thr

<210> 88
<211> 582
<212> PRT
<213> Avian paramyxovirus 2

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 88

Met Asp Ala Leu Ser Arg Glu Asn Leu Thr Glu Ile Ser Gln Gly Gly
1 5 10 15

Arg Arg Thr Trp Arg Met Leu Phe Arg Ile Leu Thr Leu Val Leu Thr
20 25 30

Leu Val Cys Leu Ala Ile Asn Ile Ala Thr Ile Ala Lys Leu Asp Ser
35 40 45

Ile Asp Thr Ser Lys Val Gln Thr Trp Thr Thr Thr Glu Ser Asp Arg
50 55 60

Val Ile Gly Ser Leu Thr Asp Thr Leu Lys Ile Pro Ile Asn Gln Val
65 70 75 80

Asn Asp Met Phe Arg Ile Val Ala Leu Asp Leu Pro Leu Gln Met Thr

85 90 95

Thr Leu Gln Lys Glu Ile Ala Ser Gln Val Gly Phe Leu Ala Glu Ser
100 105 110

Ile Asn Asn Phe Leu Ser Lys Asn Gly Ser Ala Gly Ser Val Leu Val
115 120 125

Asn Asp Pro Glu Tyr Ala Gly Gly Ile Gly Thr Ser Leu Phe His Gly
130 135 140

Asp Ser Ala Ser Gly Leu Asp Phe Glu Ala Pro Ser Leu Ile Glu His
145 150 155 160

Pro Ser Phe Ile Pro Gly Pro Thr Thr Ala Lys Gly Cys Ile Arg Ile
165 170 175

Pro Thr Phe His Met Ser Ala Ser His Trp Cys Tyr Ser His Asn Ile
180 185 190

Ile Ala Ser Gly Cys Gln Asp Ala Gly His Ser Ser Met Tyr Ile Ser
195 200 205

Met Gly Val Leu Lys Ala Thr Gln Ala Gly Ser Pro Ser Phe Leu Thr
210 215 220

Thr Ala Ser Gln Leu Val Asp Asp Lys Leu Asn Arg Lys Ser Cys Ser
225 230 235 240

Ile Ile Ser Thr Thr Tyr Gly Cys Asp Ile Leu Cys Ser Leu Val Val
245 250 255

Glu Asn Glu Asp Ala Asp Tyr Arg Ser Asp Pro Pro Thr Asp Met Ile
260 265 270

Leu Gly Arg Leu Phe Phe Asn Gly Thr Tyr Ser Glu Ser Lys Leu Asn
275 280 285

Thr Ser Ala Ile Phe Gln Leu Phe Ser Ala Asn Tyr Pro Ala Val Gly
290 295 300

Ser Gly Ile Val Leu Gly Asp Glu Ile Ala Phe Pro Val Tyr Gly Gly
305 310 315 320

Val Lys Gln Asn Thr Trp Leu Phe Asn Gln Leu Lys Asp Tyr Gly Tyr
325 330 335

Phe Ala His Asn Asn Val Tyr Lys Cys Asn Asn Ser Asn Ile His Gln
340 345 350

Thr Val Leu Asn Ala Tyr Arg Pro Pro Lys Ile Ser Gly Arg Leu Trp
355 360 365

```
Ser Gln Val Val Leu Ile Cys Pro Met Arg Leu Phe Ile Asn Thr Asp
    370                 375                 380

Cys Arg Ile Lys Val Phe Asn Thr Ser Thr Val Met Met Gly Ala Glu
385                 390                 395                 400

Ala Arg Leu Ile Gln Val Gly Ser Asp Ile Tyr Leu Tyr Gln Arg Ser
                405                 410                 415

Ser Ser Trp Trp Val Val Gly Leu Thr Tyr Lys Leu Asp Phe Gln Glu
            420                 425                 430

Leu Ser Ser Lys Thr Gly Asn Ile Leu Asn Asn Val Ser Pro Ile Ala
        435                 440                 445

His Ala Lys Phe Pro Arg Pro Ser Tyr Ser Arg Asp Ala Cys Ala Arg
    450                 455                 460

Pro Asn Ile Cys Pro Ala Val Cys Val Ser Gly Val Tyr Gln Asp Ile
465                 470                 475                 480

Trp Pro Ile Ser Thr Ala His Asn Leu Ser Gln Val Val Trp Val Gly
                485                 490                 495

Gln Tyr Leu Glu Ala Phe Tyr Ala Arg Lys Asp Pro Trp Ile Gly Ile
            500                 505                 510

Ala Thr Gln Tyr Asp Trp Lys Lys Asn Val Arg Leu Phe Asn Ala Asn
        515                 520                 525

Thr Glu Gly Gly Tyr Ser Thr Thr Thr Cys Phe Arg Asn Thr Lys Arg
    530                 535                 540

Asp Lys Ala Phe Cys Val Ile Ile Ser Glu Tyr Ala Asp Gly Val Phe
545                 550                 555                 560

Gly Ser Tyr Arg Ile Val Pro Gln Leu Ile Glu Ile Arg Thr Thr Ser
                565                 570                 575

Lys Lys Gly Leu Pro Ser
            580
```

<210> 89
<211> 610
<212> PRT
<213> Avian paramyxovirus goose/Shimane/67/2000

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase

<400> 89

```
Met Gln Pro Gly Ile Ser Glu Val Ser Phe Val Asn Asp Glu Arg Ser
1               5                   10                  15

Glu Arg Gly Thr Trp Arg Leu Leu Phe Arg Ile Leu Thr Ile Val Leu
            20                  25                  30

Cys Leu Thr Ser Ile Gly Ile Gly Ile Pro Ala Leu Ile Tyr Ser Lys
        35                  40                  45

Glu Ala Ala Thr Ser Gly Asp Ile Asp Lys Ser Leu Glu Ala Val Lys
    50                  55                  60

Thr Gly Met Ser Thr Leu Ser Ser Lys Ile Asp Glu Ser Ile Asn Thr
65                  70                  75                  80

Glu Gln Lys Ile Tyr Arg Gln Val Ile Leu Glu Ala Pro Val Ser Gln
                85                  90                  95

Leu Asn Met Glu Ser Asn Ile Leu Ser Ala Ile Thr Ser Leu Ser Tyr
            100                 105                 110

Gln Ile Asp Gly Thr Ser Asn Ser Ser Gly Cys Gly Ser Pro Met His
        115                 120                 125

Asp Gln Asp Phe Val Gly Gly Ile Asn Lys Glu Ile Trp Thr Thr Asp
    130                 135                 140

Asn Val Asn Leu Gly Glu Ile Thr Leu Thr Pro Phe Leu Glu His Leu
145                 150                 155                 160

Asn Phe Ile Pro Ala Pro Thr Thr Gly Asn Gly Cys Thr Arg Ile Pro
                165                 170                 175

Ser Phe Asp Leu Gly Leu Thr His Trp Cys Tyr Thr His Asn Val Ile
            180                 185                 190

Leu Ser Gly Cys Gln Asp Tyr Ser Ser Ser Phe Gln Tyr Ile Ala Leu
        195                 200                 205

Gly Val Leu Lys Ile Ser Ala Thr Gly His Val Phe Leu Ser Thr Met
    210                 215                 220

Arg Ser Ile Asn Leu Asp Asp Glu Arg Asn Arg Lys Ser Cys Ser Ile
225                 230                 235                 240

Ser Ala Thr Ser Ile Gly Cys Asp Ile Ile Cys Ser Leu Val Thr Glu
                245                 250                 255

Arg Glu Val Asp Asp Tyr Asn Ser Pro Ala Ala Thr Pro Met Ile His
            260                 265                 270
```

```
Gly Arg Leu Asp Phe Ser Gly Lys Tyr Asn Glu Val Asp Leu Asn Val
        275                 280                 285

Gly Gln Leu Phe Gly Asp Trp Ser Ala Asn Tyr Pro Gly Val Gly Gly
    290                 295                 300

Gly Ser Phe Leu Asn Gly Arg Val Trp Phe Pro Ile Tyr Gly Gly Val
305                 310                 315                 320

Lys Glu Gly Thr Pro Thr Phe Lys Glu Asn Asp Gly Arg Tyr Ala Ile
                325                 330                 335

Tyr Thr Arg Tyr Asn Asp Thr Cys Pro Asp Ser Glu Ser Glu Gln Val
            340                 345                 350

Ser Arg Ala Lys Ser Ser Tyr Arg Pro Ser Tyr Phe Gly Gly Lys Leu
        355                 360                 365

Val Gln Gln Ala Val Leu Ser Ile Lys Ile Asp Asp Thr Leu Gly Leu
    370                 375                 380

Asp Pro Val Leu Thr Ile Ser Asn Asn Ser Ile Thr Leu Met Gly Ala
385                 390                 395                 400

Glu Ser Arg Val Leu Gln Ile Glu Glu Lys Leu Tyr Phe Tyr Gln Arg
                405                 410                 415

Gly Thr Ser Trp Phe Pro Ser Leu Ile Met Tyr Pro Leu Thr Val Asp
            420                 425                 430

Asp Lys Met Val Arg Phe Glu Pro Pro Thr Ile Phe Asp Gln Phe Thr
        435                 440                 445

Arg Pro Gly Asn His Pro Cys Ser Ala Asp Ser Arg Cys Pro Asn Ala
    450                 455                 460

Cys Val Thr Gly Val Tyr Thr Asp Gly Tyr Pro Ile Val Phe His Asn
465                 470                 475                 480

Asn His Ser Ile Ala Ala Val Tyr Gly Met Gln Leu Asn Asp Val Thr
                485                 490                 495

Asn Arg Leu Asn Pro Arg Ser Ala Val Trp Tyr Gly Val Ser Met Ser
            500                 505                 510

Asn Val Ile Arg Val Ser Ser Ser Thr Thr Lys Ala Ala Tyr Thr Thr
        515                 520                 525

Ser Thr Cys Phe Lys Val Lys Lys Thr Gln Arg Val Tyr Cys Leu Ser
    530                 535                 540

Ile Gly Glu Ile Gly Asn Thr Leu Phe Gly Glu Phe Arg Ile Val Pro
```

545 550 555 560

Leu Leu Leu Glu Val Tyr Ser Glu Lys Gly Lys Ser Leu Lys Ser Ser
565 570 575

Phe Asp Gly Trp Glu Asp Ile Ser Ile Asn Asn Pro Leu Arg Pro Leu
580 585 590

Asp Asn His Arg Val Asp Pro Ile Leu Ile Ser Asn Tyr Thr Ser Ser
595 600 605

Trp Pro
610

<210> 90
<211> 257
<212> PRT
<213> Bovine respiratory syncytial virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 90

Met Ser Asn His Thr His His Leu Lys Phe Lys Thr Leu Lys Arg Ala
1 5 10 15

Trp Lys Ala Ser Lys Tyr Phe Ile Val Gly Leu Ser Cys Leu Tyr Lys
20 25 30

Phe Asn Leu Lys Ser Leu Val Gln Thr Ala Leu Thr Thr Leu Ala Met
35 40 45

Ile Thr Leu Thr Ser Leu Val Ile Thr Ala Ile Ile Tyr Ile Ser Val
50 55 60

Gly Asn Ala Lys Ala Lys Pro Thr Ser Lys Pro Thr Ile Gln Gln Thr
65 70 75 80

Gln Gln Pro Gln Asn His Thr Ser Pro Phe Phe Thr Glu His Asn Tyr
85 90 95

Lys Ser Thr His Thr Ser Ile Gln Ser Thr Thr Leu Ser Gln Leu Pro
100 105 110

Asn Thr Asp Thr Thr Arg Glu Thr Thr Tyr Ser His Ser Ile Asn Glu
115 120 125

Thr Gln Asn Arg Lys Ile Lys Ser Gln Ser Thr Leu Pro Ala Thr Arg
130 135 140

Lys Pro Pro Ile Asn Pro Ser Gly Ser Asn Pro Pro Glu Asn His Gln
145 150 155 160

```
Asp His Asn Asn Ser Gln Thr Leu Pro Tyr Val Pro Cys Ser Thr Cys
                165                 170                 175


Glu Gly Asn Leu Ala Cys Leu Ser Leu Cys Gln Ile Gly Pro Glu Arg
            180                 185                 190


Ala Pro Ser Arg Ala Pro Thr Ile Thr Leu Lys Lys Thr Pro Lys Pro
        195                 200                 205


Lys Thr Thr Lys Lys Pro Thr Lys Thr Thr Ile His His Arg Thr Ser
    210                 215                 220


Pro Glu Ala Lys Leu Gln Pro Lys Asn Asn Thr Ala Ala Pro Gln Gln
225                 230                 235                 240


Gly Ile Leu Ser Ser Pro Glu His His Thr Asn Gln Ser Thr Thr Gln
                245                 250                 255


Ile



<210>  91
<211>  595
<212>  PRT
<213>  Menangle virus


<220>
<221>  MISC_FEATURE
<223>  attachemnt protein

<220>
<221>  MISC_FEATURE
<223>  attachement protein

<400>  91

Met Trp Asn Ser Ile Pro Gln Leu Val Ser Asp His Glu Glu Ala Lys
1               5                   10                  15


Gly Lys Phe Thr Asp Ile Pro Leu Gln Asp Asp Thr Asp Ser Gln His
            20                  25                  30


Pro Ser Gly Ser Lys Ser Thr Cys Arg Thr Leu Phe Arg Thr Val Ser
        35                  40                  45


Ile Ile Leu Ser Leu Val Ile Leu Val Leu Gly Val Thr Ser Thr Met
    50                  55                  60


Phe Ser Ala Lys Tyr Ser Gly Gly Cys Ala Thr Asn Ser Gln Leu Leu
65                  70                  75                  80


Gly Val Ser Asn Leu Ile Asn Gln Ile Gln Lys Ser Ile Asp Ser Leu
                85                  90                  95
```

```
Ile Ser Glu Val Asn Gln Val Ser Ile Thr Thr Ala Val Thr Leu Pro
            100                 105                 110

Ile Lys Ile Met Asp Phe Gly Lys Ser Val Thr Asp Gln Val Thr Gln
        115                 120                 125

Met Ile Arg Gln Cys Asn Thr Val Cys Lys Gly Pro Gly Gln Lys Pro
    130                 135                 140

Gly Ser Gln Asn Val Arg Ile Met Pro Ser Asn Asn Leu Ser Thr Phe
145                 150                 155                 160

Gln Asn Ile Asn Met Ser Ala Arg Gly Ile Ala Tyr Gln Asp Val Pro
                165                 170                 175

Leu Thr Phe Val Arg Pro Ile Lys Asn Pro Gln Ser Cys Ser Arg Phe
            180                 185                 190

Pro Ser Tyr Ser Val Ser Phe Gly Val His Cys Phe Ala Asn Ala Val
        195                 200                 205

Thr Asp Gln Thr Cys Glu Leu Asn Gln Asn Thr Phe Tyr Arg Val Val
    210                 215                 220

Leu Ser Val Ser Lys Gly Asn Ile Ser Asp Pro Ser Ser Leu Glu Thr
225                 230                 235                 240

Lys Ala Glu Thr Arg Thr Pro Lys Gly Thr Pro Val Arg Thr Cys Ser
                245                 250                 255

Ile Ile Ser Ser Val Tyr Gly Cys Tyr Leu Leu Cys Ser Lys Ala Thr
            260                 265                 270

Val Pro Glu Ser Glu Glu Met Lys Thr Ile Gly Phe Ser Gln Met Phe
        275                 280                 285

Ile Leu Tyr Leu Ser Met Asp Ser Lys Arg Ile Ile Tyr Asp Asn Ile
    290                 295                 300

Val Ser Ser Thr Ser Ala Ile Trp Ser Gly Leu Tyr Pro Gly Glu Gly
305                 310                 315                 320

Ala Gly Ile Trp His Met Gly Gln Leu Phe Phe Pro Leu Trp Gly Gly
                325                 330                 335

Ile Pro Phe Leu Thr Pro Leu Gly Gln Lys Ile Leu Asn Ser Thr Leu
            340                 345                 350

Asp Ile Pro Glu Val Gly Ser Lys Cys Lys Ser Asp Leu Thr Ser Asn
        355                 360                 365
```

```
Pro Ala Lys Thr Lys Asp Met Leu Phe Ser Pro Tyr Tyr Gly Glu Asn
    370                 375                 380


Val Met Val Phe Gly Phe Leu Thr Cys Tyr Leu Leu Ser Asn Val Pro
385                 390                 395                 400


Thr Asn Cys His Ala Asp Tyr Leu Asn Ser Thr Val Leu Gly Phe Gly
                405                 410                 415


Ser Lys Ala Gln Phe Tyr Asp Tyr Arg Gly Ile Val Tyr Met Tyr Ile
            420                 425                 430


Gln Ser Ala Gly Trp Tyr Pro Phe Thr Gln Ile Phe Arg Ile Thr Leu
        435                 440                 445


Gln Leu Lys Gln Asn Arg Leu Gln Ala Lys Ser Ile Lys Arg Ile Glu
    450                 455                 460


Val Thr Ser Thr Thr Arg Pro Gly Asn Arg Glu Cys Ser Val Leu Arg
465                 470                 475                 480


Asn Cys Pro Tyr Ile Cys Ala Thr Gly Leu Phe Gln Val Pro Trp Ile
                485                 490                 495


Val Asn Ser Asp Ala Ile Thr Ser Lys Glu Val Asp Asn Met Val Phe
            500                 505                 510


Val Gln Ala Trp Ala Ala Asp Phe Thr Glu Phe Arg Lys Gly Ile Leu
        515                 520                 525


Ser Leu Cys Ser Gln Val Ser Cys Pro Ile Asn Asp Leu Leu Ser Lys
    530                 535                 540


Asp Asn Ser Tyr Met Arg Asp Thr Thr Thr Tyr Cys Phe Pro Gln Thr
545                 550                 555                 560


Val Pro Asn Ile Leu Ser Cys Thr Ser Phe Val Glu Trp Gly Gly Asp
                565                 570                 575


Ser Gly Asn Pro Ile Asn Ile Leu Glu Ile His Tyr Glu Val Ile Phe
            580                 585                 590


Val Ala Ser
        595


<210>  92
<211>  574
<212>  PRT
<213>  avian paramyxovirus 5


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase protein
```

<400> 92

```
Met Asp Lys Ser Tyr Tyr Thr Glu Pro Glu Asp Gln Arg Gly Asn Ser
1               5                   10                  15


Arg Thr Trp Arg Leu Leu Phe Arg Leu Ile Val Leu Thr Leu Leu Cys
            20                  25                  30


Leu Ile Ala Cys Thr Ser Val Ser Gln Leu Phe Tyr Pro Trp Leu Pro
        35                  40                  45


Gln Val Leu Ser Thr Leu Ile Ser Leu Asn Ser Ser Ile Ile Thr Ser
    50                  55                  60


Ser Asn Gly Leu Lys Lys Glu Ile Leu Asn Gln Asn Ile Lys Glu Asp
65                  70                  75                  80


Leu Ile Tyr Arg Glu Val Ala Ile Asn Ile Pro Leu Thr Leu Asp Arg
                85                  90                  95


Val Thr Val Glu Val Gly Thr Ala Val Asn Gln Ile Thr Asp Ala Leu
            100                 105                 110


Arg Gln Leu Gln Ser Val Asn Gly Ser Ala Ala Phe Ala Leu Ser Asn
        115                 120                 125


Ser Pro Asp Tyr Ser Gly Gly Ile Glu His Leu Val Phe Gln Arg Asn
    130                 135                 140


Thr Leu Ile Asn Arg Ser Val Ser Val Ser Asp Leu Ile Glu His Pro
145                 150                 155                 160


Ser Phe Ile Pro Thr Pro Thr Thr Gln His Gly Cys Thr Arg Ile Pro
                165                 170                 175


Thr Phe His Leu Gly Thr Arg His Trp Cys Tyr Ser His Asn Ile Ile
            180                 185                 190


Gly Gln Gly Cys Ala Asp Ser Gly Ala Ser Met Met Tyr Ile Ser Met
        195                 200                 205


Gly Ala Leu Gly Val Ser Ser Leu Gly Thr Pro Thr Phe Thr Thr Ser
    210                 215                 220


Ala Thr Ser Ile Leu Ser Asp Ser Leu Asn Arg Lys Ser Cys Ser Ile
225                 230                 235                 240


Val Ala Thr Thr Glu Gly Cys Asp Val Leu Cys Ser Ile Val Thr Gln
                245                 250                 255


Thr Glu Asp Gln Asp Tyr Ala Asp His Thr Pro Thr Pro Met Ile His
            260                 265                 270
```

```
Gly Arg Leu Trp Phe Asn Gly Thr Tyr Thr Glu Arg Ser Leu Ser Gln
        275                 280                 285

Ser Leu Phe Leu Gly Thr Trp Ala Ala Gln Tyr Pro Ala Val Gly Ser
    290                 295                 300

Gly Ile Met Thr Pro Gly Arg Val Ile Phe Pro Phe Tyr Gly Gly Val
305                 310                 315                 320

Ile Pro Asn Ser Pro Leu Phe Leu Asp Leu Glu Arg Phe Ala Leu Phe
                325                 330                 335

Thr His Asn Gly Asp Leu Glu Cys Arg Asn Leu Thr Gln Tyr Gln Lys
            340                 345                 350

Glu Ala Ile Tyr Ser Ala Tyr Lys Pro Pro Lys Ile Arg Gly Ser Leu
        355                 360                 365

Trp Ala Gln Gly Phe Ile Val Cys Ser Val Gly Asp Met Gly Asn Cys
    370                 375                 380

Ser Leu Lys Val Ile Asn Thr Ser Thr Val Met Met Gly Ala Glu Gly
385                 390                 395                 400

Arg Leu Gln Leu Val Gly Asp Ser Val Met Tyr Tyr Gln Arg Ser Ser
                405                 410                 415

Ser Trp Trp Pro Val Gly Ile Leu Tyr Arg Leu Ser Leu Val Asp Ile
            420                 425                 430

Ile Ala Arg Asp Ile Gln Val Val Ile Asn Ser Glu Pro Leu Pro Leu
        435                 440                 445

Ser Lys Phe Pro Arg Pro Thr Trp Thr Pro Gly Val Cys Gln Lys Pro
    450                 455                 460

Asn Val Cys Pro Ala Val Cys Val Thr Gly Val Tyr Gln Asp Leu Trp
465                 470                 475                 480

Ala Ile Ser Ala Gly Glu Thr Leu Ser Glu Met Thr Phe Phe Gly Gly
                485                 490                 495

Tyr Leu Glu Ala Ser Thr Gln Arg Lys Asp Pro Trp Ile Gly Val Ala
            500                 505                 510

Asn Gln Tyr Ser Trp Phe Met Arg Arg Arg Leu Phe Lys Thr Ser Thr
        515                 520                 525

Glu Ala Ala Tyr Ser Ser Ser Thr Cys Phe Arg Asn Thr Arg Leu Asp
    530                 535                 540
```

```
Arg Asn Phe Cys Leu Leu Ile Phe Glu Leu Thr Asp Asn Leu Leu Gly
545                 550                 555                 560


Asp Trp Arg Ile Val Pro Leu Leu Phe Glu Leu Thr Ile Val
                565                 570


<210>  93
<211>  622
<212>  PRT
<213>  cedar virus


<220>
<221>  MISC_FEATURE
<223>  attachment glycoprotein

<400>  93

Met Leu Ser Gln Leu Gln Lys Asn Tyr Leu Asp Asn Ser Asn Gln Gln
1               5                   10                  15


Gly Asp Lys Met Asn Asn Pro Asp Lys Lys Leu Ser Val Asn Phe Asn
            20                  25                  30


Pro Leu Glu Leu Asp Lys Gly Gln Lys Asp Leu Asn Lys Ser Tyr Tyr
        35                  40                  45


Val Lys Asn Lys Asn Tyr Asn Val Ser Asn Leu Leu Asn Glu Ser Leu
    50                  55                  60


His Asp Ile Lys Phe Cys Ile Tyr Cys Ile Phe Ser Leu Leu Ile Ile
65                  70                  75                  80


Ile Thr Ile Ile Asn Ile Ile Thr Ile Ser Ile Val Ile Thr Arg Leu
                85                  90                  95


Lys Val His Glu Glu Asn Asn Gly Met Glu Ser Pro Asn Leu Gln Ser
            100                 105                 110


Ile Gln Asp Ser Leu Ser Ser Leu Thr Asn Met Ile Asn Thr Glu Ile
        115                 120                 125


Thr Pro Arg Ile Gly Ile Leu Val Thr Ala Thr Ser Val Thr Leu Ser
    130                 135                 140


Ser Ser Ile Asn Tyr Val Gly Thr Lys Thr Asn Gln Leu Val Asn Glu
145                 150                 155                 160


Leu Lys Asp Tyr Ile Thr Lys Ser Cys Gly Phe Lys Val Pro Glu Leu
                165                 170                 175


Lys Leu His Glu Cys Asn Ile Ser Cys Ala Asp Pro Lys Ile Ser Lys
            180                 185                 190
```

```
Ser Ala Met Tyr Ser Thr Asn Ala Tyr Ala Glu Leu Ala Gly Pro Pro
        195                 200                 205

Lys Ile Phe Cys Lys Ser Val Ser Lys Asp Pro Asp Phe Arg Leu Lys
    210                 215                 220

Gln Ile Asp Tyr Val Ile Pro Val Gln Gln Asp Arg Ser Ile Cys Met
225                 230                 235                 240

Asn Asn Pro Leu Leu Asp Ile Ser Asp Gly Phe Phe Thr Tyr Ile His
                245                 250                 255

Tyr Glu Gly Ile Asn Ser Cys Lys Lys Ser Asp Ser Phe Lys Val Leu
            260                 265                 270

Leu Ser His Gly Glu Ile Val Asp Arg Gly Asp Tyr Arg Pro Ser Leu
        275                 280                 285

Tyr Leu Leu Ser Ser His Tyr His Pro Tyr Ser Met Gln Val Ile Asn
    290                 295                 300

Cys Val Pro Val Thr Cys Asn Gln Ser Ser Phe Val Phe Cys His Ile
305                 310                 315                 320

Ser Asn Asn Thr Lys Thr Leu Asp Asn Ser Asp Tyr Ser Ser Asp Glu
                325                 330                 335

Tyr Tyr Ile Thr Tyr Phe Asn Gly Ile Asp Arg Pro Lys Thr Lys Lys
            340                 345                 350

Ile Pro Ile Asn Asn Met Thr Ala Asp Asn Arg Tyr Ile His Phe Thr
        355                 360                 365

Phe Ser Gly Gly Gly Gly Val Cys Leu Gly Glu Glu Phe Ile Ile Pro
    370                 375                 380

Val Thr Thr Val Ile Asn Thr Asp Val Phe Thr His Asp Tyr Cys Glu
385                 390                 395                 400

Ser Phe Asn Cys Ser Val Gln Thr Gly Lys Ser Leu Lys Glu Ile Cys
                405                 410                 415

Ser Glu Ser Leu Arg Ser Pro Thr Asn Ser Ser Arg Tyr Asn Leu Asn
            420                 425                 430

Gly Ile Met Ile Ile Ser Gln Asn Asn Met Thr Asp Phe Lys Ile Gln
        435                 440                 445

Leu Asn Gly Ile Thr Tyr Asn Lys Leu Ser Phe Gly Ser Pro Gly Arg
    450                 455                 460

Leu Ser Lys Thr Leu Gly Gln Val Leu Tyr Tyr Gln Ser Ser Met Ser
```

465 470 475 480

Trp Asp Thr Tyr Leu Lys Ala Gly Phe Val Glu Lys Trp Lys Pro Phe
485 490 495

Thr Pro Asn Trp Met Asn Asn Thr Val Ile Ser Arg Pro Asn Gln Gly
500 505 510

Asn Cys Pro Arg Tyr His Lys Cys Pro Glu Ile Cys Tyr Gly Gly Thr
515 520 525

Tyr Asn Asp Ile Ala Pro Leu Asp Leu Gly Lys Asp Met Tyr Val Ser
530 535 540

Val Ile Leu Asp Ser Asp Gln Leu Ala Glu Asn Pro Glu Ile Thr Val
545 550 555 560

Phe Asn Ser Thr Thr Ile Leu Tyr Lys Glu Arg Val Ser Lys Asp Glu
565 570 575

Leu Asn Thr Arg Ser Thr Thr Thr Ser Cys Phe Leu Phe Leu Asp Glu
580 585 590

Pro Trp Cys Ile Ser Val Leu Glu Thr Asn Arg Phe Asn Gly Lys Ser
595 600 605

Ile Arg Pro Glu Ile Tyr Ser Tyr Lys Ile Pro Lys Tyr Cys
610 615 620

<210> 94
<211> 595
<212> PRT
<213> teviot virus

<220>
<221> MISC_FEATURE
<223> attachment protein

<400> 94

Met Trp Ser Thr Gln Ala Ser Lys His Pro Ala Met Val Asn Ser Ala
1 5 10 15

Thr Asn Leu Val Asp Ile Pro Leu Asp His Pro Ser Ser Ala Gln Phe
20 25 30

Pro Ile Asn Arg Lys Arg Thr Gly Arg Leu Ile Tyr Arg Leu Phe Ser
35 40 45

Ile Leu Cys Asn Leu Ile Leu Ile Ser Ile Leu Ile Ser Leu Val Val
50 55 60

Ile Trp Ser Arg Ser Ser Arg Asp Cys Ala Lys Ser Asp Gly Leu Ser
65 70 75 80

```
Ser Val Asp Asn Gln Leu Ser Ser Leu Ser Arg Ser Ile Asn Ser Leu
                85                  90                  95

Ile Thr Glu Val Asn Gln Ile Ser Val Thr Thr Ala Ile Asn Leu Pro
            100                 105                 110

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

Met Ile Arg Gln Cys Asn Ala Ala Cys Lys Gly Pro Gly Glu Lys Pro
    130                 135                 140

Gly Ile Gln Asn Val Arg Ile Asn Ile Pro Asn Asn Phe Ser Thr Tyr
145                 150                 155                 160

Ser Glu Leu Asn Arg Thr Ala Asn Ser Leu Asn Phe Gln Ser Arg Thr
                165                 170                 175

Ala Leu Phe Ala Arg Pro Asn Pro Tyr Pro Lys Thr Cys Ser Arg Phe
            180                 185                 190

Pro Ser Tyr Ser Val Tyr Phe Gly Ile His Cys Phe Ser His Ala Val
        195                 200                 205

Thr Asp Ser Ser Cys Glu Leu Ser Asp Ser Thr Tyr Tyr Arg Leu Val
    210                 215                 220

Ile Gly Val Ala Asp Lys Asn Leu Ser Asp Pro Ala Asp Val Lys Tyr
225                 230                 235                 240

Ile Gly Glu Thr Thr Thr Pro Val Arg Val Gln Thr Arg Gly Cys Ser
                245                 250                 255

Val Val Ser Ser Ile Tyr Gly Cys Tyr Leu Leu Cys Ser Lys Ser Asn
            260                 265                 270

Gln Asp Tyr Gln Asp Asp Phe Arg Glu Gln Gly Phe His Gln Met Phe
        275                 280                 285

Ile Leu Phe Leu Ser Arg Glu Leu Lys Thr Thr Phe Phe Asp Asp Met
    290                 295                 300

Val Ser Ser Thr Thr Val Thr Trp Asn Gly Leu Tyr Pro Gly Glu Gly
305                 310                 315                 320

Ser Gly Ile Trp His Met Gly His Leu Val Phe Pro Leu Trp Gly Gly
                325                 330                 335

Ile Arg Phe Gly Thr His Ala Ser Glu Gly Ile Leu Asn Ser Thr Leu
            340                 345                 350
```

```
Glu Leu Pro Pro Val Gly Pro Ser Cys Lys Arg Ser Leu Ala Asp Asn
        355                 360                 365

Gly Leu Ile Asn Lys Asp Val Leu Phe Ser Pro Tyr Phe Gly Asp Ser
    370                 375                 380

Val Met Val Phe Ala Tyr Leu Ser Cys Tyr Met Leu Ser Asn Val Pro
385                 390                 395                 400

Thr His Cys Gln Val Glu Thr Met Asn Ser Ser Val Leu Gly Phe Gly
                405                 410                 415

Ser Arg Ala Gln Phe Tyr Asp Leu Lys Gly Ile Val Tyr Leu Tyr Ile
            420                 425                 430

Gln Ser Ala Gly Trp Phe Ser Tyr Thr Gln Leu Phe Arg Leu Ser Leu
        435                 440                 445

Gln Ser Lys Gly Tyr Lys Leu Ser Val Lys Gln Ile Lys Arg Ile Pro
    450                 455                 460

Ile Ser Ser Thr Ser Arg Pro Gly Thr Glu Pro Cys Asp Ile Ile His
465                 470                 475                 480

Asn Cys Pro Tyr Thr Cys Ala Thr Gly Leu Phe Gln Ala Pro Trp Ile
                485                 490                 495

Val Asn Gly Asp Ser Ile Arg Asp Arg Asp Val Arg Asn Met Ala Phe
            500                 505                 510

Val Gln Ala Trp Ser Gly Ala Ile Asn Thr Phe Gln Arg Pro Phe Met
        515                 520                 525

Ser Ile Cys Ser Gln Tyr Ser Cys Pro Leu Ser Glu Leu Leu Asp Ser
    530                 535                 540

Glu Ser Ser Ile Met Arg Ser Thr Thr Thr Tyr Cys Phe Pro Ser Leu
545                 550                 555                 560

Thr Glu Ser Ile Leu Gln Cys Val Ser Phe Ile Glu Trp Gly Gly Pro
                565                 570                 575

Val Gly Asn Pro Ile Ser Ile Asn Glu Val Tyr Ser Ser Ile Ser Phe
            580                 585                 590

Arg Pro Asp
        595
```

<210> 95
<211> 632
<212> PRT
<213> Mossman virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 95

```
Met Val Asp Pro Pro Ala Val Ser Tyr Tyr Thr Gly Thr Gly Arg Asn
1               5                   10                  15


Asp Arg Val Lys Val Val Thr Thr Gln Ser Thr Asn Pro Tyr Trp Ala
            20                  25                  30


His Asn Pro Asn Gln Gly Leu Arg Arg Leu Ile Asp Met Val Val Asn
        35                  40                  45


Val Ile Met Val Thr Gly Val Ile Phe Ala Leu Ile Asn Ile Ile Leu
    50                  55                  60


Gly Ile Val Ile Ile Ser Gln Ser Ala Gly Ser Arg Gln Asp Thr Ser
65                  70                  75                  80


Lys Ser Leu Asp Ile Ile Gln His Val Asp Ser Ser Val Ala Ile Thr
                85                  90                  95


Lys Gln Ile Val Met Glu Asn Leu Glu Pro Lys Ile Arg Ser Ile Leu
            100                 105                 110


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


Pro Gly Lys Thr Asp Pro Pro Ile Ala Leu Pro Thr Lys Ala Ser Thr
    130                 135                 140


Pro Val Ile Pro Thr Glu Tyr Pro Ser Leu Asn Thr Thr Thr Cys Leu
145                 150                 155                 160


Arg Ile Glu Glu Ser Val Thr Gln Asn Ala Ala Ala Leu Phe Asn Ile
                165                 170                 175


Ser Phe Asp Leu Lys Thr Val Met Tyr Glu Leu Val Thr Arg Thr Gly
            180                 185                 190


Gly Cys Val Thr Leu Pro Ser Tyr Ser Glu Leu Tyr Thr Arg Val Arg
        195                 200                 205


Thr Phe Ser Thr Ala Ile Arg Asn Pro Lys Thr Cys Gln Arg Ala Gly
    210                 215                 220


Gln Glu Thr Asp Leu Asn Leu Ile Pro Ala Phe Ile Gly Thr Asp Thr
225                 230                 235                 240


Gly Ile Leu Ile Asn Ser Cys Val Arg Gln Pro Val Ile Ala Thr Gly
```

245 250 255

Asp Gly Ile Tyr Ala Leu Thr Tyr Leu Thr Met Arg Gly Thr Cys Gln
260 265 270

Asp His Arg His Ala Val Arg His Phe Glu Ile Gly Leu Val Arg Arg
275 280 285

Asp Ala Trp Trp Asp Pro Val Leu Thr Pro Ile His His Phe Thr Glu
290 295 300

Pro Gly Thr Pro Val Phe Asp Gly Cys Ser Leu Thr Val Gln Asn Gln
305 310 315 320

Thr Ala Leu Ala Leu Cys Thr Leu Thr Thr Asp Gly Pro Glu Thr Asp
325 330 335

Ile His Asn Gly Ala Ser Leu Gly Leu Ala Leu Val His Phe Asn Ile
340 345 350

Arg Gly Glu Phe Ser Lys His Lys Val Asp Pro Arg Asn Ile Asp Thr
355 360 365

Gln Asn Gln Gly Leu His Leu Val Thr Thr Ala Gly Lys Ser Ala Val
370 375 380

Lys Lys Gly Ile Leu Tyr Ser Phe Gly Tyr Met Val Thr Arg Ser Pro
385 390 395 400

Glu Pro Gly Asp Ser Lys Cys Val Thr Glu Glu Cys Asn Gln Asn Asn
405 410 415

Gln Glu Lys Cys Asn Ala Tyr Ser Lys Thr Thr Leu Asp Pro Asp Lys
420 425 430

Pro Arg Ser Met Ile Ile Phe Gln Ile Asp Val Gly Ala Glu Tyr Phe
435 440 445

Thr Val Asp Lys Val Val Val Val Pro Arg Thr Gln Tyr Tyr Gln Leu
450 455 460

Thr Ser Gly Asp Leu Phe Tyr Thr Gly Glu Glu Asn Asp Leu Leu Tyr
465 470 475 480

Gln Leu His Asn Lys Gly Trp Tyr Asn Lys Pro Ile Arg Gly Arg Val
485 490 495

Thr Phe Asp Gly Gln Val Thr Leu His Glu His Ser Arg Thr Tyr Asp
500 505 510

Ser Leu Ser Asn Gln Arg Ala Cys Asn Pro Arg Leu Gly Cys Pro Ser
515 520 525

```
Thr Cys Glu Leu Thr Ser Met Ala Ser Tyr Phe Pro Leu Asp Lys Asp
    530                 535                 540

Phe Lys Ala Ala Val Gly Val Ile Ala Leu Arg Asn Gly Met Thr Pro
545                 550                 555                 560

Ile Ile Thr Tyr Ser Thr Asp Asp Trp Arg Asn His Trp Lys Tyr Ile
                565                 570                 575

Lys Asn Ala Asp Leu Glu Phe Ser Glu Ser Ser Leu Ser Cys Tyr Ser
            580                 585                 590

Pro Asn Pro Pro Leu Asp Asp Tyr Val Leu Cys Thr Ala Val Ile Thr
        595                 600                 605

Ala Lys Val Met Ser Asn Thr Asn Pro Gln Leu Leu Ala Thr Ser Trp
    610                 615                 620

Tyr Gln Tyr Asp Lys Cys His Thr
625                 630


<210>  96
<211>  709
<212>  PRT
<213>  J-virus


<220>
<221>  MISC_FEATURE
<223>  attachment glycoprotein

<400>  96

Met Asn Pro Val Ala Met Ser Asn Phe Tyr Gly Ile Asn Gln Ala Asp
1               5                   10                  15

His Leu Arg Glu Lys Gly Asp Gln Pro Glu Lys Gly Pro Ser Val Leu
            20                  25                  30

Thr Tyr Val Ser Leu Ile Thr Gly Leu Leu Ser Leu Phe Thr Ile Ile
        35                  40                  45

Ala Leu Asn Val Thr Asn Ile Ile Tyr Leu Thr Gly Ser Gly Gly Thr
    50                  55                  60

Met Ala Thr Ile Lys Asp Asn Gln Gln Ser Met Ser Gly Ser Met Arg
65                  70                  75                  80

Asp Ile Ser Gly Met Leu Val Glu Asp Leu Lys Pro Lys Thr Asp Leu
                85                  90                  95

Ile Asn Ser Met Val Ser Tyr Thr Ile Pro Ser Gln Ile Ser Ala Met
            100                 105                 110
```

```
Ser Ala Met Ile Lys Asn Glu Val Leu Arg Gln Cys Thr Pro Ser Phe
        115                 120                 125

Met Phe Asn Asn Thr Ile Cys Pro Ile Ala Glu His Pro Val His Thr
    130                 135                 140

Ser Tyr Phe Glu Glu Val Gly Ile Glu Ala Ile Ser Met Cys Thr Gly
145                 150                 155                 160

Thr Asn Arg Lys Leu Val Val Asn Gln Gly Ile Asn Phe Val Glu Tyr
                165                 170                 175

Pro Ser Phe Ile Pro Gly Ser Thr Lys Pro Gly Gly Cys Val Arg Leu
            180                 185                 190

Pro Ser Phe Ser Leu Gly Leu Glu Val Phe Ala Tyr Ala His Ala Ile
        195                 200                 205

Thr Gln Asp Asp Cys Thr Ser Ser Ser Thr Pro Asp Tyr Tyr Phe Ser
    210                 215                 220

Val Gly Arg Ile Ala Asp His Gly Thr Asp Val Pro Val Phe Glu Thr
225                 230                 235                 240

Leu Ala Glu Trp Phe Leu Asp Asp Lys Met Asn Arg Arg Ser Cys Ser
                245                 250                 255

Val Thr Ala Ala Gly Lys Gly Gly Trp Leu Gly Cys Ser Ile Leu Val
            260                 265                 270

Gly Ser Phe Thr Asp Glu Leu Thr Ser Pro Glu Val Asn Arg Ile Ser
        275                 280                 285

Leu Ser Tyr Met Asp Thr Phe Gly Lys Lys Lys Asp Trp Leu Tyr Thr
    290                 295                 300

Gly Ser Glu Val Arg Ala Asp Gln Ser Trp Ser Ala Leu Phe Phe Ser
305                 310                 315                 320

Val Gly Ser Gly Val Val Ile Gly Asp Thr Val Tyr Phe Leu Val Trp
                325                 330                 335

Gly Gly Leu Asn His Pro Ile Asn Val Asp Ala Met Cys Arg Ala Pro
            340                 345                 350

Gly Cys Gln Ser Pro Thr Gln Ser Leu Cys Asn Tyr Ala Ile Lys Pro
        355                 360                 365

Gln Glu Trp Gly Gly Asn Gln Ile Val Asn Gly Ile Leu His Phe Lys
    370                 375                 380
```

```
His Asp Thr Asn Glu Lys Pro Thr Leu His Val Arg Thr Leu Ser Pro
385                 390                 395                 400

Asp Asn Asn Trp Met Gly Ala Glu Gly Arg Leu Phe His Phe His Asn
                405                 410                 415

Ser Gly Lys Thr Phe Ile Tyr Thr Arg Ser Ser Thr Trp His Thr Leu
            420                 425                 430

Pro Gln Val Gly Ile Leu Thr Leu Gly Trp Pro Leu Ser Val Gln Trp
        435                 440                 445

Val Asp Ile Thr Ser Ile Ser Arg Pro Gly Gln Ser Pro Cys Glu Tyr
    450                 455                 460

Asp Asn Arg Cys Pro His Gln Cys Val Thr Gly Val Tyr Thr Asp Leu
465                 470                 475                 480

Phe Pro Leu Gly Val Ser Tyr Glu Tyr Ser Val Thr Ala Tyr Leu Asp
                485                 490                 495

Gln Val Gln Ser Arg Met Asn Pro Lys Ile Ala Leu Val Gly Ala Gln
            500                 505                 510

Glu Lys Ile Tyr Glu Lys Thr Ile Thr Thr Asn Thr Gln His Ala Asp
        515                 520                 525

Tyr Thr Thr Thr Ser Cys Phe Ala Tyr Lys Leu Arg Val Trp Cys Val
    530                 535                 540

Ser Ile Val Glu Met Ser Pro Gly Val Ile Thr Thr Arg Gln Pro Val
545                 550                 555                 560

Pro Phe Leu Tyr His Leu Asn Leu Gly Cys Gln Asp Thr Ser Thr Gly
                565                 570                 575

Ser Leu Thr Pro Leu Asp Ala His Gly Gly Thr Tyr Leu Asn Thr Asp
            580                 585                 590

Pro Val Gly Asn Lys Val Asp Cys Tyr Phe Val Leu His Glu Gly Gln
        595                 600                 605

Ile Tyr Phe Gly Met Ser Val Gly Pro Ile Asn Tyr Thr Tyr Ser Ile
    610                 615                 620

Val Gly Arg Ser Arg Glu Ile Gly Ala Asn Met Asn Val Ser Leu Asn
625                 630                 635                 640

Gln Leu Cys His Ser Val Tyr Thr Glu Phe Leu Lys Glu Lys Glu His
                645                 650                 655

Pro Gly Thr Arg Asn Asn Ile Asp Val Glu Gly Trp Leu Leu Lys Arg
```

660 665 670

Ile Glu Thr Leu Asn Gly Thr Lys Ile Phe Gly Leu Asp Asp Leu Glu
675 680 685

Gly Ser Gly Pro Gly His Gln Ser Gly Pro Glu Asp Pro Ser Ile Ala
690 695 700

Pro Ile Gly His Asn
705

<210> 97
<211> 264
<212> PRT
<213> Avian metapneumovirus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 97

Met Glu Val Lys Val Glu Asn Val Gly Lys Ser Gln Glu Leu Lys Val
1 5 10 15

Lys Val Lys Asn Phe Ile Lys Arg Ser Asp Cys Lys Lys Lys Leu Phe
20 25 30

Ala Leu Ile Leu Gly Leu Val Ser Phe Glu Leu Thr Met Asn Ile Met
35 40 45

Leu Ser Val Met Tyr Val Glu Ser Asn Glu Ala Leu Ser Leu Cys Arg
50 55 60

Ile Gln Gly Thr Pro Ala Pro Arg Asp Asn Lys Thr Asn Thr Glu Asn
65 70 75 80

Ala Thr Lys Glu Thr Thr Leu His Thr Thr Thr Thr Thr Arg Asp Pro
85 90 95

Glu Val Arg Glu Thr Lys Thr Thr Lys Pro Gln Ala Asn Glu Gly Ala
100 105 110

Thr Asn Pro Ser Arg Asn Leu Thr Thr Lys Gly Asp Lys His Gln Thr
115 120 125

Thr Arg Ala Thr Thr Glu Ala Glu Leu Glu Lys Gln Ser Lys Gln Thr
130 135 140

Thr Glu Pro Gly Thr Ser Thr Gln Lys His Thr Pro Ala Arg Pro Ser
145 150 155 160

Ser Lys Ser Pro Thr Thr Thr Gln Ala Thr Ala Gln Pro Thr Thr Pro
165 170 175

```
Thr Ala Pro Lys Ala Ser Thr Ala Pro Lys Asn Arg Gln Ala Thr Thr
            180                 185                 190


Lys Lys Thr Glu Thr Asp Thr Thr Thr Ala Ser Arg Ala Arg Asn Thr
        195                 200                 205


Asn Asn Pro Thr Glu Thr Ala Thr Thr Thr Pro Lys Ala Thr Thr Glu
    210                 215                 220


Thr Gly Lys Gly Lys Glu Gly Pro Thr Gln His Thr Thr Lys Glu Gln
225                 230                 235                 240


Pro Glu Thr Thr Ala Arg Glu Thr Thr Thr Pro Gln Pro Arg Arg Thr
                245                 250                 255


Ala Gly Ala Ser Pro Arg Ala Ser
            260


<210>  98
<211>  391
<212>  PRT
<213>  Avian metapneumovirus


<220>
<221>  MISC_FEATURE
<223>  attachment protein

<400>  98

Met Gly Ser Lys Leu Tyr Met Val Gln Gly Thr Ser Ala Tyr Gln Thr
1               5                   10                  15


Ala Val Gly Phe Trp Leu Asp Ile Gly Arg Arg Tyr Ile Leu Ala Ile
            20                  25                  30


Val Leu Ser Ala Phe Gly Leu Thr Cys Thr Val Thr Ile Ala Leu Thr
        35                  40                  45


Val Ser Val Ile Val Glu Gln Ser Val Leu Glu Glu Cys Arg Asn Tyr
    50                  55                  60


Asn Gly Gly Asp Arg Asp Trp Trp Ser Thr Thr Gln Glu Gln Pro Thr
65                  70                  75                  80


Thr Ala Pro Ser Ala Thr Pro Ala Gly Asn Tyr Gly Gly Leu Gln Thr
                85                  90                  95


Ala Arg Thr Arg Lys Ser Glu Ser Cys Leu His Val Gln Ile Ser Tyr
            100                 105                 110


Gly Asp Met Tyr Ser Arg Ser Asp Thr Val Leu Gly Gly Phe Asp Cys
        115                 120                 125
```

```
Met Gly Leu Leu Val Leu Cys Lys Ser Gly Pro Ile Cys Gln Arg Asp
    130                 135                 140

Asn Gln Val Asp Pro Thr Ala Leu Cys His Cys Arg Val Asp Leu Ser
145                 150                 155                 160

Ser Val Asp Cys Cys Lys Val Asn Lys Ile Ser Thr Asn Ser Ser Thr
                165                 170                 175

Thr Ser Glu Pro Gln Lys Thr Asn Pro Ala Trp Pro Ser Gln Asp Asn
            180                 185                 190

Thr Asp Ser Asp Pro Asn Pro Gln Gly Ile Thr Thr Ser Thr Ala Thr
        195                 200                 205

Leu Leu Ser Thr Ser Leu Gly Leu Met Leu Thr Ser Lys Thr Gly Thr
    210                 215                 220

His Lys Ser Gly Pro Pro Gln Ala Leu Pro Gly Ser Asn Thr Asn Gly
225                 230                 235                 240

Lys Thr Thr Thr Asp Arg Glu Leu Gly Ser Thr Asn Gln Pro Asn Ser
                245                 250                 255

Thr Thr Asn Gly Gln His Asn Lys His Thr Gln Arg Met Thr Leu Pro
            260                 265                 270

Pro Ser Tyr Asp Asn Thr Arg Thr Ile Leu Gln His Thr Thr Pro Trp
        275                 280                 285

Glu Lys Thr Phe Ser Thr Tyr Lys Pro Thr His Ser Pro Thr Asn Glu
    290                 295                 300

Ser Asp Gln Ser Leu Pro Thr Thr Gln Asn Ser Ile Asn Cys Glu His
305                 310                 315                 320

Phe Asp Pro Gln Gly Lys Glu Lys Ile Cys Tyr Arg Val Gly Ser Tyr
                325                 330                 335

Asn Ser Asn Ile Thr Lys Gln Cys Arg Ile Asp Val Pro Leu Cys Ser
            340                 345                 350

Thr Tyr Asn Thr Val Cys Met Lys Thr Tyr Tyr Thr Glu Pro Phe Asn
        355                 360                 365

Cys Trp Arg Arg Ile Trp Arg Cys Leu Cys Asp Asp Gly Val Gly Leu
    370                 375                 380

Val Glu Trp Cys Cys Thr Ser
385                 390
```

<210> 99
<211> 1052
<212> PRT
<213> Tailam virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 99

Met Ser Gln Leu Ala Ala His Asn Leu Ala Met Ser Asn Phe Tyr Gly
1 5 10 15

Ile His Gln Gly Gly Gln Ser Thr Ser Gln Lys Glu Glu Glu Gln Pro
20 25 30

Val Gln Gly Val Ile Arg Tyr Ala Ser Met Ile Val Gly Leu Leu Ser
35 40 45

Leu Phe Thr Ile Ile Ala Leu Asn Val Thr Asn Ile Ile Tyr Met Thr
50 55 60

Glu Ser Gly Gly Thr Met Gln Ser Ile Lys Asn Ala Gln Gly Ser Ile
65 70 75 80

Asp Gly Ser Met Lys Asp Leu Ser Gly Thr Ile Met Glu Asp Ile Lys
85 90 95

Pro Lys Thr Asp Leu Ile Asn Ser Met Val Ser Tyr Asn Ile Pro Ala
100 105 110

Gln Leu Ser Met Ile His Gln Ile Ile Lys Asn Asp Val Leu Lys Gln
115 120 125

Cys Thr Pro Ser Phe Met Phe Asn Asn Thr Ile Cys Pro Leu Ala Glu
130 135 140

Asn Pro Thr His Ser Arg Tyr Phe Glu Glu Val Asn Leu Asp Ser Ile
145 150 155 160

Ser Glu Cys Ser Gly Asn Glu Met Ser Leu Glu Leu Gly Thr Glu Pro
165 170 175

Glu Phe Ile Glu Tyr Pro Ser Phe Ala Pro Gly Ser Thr Lys Pro Gly
180 185 190

Ser Cys Val Arg Leu Pro Ser Phe Ser Leu Ser Ser Thr Val Phe Ala
195 200 205

Tyr Thr His Thr Ile Met Gly His Gly Cys Ser Glu Leu Asp Val Gly
210 215 220

Asp His Tyr Leu Ala Ile Gly Arg Ile Ala Asp Ala Gly His Glu Ile

225 230 235 240

Pro Gln Phe Glu Thr Ile Ser Ser Trp Phe Ile Asn Asp Lys Ile Asn
245 250 255

Arg Arg Ser Cys Thr Val Ala Ala Gly Val Met Glu Thr Trp Met Gly
260 265 270

Cys Val Ile Met Thr Glu Thr Phe Tyr Asp Asp Leu Asp Ser Leu Asp
275 280 285

Thr Gly Lys Ile Thr Ile Ser Tyr Leu Asp Val Phe Gly Arg Lys Lys
290 295 300

Glu Trp Ile Tyr Thr Arg Ser Glu Ile Leu Tyr Asp Tyr Thr Tyr Thr
305 310 315 320

Ser Val Tyr Phe Ser Ile Gly Ser Gly Val Val Val Gly Asp Thr Val
325 330 335

Tyr Phe Leu Leu Trp Gly Ser Leu Ser Ser Pro Ile Glu Glu Thr Ala
340 345 350

Tyr Cys Tyr Ala Pro Gly Cys Ser Asn Tyr Asn Gln Arg Met Cys Asn
355 360 365

Glu Ala Gln Arg Pro Ala Lys Phe Gly His Arg Gln Met Ala Asn Ala
370 375 380

Ile Leu Arg Phe Lys Thr Asn Ser Met Gly Lys Pro Ser Ile Ser Val
385 390 395 400

Arg Thr Leu Ser Pro Thr Val Ile Pro Phe Gly Thr Glu Gly Arg Leu
405 410 415

Ile Tyr Ser Asp Phe Thr Lys Ile Ile Tyr Leu Tyr Leu Arg Ser Thr
420 425 430

Ser Trp Tyr Val Leu Pro Leu Thr Gly Leu Leu Ile Leu Gly Pro Pro
435 440 445

Val Ser Ile Ser Trp Val Thr Gln Glu Ala Val Ser Arg Pro Gly Glu
450 455 460

Tyr Pro Cys Gly Ala Ser Asn Arg Cys Pro Lys Asp Cys Ile Thr Gly
465 470 475 480

Val Tyr Thr Asp Leu Phe Pro Leu Gly Ala Arg Tyr Glu Tyr Ala Val
485 490 495

Thr Val Tyr Leu Asn Ala Glu Thr Tyr Arg Val Asn Pro Thr Leu Ala
500 505 510

```
Leu Ile Asp Arg Ser Lys Ile Ile Ala Arg Lys Lys Ile Thr Thr Glu
        515                 520                 525

Ser Gln Lys Ala Gly Tyr Thr Thr Thr Thr Cys Phe Val Phe Lys Leu
    530                 535                 540

Arg Ile Trp Cys Met Ser Val Val Glu Leu Ala Pro Ala Thr Met Thr
545                 550                 555                 560

Ala Phe Glu Pro Val Pro Phe Leu Tyr Gln Leu Asp Leu Thr Cys Lys
                565                 570                 575

Arg Asn Asn Gly Thr Thr Ala Met Gln Phe Ser Gly Gln Asp Gly Met
            580                 585                 590

Tyr Lys Ser Gly Arg Tyr Lys Ser Pro Arg Asn Glu Cys Phe Phe Glu
        595                 600                 605

Lys Val Ser Asn Lys Tyr Tyr Phe Val Val Ser Thr Pro Glu Gly Ile
    610                 615                 620

Gln Pro Tyr Glu Val Arg Asp Leu Thr Pro Glu Arg Val Ser His Val
625                 630                 635                 640

Ile Met Tyr Ile Ser Asp Val Cys Ala Pro Ala Leu Ser Ala Phe Lys
                645                 650                 655

Lys Leu Ile Pro Ala Met Arg Pro Ile Thr Thr Leu Thr Ile Gly Asn
            660                 665                 670

Trp Gln Phe Arg Pro Val Asp Ile Ser Gly Gly Leu Arg Val Asn Ile
        675                 680                 685

Tyr Arg Asn Leu Thr Arg Tyr Gly Asp Leu Ser Met Ser Ala Pro Glu
    690                 695                 700

Asp Pro Gly Thr Asp Thr Phe Pro Gly Thr His Ala Pro Ser Lys Gly
705                 710                 715                 720

His Glu Glu Val Gly His Tyr Thr Leu Pro Asn Glu Lys Leu Ser Glu
                725                 730                 735

Val Thr Thr Ala Ala Val Lys Thr Lys Glu Ser Leu Asn Leu Ile Pro
            740                 745                 750

Asp Thr Lys Asp Thr Arg Gly Glu Glu Glu Asn Gly Ser Gly Leu Asn
        755                 760                 765

Glu Ile Ile Thr Gly His Thr Thr Pro Gly His Ile Lys Thr His Pro
    770                 775                 780
```

```
Ala Glu Thr Lys Val Thr Lys His Thr Val Ile Ile Pro Gln Ile Glu
785                 790                 795                 800


Glu Asp Gly Ser Gly Ala Thr Thr Ser Thr Glu Leu Gln Asp Glu Thr
                805                 810                 815


Gly Tyr His Thr Glu Asp Tyr Asn Thr Thr Asn Thr Asn Gly Ser Leu
            820                 825                 830


Thr Ala Pro Asn Glu Arg Asn Asn Tyr Thr Ser Gly Asp His Thr Val
        835                 840                 845


Ser Gly Glu Asp Ile Thr His Thr Ile Thr Val Ser Asp Arg Thr Lys
    850                 855                 860


Thr Thr Gln Thr Leu Pro Thr Asp Asn Thr Phe Asn Gln Thr Pro Thr
865                 870                 875                 880


Lys Ile Gln Glu Gly Ser Pro Lys Ser Glu Ser Thr Pro Lys Asp Tyr
                885                 890                 895


Thr Ala Ile Glu Ser Glu Asp Ser His Phe Thr Asp Pro Thr Leu Ile
            900                 905                 910


Arg Ser Thr Pro Glu Gly Thr Ile Val Gln Val Ile Gly Asp Gln Phe
        915                 920                 925


His Ser Ala Val Thr Gln Leu Gly Glu Ser Asn Ala Ile Gly Asn Ser
    930                 935                 940


Glu Pro Ile Asp Gln Gly Asn Asn Leu Ile Pro Thr Thr Asp Arg Gly
945                 950                 955                 960


Thr Met Asp Asn Thr Ser Ser Gln Ser His Ser Ser Thr Thr Ser Thr
                965                 970                 975


Gln Gly Ser His Ser Ala Gly His Gly Ser Gln Ser Asn Met Asn Leu
            980                 985                 990


Thr Ala Leu Ala Asp Thr Asp Ser  Val Thr Asp Gln Ser  Thr Ser Thr
        995                 1000                1005


Gln Glu  Ile Asp His Glu His  Glu Asn Val Ser Ser  Ile Leu Asn
    1010                1015                1020


Pro Leu  Ser Arg His Thr Arg  Val Met Arg Asp Thr  Val Gln Glu
    1025                1030                1035


Ala Leu  Thr Gly Ala Trp Gly  Phe Ile Arg Gly Met  Ile Pro
    1040                1045                1050
```

<210> 100
<211> 242
<212> PRT
<213> Human metapneumovirus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 100

Met Glu Val Arg Val Glu Asn Ile Arg Ala Ile Asp Met Phe Lys Ala
1 5 10 15

Lys Ile Lys Asn Arg Ile Arg Asn Ser Arg Cys Tyr Arg Asn Ala Thr
20 25 30

Leu Ile Leu Ile Gly Leu Thr Ala Leu Ser Met Ala Leu Asn Ile Phe
35 40 45

Leu Ile Ile Asp His Ala Thr Leu Arg Asn Met Ile Lys Thr Glu Asn
50 55 60

Cys Ala Asn Met Pro Ser Ala Glu Pro Ser Lys Lys Thr Pro Met Thr
65 70 75 80

Ser Ile Ala Gly Pro Ser Thr Lys Pro Asn Pro Gln Gln Ala Thr Gln
85 90 95

Trp Thr Thr Glu Asn Ser Thr Ser Pro Ala Ala Thr Leu Glu Gly His
100 105 110

Pro Tyr Thr Gly Thr Thr Gln Thr Pro Asp Thr Thr Ala Pro Gln Gln
115 120 125

Thr Thr Asp Lys His Thr Ala Leu Pro Lys Ser Thr Asn Glu Gln Ile
130 135 140

Thr Gln Thr Thr Thr Glu Lys Lys Thr Thr Arg Ala Thr Thr Gln Lys
145 150 155 160

Arg Lys Lys Glu Lys Lys Thr Gln Thr Lys Pro Gln Val Gln Leu Gln
165 170 175

Pro Lys Gln Pro Thr Pro Pro Thr Lys Ser Glu Met Gln Val Arg Gln
180 185 190

Ser Gln His Pro Thr Asp Pro Glu Leu Thr Pro Leu Pro Lys Ala Val
195 200 205

Asn Arg Gln Pro Gly Gln Gln Asn Gln Ala Pro His His Ile Met His
210 215 220

Gly Glu Val Gln Asp Pro Gly Glu Arg Asn Thr Gln Val Ser His Pro

225 230 235 240

Ser Ser

<210> 101
<211> 585
<212> PRT
<213> Avian metapneumovirus type C

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 101

Met Glu Val Lys Ile Glu Asn Val Gly Lys Ser Gln Glu Leu Arg Val
1 5 10 15

Lys Val Lys Asn Phe Ile Lys Arg Ser Asp Cys Lys Lys Lys Leu Phe
20 25 30

Ala Leu Ile Leu Gly Leu Ile Ser Phe Asp Ile Thr Met Asn Ile Met
35 40 45

Leu Ser Val Met Tyr Val Glu Ser Asn Glu Ala Leu Ser Ser Cys Arg
50 55 60

Val Gln Gly Thr Pro Ala Pro Arg Asp Asn Arg Thr Asn Thr Glu Asn
65 70 75 80

Thr Ala Lys Glu Thr Thr Leu His Thr Met Thr Thr Thr Arg Asn Thr
85 90 95

Glu Ala Gly Gly Thr Lys Thr Thr Lys Pro Gln Ala Asp Glu Arg Ala
100 105 110

Thr Ser Pro Ser Lys Asn Pro Thr Ile Gly Ala Asp Lys His Lys Thr
115 120 125

Thr Arg Ala Thr Thr Glu Ala Glu Gln Glu Lys Gln Ser Lys Gln Thr
130 135 140

Thr Glu Pro Gly Thr Ser Thr Pro Lys His Ile Pro Ala Arg Pro Ser
145 150 155 160

Ser Lys Ser Pro Ala Thr Thr Lys Thr Thr Thr Gln Pro Thr Thr Pro
165 170 175

Thr Val Ala Lys Gly Gly Thr Ala Pro Lys Asn Arg Gln Thr Thr Thr
180 185 190

Lys Lys Thr Glu Ala Asp Thr Pro Thr Thr Ser Arg Ala Lys Gln Thr
195 200 205

```
Asn Lys Pro Thr Gly Thr Glu Thr Thr Pro Pro Arg Ala Thr Thr Glu
    210                 215                 220

Thr Asp Lys Asp Lys Glu Gly Pro Thr Gln His Thr Thr Lys Glu Gln
225                 230                 235                 240

Pro Glu Thr Thr Ala Gly Gly Thr Thr Thr Pro Gln Pro Arg Arg Thr
                245                 250                 255

Thr Ser Arg Pro Ala Pro Thr Thr Asn Thr Lys Glu Gly Ala Glu Thr
            260                 265                 270

Thr Gly Thr Arg Thr Thr Lys Ser Thr Gln Thr Ser Ala Ser Pro Pro
        275                 280                 285

Arg Pro Thr Arg Ser Thr Pro Ser Lys Thr Ala Thr Gly Thr Asn Lys
    290                 295                 300

Arg Ala Thr Thr Thr Lys Gly Pro Asn Thr Ala Ser Thr Asp Arg Arg
305                 310                 315                 320

Gln Gln Thr Arg Thr Thr Pro Lys Gln Asp Gln Gln Thr Gln Thr Lys
                325                 330                 335

Ala Lys Thr Thr Thr Asn Lys Ala His Ala Lys Ala Ala Thr Thr Pro
            340                 345                 350

Glu His Asn Thr Asp Thr Thr Asp Ser Met Lys Glu Asn Ser Lys Glu
        355                 360                 365

Asp Lys Thr Thr Arg Asp Pro Ser Ser Lys Ala Thr Thr Lys Gln Glu
    370                 375                 380

Asn Thr Ser Lys Gly Thr Thr Ala Thr Asn Leu Gly Asn Asn Thr Glu
385                 390                 395                 400

Ala Gly Ala Arg Thr Pro Pro Thr Thr Thr Pro Thr Arg His Thr Thr
                405                 410                 415

Glu Pro Ala Thr Ser Thr Ala Gly Gly His Thr Lys Ala Arg Thr Thr
            420                 425                 430

Arg Trp Lys Ser Thr Ala Ala Arg Gln Pro Thr Arg Asn Asn Thr Thr
        435                 440                 445

Ala Asp Thr Lys Thr Ala Gln Ser Lys Gln Thr Thr Pro Ala Gln Leu
    450                 455                 460

Gly Asn Asn Thr Thr Pro Glu Asn Thr Thr Pro Pro Asp Asn Lys Ser
465                 470                 475                 480
```

Asn Ser Gln Thr Asn Val Ala Pro Thr Glu Glu Ile Glu Ile Gly Ser
485 490 495

Ser Leu Trp Arg Arg Arg Tyr Val Tyr Gly Pro Cys Arg Glu Asn Ala
500 505 510

Leu Glu His Pro Met Asn Pro Cys Leu Lys Asp Asn Thr Thr Trp Ile
515 520 525

Tyr Leu Asp Asn Gly Arg Asn Leu Pro Ala Gly Tyr Tyr Asp Ser Lys
530 535 540

Thr Asp Lys Ile Ile Cys Tyr Gly Ile Tyr Arg Gly Asn Ser Tyr Cys
545 550 555 560

Tyr Gly Arg Ile Glu Cys Thr Cys Lys Asn Gly Thr Gly Leu Leu Ser
565 570 575

Tyr Cys Cys Asn Ser Tyr Asn Trp Ser
580 585

<210> 102
<211> 609
<212> PRT
<213> Rinderpest morbillivirus

<220>
<221> MISC_FEATURE
<223> H protein

<400> 102

Met Ser Ser Pro Arg Asp Arg Val Asn Ala Phe Tyr Lys Asp Asn Leu
1 5 10 15

Gln Phe Lys Asn Thr Arg Val Val Leu Asn Lys Glu Gln Leu Leu Ile
20 25 30

Glu Arg Pro Tyr Met Leu Leu Ala Val Leu Phe Val Met Phe Leu Ser
35 40 45

Leu Val Gly Leu Leu Ala Ile Ala Gly Ile Arg Leu His Arg Ala Ala
50 55 60

Val Asn Thr Ala Glu Ile Asn Ser Gly Leu Thr Thr Ser Ile Asp Ile
65 70 75 80

Thr Lys Ser Ile Glu Tyr Gln Val Lys Asp Val Leu Thr Pro Leu Phe
85 90 95

Lys Ile Ile Gly Asp Glu Val Gly Leu Arg Thr Pro Gln Arg Phe Thr
100 105 110

```
Asp Leu Thr Lys Phe Ile Ser Asp Lys Ile Lys Phe Leu Asn Pro Asp
        115                 120                 125

Lys Glu Tyr Asp Phe Arg Asp Ile Asn Trp Cys Ile Ser Pro Pro Glu
    130                 135                 140

Arg Ile Lys Ile Asn Tyr Asp Gln Tyr Cys Ala His Thr Ala Ala Glu
145                 150                 155                 160

Glu Leu Ile Thr Met Leu Val Asn Ser Ser Leu Ala Gly Thr Ala Val
                165                 170                 175

Leu Arg Thr Ser Leu Val Asn Leu Gly Arg Ser Cys Thr Gly Ser Thr
            180                 185                 190

Thr Thr Lys Gly Gln Phe Ser Asn Met Ser Leu Ala Leu Ser Gly Ile
        195                 200                 205

Tyr Ser Gly Arg Gly Tyr Asn Ile Ser Ser Met Ile Thr Ile Thr Glu
    210                 215                 220

Lys Gly Met Tyr Gly Ser Thr Tyr Leu Val Gly Lys His Asn Gln Gly
225                 230                 235                 240

Ala Arg Arg Pro Ser Thr Ala Trp Gln Arg Asp Tyr Arg Val Phe Glu
                245                 250                 255

Val Gly Ile Ile Arg Glu Leu Gly Val Gly Thr Pro Val Phe His Met
            260                 265                 270

Thr Asn Tyr Leu Glu Leu Pro Arg Gln Pro Glu Leu Glu Ile Cys Met
        275                 280                 285

Leu Ala Leu Gly Glu Phe Lys Leu Ala Ala Leu Cys Leu Ala Asp Asn
    290                 295                 300

Ser Val Ala Leu His Tyr Gly Gly Leu Arg Asp Asp His Lys Ile Arg
305                 310                 315                 320

Phe Val Lys Leu Gly Val Trp Pro Ser Pro Ala Asp Ser Asp Thr Leu
                325                 330                 335

Ala Thr Leu Ser Ala Val Asp Pro Thr Leu Asp Gly Leu Tyr Ile Thr
            340                 345                 350

Thr His Arg Gly Ile Ile Ala Ala Gly Lys Ala Val Trp Ala Val Pro
        355                 360                 365

Val Thr Arg Thr Asp Asp Gln Arg Lys Met Gly Gln Cys Arg Arg Glu
    370                 375                 380

Ala Cys Arg Glu Lys Pro Pro Pro Phe Cys Asn Ser Thr Asp Trp Glu
```

385 390 395 400

Pro Leu Glu Ala Gly Arg Ile Pro Ala Tyr Gly Ile Leu Thr Ile Arg
405 410 415

Leu Gly Leu Ala Asp Lys Pro Glu Ile Asp Ile Ile Ser Glu Phe Gly
420 425 430

Pro Leu Ile Thr His Asp Ser Gly Met Asp Leu Tyr Thr Pro Leu Asp
435 440 445

Gly Asn Glu Tyr Trp Leu Thr Ile Pro Pro Leu Gln Asn Ser Ala Leu
450 455 460

Gly Thr Val Asn Thr Leu Val Leu Glu Pro Ser Leu Lys Ile Ser Pro
465 470 475 480

Asn Ile Leu Thr Leu Pro Ile Arg Ser Gly Gly Gly Asp Cys Tyr Thr
485 490 495

Pro Thr Tyr Leu Ser Asp Leu Ala Asp Asp Asp Val Lys Leu Ser Ser
500 505 510

Asn Leu Val Ile Leu Pro Ser Arg Asn Leu Gln Tyr Val Ser Ala Thr
515 520 525

Tyr Asp Thr Ser Arg Val Glu His Ala Ile Val Tyr Tyr Ile Tyr Ser
530 535 540

Thr Gly Arg Leu Ser Ser Tyr Tyr Tyr Pro Val Lys Leu Pro Ile Lys
545 550 555 560

Gly Asp Pro Val Ser Leu Gln Ile Gly Cys Phe Pro Trp Gly Leu Lys
565 570 575

Leu Trp Cys His His Phe Cys Ser Val Ile Asp Ser Gly Thr Gly Lys
580 585 590

Gln Val Thr His Thr Gly Ala Val Gly Ile Glu Ile Thr Cys Asn Ser
595 600 605

Arg

<210> 103
<211> 575
<212> PRT
<213> Avian paramyxovirus 10

<220>
<221> MISC_FEATURE
<223> hemagglutinin neuraminidase

<400> 103

Met Asp Ser Ser Gln Met Asn Ile Leu Asp Ala Met Asp Arg Glu Ser
1 5 10 15

Ser Lys Arg Thr Trp Arg Gly Val Phe Arg Val Thr Thr Ile Ile Met
20 25 30

Val Val Thr Cys Val Val Leu Ser Ala Ile Thr Leu Ser Lys Val Ala
35 40 45

His Pro Gln Gly Phe Asp Thr Asn Glu Leu Gly Asn Gly Ile Val Asp
50 55 60

Arg Val Ser Asp Lys Ile Thr Glu Ala Leu Thr Val Pro Asn Asn Gln
65 70 75 80

Ile Gly Glu Ile Phe Lys Ile Val Ala Leu Asp Leu His Val Leu Val
85 90 95

Ser Ser Ser Gln Gln Ala Ile Ala Gly Gln Ile Gly Met Leu Ala Glu
100 105 110

Ser Ile Asn Ser Ile Leu Ser Gln Asn Gly Ser Ala Ser Thr Ile Leu
115 120 125

Ser Ser Ser Pro Glu Tyr Ala Gly Gly Ile Gly Val Pro Leu Phe Ser
130 135 140

Asn Lys Leu Thr Asn Gly Thr Val Ile Lys Pro Ile Thr Leu Ile Glu
145 150 155 160

His Pro Ser Phe Ile Pro Gly Pro Thr Thr Ile Gly Gly Cys Thr Arg
165 170 175

Ile Pro Thr Phe His Met Ala Ser Ser His Trp Cys Tyr Ser His Asn
180 185 190

Ile Ile Glu Lys Gly Cys Lys Asp Ser Gly Ile Ser Ser Met Tyr Ile
195 200 205

Ser Leu Gly Val Leu Gln Val Leu Lys Lys Gly Thr Pro Val Phe Leu
210 215 220

Val Thr Ala Ser Ala Val Leu Ser Asp Asp Arg Asn Arg Lys Ser Cys
225 230 235 240

Ser Ile Ile Ser Ser Arg Phe Gly Cys Glu Ile Leu Cys Ser Leu Val
245 250 255

Thr Glu Ala Glu Ser Asp Asp Tyr Lys Ser Asp Thr Pro Thr Gly Met
260 265 270

```
Val His Gly Arg Leu Tyr Phe Asn Gly Thr Tyr Arg Glu Gly Leu Val
        275                 280                 285

Asp Thr Glu Thr Ile Phe Arg Asp Phe Ser Ala Asn Tyr Pro Gly Val
    290                 295                 300

Gly Ser Gly Glu Ile Val Glu Gly His Ile His Phe Pro Ile Tyr Gly
305                 310                 315                 320

Gly Val Lys Gln Asn Thr Gly Leu Tyr Asn Ser Leu Thr Pro Tyr Trp
                325                 330                 335

Leu Asp Ala Lys Asn Lys Tyr Asp Tyr Cys Lys Leu Pro Tyr Thr Asn
            340                 345                 350

Gln Thr Ile Gln Asn Ser Tyr Lys Pro Pro Phe Ile His Gly Arg Phe
        355                 360                 365

Trp Ala Gln Gly Ile Leu Ser Cys Glu Leu Asp Leu Phe Asn Leu Gly
    370                 375                 380

Asn Cys Asn Leu Lys Ile Ile Arg Ser Asp Lys Val Met Met Gly Ala
385                 390                 395                 400

Glu Ser Arg Leu Met Leu Val Gly Ser Lys Leu Leu Met Tyr Gln Arg
                405                 410                 415

Ala Ser Ser Trp Trp Pro Leu Gly Ile Thr Gln Glu Ile Asp Ile Ala
            420                 425                 430

Glu Leu His Ser Ser Asn Thr Thr Ile Leu Arg Glu Val Lys Pro Ile
        435                 440                 445

Leu Ser Ser Lys Phe Pro Arg Pro Ser Tyr Gln Pro Asn Tyr Cys Thr
    450                 455                 460

Lys Pro Ser Val Cys Pro Ala Val Cys Val Thr Gly Val Tyr Thr Asp
465                 470                 475                 480

Met Trp Pro Ile Ser Ile Thr Gly Asn Ile Ser Asp Tyr Ala Trp Ile
                485                 490                 495

Ser His Tyr Leu Asp Ala Pro Thr Ser Arg Gln Gln Pro Arg Ile Gly
            500                 505                 510

Ile Ala Asn Gln Tyr Phe Trp Ile His Gln Thr Thr Ile Phe Pro Thr
        515                 520                 525

Asn Thr Gln Ser Ser Tyr Ser Thr Thr Thr Cys Phe Arg Asn Gln Val
    530                 535                 540
```

```
Arg Ser Arg Met Phe Cys Leu Ser Ile Ala Glu Phe Ala Asp Gly Val
545                 550                 555                 560


Phe Gly Glu Phe Arg Ile Val Pro Leu Leu Tyr Glu Leu Arg Val
                565                 570                 575


<210>  104
<211>  593
<212>  PRT
<213>  Tioman virus


<220>
<221>  MISC_FEATURE
<223>  attachment protein

<400>  104

Met Trp Ala Thr Ser Glu Ser Lys Ala Pro Ile Pro Ala Asn Ser Thr
1               5                   10                  15


Leu Asn Leu Val Asp Val Pro Leu Asp Glu Pro Gln Thr Ile Thr Lys
            20                  25                  30


His Arg Lys Gln Lys Arg Thr Gly Arg Leu Val Phe Arg Leu Leu Ser
        35                  40                  45


Leu Val Leu Ser Leu Met Thr Val Ile Leu Val Leu Val Ile Leu Ala
    50                  55                  60


Ser Trp Ser Gln Lys Ile Asn Ala Cys Ala Thr Lys Glu Gly Phe Asn
65                  70                  75                  80


Ser Leu Asp Leu Gln Ile Ser Gly Leu Val Lys Ser Ile Asn Ser Leu
                85                  90                  95


Ile Thr Glu Val Asn Gln Ile Ser Ile Thr Thr Ala Ile Asn Leu Pro
            100                 105                 110


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


Met Ile Arg Gln Cys Asn Ala Val Cys Lys Gly Pro Gly Glu Lys Pro
    130                 135                 140


Gly Ile Gln Asn Ile Arg Ile Asn Ile Pro Asn Asn Phe Ser Thr Tyr
145                 150                 155                 160


Leu Glu Leu Asn Asn Thr Val Lys Ser Ile Glu Leu Gln Arg Arg Pro
                165                 170                 175


Ala Leu Leu Ala Arg Pro Asn Pro Ile Pro Lys Ser Cys Ser Arg Phe
            180                 185                 190


Pro Ser Tyr Ser Val Asn Phe Gly Ile His Cys Phe Ala His Ala Ile
```

195 200 205

Thr Asp Gln Ser Cys Glu Leu Ser Asp Lys Thr Tyr Tyr Arg Leu Ala
210 215 220

Ile Gly Ile Ser Asp Lys Asn Leu Ser Asp Pro Ser Asp Val Lys Tyr
225 230 235 240

Ile Gly Glu Ala Phe Thr Pro Met Gly Leu Gln Ala Arg Gly Cys Ser
245 250 255

Val Ile Ser Ser Ile Tyr Gly Cys Tyr Leu Leu Cys Ser Lys Ser Asn
260 265 270

Gln Gly Tyr Glu Ala Asp Phe Gln Thr Gln Gly Phe His Gln Met Tyr
275 280 285

Ile Leu Phe Leu Ser Arg Asp Leu Lys Thr Thr Leu Phe Asn Asp Met
290 295 300

Ile Ser Ser Thr Thr Val Val Trp Asn Gly Leu Tyr Pro Gly Glu Gly
305 310 315 320

Ala Gly Ile Trp His Met Gly Tyr Leu Ile Phe Pro Leu Trp Gly Gly
325 330 335

Ile Lys Ile Gly Thr Pro Ala Ser Thr Ser Ile Leu Asn Ser Thr Leu
340 345 350

Asp Leu Pro Leu Val Gly Pro Ser Cys Lys Ser Thr Leu Glu Glu Asn
355 360 365

Asn Leu Ile Asn Lys Asp Val Leu Phe Ser Pro Tyr Phe Gly Glu Ser
370 375 380

Val Met Val Phe Gly Phe Leu Ser Cys Tyr Met Leu Ser Asn Val Pro
385 390 395 400

Thr His Cys Gln Val Glu Val Leu Asn Ser Ser Val Leu Gly Phe Gly
405 410 415

Ser Arg Ser Gln Leu Met Asp Leu Lys Gly Ile Val Tyr Leu Tyr Ile
420 425 430

Gln Ser Ala Gly Trp Tyr Ser Tyr Thr Gln Leu Phe Arg Leu Ser Leu
435 440 445

Gln Ser Arg Gly Tyr Lys Leu Thr Val Lys Gln Ile Arg Arg Ile Pro
450 455 460

Ile Ser Ser Thr Thr Arg Pro Gly Thr Ala Pro Cys Asp Val Val His
465 470 475 480

```
Asn Cys Pro Tyr Thr Cys Ala Thr Gly Leu Phe Gln Ala Pro Trp Ile
                485                 490                 495

Val Asn Gly Asp Ser Ile Leu Asp Arg Asp Val Arg Asn Leu Val Phe
            500                 505                 510

Val Gln Ala Trp Ser Gly Asn Phe Asn Thr Phe Gln Lys Gly Leu Ile
        515                 520                 525

Ser Ile Cys Asn Gln Tyr Thr Cys Pro Leu Thr Thr Leu Leu Asp Asn
    530                 535                 540

Asp Asn Ser Ile Met Arg Ser Thr Thr Thr Tyr Cys Tyr Pro Ser Leu
545                 550                 555                 560

Ser Glu Tyr Asn Leu Gln Cys Gln Ser Phe Ile Glu Trp Gly Gly Pro
                565                 570                 575

Val Gly Asn Pro Ile Gly Ile Leu Glu Val His Tyr Ile Ile Lys Phe
            580                 585                 590

Lys
```

<210> 105
<211> 604
<212> PRT
<213> Dolphin morbillivirus

<220>
<221> MISC_FEATURE
<223> haemagglutinin protein

<400> 105

```
Met Ser Ser Pro Arg Asp Lys Val Asp Ala Phe Tyr Lys Asp Ile Pro
1               5                   10                  15

Arg Pro Arg Asn Asn Arg Val Leu Leu Asp Asn Glu Arg Val Ile Ile
            20                  25                  30

Glu Arg Pro Leu Ile Leu Val Gly Val Leu Ala Val Met Phe Leu Ser
        35                  40                  45

Leu Val Gly Leu Leu Ala Ile Ala Gly Val Arg Leu Gln Lys Ala Thr
    50                  55                  60

Thr Asn Ser Ile Glu Val Asn Arg Lys Leu Ser Thr Asn Leu Glu Thr
65                  70                  75                  80

Thr Val Ser Ile Glu His His Val Lys Asp Val Leu Thr Pro Leu Phe
                85                  90                  95
```

```
Lys Ile Ile Gly Asp Glu Val Gly Leu Arg Met Pro Gln Lys Leu Thr
            100                 105                 110

Glu Ile Met Gln Phe Ile Ser Asn Lys Ile Lys Phe Leu Asn Pro Asp
        115                 120                 125

Arg Glu Tyr Asp Phe Asn Asp Leu His Trp Cys Val Asn Pro Pro Asp
    130                 135                 140

Gln Val Lys Ile Asp Tyr Ala Gln Tyr Cys Asn His Ile Ala Ala Glu
145                 150                 155                 160

Glu Leu Ile Val Thr Lys Phe Lys Glu Leu Met Asn His Ser Leu Asp
                165                 170                 175

Met Ser Lys Gly Arg Ile Phe Pro Pro Lys Asn Cys Ser Gly Ser Val
            180                 185                 190

Ile Thr Arg Gly Gln Thr Ile Lys Pro Gly Leu Thr Leu Val Asn Ile
        195                 200                 205

Tyr Thr Thr Arg Asn Phe Glu Val Ser Phe Met Val Thr Val Ile Ser
    210                 215                 220

Gly Gly Met Tyr Gly Lys Thr Tyr Phe Leu Lys Pro Pro Glu Pro Asp
225                 230                 235                 240

Asp Pro Phe Glu Phe Gln Ala Phe Arg Ile Phe Glu Val Gly Leu Val
                245                 250                 255

Arg Asp Val Gly Ser Arg Glu Pro Val Leu Gln Met Thr Asn Phe Met
            260                 265                 270

Val Ile Asp Glu Asp Glu Gly Leu Asn Phe Cys Leu Leu Ser Val Gly
        275                 280                 285

Glu Leu Arg Leu Ala Ala Val Cys Val Arg Gly Arg Pro Val Val Thr
    290                 295                 300

Lys Asp Ile Gly Gly Tyr Lys Asp Glu Pro Phe Lys Val Val Thr Leu
305                 310                 315                 320

Gly Ile Ile Gly Gly Gly Leu Ser Asn Gln Lys Thr Glu Ile Tyr Pro
                325                 330                 335

Thr Ile Asp Ser Ser Ile Glu Lys Leu Tyr Ile Thr Ser His Arg Gly
            340                 345                 350

Ile Ile Arg Asn Ser Lys Ala Arg Trp Ser Val Pro Ala Ile Arg Ser
        355                 360                 365
```

```
Asp Asp Lys Asp Lys Met Glu Lys Cys Thr Gln Ala Leu Cys Lys Ser
    370                 375                 380


Arg Pro Pro Pro Ser Cys Asn Ser Ser Asp Trp Glu Pro Leu Thr Ser
385                 390                 395                 400


Asn Arg Ile Pro Ala Tyr Ala Tyr Ile Ala Leu Glu Ile Lys Glu Asp
                405                 410                 415


Ser Gly Leu Glu Leu Asp Ile Thr Ser Asn Tyr Gly Pro Leu Ile Ile
            420                 425                 430


His Gly Ala Gly Met Asp Ile Tyr Glu Gly Pro Ser Ser Asn Gln Asp
        435                 440                 445


Trp Leu Ala Ile Pro Pro Leu Ser Gln Ser Val Leu Gly Val Ile Asn
    450                 455                 460


Lys Val Asp Phe Thr Ala Gly Phe Asp Ile Lys Pro His Thr Leu Thr
465                 470                 475                 480


Thr Ala Val Asp Tyr Glu Ser Gly Lys Cys Tyr Val Pro Val Glu Leu
                485                 490                 495


Ser Gly Ala Lys Asp Gln Asp Leu Lys Leu Glu Ser Asn Leu Val Val
            500                 505                 510


Leu Pro Thr Lys Asp Phe Gly Tyr Val Thr Ala Thr Tyr Asp Thr Ser
        515                 520                 525


Arg Ser Glu His Ala Ile Val Tyr Tyr Val Tyr Asp Thr Ala Arg Ser
    530                 535                 540


Ser Ser Tyr Phe Phe Pro Phe Arg Ile Lys Ala Arg Gly Glu Pro Ile
545                 550                 555                 560


Tyr Leu Arg Ile Glu Cys Phe Pro Trp Ser Arg Gln Leu Trp Cys His
                565                 570                 575


His Tyr Cys Met Ile Asn Ser Thr Val Ser Asn Glu Ile Val Val Val
            580                 585                 590


Asp Asn Leu Val Ser Ile Asn Met Ser Cys Ser Arg
        595                 600


<210>  106
<211>  734
<212>  PRT
<213>  Beilong virus


<220>
<221>  MISC_FEATURE
<223>  attachment glycoprotein
```

<400> 106

Met Ser Gln Leu Ala Ala His Asn Leu Ala Met Ser Asn Phe Tyr Gly
1 5 10 15

Thr His Gln Gly Asp Leu Ser Gly Ser Gln Lys Gly Glu Glu Gln Gln
20 25 30

Val Gln Gly Val Ile Arg Tyr Val Ser Met Ile Val Ser Leu Leu Ser
35 40 45

Leu Phe Thr Ile Ile Ala Leu Asn Val Thr Asn Ile Ile Tyr Met Thr
50 55 60

Glu Ser Gly Gly Thr Met Gln Ser Ile Lys Thr Ala Gln Gly Ser Ile
65 70 75 80

Asp Gly Ser Met Arg Glu Ile Ser Gly Val Ile Met Glu Asp Val Lys
85 90 95

Pro Lys Thr Asp Leu Ile Asn Ser Met Val Ser Tyr Asn Ile Pro Ala
100 105 110

Gln Leu Ser Met Ile His Gln Ile Ile Lys Asn Asp Val Pro Lys Gln
115 120 125

Cys Thr Pro Ser Phe Met Phe Asn Asn Thr Ile Cys Pro Leu Ala Glu
130 135 140

Asn Pro Thr His Ser Arg Tyr Phe Glu Glu Val Asn Leu Asp Ser Ile
145 150 155 160

Ser Glu Cys Ser Gly Pro Asp Met His Leu Gly Leu Gly Val Asn Pro
165 170 175

Glu Phe Ile Glu Phe Pro Ser Phe Ala Pro Gly Ser Thr Lys Pro Gly
180 185 190

Ser Cys Val Arg Leu Pro Ser Phe Ser Leu Ser Thr Thr Val Phe Ala
195 200 205

Tyr Thr His Thr Ile Met Gly His Gly Cys Ser Glu Leu Asp Val Gly
210 215 220

Asp His Tyr Phe Ser Val Gly Arg Ile Ala Asp Ala Gly His Glu Ile
225 230 235 240

Pro Gln Phe Glu Thr Ile Ser Ser Trp Phe Ile Asn Asp Lys Ile Asn
245 250 255

Arg Arg Ser Cys Thr Val Ala Ala Gly Ala Met Glu Ala Trp Met Gly
260 265 270

```
Cys Val Ile Met Thr Glu Thr Phe Tyr Asp Asp Arg Asn Ser Leu Asp
        275                 280                 285

Thr Gly Lys Leu Thr Ile Ser Tyr Leu Asp Val Phe Gly Arg Lys Lys
    290                 295                 300

Glu Trp Ile Tyr Thr Arg Ser Glu Ile Leu Tyr Asp Tyr Thr Tyr Thr
305                 310                 315                 320

Ser Val Tyr Phe Ser Val Gly Ser Gly Val Val Val Gly Asp Thr Val
                325                 330                 335

Tyr Phe Leu Ile Trp Gly Ser Leu Ser Ser Pro Ile Glu Glu Thr Ala
            340                 345                 350

Tyr Cys Phe Ala Pro Asp Cys Ser Asn Tyr Asn Gln Arg Met Cys Asn
        355                 360                 365

Glu Ala Gln Arg Pro Ser Lys Phe Gly His Arg Gln Met Val Asn Gly
    370                 375                 380

Ile Leu Lys Phe Lys Thr Thr Ser Thr Gly Lys Pro Leu Leu Ser Val
385                 390                 395                 400

Gly Thr Leu Ser Pro Ser Val Val Pro Phe Gly Ser Glu Gly Arg Leu
                405                 410                 415

Met Tyr Ser Glu Ile Thr Lys Ile Ile Tyr Leu Tyr Leu Arg Ser Thr
            420                 425                 430

Ser Trp His Ala Leu Pro Leu Thr Gly Leu Phe Val Leu Gly Pro Pro
        435                 440                 445

Thr Ser Ile Ser Trp Ile Val Gln Arg Ala Val Ser Arg Pro Gly Glu
    450                 455                 460

Phe Pro Cys Gly Ala Ser Asn Arg Cys Pro Lys Asp Cys Val Thr Gly
465                 470                 475                 480

Val Tyr Thr Asp Leu Phe Pro Leu Gly Ser Arg Tyr Glu Tyr Ala Ala
                485                 490                 495

Thr Val Tyr Leu Asn Ser Glu Thr Tyr Arg Val Asn Pro Thr Leu Ala
            500                 505                 510

Leu Ile Asn Gln Thr Asn Ile Ile Ala Ser Lys Lys Val Thr Thr Glu
        515                 520                 525

Ser Gln Arg Ala Gly Tyr Thr Thr Thr Thr Cys Phe Val Phe Lys Leu
    530                 535                 540
```

```
Arg Val Trp Cys Ile Ser Val Val Glu Leu Ala Pro Ser Thr Met Thr
545                 550                 555                 560


Ala Tyr Glu Pro Ile Pro Phe Leu Tyr Gln Leu Asp Leu Thr Cys Lys
                565                 570                 575


Gly Lys Asn Gly Ser Leu Ala Met Arg Phe Ala Gly Lys Glu Gly Thr
            580                 585                 590


Tyr Lys Ser Gly Arg Tyr Lys Ser Pro Arg Asn Glu Cys Phe Phe Glu
        595                 600                 605


Lys Val Ser Asn Lys Tyr Tyr Phe Ile Val Ser Thr Pro Glu Gly Ile
    610                 615                 620


Gln Pro Tyr Glu Ile Arg Asp Leu Thr Pro Asp Arg Met Pro His Ile
625                 630                 635                 640


Ile Met Tyr Ile Ser Asp Val Cys Ala Pro Ala Leu Ser Ala Phe Lys
                645                 650                 655


Lys Leu Leu Pro Ala Met Arg Pro Ile Thr Thr Leu Thr Ile Gly Asn
            660                 665                 670


Trp Gln Phe Arg Pro Val Glu Val Ser Gly Gly Leu Arg Val Asn Ile
        675                 680                 685


Gly Arg Asn Leu Thr Lys Glu Gly Asp Leu Thr Met Ser Ala Pro Glu
    690                 695                 700


Asp Pro Gly Ser Asn Thr Phe Pro Gly Asn His Ile Pro Gly Asn Gly
705                 710                 715                 720


Ile Leu Asp Ala Gly Tyr Tyr Thr Val Glu Tyr Pro Lys Glu
                725                 730


<210>  107
<211>  582
<212>  PRT
<213>  Mapuera virus


<220>
<221>  MISC_FEATURE
<223>  attachment protein

<400>  107

Met Ala Ser Leu Gln Ser Glu Pro Gly Ser Gln Lys Pro His Tyr Gln
1               5                   10                  15


Ser Asp Asp Gln Leu Val Lys Arg Thr Trp Arg Ser Phe Phe Arg Phe
            20                  25                  30
```

```
Ser Val Leu Val Val Thr Ile Thr Ser Leu Ala Leu Ser Ile Ile Thr
        35                  40                  45

Leu Ile Gly Val Asn Arg Ile Ser Thr Ala Lys Gln Ile Ser Asn Ala
    50                  55                  60

Phe Ala Ala Ile Gln Ala Asn Ile Leu Ser Ser Ile Pro Asp Ile Arg
65                  70                  75                  80

Pro Ile Asn Ser Leu Leu Asn Gln Leu Val Tyr Thr Ser Ser Val Thr
                85                  90                  95

Leu Pro Leu Arg Ile Ser Ser Leu Glu Ser Asn Val Leu Ala Ala Ile
            100                 105                 110

Gln Glu Ala Cys Thr Tyr Arg Asp Ser Gln Ser Ser Cys Ser Ala Thr
        115                 120                 125

Met Ser Val Met Asn Asp Gln Arg Tyr Ile Glu Gly Ile Gln Val Tyr
    130                 135                 140

Ser Gly Ser Phe Leu Asp Leu Gln Lys His Thr Leu Ser Pro Pro Ile
145                 150                 155                 160

Ala Phe Pro Ser Phe Ile Pro Thr Ser Thr Thr Thr Val Gly Cys Thr
                165                 170                 175

Arg Ile Pro Ser Phe Ser Leu Thr Lys Thr His Trp Cys Tyr Thr His
            180                 185                 190

Asn Tyr Ile Lys Thr Gly Cys Arg Asp Ala Thr Gln Ser Asn Gln Tyr
        195                 200                 205

Ile Ala Leu Gly Thr Ile Tyr Thr Asp Pro Asp Gly Thr Pro Gly Phe
    210                 215                 220

Ser Thr Ser Arg Ser Gln Tyr Leu Asn Asp Gly Val Asn Arg Lys Ser
225                 230                 235                 240

Cys Ser Ile Ser Ala Val Pro Met Gly Cys Ala Leu Tyr Cys Phe Ile
                245                 250                 255

Ser Val Lys Glu Glu Val Asp Tyr Tyr Lys Gly Thr Val Pro Pro Ala
            260                 265                 270

Gln Thr Leu Ile Leu Phe Phe Phe Asn Gly Thr Val His Glu His Arg
        275                 280                 285

Ile Val Pro Ser Ser Met Asn Ser Glu Trp Val Met Leu Ser Pro Gly
    290                 295                 300

Val Gly Ser Gly Val Phe Tyr Asn Asn Tyr Ile Ile Phe Pro Leu Tyr
```

305 310 315 320

Gly Gly Met Thr Lys Asp Lys Ala Glu Lys Arg Gly Glu Leu Thr Arg
325 330 335

Phe Phe Thr Pro Lys Asn Ser Arg Ser Leu Cys Lys Met Asn Asp Ser
340 345 350

Val Phe Ser Asn Ala Ala Gln Ser Ala Tyr Tyr Pro Pro Tyr Phe Ser
355 360 365

Ser Arg Trp Ile Arg Ser Gly Leu Leu Ala Cys Asn Trp Asn Gln Ile
370 375 380

Ile Thr Thr Asn Cys Glu Ile Leu Thr Phe Ser Asn Gln Val Met Met
385 390 395 400

Met Gly Ala Glu Gly Arg Leu Ile Leu Ile Asn Asp Asp Leu Phe Tyr
405 410 415

Tyr Gln Arg Ser Thr Ser Trp Trp Pro Arg Pro Leu Val Tyr Lys Leu
420 425 430

Asp Ile Glu Leu Asn Tyr Pro Asp Ser His Ile Gln Arg Val Asp Gln
435 440 445

Val Glu Val Thr Phe Pro Thr Arg Pro Gly Trp Gly Gly Cys Val Gly
450 455 460

Asn Asn Phe Cys Pro Met Ile Cys Val Ser Gly Val Tyr Gln Asp Val
465 470 475 480

Trp Pro Val Thr Asn Pro Val Asn Thr Thr Asp Ser Arg Thr Leu Trp
485 490 495

Val Gly Gly Thr Leu Leu Ser Asn Thr Thr Arg Glu Asn Pro Ala Ser
500 505 510

Val Val Thr Ser Gly Gly Ser Ile Ser Gln Thr Val Ser Trp Phe Asn
515 520 525

Gln Thr Val Pro Gly Ala Tyr Ser Thr Thr Thr Cys Phe Asn Asp Gln
530 535 540

Val Gln Gly Arg Ile Phe Cys Leu Ile Ile Phe Glu Val Gly Gly Gly
545 550 555 560

Leu Leu Gly Glu Tyr Gln Ile Val Pro Phe Leu Lys Glu Leu Lys Tyr
565 570 575

Gln Gly Ala Val His Ala
580

<210> 108
<211> 657
<212> PRT
<213> Nariva virus

<220>
<221> MISC_FEATURE
<223> attachment protein

<400> 108

Met Ala Pro Ile Asn Tyr Pro Ala Ser Tyr Tyr Thr Asn Asn Ala Glu
1 5 10 15

Arg Pro Val Val Ile Thr Thr Lys Ser Thr Glu Ser Lys Gly Gln Arg
20 25 30

Pro Leu Pro Leu Gly Asn Ala Arg Phe Trp Glu Tyr Phe Gly His Val
35 40 45

Cys Gly Thr Leu Thr Phe Cys Met Ser Leu Ile Gly Ile Ile Val Gly
50 55 60

Ile Ile Ala Leu Ala Asn Tyr Ser Ser Asp Lys Asp Trp Lys Gly Arg
65 70 75 80

Ile Gly Gly Asp Ile Gln Val Thr Arg Met Ala Thr Glu Lys Thr Val
85 90 95

Lys Leu Ile Leu Glu Asp Thr Thr Pro Lys Leu Arg Asn Ile Leu Asp
100 105 110

Ser Val Leu Phe Gln Leu Pro Lys Met Leu Ala Ser Ile Ala Ser Lys
115 120 125

Ile Asn Thr Gln Thr Pro Pro Pro Pro Thr Thr Ser Gly His Ser Thr
130 135 140

Ala Leu Ala Thr Gln Cys Ser Ser Asn Cys Glu Asn Arg Pro Glu Ile
145 150 155 160

Gly Tyr Asp Tyr Leu Arg Gln Val Glu Gln Ser Leu Gln Arg Ile Thr
165 170 175

Asn Ile Ser Ile Gln Leu Leu Glu Ala Ser Glu Ile His Ser Met Ala
180 185 190

Gly Ala Tyr Pro Asn Ala Leu Tyr Lys Ile Arg Thr Gln Asp Ser Trp
195 200 205

Ser Val Thr Ala Lys Glu Cys Pro Leu Gln Ala Phe Gln Pro Asn Leu
210 215 220

```
Asn Leu Ile Pro Ala Met Ile Gly Thr Ala Thr Gly Ala Leu Ile Arg
225                 230                 235                 240

Asn Cys Val Arg Gln Pro Val Ile Val Val Asp Asp Gly Val Tyr Met
                245                 250                 255

Leu Thr Tyr Leu Ala Met Arg Gly Ser Cys Gln Asp His Gln Lys Ser
            260                 265                 270

Val Arg His Phe Glu Met Gly Val Ile Thr Ser Asp Pro Phe Gly Asp
        275                 280                 285

Pro Val Pro Thr Pro Leu Arg His Trp Thr Lys Arg Ala Leu Pro Ala
    290                 295                 300

Tyr Asp Gly Cys Ala Leu Ala Val Lys Gly His Ala Gly Phe Ala Leu
305                 310                 315                 320

Cys Thr Glu Thr Ser Val Gly Pro Leu Arg Asp Arg Thr Ala Lys Arg
                325                 330                 335

Lys Pro Asn Ile Val Leu Phe Lys Ala Ser Leu Val Gly Glu Leu Ser
            340                 345                 350

Glu Arg Val Ile Pro Pro Gln Ser Trp Leu Ser Gly Phe Ser Phe Phe
        355                 360                 365

Ser Val Tyr Thr Val Ala Gly Lys Gly Tyr Ala Tyr His Ser Lys Phe
    370                 375                 380

His Ala Phe Gly Asn Val Val Arg Val Gly Gln Ser Glu Tyr Gln Ala
385                 390                 395                 400

Lys Cys Arg Gly Thr Gly Cys Pro Thr Ala Asn Gln Asp Asp Cys Asn
                405                 410                 415

Thr Ala Gln Arg Val Ser Gln Glu Asp Asn Thr Tyr Leu His Gln Ala
            420                 425                 430

Ile Leu Ser Val Asp Ile Asp Ser Val Ile Asp Pro Glu Asp Val Val
        435                 440                 445

Tyr Val Ile Glu Arg Asp Gln Tyr Tyr Gln Ala Ser Ala Gly Asp Leu
    450                 455                 460

Tyr Arg Val Pro Glu Thr Gly Glu Ile Leu Tyr Asn Leu His Asn Gly
465                 470                 475                 480

Gly Trp Ser Asn Glu Val Gln Val Gly Arg Ile Gln Pro Ser Asp Arg
                485                 490                 495
```

Phe Tyr Met Arg Glu Ile Gln Leu Thr Ser Thr Arg Val Pro Ala Pro
500 505 510

Asn Gly Cys Asn Arg Val Lys Gly Cys Pro Gly Gly Cys Val Ala Val
515 520 525

Ile Ser Pro Ala Phe Thr Pro Met His Pro Glu Phe Asn Val Gly Val
530 535 540

Gly Ile Phe Pro Met Asn Gln Pro His Asn Pro Ser Ile Met His Val
545 550 555 560

Gln Gln Gln Thr Glu Leu Phe Trp Lys Pro Ile Val Gly Gly Asn Ile
565 570 575

Thr Leu His Glu Ser Ser Ile Ala Cys Tyr Ser Thr Val Pro Pro Asn
580 585 590

Pro Ser Tyr Asp Leu Cys Ile Gly Val Met Thr Leu Leu Leu His Gln
595 600 605

Gly Gln Leu Pro Gln Phe Gln Ala Leu Ser Trp Tyr Gln Pro Thr Met
610 615 620

Cys Asn Gly Asn Ala Pro Gln Asn Arg Arg Ala Leu Ile Pro Val Ile
625 630 635 640

Val Glu Asp Ser Lys Ala Met Ser Val Ser Ser Asp Ala Pro Arg Thr
645 650 655

Pro

<210> 109
<211> 632
<212> PRT
<213> Bat paramyxovirus Eid_hel/GH-M74a/GHA/2009

<220>
<221> MISC_FEATURE
<223> glycoprotein

<400> 109

Met Pro Gln Lys Thr Val Glu Phe Ile Asn Met Asn Ser Pro Leu Glu
1 5 10 15

Arg Gly Val Ser Thr Leu Ser Asp Lys Lys Thr Leu Asn Gln Ser Lys
20 25 30

Ile Thr Lys Gln Gly Tyr Phe Gly Leu Gly Ser His Ser Glu Arg Asn
35 40 45

Trp Lys Lys Gln Lys Asn Gln Asn Asp His Tyr Met Thr Val Ser Thr

50 55 60

Met Ile Leu Glu Ile Leu Val Val Leu Gly Ile Met Phe Asn Leu Ile
65 70 75 80

Val Leu Thr Met Val Tyr Tyr Gln Asn Asp Asn Ile Asn Gln Arg Met
85 90 95

Ala Glu Leu Thr Ser Asn Ile Thr Val Leu Asn Leu Asn Leu Asn Gln
100 105 110

Leu Thr Asn Lys Ile Gln Arg Glu Ile Ile Pro Arg Ile Thr Leu Ile
115 120 125

Asp Thr Ala Thr Thr Ile Thr Ile Pro Ser Ala Ile Thr Tyr Ile Leu
130 135 140

Ala Thr Leu Thr Thr Arg Ile Ser Glu Leu Leu Pro Ser Ile Asn Gln
145 150 155 160

Lys Cys Glu Phe Lys Thr Pro Thr Leu Val Leu Asn Asp Cys Arg Ile
165 170 175

Asn Cys Thr Pro Pro Leu Asn Pro Ser Asp Gly Val Lys Met Ser Ser
180 185 190

Leu Ala Thr Asn Leu Val Ala His Gly Pro Ser Pro Cys Arg Asn Phe
195 200 205

Ser Ser Val Pro Thr Ile Tyr Tyr Tyr Arg Ile Pro Gly Leu Tyr Asn
210 215 220

Arg Thr Ala Leu Asp Glu Arg Cys Ile Leu Asn Pro Arg Leu Thr Ile
225 230 235 240

Ser Ser Thr Lys Phe Ala Tyr Val His Ser Glu Tyr Asp Lys Asn Cys
245 250 255

Thr Arg Gly Phe Lys Tyr Tyr Glu Leu Met Thr Phe Gly Glu Ile Leu
260 265 270

Glu Gly Pro Glu Lys Glu Pro Arg Met Phe Ser Arg Ser Phe Tyr Ser
275 280 285

Pro Thr Asn Ala Val Asn Tyr His Ser Cys Thr Pro Ile Val Thr Val
290 295 300

Asn Glu Gly Tyr Phe Leu Cys Leu Glu Cys Thr Ser Ser Asp Pro Leu
305 310 315 320

Tyr Lys Ala Asn Leu Ser Asn Ser Thr Phe His Leu Val Ile Leu Arg
325 330 335

```
His Asn Lys Asp Glu Lys Ile Val Ser Met Pro Ser Phe Asn Leu Ser
            340                 345                 350

Thr Asp Gln Glu Tyr Val Gln Ile Ile Pro Ala Glu Gly Gly Gly Thr
        355                 360                 365

Ala Glu Ser Gly Asn Leu Tyr Phe Pro Cys Ile Gly Arg Leu Leu His
    370                 375                 380

Lys Arg Val Thr His Pro Leu Cys Lys Lys Ser Asn Cys Ser Arg Thr
385                 390                 395                 400

Asp Asp Glu Ser Cys Leu Lys Ser Tyr Tyr Asn Gln Gly Ser Pro Gln
                405                 410                 415

His Gln Val Val Asn Cys Leu Ile Arg Ile Arg Asn Ala Gln Arg Asp
            420                 425                 430

Asn Pro Thr Trp Asp Val Ile Thr Val Asp Leu Thr Asn Thr Tyr Pro
        435                 440                 445

Gly Ser Arg Ser Arg Ile Phe Gly Ser Phe Ser Lys Pro Met Leu Tyr
    450                 455                 460

Gln Ser Ser Val Ser Trp His Thr Leu Leu Gln Val Ala Glu Ile Thr
465                 470                 475                 480

Asp Leu Asp Lys Tyr Gln Leu Asp Trp Leu Asp Thr Pro Tyr Ile Ser
                485                 490                 495

Arg Pro Gly Gly Ser Glu Cys Pro Phe Gly Asn Tyr Cys Pro Thr Val
            500                 505                 510

Cys Trp Glu Gly Thr Tyr Asn Asp Val Tyr Ser Leu Thr Pro Asn Asn
        515                 520                 525

Asp Leu Phe Val Thr Val Tyr Leu Lys Ser Glu Gln Val Ala Glu Asn
    530                 535                 540

Pro Tyr Phe Ala Ile Phe Ser Arg Asp Gln Ile Leu Lys Glu Phe Pro
545                 550                 555                 560

Leu Asp Ala Trp Ile Ser Ser Ala Arg Thr Thr Thr Ile Ser Cys Phe
                565                 570                 575

Met Phe Asn Asn Glu Ile Trp Cys Ile Ala Ala Leu Glu Ile Thr Arg
            580                 585                 590

Leu Asn Asp Asp Ile Ile Arg Pro Ile Tyr Tyr Ser Phe Trp Leu Pro
        595                 600                 605
```

```
Thr Asp Cys Arg Thr Pro Tyr Pro His Thr Gly Lys Met Thr Arg Val
    610                 615                 620


Pro Leu Arg Ser Thr Tyr Asn Tyr
625                 630


<210>  110
<211>  569
<212>  PRT
<213>  Avian paramyxovirus 7


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase

<400>  110

Met Glu Ser Ile Gly Lys Gly Thr Trp Arg Thr Val Tyr Arg Val Leu
1               5                   10                  15


Thr Ile Leu Leu Asp Val Val Ile Ile Ile Leu Ser Val Ile Ala Leu
            20                  25                  30


Ile Ser Leu Gly Leu Lys Pro Gly Glu Arg Ile Ile Asn Glu Val Asn
        35                  40                  45


Gly Ser Ile His Asn Gln Leu Val Pro Leu Ser Gly Ile Thr Ser Asp
    50                  55                  60


Ile Gln Ala Lys Val Ser Ser Ile Tyr Arg Ser Asn Leu Leu Ser Ile
65                  70                  75                  80


Pro Leu Gln Leu Asp Gln Ile Asn Gln Ala Ile Ser Ser Ser Ala Arg
                85                  90                  95


Gln Ile Ala Asp Thr Ile Asn Ser Phe Leu Ala Leu Asn Gly Ser Gly
            100                 105                 110


Thr Phe Ile Tyr Thr Asn Ser Pro Glu Phe Ala Asn Gly Phe Asn Arg
        115                 120                 125


Ala Met Phe Pro Thr Leu Asn Gln Ser Leu Asn Met Leu Thr Pro Gly
    130                 135                 140


Asn Leu Ile Glu Phe Thr Asn Phe Ile Pro Thr Pro Thr Thr Lys Ser
145                 150                 155                 160


Gly Cys Ile Arg Ile Pro Ser Phe Ser Met Ser Ser Ser His Trp Cys
                165                 170                 175


Tyr Thr His Asn Ile Ile Ala Ser Gly Cys Gln Asp His Ser Thr Ser
            180                 185                 190
```

```
Ser Glu Tyr Ile Ser Met Gly Val Val Glu Val Thr Asp Gln Ala Tyr
        195                 200                 205

Pro Asn Phe Arg Thr Thr Leu Ser Ile Thr Leu Ala Asp Asn Leu Asn
    210                 215                 220

Arg Lys Ser Cys Ser Ile Ala Ala Thr Gly Phe Gly Cys Asp Ile Leu
225                 230                 235                 240

Cys Ser Val Val Thr Glu Thr Glu Asn Asp Asp Tyr Gln Ser Pro Glu
                245                 250                 255

Pro Thr Gln Met Ile Tyr Gly Arg Leu Phe Phe Asn Gly Thr Tyr Ser
            260                 265                 270

Glu Met Ser Leu Asn Val Asn Gln Met Phe Ala Asp Trp Val Ala Asn
        275                 280                 285

Tyr Pro Ala Val Gly Ser Gly Val Glu Leu Ala Asp Phe Val Ile Phe
    290                 295                 300

Pro Leu Tyr Gly Gly Val Lys Ile Thr Ser Thr Leu Gly Ala Ser Leu
305                 310                 315                 320

Ser Gln Tyr Tyr Tyr Ile Pro Lys Val Pro Thr Val Asn Cys Ser Glu
                325                 330                 335

Thr Asp Ala Gln Gln Ile Glu Lys Ala Lys Ala Ser Tyr Ser Pro Pro
            340                 345                 350

Lys Val Ala Pro Asn Ile Trp Ala Gln Ala Val Val Arg Cys Asn Lys
        355                 360                 365

Ser Val Asn Leu Ala Asn Ser Cys Glu Ile Leu Thr Phe Asn Thr Ser
    370                 375                 380

Thr Met Met Met Gly Ala Glu Gly Arg Leu Leu Met Ile Gly Lys Asn
385                 390                 395                 400

Val Tyr Phe Tyr Gln Arg Ser Ser Ser Tyr Trp Pro Val Gly Ile Ile
                405                 410                 415

Tyr Lys Leu Asp Leu Gln Glu Leu Thr Thr Phe Ser Ser Asn Gln Leu
            420                 425                 430

Leu Ser Thr Ile Pro Ile Pro Phe Glu Lys Phe Pro Arg Pro Ala Ser
        435                 440                 445

Thr Ala Gly Val Cys Ser Lys Pro Asn Val Cys Pro Ala Val Cys Gln
    450                 455                 460

Thr Gly Val Tyr Gln Asp Leu Trp Val Leu Tyr Asp Leu Gly Lys Leu
```

465 470 475 480

Glu Asn Thr Thr Ala Val Gly Leu Tyr Leu Asn Ser Ala Val Gly Arg
485 490 495

Met Asn Pro Phe Ile Gly Ile Ala Asn Thr Leu Ser Trp Tyr Asn Thr
500 505 510

Thr Arg Leu Phe Ala Gln Gly Thr Pro Ala Ser Tyr Ser Thr Thr Thr
515 520 525

Cys Phe Lys Asn Thr Lys Ile Asp Thr Ala Tyr Cys Leu Ser Ile Leu
530 535 540

Glu Leu Ser Asp Ser Leu Leu Gly Ser Trp Arg Ile Thr Pro Leu Leu
545 550 555 560

Tyr Asn Ile Thr Leu Ser Ile Met Ser
565

<210> 111
<211> 588
<212> PRT
<213> Tuhoko virus

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase

<400> 111

Met Pro Pro Val Pro Thr Val Ser Gln Ser Ile Asp Glu Gly Ser Phe
1 5 10 15

Thr Asp Ile Pro Leu Ser Pro Asp Asp Ile Lys His Pro Leu Ser Lys
20 25 30

Lys Thr Cys Arg Lys Leu Phe Arg Ile Val Thr Leu Ile Gly Val Gly
35 40 45

Leu Ile Ser Ile Leu Thr Ile Ile Ser Leu Ala Gln Gln Thr Gly Ile
50 55 60

Leu Arg Lys Val Asp Ser Ser Asp Phe Gln Ser Tyr Val Gln Glu Ser
65 70 75 80

Phe Lys Gln Val Leu Asn Leu Met Lys Gln Phe Ser Ser Asn Leu Asn
85 90 95

Ser Leu Ile Glu Ile Thr Ser Val Thr Leu Pro Phe Arg Ile Asp Gln
100 105 110

Phe Gly Thr Asp Ile Lys Thr Gln Val Ala Gln Leu Val Arg Gln Cys
115 120 125

```
Asn Ala Val Cys Arg Gly Pro Ile Lys Gly Pro Thr Thr Gln Asn Ile
    130                 135                 140

Val Tyr Pro Ala Leu Tyr Glu Thr Ser Leu Asn Lys Thr Leu Glu Thr
145                 150                 155                 160

Lys Asn Val Arg Ile Gln Glu Val Arg Gln Glu Val Asp Pro Val Pro
                165                 170                 175

Gly Pro Gly Leu Ser Asn Gly Cys Thr Arg Asn Pro Ser Phe Ser Val
            180                 185                 190

Tyr His Gly Val Trp Cys Tyr Thr His Ala Thr Ser Ile Gly Asn Cys
        195                 200                 205

Asn Gly Ser Leu Gly Thr Ser Gln Leu Phe Arg Ile Gly Asn Val Leu
    210                 215                 220

Glu Gly Asp Gly Gly Ala Pro Tyr His Lys Ser Leu Ala Thr His Leu
225                 230                 235                 240

Leu Thr Thr Arg Asn Val Ser Arg Gln Cys Ser Ala Thr Ala Ser Tyr
                245                 250                 255

Tyr Gly Cys Tyr Phe Ile Cys Ser Glu Pro Val Leu Thr Glu Arg Asp
            260                 265                 270

Asp Tyr Glu Thr Pro Gly Ile Glu Pro Ile Thr Ile Phe Arg Leu Asp
        275                 280                 285

Pro Asp Gly Asn Trp Val Val Phe Pro Asn Ile Asn Arg Phe Thr Glu
    290                 295                 300

Tyr Ser Leu Lys Ala Leu Tyr Pro Gly Ile Gly Ser Gly Val Leu Phe
305                 310                 315                 320

Gln Gly Lys Leu Ile Phe Pro Met Tyr Gly Gly Ile Asp Lys Glu Arg
                325                 330                 335

Leu Ser Ala Leu Gly Leu Gly Asn Ile Gly Leu Ile Glu Arg Arg Met
            340                 345                 350

Ala Asp Thr Cys Asn His Thr Glu Lys Glu Leu Gly Arg Ser Phe Pro
        355                 360                 365

Gly Ala Phe Ser Ser Pro Tyr Tyr His Asp Ala Val Met Leu Asn Phe
    370                 375                 380

Leu Leu Ile Cys Glu Met Ile Glu Asn Leu Pro Gly Asp Cys Asp Leu
385                 390                 395                 400
```

Gln Ile Leu Asn Pro Thr Asn Met Ser Met Gly Ser Glu Ser Gln Leu
405 410 415

Ser Val Leu Asp Asn Glu Leu Phe Leu Tyr Gln Arg Ser Ala Ser Trp
420 425 430

Trp Pro Tyr Thr Leu Ile Tyr Arg Leu Asn Met Arg Tyr Thr Gly Lys
435 440 445

Tyr Leu Lys Pro Lys Ser Ile Ile Pro Met Val Ile Lys Ser Asn Thr
450 455 460

Arg Pro Gly Tyr Glu Gly Cys Asn His Glu Arg Val Cys Pro Lys Val
465 470 475 480

Cys Val Thr Gly Val Phe Gln Ala Pro Trp Ile Leu Ser Ile Gly Arg
485 490 495

Asp His Lys Glu Arg Val Ser Asn Val Thr Tyr Met Val Ala Trp Ser
500 505 510

Met Asp Lys Ser Asp Arg Thr Tyr Pro Ala Val Ser Val Cys Gly Ser
515 520 525

Asp Thr Cys Lys Leu Thr Val Pro Leu Gly Asp Ser Lys Val His Ser
530 535 540

Ala Tyr Ser Val Thr Arg Cys Tyr Leu Ser Arg Asp His Met Ser Ala
545 550 555 560

Tyr Cys Leu Val Ile Phe Glu Leu Asp Ala Arg Pro Trp Ala Glu Met
565 570 575

Arg Ile Gln Ser Phe Leu Tyr Lys Leu Ile Leu Thr
580 585

<210> 112
<211> 582
<212> PRT
<213> Tuhoko virus 3

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase

<400> 112

Met His Asn Arg Thr Gln Ser Val Ser Ser Ile Asp Thr Ser Ser Asp
1 5 10 15

Val Tyr Leu Pro Arg Arg Lys Lys Ala Val Thr Lys Phe Thr Phe Lys
20 25 30

```
Lys Ile Phe Arg Val Leu Ile Leu Thr Leu Leu Leu Ser Ile Ile Ile
        35                  40                  45

Ile Ile Ala Val Ile Phe Pro Lys Ile Asp His Ile Arg Glu Thr Cys
    50                  55                  60

Asp Asn Ser Gln Ile Leu Glu Thr Ile Thr Asn Gln Asn Ser Glu Ile
65                  70                  75                  80

Lys Asn Leu Ile Asn Ser Ala Ile Thr Asn Leu Asn Val Leu Leu Thr
                85                  90                  95

Ser Thr Thr Val Asp Leu Pro Ile Lys Leu Asn Asn Phe Gly Lys Ser
            100                 105                 110

Ile Val Asp Gln Val Thr Met Met Val Arg Gln Cys Asn Ala Val Cys
        115                 120                 125

Arg Gly Pro Gly Asp Arg Pro Thr Gln Asn Ile Glu Leu Phe Lys Gly
    130                 135                 140

Leu Tyr His Thr Ser Pro Pro Ser Asn Thr Ser Thr Lys Leu Ser Met
145                 150                 155                 160

Ile Thr Glu Ala Ser Asn Pro Asp Asp Ile Val Pro Arg Pro Gly Lys
                165                 170                 175

Leu Leu Gly Cys Thr Arg Phe Pro Ser Phe Ser Val His Tyr Gly Leu
            180                 185                 190

Trp Cys Tyr Gly His Met Ala Ser Thr Gly Asn Cys Ser Gly Ser Ser
        195                 200                 205

Pro Ser Val Gln Ile Ile Arg Ile Gly Ser Ile Gly Thr Asn Lys Asp
    210                 215                 220

Gly Thr Pro Lys Tyr Val Ile Ile Ala Ser Ala Ser Leu Pro Glu Thr
225                 230                 235                 240

Thr Arg Leu Tyr His Cys Ser Val Thr Met Thr Ser Ile Gly Cys Tyr
                245                 250                 255

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

Met Gly Ile Glu Lys Met Ser Ile Ser Phe Leu Ser Leu Asp Gly Tyr
        275                 280                 285

Leu Thr Gln Leu Gly Gln Pro Thr Gly Leu Asp Asn Gln Asn Leu Tyr
    290                 295                 300

Ala Leu Tyr Pro Gly Pro Gly Ser Gly Val Ile Phe Arg Asp Phe Leu
```

305 310 315 320

Ile Phe Pro Met Met Gly Gly Ile Arg Leu Met Asp Ala Gln Lys Met
325 330 335

Leu Asn Arg Asn Ile Thr Tyr Arg Gly Phe Pro Pro Ser Glu Thr Cys
340 345 350

Thr Glu Ser Glu Leu Lys Leu Lys Gln Glu Val Ala Asn Met Leu Thr
355 360 365

Ser Pro Tyr Tyr Gly Glu Val Leu Val Leu Asn Phe Leu Tyr Val Cys
370 375 380

Ser Leu Leu Asp Asn Ile Pro Gly Asp Cys Ser Val Gln Leu Ile Pro
385 390 395 400

Pro Asp Asn Met Thr Leu Gly Ala Glu Ser Arg Leu Tyr Val Leu Asn
405 410 415

Gly Ser Leu Ile Met Tyr Lys Arg Gly Ser Ser Trp Trp Pro Tyr Thr
420 425 430

Glu Leu Tyr Gln Ile Asn Tyr Arg Val Asn Asn Arg Ala Phe Arg Val
435 440 445

Arg Glu Ser Val Arg Ile Asn Thr Thr Ser Thr Ser Arg Pro Gly Val
450 455 460

Gln Gly Cys Asn Leu Glu Lys Val Cys Pro Lys Val Cys Val Ser Gly
465 470 475 480

Ile Tyr Gln Ser Pro Gly Ile Ile Ser Ala Pro Val Asn Pro Thr Arg
485 490 495

Gln Glu Glu Gly Leu Leu Tyr Phe Leu Val Trp Thr Ser Ser Met Ser
500 505 510

Ser Arg Thr Gly Pro Leu Ser Ser Leu Cys Asp His Ser Thr Cys Arg
515 520 525

Ile Thr Tyr Pro Ile Gly Asp Asp Thr Ile Phe Ile Gly Tyr Thr Asp
530 535 540

Ser Ser Cys Phe Met Ser Ser Ile Lys Glu Gly Ile Tyr Cys Ile Ala
545 550 555 560

Phe Leu Glu Leu Asp Asn Gln Pro Tyr Ser Met Met Ala Ile Arg Ser
565 570 575

Leu Ser Tyr Ile Ile Asn
580

<210> 113
<211> 625
<212> PRT
<213> Mojiang virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 113

Met Ala Thr Asn Arg Asp Asn Thr Ile Thr Ser Ala Glu Val Ser Gln
1 5 10 15

Glu Asp Lys Val Lys Lys Tyr Tyr Gly Val Glu Thr Ala Glu Lys Val
20 25 30

Ala Asp Ser Ile Ser Gly Asn Lys Val Phe Ile Leu Met Asn Thr Leu
35 40 45

Leu Ile Leu Thr Gly Ala Ile Ile Thr Ile Thr Leu Asn Ile Thr Asn
50 55 60

Leu Thr Ala Ala Lys Ser Gln Gln Asn Met Leu Lys Ile Ile Gln Asp
65 70 75 80

Asp Val Asn Ala Lys Leu Glu Met Phe Val Asn Leu Asp Gln Leu Val
85 90 95

Lys Gly Glu Ile Lys Pro Lys Val Ser Leu Ile Asn Thr Ala Val Ser
100 105 110

Val Ser Ile Pro Gly Gln Ile Ser Asn Leu Gln Thr Lys Phe Leu Gln
115 120 125

Lys Tyr Val Tyr Leu Glu Glu Ser Ile Thr Lys Gln Cys Thr Cys Asn
130 135 140

Pro Leu Ser Gly Ile Phe Pro Thr Ser Gly Pro Thr Tyr Pro Pro Thr
145 150 155 160

Asp Lys Pro Asp Asp Asp Thr Thr Asp Asp Asp Lys Val Asp Thr Thr
165 170 175

Ile Lys Pro Ile Glu Tyr Pro Lys Pro Asp Gly Cys Asn Arg Thr Gly
180 185 190

Asp His Phe Thr Met Glu Pro Gly Ala Asn Phe Tyr Thr Val Pro Asn
195 200 205

Leu Gly Pro Ala Ser Ser Asn Ser Asp Glu Cys Tyr Thr Asn Pro Ser
210 215 220

```
Phe Ser Ile Gly Ser Ser Ile Tyr Met Phe Ser Gln Glu Ile Arg Lys
225                 230                 235                 240

Thr Asp Cys Thr Ala Gly Glu Ile Leu Ser Ile Gln Ile Val Leu Gly
                245                 250                 255

Arg Ile Val Asp Lys Gly Gln Gln Gly Pro Gln Ala Ser Pro Leu Leu
            260                 265                 270

Val Trp Ala Val Pro Asn Pro Lys Ile Ile Asn Ser Cys Ala Val Ala
        275                 280                 285

Ala Gly Asp Glu Met Gly Trp Val Leu Cys Ser Val Thr Leu Thr Ala
    290                 295                 300

Ala Ser Gly Glu Pro Ile Pro His Met Phe Asp Gly Phe Trp Leu Tyr
305                 310                 315                 320

Lys Leu Glu Pro Asp Thr Glu Val Val Ser Tyr Arg Ile Thr Gly Tyr
                325                 330                 335

Ala Tyr Leu Leu Asp Lys Gln Tyr Asp Ser Val Phe Ile Gly Lys Gly
            340                 345                 350

Gly Gly Ile Gln Lys Gly Asn Asp Leu Tyr Phe Gln Met Tyr Gly Leu
        355                 360                 365

Ser Arg Asn Arg Gln Ser Phe Lys Ala Leu Cys Glu His Gly Ser Cys
    370                 375                 380

Leu Gly Thr Gly Gly Gly Gly Tyr Gln Val Leu Cys Asp Arg Ala Val
385                 390                 395                 400

Met Ser Phe Gly Ser Glu Glu Ser Leu Ile Thr Asn Ala Tyr Leu Lys
                405                 410                 415

Val Asn Asp Leu Ala Ser Gly Lys Pro Val Ile Ile Gly Gln Thr Phe
            420                 425                 430

Pro Pro Ser Asp Ser Tyr Lys Gly Ser Asn Gly Arg Met Tyr Thr Ile
        435                 440                 445

Gly Asp Lys Tyr Gly Leu Tyr Leu Ala Pro Ser Ser Trp Asn Arg Tyr
    450                 455                 460

Leu Arg Phe Gly Ile Thr Pro Asp Ile Ser Val Arg Ser Thr Thr Trp
465                 470                 475                 480

Leu Lys Ser Gln Asp Pro Ile Met Lys Ile Leu Ser Thr Cys Thr Asn
                485                 490                 495
```

Thr Asp Arg Asp Met Cys Pro Glu Ile Cys Asn Thr Arg Gly Tyr Gln
500 505 510

Asp Ile Phe Pro Leu Ser Glu Asp Ser Glu Tyr Tyr Thr Tyr Ile Gly
515 520 525

Ile Thr Pro Asn Asn Gly Gly Thr Lys Asn Phe Val Ala Val Arg Asp
530 535 540

Ser Asp Gly His Ile Ala Ser Ile Asp Ile Leu Gln Asn Tyr Tyr Ser
545 550 555 560

Ile Thr Ser Ala Thr Ile Ser Cys Phe Met Tyr Lys Asp Glu Ile Trp
565 570 575

Cys Ile Ala Ile Thr Glu Gly Lys Lys Gln Lys Asp Asn Pro Gln Arg
580 585 590

Ile Tyr Ala His Ser Tyr Lys Ile Arg Gln Met Cys Tyr Asn Met Lys
595 600 605

Ser Ala Thr Val Thr Val Gly Asn Ala Lys Asn Ile Thr Ile Arg Arg
610 615 620

Tyr
625

<210> 114
<211> 614
<212> PRT
<213> Avian paramyxovirus 12

<220>
<221> MISC_FEATURE
<223> hemagglutinin neuraminidase

<400> 114

Met Glu Ser Ala Thr Ser Gln Val Ser Phe Glu Asn Asp Lys Thr Ser
1 5 10 15

Asp Arg Arg Thr Trp Arg Ala Val Phe Arg Val Leu Met Ile Ile Leu
20 25 30

Ala Leu Ser Ser Leu Cys Val Thr Val Ala Ala Leu Ile Tyr Ser Ala
35 40 45

Lys Ala Ala Ile Pro Gly Asn Ile Asp Ala Ser Glu Gln Arg Ile Leu
50 55 60

Ser Ser Val Glu Ala Val Gln Val Pro Val Ser Arg Leu Glu Asp Thr
65 70 75 80

Ser Gln Lys Ile Tyr Arg Gln Val Ile Leu Glu Ala Pro Val Thr Gln

```
                85                  90                  95

Leu Asn Met Glu Thr Asn Ile Leu Asn Ala Ile Thr Ser Leu Ser Tyr
            100                 105                 110

Gln Ile Asp Ala Ser Ala Asn Ser Ser Gly Cys Gly Ala Pro Val His
        115                 120                 125

Asp Ser Asp Phe Thr Gly Gly Val Gly Arg Glu Leu Leu Gln Glu Ala
    130                 135                 140

Glu Val Asn Leu Thr Ile Ile Arg Pro Ser Lys Phe Leu Glu His Leu
145                 150                 155                 160

Asn Phe Ile Pro Ala Pro Thr Thr Gly Asn Gly Cys Thr Arg Ile Pro
                165                 170                 175

Ser Phe Asp Leu Gly Gln Thr His Trp Cys Tyr Thr His Asn Val Val
            180                 185                 190

Leu Asn Gly Cys Arg Asp Arg Gly His Ser Phe Gln Tyr Val Ala Leu
        195                 200                 205

Gly Ile Leu Arg Thr Ser Ala Thr Gly Ser Val Phe Leu Ser Thr Leu
    210                 215                 220

Arg Ser Val Asn Leu Asp Asp Asp Arg Asn Arg Lys Ser Cys Ser Val
225                 230                 235                 240

Ser Ala Thr Pro Ile Gly Cys Glu Met Leu Cys Ser Leu Val Thr Glu
                245                 250                 255

Thr Glu Glu Gly Asp Tyr Asp Ser Ile Asp Pro Thr Pro Met Val His
            260                 265                 270

Gly Arg Leu Gly Phe Asp Gly Lys Tyr Arg Glu Val Asp Leu Ser Glu
        275                 280                 285

Lys Glu Ile Phe Ala Asp Trp Arg Ala Asn Tyr Pro Ala Val Gly Gly
    290                 295                 300

Gly Ala Phe Phe Gly Asn Arg Val Trp Phe Pro Val Tyr Gly Gly Leu
305                 310                 315                 320

Lys Glu Gly Thr Gln Ser Glu Arg Asp Ala Glu Lys Gly Tyr Ala Ile
                325                 330                 335

Tyr Lys Arg Phe Asn Asn Thr Cys Pro Asp Asp Asn Thr Thr Gln Ile
            340                 345                 350

Ala Asn Ala Lys Ala Ser Tyr Arg Pro Ser Arg Phe Gly Gly Arg Phe
        355                 360                 365
```

Ile Gln Gln Gly Ile Leu Ser Phe Lys Val Glu Gly Asn Leu Gly Ser
370 375 380

Asp Pro Ile Leu Ser Leu Thr Asp Asn Ser Ile Thr Leu Met Gly Ala
385 390 395 400

Glu Ala Arg Val Met Asn Ile Glu Asn Lys Leu Tyr Leu Tyr Gln Arg
405 410 415

Gly Thr Ser Trp Phe Pro Ser Ala Leu Val Tyr Pro Leu Asp Val Ala
420 425 430

Asn Thr Ala Val Lys Val Arg Ala Pro Tyr Ile Phe Asp Lys Phe Thr
435 440 445

Arg Pro Gly Gly His Pro Cys Ser Ala Ser Ser Arg Cys Pro Asn Val
450 455 460

Cys Val Thr Gly Val Tyr Thr Asp Ala Tyr Pro Leu Val Phe Ser Arg
465 470 475 480

Ser His Asp Ile Val Ala Val Tyr Gly Met Gln Leu Ala Ala Gly Thr
485 490 495

Ala Arg Leu Asp Pro Gln Ala Ala Ile Trp Tyr Gly Asn Glu Met Ser
500 505 510

Thr Pro Thr Lys Val Ser Ser Ser Thr Thr Lys Ala Ala Tyr Thr Thr
515 520 525

Ser Thr Cys Phe Lys Val Thr Lys Thr Lys Arg Ile Tyr Cys Ile Ser
530 535 540

Ile Ala Glu Ile Gly Asn Thr Leu Phe Gly Glu Phe Arg Ile Val Pro
545 550 555 560

Leu Leu Ile Glu Val Gln Lys Thr Pro Leu Thr Arg Arg Ser Glu Leu
565 570 575

Arg Gln Gln Met Pro Gln Pro Pro Ile Asp Leu Val Ile Asp Asn Pro
580 585 590

Phe Cys Ala Pro Ser Gly Asn Leu Ser Arg Lys Asn Ala Ile Asp Glu
595 600 605

Tyr Ala Asn Ser Trp Pro
610

<210> 115
<211> 577
<212> PRT

<213> avian-paramyxovirus 3

<220>
<221> MISC_FEATURE
<223> hemagglutinin

<400> 115

```
Met Glu Pro Thr Gly Ser Lys Val Asp Ile Val Pro Ser Gln Gly Thr
1               5                   10                  15

Lys Arg Thr Cys Arg Thr Phe Tyr Arg Leu Leu Ile Leu Ile Leu Asn
            20                  25                  30

Leu Ile Ile Ile Ile Leu Thr Ile Ile Ser Ile Tyr Val Ser Ile Ser
        35                  40                  45

Thr Asp Gln His Lys Leu Cys Asn Asn Glu Ala Asp Ser Leu Leu His
    50                  55                  60

Ser Ile Val Glu Pro Ile Thr Val Pro Leu Gly Thr Asp Ser Asp Val
65                  70                  75                  80

Glu Asp Glu Leu Arg Glu Ile Arg Arg Asp Thr Gly Ile Asn Ile Pro
                85                  90                  95

Ile Gln Ile Asp Asn Thr Glu Asn Ile Ile Leu Thr Thr Leu Ala Ser
            100                 105                 110

Ile Asn Ser Asn Ile Ala Arg Leu His Asn Ala Thr Asp Glu Ser Pro
        115                 120                 125

Thr Cys Leu Ser Pro Val Asn Asp Pro Arg Phe Ile Ala Gly Ile Asn
    130                 135                 140

Lys Ile Thr Lys Gly Ser Met Ile Tyr Arg Asn Phe Ser Asn Leu Ile
145                 150                 155                 160

Glu His Val Asn Phe Ile Pro Ser Pro Thr Thr Leu Ser Gly Cys Thr
                165                 170                 175

Arg Ile Pro Ser Phe Ser Leu Ser Lys Thr His Trp Cys Tyr Ser His
            180                 185                 190

Asn Val Ile Ser Thr Gly Cys Gln Asp His Ala Ala Ser Ser Gln Tyr
        195                 200                 205

Ile Ser Ile Gly Ile Val Asp Thr Gly Leu Asn Asn Glu Pro Tyr Leu
    210                 215                 220

Arg Thr Met Ser Ser Arg Leu Leu Asn Asp Gly Leu Asn Arg Lys Ser
225                 230                 235                 240
```

```
Cys Ser Val Thr Ala Gly Ala Gly Val Cys Trp Leu Leu Cys Ser Val
                245                 250                 255

Val Thr Glu Ser Glu Ser Ala Asp Tyr Arg Ser Arg Ala Pro Thr Ala
            260                 265                 270

Met Ile Leu Gly Arg Phe Asn Phe Tyr Gly Asp Tyr Thr Glu Ser Pro
        275                 280                 285

Val Pro Ala Ser Leu Phe Ser Gly Arg Phe Thr Ala Asn Tyr Pro Gly
    290                 295                 300

Val Gly Ser Gly Thr Gln Leu Asn Gly Thr Leu Tyr Phe Pro Ile Tyr
305                 310                 315                 320

Gly Gly Val Val Asn Asp Ser Asp Ile Glu Leu Ser Asn Arg Gly Lys
                325                 330                 335

Ser Phe Arg Pro Arg Asn Pro Thr Asn Pro Cys Pro Asp Pro Glu Val
            340                 345                 350

Thr Gln Ser Gln Arg Ala Gln Ala Ser Tyr Tyr Pro Thr Arg Phe Gly
        355                 360                 365

Arg Leu Leu Ile Gln Gln Ala Ile Leu Ala Cys Arg Ile Ser Asp Thr
    370                 375                 380

Thr Cys Thr Asp Tyr Tyr Leu Leu Tyr Phe Asp Asn Asn Gln Val Met
385                 390                 395                 400

Met Gly Ala Glu Ala Arg Ile Tyr Tyr Leu Asn Asn Gln Met Tyr Leu
                405                 410                 415

Tyr Gln Arg Ser Ser Ser Trp Trp Pro His Pro Leu Phe Tyr Arg Phe
            420                 425                 430

Ser Leu Pro His Cys Glu Pro Met Ser Val Cys Met Ile Thr Asp Thr
        435                 440                 445

His Leu Ile Leu Thr Tyr Ala Thr Ser Arg Pro Gly Thr Ser Ile Cys
    450                 455                 460

Thr Gly Ala Ser Arg Cys Pro Asn Asn Cys Val Asp Gly Val Tyr Thr
465                 470                 475                 480

Asp Val Trp Pro Leu Thr Glu Gly Thr Thr Gln Asp Pro Asp Ser Tyr
                485                 490                 495

Tyr Thr Val Phe Leu Asn Ser Pro Asn Arg Arg Ile Ser Pro Thr Ile
            500                 505                 510

Ser Ile Tyr Ser Tyr Asn Gln Lys Ile Ser Ser Arg Leu Ala Val Gly
```

515 520 525

Ser Glu Ile Gly Ala Ala Tyr Thr Thr Ser Thr Cys Phe Ser Arg Thr
530 535 540

Asp Thr Gly Ala Leu Tyr Cys Ile Thr Ile Ile Glu Ala Val Asn Thr
545 550 555 560

Ile Phe Gly Gln Tyr Arg Ile Val Pro Ile Leu Val Gln Leu Ile Ser
565 570 575

Asp

<210> 116
<211> 620
<212> PRT
<213> Salem virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 116

Met Lys Ala Met His Tyr Tyr Lys Asn Asp Phe Ala Asp Pro Gly Thr
1 5 10 15

Asn Asp Asn Ser Ser Asp Leu Thr Thr Asn Pro Phe Ile Ser Asn Gln
20 25 30

Ile Lys Ser Asn Leu Ser Pro Pro Val Leu Ala Glu Gly His Leu Ser
35 40 45

Pro Ser Pro Ile Pro Lys Phe Arg Lys Ile Leu Leu Thr Ile Ser Phe
50 55 60

Val Ser Thr Ile Val Val Leu Thr Val Ile Leu Leu Val Leu Thr Ile
65 70 75 80

Arg Ile Leu Thr Ile Ile Glu Ala Ser Ala Gly Asp Glu Lys Asp Ile
85 90 95

His Thr Ile Leu Ser Ser Leu Leu Asn Thr Phe Met Asn Glu Tyr Ile
100 105 110

Pro Val Phe Lys Asn Leu Val Ser Ile Ile Ser Leu Gln Ile Pro Gln
115 120 125

Met Leu Ile Asp Leu Lys Thr Ser Ser Thr Gln Met Met Gln Ser Leu
130 135 140

Lys Thr Phe Pro Arg Asp Leu Glu Thr Leu Ser Thr Val Thr Gln Ser
145 150 155 160

```
Val Ala Val Leu Leu Glu Lys Ala Lys Ser Thr Ile Pro Asp Ile Asn
                165                 170                 175

Lys Phe Tyr Lys Asn Val Gly Lys Val Thr Phe Asn Asp Pro Asn Ile
            180                 185                 190

Lys Val Leu Thr Leu Glu Val Pro Ala Trp Leu Pro Ile Val Arg Gln
        195                 200                 205

Cys Leu Lys Gln Asp Phe Arg Gln Val Ile Ser Asn Ser Thr Gly Phe
    210                 215                 220

Ala Leu Ile Gly Ala Leu Pro Ser Gln Leu Phe Asn Glu Phe Glu Gly
225                 230                 235                 240

Tyr Pro Ser Leu Ala Ile Val Ser Glu Val Tyr Ala Ile Thr Tyr Leu
                245                 250                 255

Lys Gly Val Met Phe Glu Asn Gln Glu Asn Phe Leu Tyr Gln Tyr Phe
            260                 265                 270

Glu Ile Gly Thr Ile Ser Pro Asp Gly Tyr Asn Lys Pro Tyr Phe Leu
        275                 280                 285

Arg His Thr Ser Val Met Leu Ser Thr Phe Lys Leu Ser Gly Lys Cys
    290                 295                 300

Thr Ala Ala Val Asp Tyr Arg Gly Gly Ile Phe Leu Cys Thr Pro Ser
305                 310                 315                 320

Pro Lys Ile Pro Lys Ile Leu Gln Asn Pro Pro Asp Leu Pro Thr Leu
                325                 330                 335

Thr Val Val Ser Ile Pro Phe Asp Gly Arg Tyr Thr Ile Arg Asn Ile
            340                 345                 350

Ser Leu Met Leu Thr Asp Glu Ala Asp Ile Ile Tyr Asp Leu Asp Thr
        355                 360                 365

Leu Gln Gly Arg Gly Val Leu Gln Ala Met Arg Phe Tyr Ala Leu Val
    370                 375                 380

Arg Val Ile Ser Ser Ser Ser Pro Arg His Phe Pro Phe Cys Lys Asn
385                 390                 395                 400

Ser Trp Cys Pro Thr Ala Asp Asp Lys Ile Cys Asp Gln Ser Arg Arg
                405                 410                 415

Leu Gly Ala Asp Gly Asn Tyr Pro Val Met Tyr Gly Leu Ile Ser Ile
            420                 425                 430
```

```
Pro Ala His Ser Ser Tyr Gln Gly Asn Val Ser Leu Lys Leu Ile Asp
        435                 440                 445

Pro Lys Tyr Tyr Ala Tyr Thr Arg Asp Ala Ser Leu Phe Tyr Asn Ser
    450                 455                 460

Met Thr Asp Thr Tyr His Tyr Ser Phe Gly Thr Arg Gly Trp Val Ser
465                 470                 475                 480

Arg Pro Ile Ile Gly Glu Leu Leu Leu Gly Asp Asp Ile Val Leu Thr
                485                 490                 495

Arg Tyr Thr Val Arg Ser Val Ser Arg Ala Thr Ala Gly Asp Cys Thr
            500                 505                 510

Thr Val Ser Met Cys Pro Gln Ala Cys Ser Gly Gly Met Asn Ser Ile
        515                 520                 525

Phe Tyr Pro Leu Asn Phe Asp Lys Pro Gln Val Thr Gly Val Ala Ile
    530                 535                 540

Arg Gln Tyr Glu Arg Gln Gln Glu Gly Ile Ile Val Val Thr Met Asn
545                 550                 555                 560

Asp His Tyr Tyr Tyr Ser Val Pro Ile Ile Lys Asn Gly Thr Leu Leu
                565                 570                 575

Ile Ser Ser Val Thr Asp Cys Phe Trp Leu Met Gly Asp Leu Trp Cys
            580                 585                 590

Met Ser Leu Met Glu Lys Asn Asn Leu Pro Leu Gly Val Arg Ser Leu
        595                 600                 605

Ala His Leu Thr Trp Asn Ile His Trp Ser Cys Ser
    610                 615                 620


<210>  117
<211>  579
<212>  PRT
<213>  Avian paramyxovirus 9


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase

<400>  117

Met Glu Ser Gly Ile Ser Gln Ala Ser Leu Val Asn Asp Asn Ile Glu
1               5                   10                  15

Leu Arg Asn Thr Trp Arg Thr Ala Phe Arg Val Val Ser Leu Leu Leu
            20                  25                  30
```

```
Gly Phe Thr Ser Leu Val Leu Thr Ala Cys Ala Leu His Phe Ala Leu
        35                  40                  45

Asn Ala Ala Thr Pro Ala Asp Leu Ser Ser Ile Pro Val Ala Val Asp
    50                  55                  60

Gln Ser His His Glu Ile Leu Gln Thr Leu Ser Leu Met Ser Asp Ile
65                  70                  75                  80

Gly Asn Lys Ile Tyr Lys Gln Val Ala Leu Asp Ser Pro Val Ala Leu
                85                  90                  95

Leu Asn Thr Glu Ser Thr Leu Met Ser Ala Ile Thr Ser Leu Ser Tyr
            100                 105                 110

Gln Ile Asn Asn Ala Ala Asn Asn Ser Gly Cys Gly Ala Pro Val His
        115                 120                 125

Asp Lys Asp Phe Ile Asn Gly Val Ala Lys Glu Leu Phe Val Gly Ser
    130                 135                 140

Gln Tyr Asn Ala Ser Asn Tyr Arg Pro Ser Arg Phe Leu Glu His Leu
145                 150                 155                 160

Asn Phe Ile Pro Ala Pro Thr Thr Gly Lys Gly Cys Thr Arg Ile Pro
                165                 170                 175

Ser Phe Asp Leu Ala Ala Thr His Trp Cys Tyr Thr His Asn Val Ile
            180                 185                 190

Leu Asn Gly Cys Asn Asp His Ala Gln Ser Tyr Gln Tyr Ile Ser Leu
        195                 200                 205

Gly Ile Leu Lys Val Ser Ala Thr Gly Asn Val Phe Leu Ser Thr Leu
    210                 215                 220

Arg Ser Ile Asn Leu Asp Asp Asp Glu Asn Arg Lys Ser Cys Ser Ile
225                 230                 235                 240

Ser Ala Thr Pro Leu Gly Cys Asp Leu Leu Cys Ala Lys Val Thr Glu
                245                 250                 255

Arg Glu Glu Ala Asp Tyr Asn Ser Asp Ala Ala Thr Arg Leu Val His
            260                 265                 270

Gly Arg Leu Gly Phe Asp Gly Val Tyr His Glu Gln Ala Leu Pro Val
        275                 280                 285

Glu Ser Leu Phe Ser Asp Trp Val Ala Asn Tyr Pro Ser Val Gly Gly
    290                 295                 300

Gly Ser Tyr Phe Asp Asn Arg Val Trp Phe Gly Val Tyr Gly Gly Ile
```

305 310 315 320

Arg Pro Gly Ser Gln Thr Asp Leu Leu Gln Ser Glu Lys Tyr Ala Ile
325 330 335

Tyr Arg Arg Tyr Asn Asn Thr Cys Pro Asp Asn Asn Pro Thr Gln Ile
340 345 350

Glu Arg Ala Lys Ser Ser Tyr Arg Pro Gln Arg Phe Gly Gln Arg Leu
355 360 365

Val Gln Gln Ala Ile Leu Ser Ile Arg Val Glu Pro Ser Leu Gly Asn
370 375 380

Asp Pro Lys Leu Ser Val Leu Asp Asn Thr Val Val Leu Met Gly Ala
385 390 395 400

Glu Ala Arg Ile Met Thr Phe Gly His Val Ala Leu Met Tyr Gln Arg
405 410 415

Gly Ser Ser Tyr Phe Pro Ser Ala Leu Leu Tyr Pro Leu Ser Leu Thr
420 425 430

Asn Gly Ser Ala Ala Ala Ser Lys Pro Phe Ile Phe Glu Gln Tyr Thr
435 440 445

Arg Pro Gly Ser Pro Pro Cys Gln Ala Thr Ala Arg Cys Pro Asn Ser
450 455 460

Cys Val Thr Gly Val Tyr Thr Asp Ala Tyr Pro Leu Phe Trp Ser Glu
465 470 475 480

Asp His Lys Val Asn Gly Val Tyr Gly Met Met Leu Asp Asp Ile Thr
485 490 495

Ser Arg Leu Asn Pro Val Ala Ala Ile Phe Asp Arg Tyr Gly Arg Ser
500 505 510

Arg Val Thr Arg Val Ser Ser Ser Ser Thr Lys Ala Ala Tyr Thr Thr
515 520 525

Asn Thr Cys Phe Lys Val Val Lys Thr Lys Arg Val Tyr Cys Leu Ser
530 535 540

Ile Ala Glu Ile Glu Asn Thr Leu Phe Gly Glu Phe Arg Ile Thr Pro
545 550 555 560

Leu Leu Ser Glu Ile Ile Phe Asp Pro Asn Leu Glu Pro Ser Asp Thr
565 570 575

Ser Arg Asn

<210> 118
<211> 595
<212> PRT
<213> Achimota virus

<220>
<221> MISC_FEATURE
<223> attachment protein

<400> 118

Met Ala Thr Asn Leu Ser Thr Ile Thr Asn Gly Lys Phe Ser Gln Asn
1 5 10 15

Ser Asp Glu Gly Ser Leu Thr Glu Leu Pro Phe Phe Glu His Asn Arg
20 25 30

Lys Val Ala Thr Thr Lys Arg Thr Cys Arg Phe Val Phe Arg Ser Val
35 40 45

Ile Thr Leu Cys Asn Leu Thr Ile Leu Ile Val Thr Val Val Val Leu
50 55 60

Phe Gln Gln Ala Gly Phe Ile Lys Arg Thr Glu Ser Asn Gln Val Cys
65 70 75 80

Glu Thr Leu Gln Asn Asp Met His Gly Val Val Thr Met Ser Lys Gly
85 90 95

Val Ile Thr Thr Leu Asn Asn Leu Ile Glu Ile Thr Ser Val Asn Leu
100 105 110

Pro Phe Gln Met Lys Gln Phe Gly Gln Gly Ile Val Thr Gln Val Thr
115 120 125

Gln Met Val Arg Gln Cys Asn Ala Val Cys Lys Gly Pro Thr Ile Gly
130 135 140

Pro Asp Ile Gln Asn Ile Val Tyr Pro Ala Ser Tyr Glu Ser Met Ile
145 150 155 160

Lys His Pro Val Asn Asn Ser Asn Ile Leu Leu Ser Glu Ile Arg Gln
165 170 175

Pro Leu Asn Phe Val Pro Asn Thr Gly Lys Leu Asn Gly Cys Thr Arg
180 185 190

Thr Pro Ser Phe Ser Val Tyr Asn Gly Phe Trp Cys Tyr Thr His Ala
195 200 205

Glu Ser Asp Trp Asn Cys Asn Gly Ser Ser Pro Tyr Met Gln Val Phe
210 215 220

```
Arg Val Gly Val Val Thr Ser Asp Tyr Asp Tyr Asn Val Ile His Lys
225                 230                 235                 240

Thr Leu His Thr Lys Thr Ser Arg Leu Ala Asn Val Thr Tyr Gln Cys
                245                 250                 255

Ser Thr Ile Ser Thr Gly Tyr Glu Cys Tyr Phe Leu Cys Ser Thr Pro
            260                 265                 270

Asn Val Asp Glu Ile Thr Asp Tyr Lys Thr Pro Gly Ile Glu Ser Leu
        275                 280                 285

Gln Ile Tyr Lys Ile Asp Asn Arg Gly Thr Phe Ala Lys Phe Pro Ile
    290                 295                 300

Thr Asp Gln Leu Asn Lys Glu Leu Leu Thr Ala Leu Tyr Pro Gly Pro
305                 310                 315                 320

Gly Asn Gly Val Leu Tyr Gln Gly Arg Leu Leu Phe Pro Met His Gly
                325                 330                 335

Gly Met Gln Ser Ser Glu Leu Asn Lys Val Asn Leu Asn Asn Thr Val
            340                 345                 350

Leu Ser Gln Phe Asn Asp Asn Lys Gly Cys Asn Ala Thr Glu Ile Lys
        355                 360                 365

Leu Glu Ser Glu Phe Pro Gly Thr Phe Thr Ser Pro Tyr Tyr Ser Asn
    370                 375                 380

Gln Val Met Leu Asn Tyr Ile Leu Ile Cys Glu Met Ile Glu Asn Leu
385                 390                 395                 400

Pro Gly Asn Cys Asp Leu Gln Ile Val Ala Pro Lys Asn Met Ser Met
                405                 410                 415

Gly Ser Glu Ser Gln Leu Tyr Ser Ile Asn Asn Lys Leu Tyr Leu Tyr
            420                 425                 430

Gln Arg Ser Ser Ser Arg Trp Pro Tyr Pro Leu Ile Tyr Glu Val Gly
        435                 440                 445

Thr Arg Leu Thr Asn Arg Gln Phe Arg Leu Arg Ala Ile Asn Arg Phe
    450                 455                 460

Leu Ile Lys Ser Thr Thr Arg Pro Gly Ser Glu Gly Cys Asn Ile Tyr
465                 470                 475                 480

Arg Val Cys Pro Lys Val Cys Val Thr Gly Val Tyr Gln Ala Pro Trp
                485                 490                 495
```

Ile Leu His Val Ser Lys Ala Gly Ser Gln Ser Ile Ala Lys Val Leu
500 505 510

Tyr Ala Val Ala Trp Ser Lys Asp His Met Ser Arg Lys Gly Pro Leu
515 520 525

Phe Ser Ile Cys Asp Asn Asp Thr Cys Phe Leu Thr Lys Ser Leu Ala
530 535 540

Ser Glu His Val His Ser Gly Tyr Ser Ile Thr Arg Cys Tyr Leu Glu
545 550 555 560

Asn Ser Glu Arg His Ile Ile Cys Val Val Ile Met Glu Leu Asp Ala
565 570 575

Ser Pro Trp Ala Glu Met Arg Ile Gln Ser Val Ile Tyr Asn Ile Thr
580 585 590

Leu Pro Ser
595

<210> 119
<211> 583
<212> PRT
<213> Achimota virus 2

<220>
<221> MISC_FEATURE
<223> attachment protein

<400> 119

Met Asp Asn Ser Met Ser Ile Ser Thr Ile Ser Leu Asp Ala Gln Pro
1 5 10 15

Arg Ile Trp Ser Arg His Glu Ser Arg Arg Thr Trp Arg Asn Ile Phe
20 25 30

Arg Ile Thr Ser Leu Val Leu Leu Gly Val Thr Val Ile Ile Cys Ile
35 40 45

Trp Leu Cys Cys Glu Val Ala Arg Glu Ser Glu Leu Glu Leu Leu Ala
50 55 60

Ser Pro Leu Gly Ala Leu Ile Met Ala Ile Asn Thr Ile Lys Ser Ser
65 70 75 80

Val Val Lys Met Thr Thr Glu Leu Asn Gln Val Thr Phe Thr Thr Ser
85 90 95

Ile Ile Leu Pro Asn Lys Val Asp Gln Phe Gly Gln Asn Val Val Ser
100 105 110

Gln Val Ala Gln Leu Val Lys Gln Cys Asn Ala Val Cys Arg Gly His

115 120 125

Gln Asp Thr Pro Glu Leu Glu Gln Phe Ile Asn Gln Lys Asn Pro Thr
130 135 140

Trp Ile Leu Gln Pro Asn Tyr Thr Thr Lys Leu Thr Asn Leu His Glu
145 150 155 160

Ile Asp Ser Ile Ile Pro Leu Val Asp Tyr Pro Gly Phe Ser Lys Ser
165 170 175

Cys Thr Arg Phe Pro Ser Phe Ser Glu Gly Ser Lys Phe Trp Cys Phe
180 185 190

Thr Tyr Ala Val Val Lys Glu Pro Cys Ser Asp Ile Ser Ser Ser Ile
195 200 205

Gln Val Val Lys Tyr Gly Ala Ile Lys Ala Asn His Ser Asp Gly Asn
210 215 220

Pro Tyr Leu Val Leu Gly Thr Lys Val Leu Asp Asp Gly Lys Phe Arg
225 230 235 240

Arg Gly Cys Ser Ile Thr Ser Ser Leu Tyr Gly Cys Tyr Leu Leu Cys
245 250 255

Ser Thr Ala Asn Val Ser Glu Val Asn Asp Tyr Ala His Thr Pro Ala
260 265 270

Tyr Pro Leu Thr Leu Glu Leu Ile Ser Lys Asp Gly Ile Thr Thr Asp
275 280 285

Leu Ser Pro Thr Tyr Thr Val Gln Leu Asp Lys Trp Ser Ala Leu Tyr
290 295 300

Pro Gly Ile Gly Ser Gly Val Ile Phe Lys Gly Tyr Leu Met Phe Pro
305 310 315 320

Val Tyr Gly Gly Leu Pro Phe Lys Ser Pro Leu Ile Ser Ala Ser Trp
325 330 335

Val Gly Pro Gly Asn Lys Trp Pro Val Asp Phe Ser Cys Ser Glu Asp
340 345 350

Gln Tyr Ser Thr Phe Asn Phe Ser Asn Pro Tyr Ser Ala Leu Tyr Ser
355 360 365

Pro His Phe Ser Asn Asn Ile Val Val Ser Ala Leu Phe Val Cys Pro
370 375 380

Leu Asn Glu Asn Leu Pro Tyr Ser Cys Glu Val Gln Val Leu Pro Gln
385 390 395 400

```
Gly Asn Leu Thr Ile Gly Ala Glu Gly Arg Leu Tyr Val Ile Asp Gln
                405                 410                 415


Asp Leu Tyr Tyr Tyr Gln Arg Ser Thr Ser Trp Trp Pro Tyr Leu Gln
            420                 425                 430


Leu Tyr Lys Leu Asn Ile Arg Ile Thr Asn Arg Val Phe Arg Val Arg
        435                 440                 445


Ser Leu Ser Leu Leu Pro Ile Lys Ser Thr Thr Arg Pro Gly Tyr Gly
    450                 455                 460


Asn Cys Thr Tyr Phe Lys Leu Cys Pro His Ile Cys Val Thr Gly Val
465                 470                 475                 480


Tyr Gln Ser Pro Trp Leu Ile Ser Ile Arg Asp Lys Arg Pro His Glu
                485                 490                 495


Glu Lys Asn Ile Leu Tyr Phe Ile Gly Trp Ser Pro Asp Glu Gln Ile
            500                 505                 510


Arg Gln Asn Pro Leu Val Ser Leu Cys His Glu Thr Ala Cys Phe Ile
        515                 520                 525


Asn Arg Ser Leu Ala Thr Asn Lys Thr His Ala Gly Tyr Ser Glu Ser
    530                 535                 540


His Cys Val Gln Ser Phe Glu Arg Asn Lys Leu Thr Cys Thr Val Phe
545                 550                 555                 560


Tyr Glu Leu Thr Ala Lys Pro Trp Ala Glu Met Arg Val Gln Ser Leu
                565                 570                 575


Leu Phe Gln Val Asp Phe Leu
            580


<210>  120
<211>  580
<212>  PRT
<213>  Tuhoko virus 1


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase

<400>  120

Met Asp Ser Arg Ser Asp Ser Phe Thr Asp Ile Pro Leu Asp Asn Arg
1               5                   10                  15


Ile Glu Arg Thr Val Thr Ser Lys Lys Thr Trp Arg Ser Ile Phe Arg
            20                  25                  30
```

```
Val Thr Ala Ile Ile Leu Leu Ile Ile Cys Val Val Val Ser Ser Ile
        35                  40                  45

Ser Leu Asn Gln His Asn Asp Ala Pro Leu Asn Gly Ala Gly Asn Gln
    50                  55                  60

Ala Thr Ser Gly Phe Met Asp Ala Ile Lys Ser Leu Glu Lys Leu Met
65                  70                  75                  80

Ser Gln Thr Ile Asn Glu Leu Asn Gln Val Val Met Thr Thr Ser Val
                85                  90                  95

Gln Leu Pro Asn Arg Ile Thr Lys Phe Gly Gln Asp Ile Leu Asp Gln
            100                 105                 110

Val Thr Gln Met Val Arg Gln Cys Asn Ala Val Cys Arg Gly Pro Gly
        115                 120                 125

Val Gly Pro Ser Ile Gln Asn Tyr Val Ile Gln Gly His Ala Pro Thr
    130                 135                 140

Val Ser Phe Asp Pro Ile Ser Ala Glu Tyr Gln Lys Phe Val Phe Gly
145                 150                 155                 160

Ile Thr Glu Lys Thr Leu Ile Thr Ala Tyr His Asn Pro Trp Glu Cys
                165                 170                 175

Leu Arg Phe Pro Ser Gln His Leu Phe Asp Thr Thr Trp Cys Val Ser
            180                 185                 190

Tyr Gln Ile Leu Thr Gln Asn Cys Ser Asp His Gly Pro Arg Ile Thr
        195                 200                 205

Val Ile Gln Leu Gly Glu Ile Met Ile Ala Asn Asn Leu Ser Thr Val
    210                 215                 220

Phe Arg Asp Pro Val Ile Lys Tyr Ile Arg His His Ile Trp Leu Arg
225                 230                 235                 240

Ser Cys Ser Val Val Ala Tyr Tyr Ser Gln Cys Thr Ile Phe Cys Thr
                245                 250                 255

Ser Thr Asn Lys Ser Glu Pro Ser Asp Tyr Ala Asp Thr Gly Tyr Glu
            260                 265                 270

Gln Leu Phe Leu Ala Thr Leu Gln Ser Asp Gly Thr Phe Thr Glu His
        275                 280                 285

Ser Met His Gly Val Asn Ile Val His Gln Trp Asn Ala Ile Tyr Gly
    290                 295                 300
```

```
Gly Val Gly Asn Gly Val Ile Ile Gly Arg Asn Met Leu Ile Pro Leu
305                 310                 315                 320

Tyr Gly Gly Ile Asn Tyr Tyr Asp His Asn Thr Thr Ile Val Gln Thr
                325                 330                 335

Val Asp Leu Arg Pro Tyr Pro Ile Pro Asp Ser Cys Ser Gln Thr Asp
            340                 345                 350

Asn Tyr Gln Thr Asn Tyr Leu Pro Ser Met Phe Thr Asn Ser Tyr Tyr
        355                 360                 365

Gly Thr Asn Leu Val Val Ser Gly Tyr Leu Ser Cys Arg Leu Met Ala
    370                 375                 380

Gly Thr Pro Thr Ser Cys Ser Ile Arg Val Ile Pro Ile Glu Asn Met
385                 390                 395                 400

Thr Met Gly Ser Glu Gly Gln Phe Tyr Leu Ile Asn Asn Gln Leu Tyr
                405                 410                 415

Tyr Tyr Lys Arg Ser Ser Asn Trp Ile Arg Asp Thr Gln Val Tyr Leu
            420                 425                 430

Leu Ser Tyr Ser Asp Lys Gly Asn Ile Ile Glu Ile Thr Ser Ala Glu
        435                 440                 445

Arg Tyr Ile Phe Lys Ser Val Thr Ser Pro Asp Glu Gly Asp Cys Val
    450                 455                 460

Thr Asn His Gly Cys Pro Ser Asn Cys Ile Gly Gly Leu Phe Gln Ala
465                 470                 475                 480

Pro Trp Ile Leu Asn Asp Phe Lys Leu Cys Gly Ser Asn Ile Thr Cys
                485                 490                 495

Pro Lys Ile Val Thr Val Trp Ala Asp Gln Pro Asp Lys Arg Ser Asn
            500                 505                 510

Pro Met Leu Ser Ile Ala Glu Thr Asp Lys Leu Leu Leu His Lys Ser
        515                 520                 525

Tyr Ile Asn Tyr His Thr Ala Val Gly Tyr Ser Thr Val Leu Cys Phe
    530                 535                 540

Asp Ser Pro Lys Leu Asn Leu Lys Thr Cys Val Val Leu Gln Glu Leu
545                 550                 555                 560

Met Ser Asp Asp Lys Leu Leu Ile Arg Ile Ser Tyr Ser Ile Val Ser
                565                 570                 575

Ile Met Val Glu
```

580

<210> 121
<211> 607
<212> PRT
<213> Phocine distemper virus

<400> 121

Met Phe Ser His Gln Asp Lys Val Gly Ala Phe Tyr Lys Asn Asn Ala
1 5 10 15

Arg Ala Asn Ser Ser Lys Leu Ser Leu Val Thr Asp Glu Val Glu Glu
20 25 30

Arg Arg Ser Pro Trp Phe Leu Ser Ile Leu Leu Ile Leu Leu Val Gly
35 40 45

Ile Leu Ile Leu Leu Ala Ile Thr Gly Ile Arg Phe His Gln Val Val
50 55 60

Lys Ser Asn Leu Glu Phe Asn Lys Leu Leu Ile Glu Asp Met Glu Lys
65 70 75 80

Thr Lys Ala Val His His Gln Val Lys Asp Val Leu Thr Pro Leu Phe
85 90 95

Lys Ile Ile Gly Asp Glu Val Gly Leu Arg Leu Pro Gln Lys Leu Asn
100 105 110

Glu Ile Lys Gln Phe Ile Val Gln Lys Thr Asn Phe Phe Asn Pro Asn
115 120 125

Arg Glu Phe Asp Phe Arg Glu Leu His Trp Cys Ile Asn Pro Pro Ser
130 135 140

Lys Val Lys Val Asn Phe Thr Gln Tyr Cys Glu Ile Thr Glu Phe Lys
145 150 155 160

Glu Ala Thr Arg Ser Val Ala Asn Ser Ile Leu Leu Leu Thr Leu Tyr
165 170 175

Arg Gly Arg Asp Asp Ile Phe Pro Pro Tyr Lys Cys Arg Gly Ala Thr
180 185 190

Thr Ser Met Gly Asn Val Phe Pro Leu Ala Val Ser Leu Ser Met Ser
195 200 205

Leu Ile Ser Lys Pro Ser Glu Val Ile Asn Met Leu Thr Ala Ile Ser
210 215 220

Glu Gly Ile Tyr Gly Lys Thr Tyr Leu Leu Val Thr Asp Asp Thr Glu
225 230 235 240

```
Glu Asn Phe Glu Thr Pro Glu Ile Arg Val Phe Glu Ile Gly Phe Ile
                245                 250                 255

Asn Arg Trp Leu Gly Asp Met Pro Leu Phe Gln Thr Thr Asn Tyr Arg
            260                 265                 270

Ile Ile Ser Asn Asn Ser Asn Thr Lys Ile Cys Thr Ile Ala Val Gly
        275                 280                 285

Glu Leu Ala Leu Ala Ser Leu Cys Thr Lys Glu Ser Thr Ile Leu Leu
    290                 295                 300

Asn Leu Gly Asp Glu Glu Ser Gln Asn Ser Val Leu Val Val Ile Leu
305                 310                 315                 320

Gly Leu Phe Gly Ala Thr His Met Asp Gln Leu Glu Glu Val Ile Pro
                325                 330                 335

Val Ala His Pro Ser Ile Glu Lys Ile His Ile Thr Asn His Arg Gly
            340                 345                 350

Phe Ile Lys Asp Ser Val Ala Thr Trp Met Val Pro Ala Leu Ala Leu
        355                 360                 365

Ser Glu Gln Gly Glu Gln Ile Asn Cys Leu Arg Ser Ala Cys Lys Arg
    370                 375                 380

Arg Thr Tyr Pro Met Cys Asn Gln Thr Ser Trp Glu Pro Phe Gly Asp
385                 390                 395                 400

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

Thr Asp Leu Ser Ile Asn Val Ser Val Ala Gln Gly Pro Ile Ile Phe
            420                 425                 430

Asn Gly Asp Gly Met Asp Tyr Tyr Glu Gly Thr Leu Leu Asn Ser Gly
        435                 440                 445

Trp Leu Thr Ile Pro Pro Lys Asn Gly Thr Ile Leu Gly Leu Ile Asn
    450                 455                 460

Gln Ala Ser Lys Gly Asp Gln Phe Ile Val Thr Pro His Ile Leu Thr
465                 470                 475                 480

Phe Ala Pro Arg Glu Ser Ser Thr Asp Cys His Leu Pro Ile Gln Thr
                485                 490                 495

Tyr Gln Ile Gln Asp Asp Asp Val Leu Leu Glu Ser Asn Leu Val Val
            500                 505                 510
```

Leu Pro Thr Gln Ser Phe Glu Tyr Val Val Ala Thr Tyr Asp Val Ser
515 520 525

Arg Ser Asp His Ala Ile Val Tyr Tyr Val Tyr Asp Pro Ala Arg Thr
530 535 540

Val Ser Tyr Thr Tyr Pro Phe Arg Leu Arg Thr Lys Gly Arg Pro Asp
545 550 555 560

Ile Leu Arg Ile Glu Cys Phe Val Trp Asp Gly His Leu Trp Cys His
565 570 575

Gln Phe Tyr Arg Phe Gln Leu Asp Ala Thr Asn Ser Thr Ser Val Val
580 585 590

Glu Asn Leu Ile Arg Ile Arg Phe Ser Cys Asp Arg Leu Asp Pro
595 600 605

<210> 122
<211> 574
<212> PRT
<213> Caprine parainfluenza virus 3

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase

<400> 122

Met Glu Tyr Trp Gly His Thr Asn Asn Pro Asp Lys Ile Asn Arg Lys
1 5 10 15

Val Gly Val Asp Gln Val Arg Asp Arg Ser Lys Thr Leu Lys Ile Ile
20 25 30

Thr Phe Ile Ile Ser Met Met Thr Ser Ile Met Ser Thr Val Ala Leu
35 40 45

Ile Leu Ile Leu Ile Met Phe Ile Gln Asn Asn Asn Asn Asn Arg Ile
50 55 60

Ile Leu Gln Glu Leu Arg Asp Glu Thr Asp Ala Ile Glu Ala Arg Ile
65 70 75 80

Gln Lys Ala Ser Asn Asp Ile Gly Val Ser Ile Gln Ser Gly Ile Asn
85 90 95

Thr Arg Leu Leu Thr Ile Gln Asn His Val Gln Asn Tyr Ile Pro Leu
100 105 110

Ala Leu Thr Gln Gln Val Ser Ser Leu Arg Glu Ser Ile Asn Asp Val
115 120 125

Ile Thr Lys Arg Glu Glu Thr Gln Ser Lys Met Pro Ile Gln Arg Met

130 135 140

Thr His Asp Asp Gly Ile Glu Pro Leu Ile Pro Asp Asn Phe Trp Lys
145 150 155 160

Cys Pro Ser Gly Ile Pro Thr Ile Ser Ala Ser Pro Lys Ile Arg Leu
165 170 175

Ile Pro Gly Pro Gly Leu Leu Ala Thr Ser Thr Thr Ile Asn Gly Cys
180 185 190

Ile Arg Leu Pro Ser Leu Val Ile Asn Asn Leu Ile Tyr Ala Tyr Thr
195 200 205

Ser Asn Leu Ile Thr Gln Gly Cys Gln Asp Ile Gly Lys Ser Tyr Gln
210 215 220

Val Leu Gln Ile Gly Ile Ile Thr Ile Asn Ser Asp Leu Val Pro Asp
225 230 235 240

Leu Asn Pro Arg Ile Thr His Thr Phe Asp Ile Asp Asp Asn Arg Lys
245 250 255

Ser Cys Ser Leu Ala Leu Arg Asn Ala Asp Val Tyr Gln Leu Cys Ser
260 265 270

Thr Pro Lys Val Asp Glu Arg Ser Asp Tyr Ser Ser Ile Gly Ile Glu
275 280 285

Asp Ile Val Leu Asp Ile Val Thr Ser Glu Gly Thr Val Ser Thr Thr
290 295 300

Arg Phe Thr Asn Asn Asn Ile Thr Phe Asp Lys Pro Tyr Ala Ala Leu
305 310 315 320

Tyr Pro Ser Val Gly Pro Gly Ile Tyr Tyr Asp Asn Lys Ile Ile Phe
325 330 335

Leu Gly Tyr Gly Gly Leu Glu His Glu Glu Asn Gly Asp Val Ile Cys
340 345 350

Asn Ile Thr Gly Cys Pro Gly Lys Thr Gln His Asp Cys Asn Gln Ala
355 360 365

Ser Tyr Ser Pro Trp Phe Ser Asn Arg Arg Met Val Asn Ala Ile Ile
370 375 380

Leu Val Asn Lys Gly Leu Asn Lys Val Pro Ser Leu Gln Val Trp Thr
385 390 395 400

Ile Pro Met Arg Gln Asn Tyr Trp Gly Ser Glu Gly Arg Leu Leu Leu
405 410 415

```
Leu Gly Asn Lys Ile Tyr Ile Tyr Thr Arg Ser Thr Ser Trp His Ser
            420                 425                 430


Lys Leu Gln Leu Gly Thr Leu Asp Ile Ser Asn Tyr Asn Asp Ile Arg
        435                 440                 445


Ile Arg Trp Thr His His Asp Val Leu Ser Arg Pro Gly Ser Glu Glu
    450                 455                 460


Cys Pro Trp Gly Asn Thr Cys Pro Arg Gly Cys Ile Thr Gly Val Tyr
465                 470                 475                 480


Asn Asp Ala Tyr Pro Leu Asn Pro Ser Gly Ser Val Val Ser Ser Val
                485                 490                 495


Ile Leu Asp Ser Arg Thr Ser Arg Glu Asn Pro Ile Ile Thr Tyr Ser
            500                 505                 510


Thr Asp Thr Ser Arg Val Asn Glu Leu Ala Ile Arg Asn Asn Thr Leu
        515                 520                 525


Ser Ala Ala Tyr Thr Thr Thr Asn Cys Val Thr His Tyr Gly Lys Gly
    530                 535                 540


Tyr Cys Phe His Ile Ile Glu Ile Asn His Lys Ser Leu Asn Thr Leu
545                 550                 555                 560


Gln Pro Met Leu Phe Lys Thr Glu Ile Pro Lys Ser Cys Asn
                565                 570


<210>  123
<211>  414
<212>  PRT
<213>  Avian metapneumovirus


<220>
<221>  MISC_FEATURE
<223>  attachment glycoprotein

<400>  123

Met Gly Ser Glu Leu Tyr Ile Ile Glu Gly Val Ser Ser Ser Glu Ile
1               5                   10                  15


Val Leu Lys Gln Val Leu Arg Arg Ser Lys Lys Ile Leu Leu Gly Leu
            20                  25                  30


Val Leu Ser Ala Leu Gly Leu Thr Leu Thr Ser Thr Ile Val Ile Ser
        35                  40                  45


Ile Cys Ile Ser Val Glu Gln Val Lys Leu Arg Gln Cys Val Asp Thr
    50                  55                  60
```

```
Tyr Trp Ala Glu Asn Gly Ser Leu His Pro Gly Gln Ser Thr Glu Asn
65                  70                  75                  80

Thr Ser Thr Arg Gly Lys Thr Thr Thr Lys Asp Pro Arg Arg Leu Gln
                85                  90                  95

Ala Thr Gly Ala Gly Lys Phe Glu Ser Cys Gly Tyr Val Gln Val Val
            100                 105                 110

Asp Gly Asp Met His Asp Arg Ser Tyr Ala Val Leu Gly Gly Val Asp
        115                 120                 125

Cys Leu Gly Leu Leu Ala Leu Cys Glu Ser Gly Pro Ile Cys Gln Gly
    130                 135                 140

Asp Thr Trp Ser Glu Asp Gly Asn Phe Cys Arg Cys Thr Phe Ser Ser
145                 150                 155                 160

His Gly Val Ser Cys Cys Lys Lys Pro Lys Ser Lys Ala Thr Thr Ala
                165                 170                 175

Gln Arg Asn Ser Lys Pro Ala Asn Ser Lys Ser Thr Pro Pro Val His
            180                 185                 190

Ser Asp Arg Ala Ser Lys Glu His Asn Pro Ser Gln Gly Glu Gln Pro
        195                 200                 205

Arg Arg Gly Pro Thr Ser Ser Lys Thr Thr Ile Ala Ser Thr Pro Ser
    210                 215                 220

Thr Glu Asp Thr Ala Lys Pro Thr Ile Ser Lys Pro Lys Leu Thr Ile
225                 230                 235                 240

Arg Pro Ser Gln Arg Gly Pro Ser Gly Ser Thr Lys Ala Ala Ser Ser
                245                 250                 255

Thr Pro Ser His Lys Thr Asn Thr Arg Gly Thr Ser Lys Thr Thr Asp
            260                 265                 270

Gln Arg Pro Arg Thr Gly Pro Thr Pro Glu Arg Pro Arg Gln Thr His
        275                 280                 285

Ser Thr Ala Thr Pro Pro Pro Thr Thr Pro Ile His Lys Gly Arg Ala
    290                 295                 300

Pro Thr Pro Lys Pro Thr Thr Asp Leu Lys Val Asn Pro Arg Glu Gly
305                 310                 315                 320

Ser Thr Ser Pro Thr Ala Ile Gln Lys Asn Pro Thr Thr Gln Ser Asn
                325                 330                 335
```

Leu Val Asp Cys Thr Leu Ser Asp Pro Asp Glu Pro Gln Arg Ile Cys
340 345 350

Tyr Gln Val Gly Thr Tyr Asn Pro Ser Gln Ser Gly Thr Cys Asn Ile
355 360 365

Glu Val Pro Lys Cys Ser Thr Tyr Gly His Ala Cys Met Ala Thr Leu
370 375 380

Tyr Asp Thr Pro Phe Asn Cys Trp Arg Arg Thr Arg Arg Cys Ile Cys
385 390 395 400

Asp Ser Gly Gly Glu Leu Ile Glu Trp Cys Cys Thr Ser Gln
405 410

<210> 124
<211> 665
<212> PRT
<213> Tupaia paramyxovirus

<220>
<221> MISC_FEATURE
<223> hemagglutinin

<400> 124

Met Asp Tyr His Ser His Thr Thr Gln Thr Gly Ser Asn Glu Thr Leu
1 5 10 15

Tyr Gln Asp Pro Leu Gln Ser Gln Ser Gly Ser Arg Asp Thr Leu Asp
20 25 30

Gly Pro Pro Ser Thr Leu Gln His Tyr Ser Asn Pro Pro Pro Tyr Ser
35 40 45

Glu Glu Asp Gln Gly Ile Asp Gly Pro Gln Arg Ser Gln Pro Leu Ser
50 55 60

Thr Pro His Gln Tyr Asp Arg Tyr Tyr Gly Val Asn Ile Gln His Thr
65 70 75 80

Arg Val Tyr Asn His Leu Gly Thr Ile Tyr Lys Gly Leu Lys Leu Ala
85 90 95

Phe Gln Ile Leu Gly Trp Val Ser Val Ile Ile Thr Met Ile Ile Thr
100 105 110

Val Thr Thr Leu Lys Lys Met Ser Asp Gly Asn Ser Gln Asp Ser Ala
115 120 125

Met Leu Lys Ser Leu Asp Glu Asn Phe Asp Ala Ile Gln Glu Val Ala
130 135 140

Asn Leu Leu Asp Asn Glu Val Arg Pro Lys Leu Gly Val Thr Met Thr

```
145                 150                 155                 160

Gln Thr Thr Phe Gln Leu Pro Lys Glu Leu Ser Glu Ile Lys Arg Tyr
                165                 170                 175

Leu Leu Arg Leu Glu Arg Asn Cys Pro Val Cys Gly Thr Glu Ala Thr
            180                 185                 190

Pro Gln Gly Ser Lys Gly Asn Ala Ser Gly Asp Thr Ala Phe Cys Pro
        195                 200                 205

Pro Cys Leu Thr Arg Gln Cys Ser Glu Asp Ser Thr His Asp Gln Gly
    210                 215                 220

Pro Gly Val Glu Gly Thr Ser Arg Asn His Lys Gly Lys Ile Asn Phe
225                 230                 235                 240

Pro His Ile Leu Gln Ser Asp Asp Cys Gly Arg Ser Asp Asn Leu Ile
                245                 250                 255

Val Tyr Ser Ile Asn Leu Val Pro Gly Leu Ser Phe Ile Gln Leu Pro
            260                 265                 270

Ser Gly Thr Lys His Cys Ile Ile Asp Val Ser Tyr Thr Phe Ser Asp
        275                 280                 285

Thr Leu Ala Gly Tyr Leu Ile Val Gly Gly Val Asp Gly Cys Gln Leu
    290                 295                 300

His Asn Lys Ala Ile Ile Tyr Leu Ser Leu Gly Tyr Tyr Lys Thr Lys
305                 310                 315                 320

Met Ile Tyr Pro Pro Asp Tyr Ile Ala Ile Ala Thr Tyr Thr Tyr Asp
                325                 330                 335

Leu Val Pro Asn Leu Arg Asp Cys Ser Ile Ala Val Asn Gln Thr Ser
            340                 345                 350

Leu Ala Ala Ile Cys Thr Ser Lys Lys Thr Lys Glu Asn Gln Asp Phe
        355                 360                 365

Ser Thr Ser Gly Val His Pro Phe Tyr Ile Phe Thr Leu Asn Thr Asp
    370                 375                 380

Gly Ile Phe Thr Val Thr Val Ile Glu Gln Ser Gln Leu Lys Leu Asp
385                 390                 395                 400

Tyr Gln Tyr Ala Ala Leu Tyr Pro Ala Thr Gly Pro Gly Ile Phe Ile
                405                 410                 415

Gly Asp His Leu Val Phe Leu Met Trp Gly Gly Leu Met Thr Lys Ala
            420                 425                 430
```

```
Glu Gly Asp Ala Tyr Cys Gln Ala Ser Gly Cys Asn Asp Ala His Arg
        435                 440                 445

Thr Ser Cys Asn Ile Ala Gln Met Pro Ser Ala Tyr Gly His Arg Gln
    450                 455                 460

Leu Val Asn Gly Leu Leu Met Leu Pro Ile Lys Glu Leu Gly Ser His
465                 470                 475                 480

Leu Ile Gln Pro Ser Leu Glu Thr Ile Ser Pro Lys Ile Asn Trp Ala
                485                 490                 495

Gly Gly His Gly Arg Leu Tyr Tyr Asn Trp Glu Ile Asn Thr Thr Tyr
            500                 505                 510

Ile Tyr Ile Glu Gly Lys Thr Trp Arg Ser Arg Pro Asn Leu Gly Ile
        515                 520                 525

Ile Ser Trp Ser Lys Pro Leu Ser Ile Arg Trp Ile Asp His Ser Val
    530                 535                 540

Ala Arg Arg Pro Gly Ala Arg Pro Cys Asp Ser Ala Asn Asp Cys Pro
545                 550                 555                 560

Glu Asp Cys Leu Val Gly Gly Tyr Tyr Asp Met Phe Pro Met Ser Ser
                565                 570                 575

Asp Tyr Lys Thr Ala Ile Thr Ile Ile Pro Thr His His Gln Trp Pro
            580                 585                 590

Ser Ser Pro Ala Leu Lys Leu Phe Asn Thr Asn Arg Glu Val Arg Val
        595                 600                 605

Val Met Ile Leu Arg Pro Pro Asn Asn Val Lys Lys Thr Thr Ile Ser
    610                 615                 620

Cys Ile Arg Ile Met Gln Thr Asn Trp Cys Leu Gly Phe Ile Ile Phe
625                 630                 635                 640

Lys Glu Gly Asn Asn Ala Trp Gly Gln Ile Tyr Ser Tyr Ile Tyr Gln
                645                 650                 655

Val Glu Ser Thr Cys Pro Asn Thr Lys
            660                 665
```

<210> 125
<211> 585
<212> PRT
<213> Avian metapneumovirus

<220>

<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 125

```
Met Glu Val Lys Val Glu Asn Val Gly Lys Ser Gln Glu Leu Lys Val
1               5                  10                  15

Lys Val Lys Asn Phe Ile Lys Arg Ser Asp Cys Lys Lys Lys Leu Phe
            20                  25                  30

Ala Leu Ile Leu Gly Leu Val Ser Phe Glu Leu Thr Met Asn Ile Met
        35                  40                  45

Leu Ser Val Met Tyr Val Glu Ser Asn Glu Ala Leu Ser Leu Cys Arg
    50                  55                  60

Ile Gln Gly Thr Pro Ala Pro Arg Asp Asn Lys Thr Asn Thr Glu Asn
65                  70                  75                  80

Ala Thr Lys Glu Thr Thr Leu His Thr Thr Thr Thr Thr Arg Asp Pro
                85                  90                  95

Glu Val Arg Glu Thr Lys Thr Thr Lys Pro Gln Ala Asn Glu Gly Ala
            100                 105                 110

Thr Asn Pro Ser Arg Asn Leu Thr Thr Lys Gly Asp Lys His Gln Thr
        115                 120                 125

Thr Arg Ala Thr Thr Glu Ala Glu Leu Glu Lys Gln Ser Lys Gln Thr
    130                 135                 140

Thr Glu Pro Gly Thr Ser Thr Gln Lys His Thr Pro Thr Arg Pro Ser
145                 150                 155                 160

Ser Lys Ser Pro Thr Thr Thr Gln Ala Ile Ala Gln Leu Thr Thr Pro
                165                 170                 175

Thr Thr Pro Lys Ala Ser Thr Ala Pro Lys Asn Arg Gln Ala Thr Thr
            180                 185                 190

Lys Lys Thr Glu Thr Asp Thr Thr Thr Ala Ser Arg Ala Arg Asn Thr
        195                 200                 205

Asn Asn Pro Thr Glu Thr Ala Thr Thr Thr Pro Lys Ala Thr Thr Glu
    210                 215                 220

Thr Gly Lys Ser Lys Glu Gly Pro Thr Gln His Thr Thr Lys Glu Gln
225                 230                 235                 240

Pro Glu Thr Thr Ala Gly Glu Thr Thr Thr Pro Gln Pro Arg Arg Thr
                245                 250                 255
```

```
Ala Ser Arg Pro Ala Pro Thr Thr Lys Ile Glu Glu Glu Ala Glu Thr
            260                 265                 270

Thr Lys Thr Arg Thr Thr Lys Ser Thr Gln Thr Ser Thr Gly Pro Pro
        275                 280                 285

Arg Pro Thr Gly Gly Ala Pro Ser Gly Ala Ala Thr Glu Gly Ser Gly
    290                 295                 300

Arg Ala Ala Ala Ala Gly Gly Pro Ser Ala Ala Ser Ala Gly Gly Arg
305                 310                 315                 320

Arg Arg Thr Glu Ala Ala Ala Glu Arg Asp Arg Arg Thr Arg Ala Gly
                325                 330                 335

Ala Gly Pro Thr Ala Gly Gly Ala Arg Ala Arg Thr Ala Ala Ala Ser
            340                 345                 350

Glu Arg Gly Ala Asp Thr Ala Gly Ser Ala Gly Gly Gly Pro Gly Gly
        355                 360                 365

Asp Gly Ala Thr Gly Gly Leu Ser Gly Gly Ala Pro Ala Glu Arg Glu
    370                 375                 380

Asp Ala Ser Gly Gly Thr Ala Ala Ala Gly Pro Gly Asp Gly Thr Glu
385                 390                 395                 400

Ala Asp Gly Arg Ala Pro Pro Ala Ala Ala Leu Ala Gly Arg Thr Thr
                405                 410                 415

Glu Ser Ala Ala Gly Ala Ala Gly Asp Ser Gly Arg Ala Gly Thr Ala
            420                 425                 430

Gly Trp Gly Ser Ala Ala Asp Gly Arg Ser Thr Gly Gly Asn Ala Ala
        435                 440                 445

Ala Glu Ala Gly Ala Ala Gln Ser Gly Arg Ala Ala Pro Arg Gln Pro
    450                 455                 460

Ser Gly Gly Thr Ala Pro Glu Ser Thr Ala Pro Pro Asn Ser Gly Gly
465                 470                 475                 480

Ser Gly Arg Ala Asp Ala Ala Pro Thr Glu Glu Val Gly Val Gly Ser
                485                 490                 495

Gly Leu Trp Arg Gly Arg Tyr Val Cys Gly Pro Cys Gly Glu Ser Val
            500                 505                 510

Pro Glu His Pro Met Asn Pro Cys Phe Gly Asp Gly Thr Ala Trp Ile
        515                 520                 525

Cys Ser Asp Asp Gly Gly Ser Leu Pro Ala Gly Cys Tyr Asp Gly Gly
```

530 535 540

Thr Asp Gly Val Val Cys Cys Gly Val Cys Gly Gly Asn Ser Cys Cys
545 550 555 560

Cys Gly Arg Val Glu Cys Thr Cys Gly Gly Gly Ala Gly Leu Leu Ser
565 570 575

Cys Cys Cys Gly Ser Tyr Ser Trp Ser
580 585

<210> 126
<211> 577
<212> PRT
<213> Avian paramyxovirus 3

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 126

Met Glu Ser Pro Pro Ser Gly Lys Asp Ala Pro Ala Phe Arg Glu Pro
1 5 10 15

Lys Arg Thr Cys Arg Leu Cys Tyr Arg Ala Thr Thr Leu Ser Leu Asn
20 25 30

Leu Thr Ile Val Val Leu Ser Ile Ile Ser Ile Tyr Val Ser Thr Gln
35 40 45

Thr Gly Ala Asn Asn Ser Cys Val Asn Pro Thr Ile Val Thr Pro Asp
50 55 60

Tyr Leu Thr Gly Ser Thr Thr Gly Ser Val Glu Asp Leu Ala Asp Leu
65 70 75 80

Glu Ser Gln Leu Arg Glu Ile Arg Arg Asp Thr Gly Ile Asn Leu Pro
85 90 95

Val Gln Ile Asp Asn Thr Glu Asn Leu Ile Leu Thr Thr Leu Ala Ser
100 105 110

Ile Asn Ser Asn Leu Arg Phe Leu Gln Asn Ala Thr Thr Glu Ser Gln
115 120 125

Thr Cys Leu Ser Pro Val Asn Asp Pro Arg Phe Val Ala Gly Ile Asn
130 135 140

Arg Ile Pro Ala Gly Ser Met Ala Tyr Asn Asp Phe Ser Asn Leu Ile
145 150 155 160

Glu His Val Asn Phe Ile Pro Ser Pro Thr Thr Leu Ser Gly Cys Thr
165 170 175

```
Arg Ile Pro Ser Phe Ser Leu Ser Lys Thr His Trp Cys Tyr Thr His
            180                 185                 190

Asn Val Ile Ser Asn Gly Cys Leu Asp His Ala Ala Ser Ser Gln Tyr
        195                 200                 205

Ile Ser Ile Gly Ile Val Asp Thr Gly Leu Asn Asn Glu Pro Tyr Phe
    210                 215                 220

Arg Thr Met Ser Ser Lys Ser Leu Asn Asp Gly Leu Asn Arg Lys Ser
225                 230                 235                 240

Cys Ser Val Thr Ala Ala Ala Asn Ala Cys Trp Leu Leu Cys Ser Val
                245                 250                 255

Val Thr Glu Tyr Glu Ala Ala Asp Tyr Arg Ser Arg Thr Pro Thr Ala
            260                 265                 270

Met Val Leu Gly Arg Phe Asp Phe Asn Gly Glu Tyr Thr Glu Ile Ala
        275                 280                 285

Val Pro Ser Ser Leu Phe Asp Gly Arg Phe Ala Ser Asn Tyr Pro Gly
    290                 295                 300

Val Gly Ser Gly Thr Gln Val Asn Gly Thr Leu Tyr Phe Pro Leu Tyr
305                 310                 315                 320

Gly Gly Val Leu Asn Gly Ser Asp Ile Glu Thr Ala Asn Lys Gly Lys
                325                 330                 335

Ser Phe Arg Pro Gln Asn Pro Lys Asn Arg Cys Pro Asp Ser Glu Ala
            340                 345                 350

Ile Gln Ser Phe Arg Ala Gln Asp Ser Tyr Tyr Pro Thr Arg Phe Gly
        355                 360                 365

Lys Val Leu Ile Gln Gln Ala Ile Ile Ala Cys Arg Ile Ser Asn Lys
    370                 375                 380

Ser Cys Thr Asp Phe Tyr Leu Leu Tyr Phe Asp Asn Asn Arg Val Met
385                 390                 395                 400

Met Gly Ala Glu Ala Arg Leu Tyr Tyr Leu Asn Asn Gln Leu Tyr Leu
                405                 410                 415

Tyr Gln Arg Ser Ser Ser Trp Trp Pro His Pro Leu Phe Tyr Ser Ile
            420                 425                 430

Ser Leu Pro Ser Cys Gln Ala Leu Ala Val Cys Gln Ile Thr Glu Ala
        435                 440                 445
```

```
His Leu Thr Leu Thr Tyr Ala Thr Ser Arg Pro Gly Met Ser Ile Cys
    450                 455                 460


Thr Gly Ala Ser Arg Cys Pro Asn Asn Cys Val Asp Gly Val Tyr Thr
465                 470                 475                 480


Asp Val Trp Pro Leu Thr Lys Asn Asp Ala Gln Asp Pro Asn Leu Phe
                485                 490                 495


Tyr Thr Val Tyr Leu Asn Asn Ser Thr Arg Arg Ile Ser Pro Thr Ile
            500                 505                 510


Ser Leu Tyr Thr Tyr Asp Arg Arg Ile Lys Ser Lys Leu Ala Val Gly
        515                 520                 525


Ser Asp Ile Gly Ala Ala Tyr Thr Thr Ser Thr Cys Phe Gly Arg Ser
    530                 535                 540


Asp Thr Gly Ala Val Tyr Cys Leu Thr Ile Met Glu Thr Val Asn Thr
545                 550                 555                 560


Ile Phe Gly Gln Tyr Arg Ile Val Pro Ile Leu Leu Arg Val Thr Ser
                565                 570                 575


Arg
```

<210> 127
<211> 585
<212> PRT
<213> Avian metapneumovirus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 127

```
Met Glu Val Lys Val Glu Asn Val Gly Lys Ser Gln Glu Leu Lys Val
1               5                   10                  15


Lys Val Lys Asn Phe Ile Lys Arg Ser Asp Cys Lys Lys Lys Leu Phe
            20                  25                  30


Ala Leu Ile Leu Gly Leu Val Ser Phe Glu Leu Thr Met Asn Ile Met
        35                  40                  45


Leu Ser Val Met Tyr Val Glu Ser Asn Glu Ala Leu Ser Leu Cys Arg
    50                  55                  60


Ile Gln Gly Thr Pro Ala Pro Arg Asp Asn Lys Thr Asn Thr Glu Asn
65                  70                  75                  80
```

```
Ala Thr Lys Glu Thr Thr Leu His Thr Thr Thr Thr Thr Arg Asp Pro
                 85                  90                  95

Glu Val Arg Glu Thr Lys Thr Thr Lys Pro Gln Ala Asn Glu Gly Ala
            100                 105                 110

Thr Asn Pro Ser Arg Asn Leu Thr Thr Lys Gly Asp Lys His Gln Thr
        115                 120                 125

Thr Arg Ala Thr Thr Glu Ala Glu Leu Glu Lys Gln Ser Lys Gln Thr
    130                 135                 140

Thr Glu Pro Gly Thr Ser Thr Gln Lys His Thr Pro Ala Arg Pro Ser
145                 150                 155                 160

Ser Lys Ser Pro Thr Thr Thr Gln Ala Thr Ala Gln Pro Thr Thr Pro
                165                 170                 175

Thr Ala Pro Lys Ala Ser Thr Ala Pro Lys Asn Arg Gln Ala Thr Thr
            180                 185                 190

Lys Lys Thr Glu Thr Asp Thr Thr Thr Ala Ser Arg Ala Arg Asn Thr
        195                 200                 205

Asn Asn Pro Thr Glu Thr Ala Thr Thr Thr Pro Lys Ala Thr Thr Glu
    210                 215                 220

Thr Gly Lys Gly Lys Glu Gly Pro Thr Gln His Thr Thr Lys Glu Gln
225                 230                 235                 240

Pro Glu Thr Thr Ala Arg Glu Thr Thr Thr Pro Gln Pro Arg Arg Thr
                245                 250                 255

Ala Ser Arg Pro Ala Pro Thr Thr Lys Ile Glu Glu Glu Ala Glu Thr
            260                 265                 270

Thr Lys Thr Arg Thr Thr Lys Asn Thr Gln Thr Ser Thr Gly Pro Pro
        275                 280                 285

Arg Pro Thr Arg Ser Thr Pro Ser Lys Thr Ala Thr Glu Asn Asn Lys
    290                 295                 300

Arg Thr Thr Thr Thr Lys Arg Pro Asn Thr Ala Ser Thr Asp Ser Arg
305                 310                 315                 320

Gln Gln Thr Arg Thr Thr Ala Glu Gln Asp Gln Gln Thr Gln Thr Arg
                325                 330                 335

Ala Lys Pro Thr Thr Asn Gly Ala His Pro Gln Thr Thr Thr Thr Pro
            340                 345                 350

Glu His Asn Thr Asp Thr Thr Asn Ser Thr Lys Gly Ser Pro Lys Glu
```

```
        355                 360                 365

Asp Lys Thr Thr Arg Asp Pro Ser Ser Lys Thr Pro Thr Glu Gln Glu
    370                 375                 380

Asp Ala Ser Lys Gly Thr Ala Ala Ala Asn Pro Gly Gly Ser Ala Glu
385                 390                 395                 400

Ala Asp Arg Arg Ala Pro Pro Ala Thr Thr Pro Thr Gly Arg Thr Thr
                405                 410                 415

Glu Ser Ala Ala Gly Thr Thr Gly Asp Asp Ser Gly Ala Glu Thr Thr
            420                 425                 430

Arg Arg Arg Ser Ala Ala Asp Arg Arg Pro Thr Gly Gly Ser Thr Ala
        435                 440                 445

Ala Glu Ala Gly Thr Ala Gln Ser Gly Arg Ala Thr Pro Lys Gln Pro
    450                 455                 460

Ser Gly Gly Thr Ala Ala Gly Asn Thr Ala Pro Pro Asn Asn Glu Ser
465                 470                 475                 480

Ser Gly Arg Ala Asp Ala Ala Pro Ala Glu Glu Ala Gly Val Gly Pro
                485                 490                 495

Ser Ile Arg Arg Gly Arg His Ala Cys Gly Pro Arg Arg Glu Ser Ala
            500                 505                 510

Pro Glu His Pro Thr Asn Pro Cys Pro Gly Asp Gly Thr Ala Trp Thr
        515                 520                 525

Arg Ser Asp Gly Gly Gly Asn Leu Pro Ala Gly Arg His Asp Ser Gly
    530                 535                 540

Ala Asp Gly Ala Ala Arg Arg Gly Ala Arg Gly Gly Asn Pro Arg Arg
545                 550                 555                 560

Arg Gly Arg Ala Glu Arg Thr Arg Gly Gly Gly Ala Gly Pro Pro Ser
                565                 570                 575

Cys Arg Cys Gly Ser His Asn Arg Ser
            580                 585


<210>  128
<211>  389
<212>  PRT
<213>  Avian metapneumovirus type D


<220>
<221>  MISC_FEATURE
<223>  attachment glycoprotein
```

<400> 128

Met Gly Ala Lys Leu Tyr Ala Ile Ser Gly Ala Ser Asp Ala Gln Leu
1 5 10 15

Met Lys Lys Thr Cys Ala Lys Leu Leu Glu Lys Val Val Pro Ile Ile
20 25 30

Ile Leu Ala Val Leu Gly Ile Thr Gly Thr Thr Thr Ile Ala Leu Ser
35 40 45

Ile Ser Ile Ser Ile Glu Arg Ala Val Leu Ser Asp Cys Thr Thr Gln
50 55 60

Leu Arg Asn Gly Thr Thr Ser Gly Ser Leu Ser Asn Pro Thr Arg Ser
65 70 75 80

Thr Thr Ser Thr Ala Val Thr Thr Arg Asp Ile Arg Gly Leu Gln Thr
85 90 95

Thr Arg Thr Arg Glu Leu Lys Ser Cys Ser Asn Val Gln Ile Ala Tyr
100 105 110

Gly Tyr Leu His Asp Ser Ser Asn Pro Val Leu Asp Ser Ile Gly Cys
115 120 125

Leu Gly Leu Leu Ala Leu Cys Glu Ser Gly Pro Phe Cys Gln Arg Asn
130 135 140

Tyr Asn Pro Arg Asp Arg Pro Lys Cys Arg Cys Thr Leu Arg Gly Lys
145 150 155 160

Asp Ile Ser Cys Cys Lys Glu Pro Pro Thr Ala Val Thr Thr Ser Lys
165 170 175

Thr Thr Pro Trp Gly Thr Glu Val His Pro Thr Tyr Pro Thr Gln Val
180 185 190

Thr Pro Gln Ser Gln Pro Ala Thr Met Ala His Gln Thr Ala Thr Ala
195 200 205

Asn Gln Arg Ser Ser Thr Thr Glu Pro Val Gly Ser Gln Gly Asn Thr
210 215 220

Thr Ser Ser Asn Pro Glu Gln Gln Thr Glu Pro Pro Pro Ser Pro Gln
225 230 235 240

His Pro Pro Thr Thr Thr Ser Gln Asp Gln Ser Thr Glu Thr Ala Asp
245 250 255

Gly Gln Glu His Thr Pro Thr Arg Lys Thr Pro Thr Ala Thr Ser Asn
260 265 270

Arg Arg Ser Pro Thr Pro Lys Arg Gln Glu Thr Gly Arg Ala Thr Pro
275 280 285

Arg Asn Thr Ala Thr Thr Gln Ser Gly Ser Ser Pro Pro His Ser Ser
290 295 300

Pro Pro Gly Val Asp Ala Asn Met Glu Gly Gln Cys Lys Glu Leu Gln
305 310 315 320

Ala Pro Lys Pro Asn Ser Val Cys Lys Gly Leu Asp Ile Tyr Arg Glu
325 330 335

Ala Leu Pro Arg Gly Cys Asp Lys Val Leu Pro Leu Cys Lys Thr Ser
340 345 350

Thr Ile Met Cys Val Asp Ala Tyr Tyr Ser Lys Pro Pro Ile Cys Phe
355 360 365

Gly Tyr Asn Gln Arg Cys Phe Cys Met Glu Thr Phe Gly Pro Ile Glu
370 375 380

Phe Cys Cys Lys Ser
385

<210> 129
<211> 197
<212> PRT
<213> Respiratory syncytial virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 129

Met Ser Lys Asn Lys Asn Gln Arg Thr Ala Arg Thr Leu Glu Lys Thr
1 5 10 15

Trp Asp Thr Leu Asn His Leu Ile Val Ile Ser Ser Cys Leu Tyr Lys
20 25 30

Leu Asn Leu Lys Ser Ile Ala Gln Ile Ala Leu Ser Val Leu Ala Met
35 40 45

Ile Ile Ser Thr Ser Leu Ile Ile Ala Ala Ile Ile Phe Ile Ile Ser
50 55 60

Ala Asn His Lys Val Thr Leu Thr Thr Val Thr Val Gln Thr Ile Lys
65 70 75 80

Asn His Thr Glu Lys Asn Ile Thr Thr Tyr Leu Thr Gln Val Ser Pro
85 90 95

Glu Arg Val Ser Pro Ser Lys Gln Pro Thr Thr Thr Pro Pro Ile His
100 105 110

Thr Asn Ser Ala Thr Ile Ser Pro Asn Thr Lys Ser Glu Ile His His
115 120 125

Thr Thr Ala Gln Thr Lys Gly Arg Thr Ser Thr Pro Thr Gln Asn Asn
130 135 140

Lys Pro Asn Thr Lys Pro Arg Pro Lys Asn Pro Pro Lys Lys Asp Asp
145 150 155 160

Tyr His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile Cys Gly Asn
165 170 175

Asn Gln Leu Cys Lys Ser Ile Cys Lys Thr Ile Pro Ser Asn Lys Pro
180 185 190

Arg Lys Asn Gln Pro
195

<210> 130
<211> 580
<212> PRT
<213> Avian paramyxovirus 14

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 130

Met Glu Gly Ser Arg Thr Val Ile Tyr Gln Gly Asp Pro Asn Glu Lys
1 5 10 15

Asn Thr Trp Arg Leu Val Phe Arg Thr Leu Thr Leu Ile Leu Asn Leu
20 25 30

Ala Ile Leu Ser Val Thr Ile Ala Ser Ile Ile Ile Thr Ser Lys Ile
35 40 45

Thr Leu Ser Glu Val Thr Thr Leu Lys Thr Glu Gly Val Glu Glu Val
50 55 60

Ile Thr Pro Leu Met Ala Thr Leu Ser Asp Ser Val Gln Gln Glu Lys
65 70 75 80

Met Ile Tyr Lys Glu Val Ala Ile Ser Ile Pro Leu Val Leu Asp Lys
85 90 95

Ile Gln Thr Asp Val Gly Thr Ser Val Ala Gln Ile Thr Asp Ala Leu
100 105 110

Arg Gln Ile Gln Gly Val Asn Gly Thr Gln Ala Phe Ala Leu Ser Asn

```
        115                 120                 125

Ala Pro Glu Tyr Ser Gly Gly Ile Glu Val Pro Leu Phe Gln Ile Asp
    130                 135                 140

Ser Phe Val Asn Lys Ser Met Ser Ile Ser Gly Leu Leu Glu His Ala
145                 150                 155                 160

Ser Phe Ile Pro Ser Pro Thr Thr Leu His Gly Cys Thr Arg Ile Pro
                165                 170                 175

Ser Phe His Leu Gly Pro Arg His Trp Cys Tyr Thr His Asn Ile Ile
            180                 185                 190

Gly Ser Arg Cys Arg Asp Glu Gly Phe Ser Ser Met Tyr Ile Ser Ile
        195                 200                 205

Gly Ala Ile Thr Val Asn Arg Asp Gly Asn Pro Leu Phe Ile Thr Thr
    210                 215                 220

Ala Ser Thr Ile Leu Ala Asp Asp Asn Asn Arg Lys Ser Cys Ser Ile
225                 230                 235                 240

Ile Ala Ser Ser Tyr Gly Cys Asp Leu Leu Cys Ser Ile Val Thr Glu
                245                 250                 255

Ser Glu Asn Asp Asp Tyr Ala Asn Pro Asn Pro Thr Lys Met Val His
            260                 265                 270

Gly Arg Phe Leu Tyr Asn Gly Ser Tyr Val Glu Gln Ala Leu Pro Asn
        275                 280                 285

Ser Leu Phe Gln Asp Lys Trp Val Ala Gln Tyr Pro Gly Val Gly Ser
    290                 295                 300

Gly Ile Thr Thr His Gly Lys Val Leu Phe Pro Ile Tyr Gly Gly Ile
305                 310                 315                 320

Lys Lys Asn Thr Gln Leu Phe Tyr Glu Leu Ser Lys Tyr Gly Phe Phe
                325                 330                 335

Ala His Asn Lys Glu Leu Glu Cys Lys Asn Met Thr Glu Glu Gln Ile
            340                 345                 350

Arg Asp Ile Lys Ala Ala Tyr Leu Pro Ser Lys Thr Ser Gly Asn Leu
        355                 360                 365

Phe Ala Gln Gly Ile Ile Tyr Cys Asn Ile Ser Lys Leu Gly Asp Cys
    370                 375                 380

Asn Val Ala Val Leu Asn Thr Ser Thr Thr Met Met Gly Ala Glu Gly
385                 390                 395                 400
```

Arg Leu Gln Met Met Gly Glu Tyr Val Tyr Tyr Tyr Gln Arg Ser Ser
405 410 415

Ser Trp Trp Pro Val Gly Ile Val Tyr Lys Lys Ser Leu Ala Glu Leu
420 425 430

Met Asn Gly Ile Asn Met Glu Val Leu Ser Phe Glu Pro Ile Pro Leu
435 440 445

Ser Lys Phe Pro Arg Pro Thr Trp Thr Ala Gly Leu Cys Gln Lys Pro
450 455 460

Ser Ile Cys Pro Asp Val Cys Val Thr Gly Val Tyr Thr Asp Leu Phe
465 470 475 480

Ser Val Thr Ile Gly Ser Thr Thr Asp Lys Asp Thr Tyr Phe Gly Val
485 490 495

Tyr Leu Asp Ser Ala Thr Glu Arg Lys Asp Pro Trp Val Ala Ala Ala
500 505 510

Asp Gln Tyr Glu Trp Arg Asn Arg Val Arg Leu Phe Glu Ser Thr Thr
515 520 525

Glu Ala Ala Tyr Thr Thr Ser Thr Cys Phe Lys Asn Thr Val Asn Asn
530 535 540

Arg Val Phe Cys Val Ser Ile Val Glu Leu Arg Glu Asn Leu Leu Gly
545 550 555 560

Asp Trp Lys Ile Val Pro Leu Leu Phe Gln Ile Gly Val Ser Gln Gly
565 570 575

Pro Pro Pro Lys
580

<210> 131
<211> 1046
<212> PRT
<213> Beilong virus

<220>
<221> MISC_FEATURE
<223> attachment glycoprotein

<400> 131

Met Ser Gln Leu Ala Ala His Asn Leu Ala Met Ser Asn Phe Tyr Gly
1 5 10 15

Thr His Gln Gly Asp Leu Ser Gly Ser Gln Lys Gly Glu Glu Gln Gln
20 25 30

```
Val Gln Gly Val Ile Arg Tyr Val Ser Met Ile Val Gly Leu Leu Ser
        35                  40                  45

Leu Phe Thr Ile Ile Ala Leu Asn Val Thr Asn Ile Ile Tyr Met Thr
    50                  55                  60

Glu Ser Gly Gly Thr Met Gln Ser Ile Lys Thr Ala Gln Gly Ser Ile
65                  70                  75                  80

Asp Gly Ser Met Arg Glu Ile Ser Gly Val Ile Met Glu Asp Val Lys
                85                  90                  95

Pro Lys Thr Asp Leu Ile Asn Ser Met Val Ser Tyr Asn Ile Pro Ala
            100                 105                 110

Gln Leu Ser Met Ile His Gln Ile Ile Lys Asn Asp Val Leu Lys Gln
        115                 120                 125

Cys Thr Pro Ser Phe Met Phe Asn Asn Thr Ile Cys Pro Leu Ala Glu
    130                 135                 140

Asn Pro Thr His Ser Arg Tyr Phe Glu Glu Val Asn Leu Asp Ser Ile
145                 150                 155                 160

Ser Glu Cys Ser Gly Pro Asp Met His Leu Gly Leu Gly Val Asn Pro
                165                 170                 175

Glu Phe Ile Glu Phe Pro Ser Phe Ala Pro Gly Ser Thr Lys Pro Gly
            180                 185                 190

Ser Cys Val Arg Leu Pro Ser Phe Ser Leu Ser Thr Thr Val Phe Ala
        195                 200                 205

Tyr Thr His Thr Ile Met Gly His Gly Cys Ser Glu Leu Asp Val Gly
    210                 215                 220

Asp His Tyr Phe Ser Val Gly Arg Ile Ala Asp Ala Gly His Glu Ile
225                 230                 235                 240

Pro Gln Phe Glu Thr Ile Ser Ser Trp Phe Ile Asn Asp Lys Ile Asn
                245                 250                 255

Arg Arg Ser Cys Thr Val Ala Ala Gly Ala Met Glu Ala Trp Met Gly
            260                 265                 270

Cys Val Ile Met Thr Glu Thr Phe Tyr Asp Asp Leu Asn Ser Leu Asp
        275                 280                 285

Thr Gly Lys Leu Thr Ile Ser Tyr Leu Asp Val Phe Gly Arg Lys Lys
    290                 295                 300
```

Glu Trp Ile Tyr Thr Arg Ser Glu Ile Leu Tyr Asp Tyr Thr Tyr Thr
305 310 315 320

Ser Val Tyr Phe Ser Val Gly Ser Gly Val Val Val Gly Asp Thr Val
325 330 335

Tyr Phe Leu Ile Trp Gly Ser Leu Ser Ser Pro Ile Glu Glu Thr Ala
340 345 350

Tyr Cys Phe Ala Pro Asp Cys Ser Asn Tyr Asn Gln Arg Met Cys Asn
355 360 365

Glu Ala Gln Arg Pro Ser Lys Phe Gly His Arg Gln Met Val Asn Gly
370 375 380

Ile Leu Lys Phe Lys Thr Thr Ser Thr Gly Lys Pro Leu Leu Ser Val
385 390 395 400

Gly Thr Leu Ser Pro Ser Val Val Pro Phe Gly Ser Glu Gly Arg Leu
405 410 415

Met Tyr Ser Glu Ile Thr Lys Ile Ile Tyr Leu Tyr Leu Arg Ser Thr
420 425 430

Ser Trp His Ala Leu Pro Leu Thr Gly Leu Phe Val Leu Gly Pro Pro
435 440 445

Thr Ser Ile Ser Trp Ile Val Gln Arg Ala Val Ser Arg Pro Gly Glu
450 455 460

Phe Pro Cys Gly Ala Ser Asn Arg Cys Pro Lys Asp Cys Val Thr Gly
465 470 475 480

Val Tyr Thr Asp Leu Phe Pro Leu Gly Ser Arg Tyr Glu Tyr Ala Ala
485 490 495

Thr Val Tyr Leu Asn Ser Glu Thr Tyr Arg Val Asn Pro Thr Leu Ala
500 505 510

Leu Ile Asn Gln Thr Asn Ile Ile Ala Ser Lys Lys Val Thr Thr Glu
515 520 525

Ser Gln Arg Ala Gly Tyr Thr Thr Thr Thr Cys Phe Val Phe Lys Leu
530 535 540

Arg Val Trp Cys Ile Ser Val Val Glu Leu Ala Pro Ser Thr Met Thr
545 550 555 560

Ala Tyr Glu Pro Ile Pro Phe Leu Tyr Gln Leu Asp Leu Thr Cys Lys
565 570 575

Gly Lys Asn Gly Ser Leu Ala Met Arg Phe Thr Gly Lys Glu Gly Thr

580 585 590

Tyr Lys Ser Gly Arg Tyr Lys Ser Pro Arg Asn Glu Cys Phe Phe Glu
595 600 605

Lys Val Ser Asn Lys Tyr Tyr Phe Ile Val Ser Thr Pro Glu Gly Ile
610 615 620

Gln Pro Tyr Glu Ile Arg Asp Leu Thr Pro Asp Arg Met Pro His Ile
625 630 635 640

Ile Met Tyr Ile Ser Asp Val Cys Ala Pro Ala Leu Ser Ala Phe Lys
645 650 655

Lys Leu Leu Pro Ala Met Arg Pro Ile Thr Thr Leu Thr Ile Gly Asn
660 665 670

Trp Gln Phe Arg Pro Val Glu Val Ser Gly Gly Leu Arg Val Ser Ile
675 680 685

Gly Arg Asn Leu Thr Lys Glu Gly Asp Leu Thr Met Ser Ala Pro Glu
690 695 700

Asp Pro Gly Ser Asn Thr Phe Pro Gly Gly His Ile Pro Gly Asn Gly
705 710 715 720

Leu Phe Asp Ala Gly Tyr Tyr Thr Val Glu Tyr Pro Lys Glu Trp Lys
725 730 735

Gln Thr Thr Pro Lys Pro Ser Glu Gly Gly Asn Ile Ile Asp Lys Asn
740 745 750

Lys Thr Pro Val Ile Pro Ser Arg Asp Asn Pro Thr Ser Asp Ser Ser
755 760 765

Ile Pro His Arg Glu Ser Ile Glu Pro Val Arg Pro Thr Arg Glu Val
770 775 780

Leu Lys Ser Ser Asp Tyr Val Thr Ile Val Ser Thr Asp Ser Gly Ser
785 790 795 800

Gly Ser Gly Asp Phe Ala Thr Gly Val Pro Trp Thr Gly Val Ser Pro
805 810 815

Lys Ala Pro Gln Asn Gly Ile Asn Leu Pro Gly Thr Glu Leu Pro His
820 825 830

Pro Thr Val Leu Asp Arg Ile Asn Thr Pro Ala Pro Ser Asp Pro Lys
835 840 845

Val Ser Ala Asp Ser Asp His Thr Arg Asp Thr Ile Asp Pro Thr Ala
850 855 860

```
Leu Ser Lys Pro Leu Asn His Asp Thr Thr Gly Asp Thr Asp Thr Arg
865                 870                 875                 880


Ile Asn Thr Gly Thr Ala Thr Tyr Gly Phe Thr Pro Gly Arg Glu Ala
                885                 890                 895


Thr Ser Ser Gly Lys Leu Ala Asn Asp Leu Thr Asn Ser Thr Ser Val
            900                 905                 910


Pro Ser Glu Ala His Pro Ser Ala Ser Thr Ser Glu Ala Ser Lys Pro
        915                 920                 925


Glu Lys Asn Thr Asp Asn Arg Val Thr Gln Asp Pro Thr Ser Gly Thr
    930                 935                 940


Ala Glu Arg Pro Thr Thr Asn Ala Pro Val Asp Gly Lys His Ser Thr
945                 950                 955                 960


Gln Leu Thr Asp Ala Arg Pro Asn Thr Ala Asp Pro Glu Arg Thr Ser
                965                 970                 975


Gln His Ser Ser Ser Thr Thr Arg Asp Glu Val Lys Pro Ser Leu Pro
            980                 985                 990


Ser Thr Thr Glu Ala Ser Thr His  Gln Arg Thr Glu Ala  Ala Thr Pro
        995                 1000                1005


Pro Glu  Leu Val Asn Asn Thr  Leu Asn Pro Pro Ser  Thr Gln Val
    1010                1015                1020


Arg Ser  Val Arg Ser Leu Met  Gln Asp Ala Ile Ala  Gln Ala Trp
    1025                1030                1035


Asn Phe  Val Arg Gly Val Thr  Pro
    1040                1045


<210>  132
<211>  618
<212>  PRT
<213>  Avian paramyxovirus UPO216


<220>
<221>  MISC_FEATURE
<223>  hemagglutinin-neuraminidase protein

<400>  132

Met Glu Arg Gly Ile Ser Glu Val Ala Leu Ala Asn Asp Arg Thr Glu
1               5                   10                  15


Glu Lys Asn Thr Trp Arg Leu Ile Phe Arg Ile Thr Val Leu Val Val
            20                  25                  30
```

```
Ser Val Ile Thr Leu Gly Leu Thr Ala Ala Ser Leu Val Tyr Ser Met
        35                  40                  45

Asn Ala Ala Gln Pro Ala Asp Phe Asp Gly Ile Ile Pro Ala Val Gln
    50                  55                  60

Gln Val Gly Thr Ser Leu Thr Asn Ser Ile Gly Gly Met Gln Asp Val
65                  70                  75                  80

Leu Asp Arg Thr Tyr Lys Gln Val Ala Leu Glu Ser Pro Leu Thr Leu
                85                  90                  95

Leu Asn Met Glu Ser Thr Ile Met Asn Ala Ile Thr Ser Leu Ser Tyr
            100                 105                 110

Lys Ile Asn Asn Gly Gly Asn Ser Ser Gly Cys Gly Ala Pro Ile His
        115                 120                 125

Asp Pro Glu Tyr Ile Gly Gly Ile Gly Lys Glu Leu Leu Ile Asp Asp
    130                 135                 140

Asn Val Asp Val Thr Ser Phe Tyr Pro Ser Ala Phe Lys Glu His Leu
145                 150                 155                 160

Asn Phe Ile Pro Ala Pro Thr Thr Gly Ala Gly Cys Thr Arg Ile Pro
                165                 170                 175

Ser Phe Asp Leu Ser Ala Thr His Tyr Cys Tyr Thr His Asn Val Ile
            180                 185                 190

Leu Ser Gly Cys Gln Asp His Ser His Ser His Gln Tyr Ile Ala Leu
        195                 200                 205

Gly Val Leu Lys Leu Ser Asp Thr Gly Asn Val Phe Phe Ser Thr Leu
    210                 215                 220

Arg Ser Ile Asn Leu Asp Asp Thr Ala Asn Arg Lys Ser Cys Ser Ile
225                 230                 235                 240

Ser Ala Thr Pro Leu Gly Cys Asp Ile Leu Cys Ser Lys Val Thr Glu
                245                 250                 255

Thr Glu Leu Glu Asp Tyr Lys Ser Glu Glu Pro Thr Pro Met Val His
            260                 265                 270

Gly Arg Leu Ser Phe Asp Gly Thr Tyr Ser Glu Lys Asp Leu Asp Val
        275                 280                 285

Asn Asn Leu Phe Ser Asp Trp Thr Ala Asn Tyr Pro Ser Val Gly Gly
    290                 295                 300
```

```
Gly Ser Tyr Ile Gly Asn Arg Val Trp Tyr Ala Val Tyr Gly Gly Leu
305                 310                 315                 320

Lys Pro Gly Ser Asn Thr Asp Gln Ser Gln Arg Asp Lys Tyr Val Ile
                325                 330                 335

Tyr Lys Arg Tyr Asn Asn Thr Cys Pro Asp Pro Glu Asp Tyr Gln Ile
            340                 345                 350

Asn Lys Ala Lys Ser Ser Tyr Thr Pro Ser Tyr Phe Gly Ser Lys Arg
        355                 360                 365

Val Gln Gln Ala Ile Leu Ser Ile Ala Val Ser Pro Thr Leu Gly Ser
    370                 375                 380

Asp Pro Val Leu Thr Pro Leu Ser Asn Asp Val Val Leu Met Gly Ala
385                 390                 395                 400

Glu Gly Arg Val Met His Ile Gly Gly Tyr Thr Tyr Leu Tyr Gln Arg
                405                 410                 415

Gly Thr Ser Tyr Tyr Ser Pro Ala Leu Leu Tyr Pro Leu Asn Ile Gln
            420                 425                 430

Asp Lys Ser Ala Thr Ala Ser Ser Pro Tyr Lys Phe Asp Ala Phe Thr
        435                 440                 445

Arg Pro Gly Ser Val Pro Cys Gln Ala Asp Ala Arg Cys Pro Gln Ser
    450                 455                 460

Cys Val Thr Gly Val Tyr Thr Asp Pro Tyr Pro Leu Ile Phe Ala Lys
465                 470                 475                 480

Asp His Ser Ile Arg Gly Val Tyr Gly Met Met Leu Asn Asp Val Thr
                485                 490                 495

Ala Arg Leu Asn Pro Ile Ala Ala Val Phe Ser Asn Ile Ser Arg Ser
            500                 505                 510

Gln Ile Thr Arg Val Ser Ser Ser Ser Thr Lys Ala Ala Tyr Thr Thr
        515                 520                 525

Ser Thr Cys Phe Lys Val Ile Lys Thr Asn Arg Ile Tyr Cys Met Ser
    530                 535                 540

Ile Ala Glu Ile Ser Asn Thr Leu Phe Gly Glu Phe Arg Ile Val Pro
545                 550                 555                 560

Leu Leu Val Glu Ile Leu Ser Asn Gly Gly Asn Thr Ala Arg Ser Ala
                565                 570                 575

Gly Gly Thr Pro Val Lys Glu Ser Pro Lys Gly Trp Ser Asp Ala Ile
```

580 585 590

Ala Glu Pro Leu Phe Cys Thr Pro Thr Asn Val Thr Arg Tyr Asn Ala
595 600 605

Asp Ile Arg Arg Tyr Ala Tyr Ser Trp Pro
610 615

<210> 133
<211> 576
<212> PRT
<213> Atlantic salmon paramyxovirus

<220>
<221> MISC_FEATURE
<223> hemagglutinin-neuraminidase protein

<400> 133

Met Pro Pro Ala Pro Ser Pro Val His Asp Pro Ser Ser Phe Tyr Gly
1 5 10 15

Ser Ser Leu Phe Asn Glu Asp Thr Ala Ser Arg Lys Gly Thr Ser Glu
20 25 30

Glu Ile His Leu Leu Gly Ile Arg Trp Asn Thr Val Leu Ile Val Leu
35 40 45

Gly Leu Ile Leu Ala Ile Ile Gly Ile Gly Ile Gly Ala Ser Ser Phe
50 55 60

Ser Ala Ser Gly Ile Thr Gly Asn Thr Thr Lys Glu Ile Arg Leu Ile
65 70 75 80

Val Glu Glu Met Ser Tyr Gly Leu Val Arg Ile Ser Asp Ser Val Arg
85 90 95

Gln Glu Ile Ser Pro Lys Val Thr Leu Leu Gln Asn Ala Val Leu Ser
100 105 110

Ser Ile Pro Ala Leu Val Thr Thr Glu Thr Asn Thr Ile Ile Asn Ala
115 120 125

Val Lys Asn His Cys Asn Ser Pro Pro Thr Pro Pro Pro Pro Thr Glu
130 135 140

Ala Pro Leu Lys Lys His Glu Thr Gly Met Ala Pro Leu Asp Pro Thr
145 150 155 160

Thr Tyr Trp Thr Cys Thr Ser Gly Thr Pro Arg Phe Tyr Ser Ser Pro
165 170 175

Asn Ala Thr Phe Ile Pro Gly Pro Ser Pro Leu Pro His Thr Ala Thr
180 185 190

```
Pro Gly Gly Cys Val Arg Ile Pro Ser Met His Ile Gly Ser Glu Ile
        195                 200                 205

Tyr Ala Tyr Thr Ser Asn Leu Ile Ala Ser Gly Cys Gln Asp Ile Gly
    210                 215                 220

Lys Ser Tyr Gln Asn Val Gln Ile Gly Val Leu Asp Arg Thr Pro Glu
225                 230                 235                 240

Gly Asn Pro Glu Met Ser Pro Met Leu Ser His Thr Phe Pro Ile Asn
                245                 250                 255

Asp Asn Arg Lys Ser Cys Ser Ile Val Thr Leu Lys Arg Ala Ala Tyr
            260                 265                 270

Ile Tyr Cys Ser Gln Pro Lys Val Thr Glu Phe Val Asp Tyr Gln Thr
        275                 280                 285

Pro Gly Ile Glu Pro Met Ser Leu Asp His Ile Asn Ala Asn Gly Thr
    290                 295                 300

Thr Lys Thr Trp Ile Tyr Ser Pro Thr Glu Val Val Thr Asp Val Pro
305                 310                 315                 320

Tyr Ala Ser Met Tyr Pro Ser Val Gly Ser Gly Val Val Ile Asp Gly
                325                 330                 335

Lys Leu Val Phe Leu Val Tyr Gly Gly Leu Leu Asn Gly Ile Gln Val
            340                 345                 350

Pro Ala Met Cys Leu Ser Pro Glu Cys Pro Gly Ile Asp Gln Ala Ala
        355                 360                 365

Cys Asn Ala Ser Gln Tyr Asn Gln Tyr Leu Ser Gly Arg Gln Val Val
    370                 375                 380

Asn Gly Ile Ala Thr Val Asp Leu Met Asn Gly Gln Lys Pro His Ile
385                 390                 395                 400

Ser Val Glu Thr Ile Ser Pro Ser Lys Asn Trp Phe Gly Ala Glu Gly
                405                 410                 415

Arg Leu Val Tyr Met Gly Gly Arg Leu Tyr Ile Tyr Ile Arg Ser Thr
            420                 425                 430

Gly Trp His Ser Pro Ile Gln Ile Gly Val Ile Tyr Thr Met Asn Pro
        435                 440                 445

Leu Ala Ile Thr Trp Val Thr Asn Thr Val Leu Ser Arg Pro Gly Ser
    450                 455                 460
```

```
Ala Gly Cys Asp Trp Asn Asn Arg Cys Pro Lys Ala Cys Leu Ser Gly
465                 470                 475                 480


Val Tyr Thr Asp Ala Tyr Pro Ile Ser Pro Asp Tyr Asn His Leu Ala
                485                 490                 495


Thr Met Ile Leu His Ser Thr Ser Thr Arg Ser Asn Pro Val Met Val
            500                 505                 510


Tyr Ser Ser Pro Thr Asn Met Val Asn Tyr Ala Gln Leu Thr Thr Thr
        515                 520                 525


Ala Gln Ile Ala Gly Tyr Thr Thr Thr Ser Cys Phe Thr Asp Asn Glu
    530                 535                 540


Val Gly Tyr Cys Ala Thr Ala Leu Glu Leu Thr Pro Gly Thr Leu Ser
545                 550                 555                 560


Ser Val Gln Pro Ile Leu Val Met Thr Lys Ile Pro Lys Glu Cys Val
                565                 570                 575


<210>  134
<211>  8797
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  SYNE1

<400>  134

Met Ala Thr Ser Arg Gly Ala Ser Arg Cys Pro Arg Asp Ile Ala Asn
1               5                   10                  15


Val Met Gln Arg Leu Gln Asp Glu Gln Glu Ile Val Gln Lys Arg Thr
            20                  25                  30


Phe Thr Lys Trp Ile Asn Ser His Leu Ala Lys Arg Lys Pro Pro Met
        35                  40                  45


Val Val Asp Asp Leu Phe Glu Asp Met Lys Asp Gly Val Lys Leu Leu
    50                  55                  60


Ala Leu Leu Glu Val Leu Ser Gly Gln Lys Leu Pro Cys Glu Gln Gly
65                  70                  75                  80


Arg Arg Met Lys Arg Ile His Ala Val Ala Asn Ile Gly Thr Ala Leu
                85                  90                  95


Lys Phe Leu Glu Gly Arg Lys Ile Lys Leu Val Asn Ile Asn Ser Thr
            100                 105                 110
```

```
Asp Ile Ala Asp Gly Arg Pro Ser Ile Val Leu Gly Leu Met Trp Thr
        115                 120                 125

Ile Ile Leu Tyr Phe Gln Ile Glu Glu Leu Thr Ser Asn Leu Pro Gln
    130                 135                 140

Leu Gln Ser Leu Ser Ser Ser Ala Ser Ser Val Asp Ser Ile Val Ser
145                 150                 155                 160

Ser Glu Thr Pro Ser Pro Pro Ser Lys Arg Lys Val Thr Thr Lys Ile
                165                 170                 175

Gln Gly Asn Ala Lys Lys Ala Leu Leu Lys Trp Val Gln Tyr Thr Ala
            180                 185                 190

Gly Lys Gln Thr Gly Ile Glu Val Lys Asp Phe Gly Lys Ser Trp Arg
        195                 200                 205

Ser Gly Val Ala Phe His Ser Val Ile His Ala Ile Arg Pro Glu Leu
    210                 215                 220

Val Asp Leu Glu Thr Val Lys Gly Arg Ser Asn Arg Glu Asn Leu Glu
225                 230                 235                 240

Asp Ala Phe Thr Ile Ala Glu Thr Glu Leu Gly Ile Pro Arg Leu Leu
                245                 250                 255

Asp Pro Glu Asp Val Asp Val Asp Lys Pro Asp Glu Lys Ser Ile Met
            260                 265                 270

Thr Tyr Val Ala Gln Phe Leu Lys His Tyr Pro Asp Ile His Asn Ala
        275                 280                 285

Ser Thr Asp Gly Gln Glu Asp Asp Glu Ile Leu Pro Gly Phe Pro Ser
    290                 295                 300

Phe Ala Asn Ser Val Gln Asn Phe Lys Arg Glu Asp Arg Val Ile Phe
305                 310                 315                 320

Lys Glu Met Lys Val Trp Ile Glu Gln Phe Glu Arg Asp Leu Thr Arg
                325                 330                 335

Ala Gln Met Val Glu Ser Asn Leu Gln Asp Lys Tyr Gln Ser Phe Lys
            340                 345                 350

His Phe Arg Val Gln Tyr Glu Met Lys Arg Lys Gln Ile Glu His Leu
        355                 360                 365

Ile Gln Pro Leu His Arg Asp Gly Lys Leu Ser Leu Asp Gln Ala Leu
    370                 375                 380

Val Lys Gln Ser Trp Asp Arg Val Thr Ser Arg Leu Phe Asp Trp His
```

385 390 395 400

Ile Gln Leu Asp Lys Ser Leu Pro Ala Pro Leu Gly Thr Ile Gly Ala
405 410 415

Trp Leu Tyr Arg Ala Glu Val Ala Leu Arg Glu Glu Ile Thr Val Gln
420 425 430

Gln Val His Glu Glu Thr Ala Asn Thr Ile Gln Arg Lys Leu Glu Gln
435 440 445

His Lys Asp Leu Leu Gln Asn Thr Asp Ala His Lys Arg Ala Phe His
450 455 460

Glu Ile Tyr Arg Thr Arg Ser Val Asn Gly Ile Pro Val Pro Pro Asp
465 470 475 480

Gln Leu Glu Asp Met Ala Glu Arg Phe His Phe Val Ser Ser Thr Ser
485 490 495

Glu Leu His Leu Met Lys Met Glu Phe Leu Glu Leu Lys Tyr Arg Leu
500 505 510

Leu Ser Leu Leu Val Leu Ala Glu Ser Lys Leu Lys Ser Trp Ile Ile
515 520 525

Lys Tyr Gly Arg Arg Glu Ser Val Glu Gln Leu Leu Gln Asn Tyr Val
530 535 540

Ser Phe Ile Glu Asn Ser Lys Phe Phe Glu Gln Tyr Glu Val Thr Tyr
545 550 555 560

Gln Ile Leu Lys Gln Thr Ala Glu Met Tyr Val Lys Ala Asp Gly Ser
565 570 575

Val Glu Glu Ala Glu Asn Val Met Lys Phe Met Asn Glu Thr Thr Ala
580 585 590

Gln Trp Arg Asn Leu Ser Val Glu Val Arg Ser Val Arg Ser Met Leu
595 600 605

Glu Glu Val Ile Ser Asn Trp Asp Arg Tyr Gly Asn Thr Val Ala Ser
610 615 620

Leu Gln Ala Trp Leu Glu Asp Ala Glu Lys Met Leu Asn Gln Ser Glu
625 630 635 640

Asn Ala Lys Lys Asp Phe Phe Arg Asn Leu Pro His Trp Ile Gln Gln
645 650 655

His Thr Ala Met Asn Asp Ala Gly Asn Phe Leu Ile Glu Thr Cys Asp
660 665 670

```
Glu Met Val Ser Arg Asp Leu Lys Gln Gln Leu Leu Leu Leu Asn Gly
        675                 680                 685

Arg Trp Arg Glu Leu Phe Met Glu Val Lys Gln Tyr Ala Gln Ala Asp
    690                 695                 700

Glu Met Asp Arg Met Lys Lys Glu Tyr Thr Asp Cys Val Val Thr Leu
705                 710                 715                 720

Ser Ala Phe Ala Thr Glu Ala His Lys Lys Leu Ser Glu Pro Leu Glu
                725                 730                 735

Val Ser Phe Met Asn Val Lys Leu Leu Ile Gln Asp Leu Glu Asp Ile
            740                 745                 750

Glu Gln Arg Val Pro Val Met Asp Ala Gln Tyr Lys Ile Ile Thr Lys
        755                 760                 765

Thr Ala His Leu Ile Thr Lys Glu Ser Pro Gln Glu Glu Gly Lys Glu
    770                 775                 780

Met Phe Ala Thr Met Ser Lys Leu Lys Glu Gln Leu Thr Lys Val Lys
785                 790                 795                 800

Glu Cys Tyr Ser Pro Leu Leu Tyr Glu Ser Gln Gln Leu Leu Ile Pro
                805                 810                 815

Leu Glu Glu Leu Glu Lys Gln Met Thr Ser Phe Tyr Asp Ser Leu Gly
            820                 825                 830

Lys Ile Asn Glu Ile Ile Thr Val Leu Glu Arg Glu Ala Gln Ser Ser
        835                 840                 845

Ala Leu Phe Lys Gln Lys His Gln Glu Leu Leu Ala Cys Gln Glu Asn
    850                 855                 860

Cys Lys Lys Thr Leu Thr Leu Ile Glu Lys Gly Ser Gln Ser Val Gln
865                 870                 875                 880

Lys Phe Val Thr Leu Ser Asn Val Leu Lys His Phe Asp Gln Thr Arg
                885                 890                 895

Leu Gln Arg Gln Ile Ala Asp Ile His Val Ala Phe Gln Ser Met Val
            900                 905                 910

Lys Lys Thr Gly Asp Trp Lys Lys His Val Glu Thr Asn Ser Arg Leu
        915                 920                 925

Met Lys Lys Phe Glu Glu Ser Arg Ala Glu Leu Glu Lys Val Leu Arg
    930                 935                 940
```

```
Ile Ala Gln Glu Gly Leu Glu Glu Lys Gly Asp Pro Glu Glu Leu Leu
945                 950                 955                 960


Arg Arg His Thr Glu Phe Phe Ser Gln Leu Asp Gln Arg Val Leu Asn
                965                 970                 975


Ala Phe Leu Lys Ala Cys Asp Glu Leu Thr Asp Ile Leu Pro Glu Gln
            980                 985                 990


Glu Gln Gln Gly Leu Gln Glu Ala  Val Arg Lys Leu His  Lys Gln Trp
        995                 1000                1005


Lys Asp  Leu Gln Gly Glu Ala  Pro Tyr His Leu Leu  His Leu Lys
    1010                1015                1020


Ile Asp  Val Glu Lys Asn Arg  Phe Leu Ala Ser Val  Glu Glu Cys
    1025                1030                1035


Arg Thr  Glu Leu Asp Arg Glu  Thr Lys Leu Met Pro  Gln Glu Gly
    1040                1045                1050


Ser Glu  Lys Ile Ile Lys Glu  His Arg Val Phe Phe  Ser Asp Lys
    1055                1060                1065


Gly Pro  His His Leu Cys Glu  Lys Arg Leu Gln Leu  Ile Glu Glu
    1070                1075                1080


Leu Cys  Val Lys Leu Pro Val  Arg Asp Pro Val Arg  Asp Thr Pro
    1085                1090                1095


Gly Thr  Cys His Val Thr Leu  Lys Glu Leu Arg Ala  Ala Ile Asp
    1100                1105                1110


Ser Thr  Tyr Arg Lys Leu Met  Glu Asp Pro Asp Lys  Trp Lys Asp
    1115                1120                1125


Tyr Thr  Ser Arg Phe Ser Glu  Phe Ser Ser Trp Ile  Ser Thr Asn
    1130                1135                1140


Glu Thr  Gln Leu Lys Gly Ile  Lys Gly Glu Ala Ile  Asp Thr Ala
    1145                1150                1155


Asn His  Gly Glu Val Lys Arg  Ala Val Glu Glu Ile  Arg Asn Gly
    1160                1165                1170


Val Thr  Lys Arg Gly Glu Thr  Leu Ser Trp Leu Lys  Ser Arg Leu
    1175                1180                1185


Lys Val  Leu Thr Glu Val Ser  Ser Glu Asn Glu Ala  Gln Lys Gln
    1190                1195                1200
```

```
Gly Asp  Glu Leu Ala Lys Leu  Ser Ser Ser Phe Lys  Ala Leu Val
    1205                 1210                 1215


Thr Leu  Leu Ser Glu Val Glu  Lys Met Leu Ser Asn  Phe Gly Asp
    1220                 1225                 1230


Cys Val  Gln Tyr Lys Glu Ile  Val Lys Asn Ser Leu  Glu Glu Leu
    1235                 1240                 1245


Ile Ser  Gly Ser Lys Glu Val  Gln Glu Gln Ala Glu  Lys Ile Leu
    1250                 1255                 1260


Asp Thr  Glu Asn Leu Phe Glu  Ala Gln Gln Leu Leu  Leu His His
    1265                 1270                 1275


Gln Gln  Lys Thr Lys Arg Ile  Ser Ala Lys Lys Arg  Asp Val Gln
    1280                 1285                 1290


Gln Gln  Ile Ala Gln Ala Gln  Gln Gly Glu Gly Gly  Leu Pro Asp
    1295                 1300                 1305


Arg Gly  His Glu Glu Leu Arg  Lys Leu Glu Ser Thr  Leu Asp Gly
    1310                 1315                 1320


Leu Glu  Arg Ser Arg Glu Arg  Gln Glu Arg Arg Ile  Gln Val Thr
    1325                 1330                 1335


Leu Arg  Lys Trp Glu Arg Phe  Glu Thr Asn Lys Glu  Thr Val Val
    1340                 1345                 1350


Arg Tyr  Leu Phe Gln Thr Gly  Ser Ser His Glu Arg  Phe Leu Ser
    1355                 1360                 1365


Phe Ser  Ser Leu Glu Ser Leu  Ser Ser Glu Leu Glu  Gln Thr Lys
    1370                 1375                 1380


Glu Phe  Ser Lys Arg Thr Glu  Ser Ile Ala Val Gln  Ala Glu Asn
    1385                 1390                 1395


Leu Val  Lys Glu Ala Ser Glu  Ile Pro Leu Gly Pro  Gln Asn Lys
    1400                 1405                 1410


Gln Leu  Leu Gln Gln Gln Ala  Lys Ser Ile Lys Glu  Gln Val Lys
    1415                 1420                 1425


Lys Leu  Glu Asp Thr Leu Glu  Glu Asp Ile Lys Thr  Met Glu Met
    1430                 1435                 1440


Val Lys  Thr Lys Trp Asp His  Phe Gly Ser Asn Phe  Glu Thr Leu
    1445                 1450                 1455


Ser Val  Trp Ile Thr Glu Lys  Glu Lys Glu Leu Asn  Ala Leu Glu
```

```
    1460                1465                1470

Thr Ser  Ser Ser Ala Met Asp  Met Gln Ile Ser Gln  Ile Lys Val
    1475                1480                1485

Thr Ile  Gln Glu Ile Glu Ser  Lys Leu Ser Ser Ile  Val Gly Leu
    1490                1495                1500

Glu Glu  Glu Ala Gln Ser Phe  Ala Gln Phe Val Thr  Thr Gly Glu
    1505                1510                1515

Ser Ala  Arg Ile Lys Ala Lys  Leu Thr Gln Ile Arg  Arg Tyr Gly
    1520                1525                1530

Glu Glu  Leu Arg Glu His Ala  Gln Cys Leu Glu Gly  Thr Ile Leu
    1535                1540                1545

Gly His  Leu Ser Gln Gln Gln  Lys Phe Glu Glu Asn  Leu Arg Lys
    1550                1555                1560

Ile Gln  Gln Ser Val Ser Glu  Phe Glu Asp Lys Leu  Ala Val Pro
    1565                1570                1575

Ile Lys  Ile Cys Ser Ser Ala  Thr Glu Thr Tyr Lys  Val Leu Gln
    1580                1585                1590

Glu His  Met Asp Leu Cys Gln  Ala Leu Glu Ser Leu  Ser Ser Ala
    1595                1600                1605

Ile Thr  Ala Phe Ser Ala Ser  Ala Arg Lys Val Val  Asn Arg Asp
    1610                1615                1620

Ser Cys  Val Gln Glu Ala Ala  Ala Leu Gln Gln Gln  Tyr Glu Asp
    1625                1630                1635

Ile Leu  Arg Arg Ala Lys Glu  Arg Gln Thr Ala Leu  Glu Asn Leu
    1640                1645                1650

Leu Ala  His Trp Gln Arg Leu  Glu Lys Glu Leu Ser  Ser Phe Leu
    1655                1660                1665

Thr Trp  Leu Glu Arg Gly Glu  Ala Lys Ala Ser Ser  Pro Glu Met
    1670                1675                1680

Asp Ile  Ser Ala Asp Arg Val  Lys Val Glu Gly Glu  Leu Gln Leu
    1685                1690                1695

Ile Gln  Ala Leu Gln Asn Glu  Val Val Ser Gln Ala  Ser Phe Tyr
    1700                1705                1710

Ser Lys  Leu Leu Gln Leu Lys  Glu Ser Leu Phe Ser  Val Ala Ser
    1715                1720                1725
```

```
Lys Asp  Asp Val Lys Met Met  Lys Leu His Leu Glu  Gln Leu Asp
    1730                1735                1740

Glu Arg  Trp Arg Asp Leu Pro  Gln Ile Ile Asn Lys  Arg Ile Asn
    1745                1750                1755

Phe Leu  Gln Ser Val Val Ala  Glu His Gln Gln Phe  Asp Glu Leu
    1760                1765                1770

Leu Leu  Ser Phe Ser Val Trp  Ile Lys Leu Phe Leu  Ser Glu Leu
    1775                1780                1785

Gln Thr  Thr Ser Glu Ile Ser  Ile Met Asp His Gln  Val Ala Leu
    1790                1795                1800

Thr Arg  His Lys Asp His Ala  Ala Glu Val Glu Ser  Lys Lys Gly
    1805                1810                1815

Glu Leu  Gln Ser Leu Gln Gly  His Leu Ala Lys Leu  Gly Ser Leu
    1820                1825                1830

Gly Arg  Ala Glu Asp Leu His  Leu Leu Gln Gly Lys  Ala Glu Asp
    1835                1840                1845

Cys Phe  Gln Leu Phe Glu Glu  Ala Ser Gln Val Val  Glu Arg Arg
    1850                1855                1860

Gln Leu  Ala Leu Ser His Leu  Ala Glu Phe Leu Gln  Ser His Ala
    1865                1870                1875

Ser Leu  Ser Gly Ile Leu Arg  Gln Leu Arg Gln Thr  Val Glu Ala
    1880                1885                1890

Thr Asn  Ser Met Asn Lys Asn  Glu Ser Asp Leu Ile  Glu Lys Asp
    1895                1900                1905

Leu Asn  Asp Ala Leu Gln Asn  Ala Lys Ala Leu Glu  Ser Ala Ala
    1910                1915                1920

Val Ser  Leu Asp Gly Ile Leu  Ser Lys Ala Gln Tyr  His Leu Lys
    1925                1930                1935

Ile Gly  Ser Ser Glu Gln Arg  Thr Ser Cys Arg Ala  Thr Ala Asp
    1940                1945                1950

Gln Leu  Cys Gly Glu Val Glu  Arg Ile Gln Asn Leu  Leu Gly Thr
    1955                1960                1965

Lys Gln  Ser Glu Ala Asp Ala  Leu Ala Val Leu Lys  Lys Ala Phe
    1970                1975                1980
```

```
Gln Asp  Gln Lys Glu Glu Leu  Leu Lys Ser Ile Glu  Asp Ile Glu
    1985                 1990                 1995

Glu Arg  Thr Asp Lys Glu Arg  Leu Lys Glu Pro Thr  Arg Gln Ala
    2000                 2005                 2010

Leu Gln  Gln Arg Leu Arg Val  Phe Asn Gln Leu Glu  Asp Glu Leu
    2015                 2020                 2025

Asn Ser  His Glu His Glu Leu  Cys Trp Leu Lys Asp  Lys Ala Lys
    2030                 2035                 2040

Gln Ile  Ala Gln Lys Asp Val  Ala Phe Ala Pro Glu  Val Asp Arg
    2045                 2050                 2055

Glu Ile  Asn Arg Leu Glu Val  Thr Trp Asp Asp Thr  Lys Arg Leu
    2060                 2065                 2070

Ile His  Glu Asn Gln Gly Gln  Cys Cys Gly Leu Ile  Asp Leu Met
    2075                 2080                 2085

Arg Glu  Tyr Gln Asn Leu Lys  Ser Ala Val Ser Lys  Val Leu Glu
    2090                 2095                 2100

Asn Ala  Ser Ser Val Ile Val  Thr Arg Thr Thr Ile  Lys Asp Gln
    2105                 2110                 2115

Glu Asp  Leu Lys Trp Ala Phe  Ser Lys His Glu Thr  Ala Lys Asn
    2120                 2125                 2130

Lys Met  Asn Tyr Lys Gln Lys  Asp Leu Asp Asn Phe  Thr Ser Lys
    2135                 2140                 2145

Gly Lys  His Leu Leu Ser Glu  Leu Lys Lys Ile His  Ser Ser Asp
    2150                 2155                 2160

Phe Ser  Leu Val Lys Thr Asp  Met Glu Ser Thr Val  Asp Lys Trp
    2165                 2170                 2175

Leu Asp  Val Ser Glu Lys Leu  Glu Glu Asn Met Asp  Arg Leu Arg
    2180                 2185                 2190

Val Ser  Leu Ser Ile Trp Asp  Asp Val Leu Ser Thr  Arg Asp Glu
    2195                 2200                 2205

Ile Glu  Gly Trp Ser Asn Asn  Cys Val Pro Gln Met  Ala Glu Asn
    2210                 2215                 2220

Ile Ser  Asn Leu Asp Asn His  Leu Arg Ala Glu Glu  Leu Leu Lys
    2225                 2230                 2235
```

```
Glu Phe  Glu Ser Glu Val Lys  Asn Lys Ala Leu Arg  Leu Glu Glu
    2240                 2245                 2250


Leu His  Ser Lys Val Asn Asp  Leu Lys Glu Leu Thr  Lys Asn Leu
    2255                 2260                 2265


Glu Thr  Pro Pro Asp Leu Gln  Phe Ile Glu Ala Asp  Leu Met Gln
    2270                 2275                 2280


Lys Leu  Glu His Ala Lys Glu  Ile Thr Glu Val Ala  Lys Gly Thr
    2285                 2290                 2295


Leu Lys  Asp Phe Thr Ala Gln  Ser Thr Gln Val Glu  Lys Phe Ile
    2300                 2305                 2310


Asn Asp  Ile Thr Thr Trp Phe  Thr Lys Val Glu Glu  Ser Leu Met
    2315                 2320                 2325


Asn Cys  Ala Gln Asn Glu Thr  Cys Glu Ala Leu Lys  Lys Val Lys
    2330                 2335                 2340


Asp Ile  Gln Lys Glu Leu Gln  Ser Gln Gln Ser Asn  Ile Ser Ser
    2345                 2350                 2355


Thr Gln  Glu Asn Leu Asn Ser  Leu Cys Arg Lys Tyr  His Ser Ala
    2360                 2365                 2370


Glu Leu  Glu Ser Leu Gly Arg  Ala Met Thr Gly Leu  Ile Lys Lys
    2375                 2380                 2385


His Glu  Ala Val Ser Gln Leu  Cys Ser Lys Thr Gln  Ala Ser Leu
    2390                 2395                 2400


Gln Glu  Ser Leu Glu Lys His  Phe Ser Glu Ser Met  Gln Glu Phe
    2405                 2410                 2415


Gln Glu  Trp Phe Leu Gly Ala  Lys Ala Ala Ala Lys  Glu Ser Ser
    2420                 2425                 2430


Asp Arg  Thr Gly Asp Ser Lys  Val Leu Glu Ala Lys  Leu His Asp
    2435                 2440                 2445


Leu Gln  Asn Ile Leu Asp Ser  Val Ser Asp Gly Gln  Ser Lys Leu
    2450                 2455                 2460


Asp Ala  Val Thr Gln Glu Gly  Gln Thr Leu Tyr Ala  His Leu Ser
    2465                 2470                 2475


Lys Gln  Ile Val Ser Ser Ile  Gln Glu Gln Ile Thr  Lys Ala Asn
    2480                 2485                 2490


Glu Glu  Phe Gln Ala Phe Leu  Lys Gln Cys Leu Lys  Asp Lys Gln
```

2495 2500 2505

Ala Leu Gln Asp Cys Ala Ser Glu Leu Gly Ser Phe Glu Asp Gln
2510 2515 2520

His Arg Lys Leu Asn Leu Trp Ile His Glu Met Glu Glu Arg Phe
2525 2530 2535

Asn Thr Glu Asn Leu Gly Glu Ser Lys Gln His Ile Pro Glu Lys
2540 2545 2550

Lys Asn Glu Val His Lys Val Glu Met Phe Leu Gly Glu Leu Leu
2555 2560 2565

Ala Ala Arg Glu Ser Leu Asp Lys Leu Ser Gln Arg Gly Gln Leu
2570 2575 2580

Leu Ser Glu Glu Gly His Gly Ala Gly Gln Glu Gly Arg Leu Cys
2585 2590 2595

Ser Gln Leu Leu Thr Ser His Gln Asn Leu Leu Arg Met Thr Lys
2600 2605 2610

Glu Lys Leu Arg Ser Cys Gln Val Ala Leu Gln Glu His Glu Ala
2615 2620 2625

Leu Glu Glu Ala Leu Gln Ser Met Trp Phe Trp Val Lys Ala Ile
2630 2635 2640

Gln Asp Arg Leu Ala Cys Ala Glu Ser Thr Leu Gly Ser Lys Asp
2645 2650 2655

Thr Leu Glu Lys Arg Leu Ser Gln Ile Gln Asp Ile Leu Leu Met
2660 2665 2670

Lys Gly Glu Gly Glu Val Lys Leu Asn Met Ala Ile Gly Lys Gly
2675 2680 2685

Glu Gln Ala Leu Arg Ser Ser Asn Lys Glu Gly Gln Arg Val Ile
2690 2695 2700

Gln Thr Gln Leu Glu Thr Leu Lys Glu Val Trp Ala Asp Ile Met
2705 2710 2715

Ser Ser Ser Val His Ala Gln Ser Thr Leu Glu Ser Val Ile Ser
2720 2725 2730

Gln Trp Asn Asp Tyr Val Glu Arg Lys Asn Gln Leu Glu Gln Trp
2735 2740 2745

Met Glu Ser Val Asp Gln Lys Ile Glu His Pro Leu Gln Pro Gln
2750 2755 2760

```
Pro Gly  Leu Lys Glu Lys Phe  Val Leu Leu Asp His  Leu Gln Ser
    2765                 2770                 2775

Ile Leu  Ser Glu Ala Glu Asp  His Thr Arg Ala Leu  His Arg Leu
    2780                 2785                 2790

Ile Ala  Lys Ser Arg Glu Leu  Tyr Glu Lys Thr Glu  Asp Glu Ser
    2795                 2800                 2805

Phe Lys  Asp Thr Ala Gln Glu  Glu Leu Lys Thr Gln  Phe Asn Asp
    2810                 2815                 2820

Ile Met  Thr Val Ala Lys Glu  Lys Met Arg Lys Val  Glu Glu Ile
    2825                 2830                 2835

Val Lys  Asp His Leu Met Tyr  Leu Asp Ala Val His  Glu Phe Thr
    2840                 2845                 2850

Asp Trp  Leu His Ser Ala Lys  Glu Glu Leu His Arg  Trp Ser Asp
    2855                 2860                 2865

Met Ser  Gly Asp Ser Ser Ala  Thr Gln Lys Lys Leu  Ser Lys Ile
    2870                 2875                 2880

Lys Glu  Leu Ile Asp Ser Arg  Glu Ile Gly Ala Ser  Arg Leu Ser
    2885                 2890                 2895

Arg Val  Glu Ser Leu Ala Pro  Glu Val Lys Gln Asn  Thr Thr Ala
    2900                 2905                 2910

Ser Gly  Cys Glu Leu Met His  Thr Glu Met Gln Ala  Leu Arg Ala
    2915                 2920                 2925

Asp Trp  Lys Gln Trp Glu Asp  Ser Val Phe Gln Thr  Gln Ser Cys
    2930                 2935                 2940

Leu Glu  Asn Leu Val Ser Gln  Met Ala Leu Ser Glu  Gln Glu Phe
    2945                 2950                 2955

Ser Gly  Gln Val Ala Gln Leu  Glu Gln Ala Leu Glu  Gln Phe Ser
    2960                 2965                 2970

Ala Leu  Leu Lys Thr Trp Ala  Gln Gln Leu Thr Leu  Leu Glu Gly
    2975                 2980                 2985

Lys Asn  Thr Asp Glu Glu Ile  Val Glu Cys Trp His  Lys Gly Gln
    2990                 2995                 3000

Glu Ile  Leu Asp Ala Leu Gln  Lys Ala Glu Pro Arg  Thr Glu Asp
    3005                 3010                 3015
```

```
Leu Lys  Ser Gln Leu Asn Glu  Leu Cys Arg Phe Ser  Arg Asp Leu
    3020                3025                3030

Ser Thr  Tyr Ser Gly Lys Val  Ser Gly Leu Ile Lys  Glu Tyr Asn
    3035                3040                3045

Cys Leu  Cys Leu Gln Ala Ser  Lys Gly Cys Gln Asn  Lys Glu Gln
    3050                3055                3060

Ile Leu  Gln Gln Arg Phe Arg  Lys Ala Phe Arg Asp  Phe Gln Gln
    3065                3070                3075

Trp Leu  Val Asn Ala Lys Ile  Thr Thr Ala Lys Cys  Phe Asp Ile
    3080                3085                3090

Pro Gln  Asn Ile Ser Glu Val  Ser Thr Ser Leu Gln  Lys Ile Gln
    3095                3100                3105

Glu Phe  Leu Ser Glu Ser Glu  Asn Gly Gln His Lys  Leu Asn Met
    3110                3115                3120

Met Leu  Ser Lys Gly Glu Leu  Leu Ser Thr Leu Leu  Thr Lys Glu
    3125                3130                3135

Lys Ala  Lys Gly Ile Gln Ala  Lys Val Thr Ala Ala  Lys Glu Asp
    3140                3145                3150

Trp Lys  Asn Phe His Ser Asn  Leu His Gln Lys Glu  Ser Ala Leu
    3155                3160                3165

Glu Asn  Leu Lys Ile Gln Met  Lys Asp Phe Glu Val  Ser Ala Glu
    3170                3175                3180

Pro Ile  Gln Asp Trp Leu Ser  Lys Thr Glu Lys Met  Val His Glu
    3185                3190                3195

Ser Ser  Asn Arg Leu Tyr Asp  Leu Pro Ala Lys Arg  Arg Glu Gln
    3200                3205                3210

Gln Lys  Leu Gln Ser Val Leu  Glu Glu Ile His Cys  Tyr Glu Pro
    3215                3220                3225

Gln Leu  Asn Arg Leu Lys Glu  Lys Ala Gln Gln Leu  Trp Glu Gly
    3230                3235                3240

Gln Ala  Ala Ser Lys Ser Phe  Arg His Arg Val Ser  Gln Leu Ser
    3245                3250                3255

Ser Gln  Tyr Leu Ala Leu Ser  Asn Leu Thr Lys Glu  Lys Val Ser
    3260                3265                3270
```

```
Arg Leu  Asp Arg Ile Val Ala  Glu His Asn Gln Phe  Ser Leu Gly
    3275                 3280                 3285


Ile Lys  Glu Leu Gln Asp Trp  Met Thr Asp Ala Ile  His Met Leu
    3290                 3295                 3300


Asp Ser  Tyr Cys His Pro Thr  Ser Asp Lys Ser Val  Leu Asp Ser
    3305                 3310                 3315


Arg Thr  Leu Lys Leu Glu Ala  Leu Leu Ser Val Lys  Gln Glu Lys
    3320                 3325                 3330


Glu Ile  Gln Met Lys Met Ile  Val Thr Arg Gly Glu  Ser Val Leu
    3335                 3340                 3345


Gln Asn  Thr Ser Pro Glu Gly  Ile Pro Thr Ile Gln  Gln Gln Leu
    3350                 3355                 3360


Gln Ser  Val Lys Asp Met Trp  Ala Ser Leu Leu Ser  Ala Gly Ile
    3365                 3370                 3375


Arg Cys  Lys Ser Gln Leu Glu  Gly Ala Leu Ser Lys  Trp Thr Ser
    3380                 3385                 3390


Tyr Gln  Asp Gly Val Arg Gln  Phe Ser Gly Trp Met  Asp Ser Met
    3395                 3400                 3405


Glu Ala  Asn Leu Asn Glu Ser  Glu Arg Gln His Ala  Glu Leu Arg
    3410                 3415                 3420


Asp Lys  Thr Thr Met Leu Gly  Lys Ala Lys Leu Leu  Asn Glu Glu
    3425                 3430                 3435


Val Leu  Ser Tyr Ser Ser Leu  Leu Glu Thr Ile Glu  Val Lys Gly
    3440                 3445                 3450


Ala Gly  Met Thr Glu His Tyr  Val Thr Gln Leu Glu  Leu Gln Asp
    3455                 3460                 3465


Leu Gln  Glu Arg Tyr Arg Ala  Ile Gln Glu Arg Ala  Lys Glu Ala
    3470                 3475                 3480


Val Thr  Lys Ser Glu Lys Leu  Val Arg Leu His Gln  Glu Tyr Gln
    3485                 3490                 3495


Arg Asp  Leu Lys Ala Phe Glu  Val Trp Leu Gly Gln  Glu Gln Glu
    3500                 3505                 3510


Lys Leu  Asp Gln Tyr Ser Val  Leu Glu Gly Asp Ala  His Thr His
    3515                 3520                 3525


Glu Thr  Thr Leu Arg Asp Leu  Gln Glu Leu Gln Val  His Cys Ala
```

```
     3530                3535                3540

Glu Gly  Gln Ala Leu Leu Asn  Ser Val Leu His Thr  Arg Glu Asp
     3545                3550                3555

Val Ile  Pro Ser Gly Ile Pro  Gln Ala Glu Asp Arg  Ala Leu Glu
     3560                3565                3570

Ser Leu  Arg Gln Asp Trp Gln  Ala Tyr Gln His Arg  Leu Ser Glu
     3575                3580                3585

Thr Arg  Thr Gln Phe Asn Asn  Val Val Asn Lys Leu  Arg Leu Met
     3590                3595                3600

Glu Gln  Lys Phe Gln Gln Val  Asp Glu Trp Leu Lys  Thr Ala Glu
     3605                3610                3615

Glu Lys  Val Ser Pro Arg Thr  Arg Arg Gln Ser Asn  Arg Ala Thr
     3620                3625                3630

Lys Glu  Ile Gln Leu His Gln  Met Lys Lys Trp His  Glu Glu Val
     3635                3640                3645

Thr Ala  Tyr Arg Asp Glu Val  Glu Glu Val Gly Ala  Arg Ala Gln
     3650                3655                3660

Glu Ile  Leu Asp Glu Ser His  Val Asn Ser Arg Met  Gly Cys Gln
     3665                3670                3675

Ala Thr  Gln Leu Thr Ser Arg  Tyr Gln Ala Leu Leu  Leu Gln Val
     3680                3685                3690

Leu Glu  Gln Ile Lys Phe Leu  Glu Glu Glu Ile Gln  Ser Leu Glu
     3695                3700                3705

Glu Ser  Glu Ser Ser Leu Ser  Ser Tyr Ser Asp Trp  Tyr Gly Ser
     3710                3715                3720

Thr His  Lys Asn Phe Lys Asn  Val Ala Thr Lys Ile  Asp Lys Val
     3725                3730                3735

Asp Thr  Val Met Met Gly Lys  Lys Leu Lys Thr Leu  Glu Val Leu
     3740                3745                3750

Leu Lys  Asp Met Glu Lys Gly  His Ser Leu Leu Lys  Ser Ala Arg
     3755                3760                3765

Glu Lys  Gly Glu Arg Ala Val  Lys Tyr Leu Glu Glu  Gly Glu Ala
     3770                3775                3780

Glu Arg  Leu Arg Lys Glu Ile  His Asp His Met Glu  Gln Leu Lys
     3785                3790                3795
```

```
Glu Leu  Thr Ser Thr Val Arg  Lys Glu His Met Thr  Leu Glu Lys
    3800                 3805                 3810

Gly Leu  His Leu Ala Lys Glu  Phe Ser Asp Lys Cys  Lys Ala Leu
    3815                 3820                 3825

Thr Gln  Trp Ile Ala Glu Tyr  Gln Glu Ile Leu His  Val Pro Glu
    3830                 3835                 3840

Glu Pro  Lys Met Glu Leu Tyr  Glu Lys Lys Ala Gln  Leu Ser Lys
    3845                 3850                 3855

Tyr Lys  Ser Leu Gln Gln Thr  Val Leu Ser His Glu  Pro Ser Val
    3860                 3865                 3870

Lys Ser  Val Arg Glu Lys Gly  Glu Ala Leu Leu Glu  Leu Val Gln
    3875                 3880                 3885

Asp Val  Thr Leu Lys Asp Lys  Ile Asp Gln Leu Gln  Ser Asp Tyr
    3890                 3895                 3900

Gln Asp  Leu Cys Ser Ile Gly  Lys Glu His Val Phe  Ser Leu Glu
    3905                 3910                 3915

Ala Lys  Val Lys Asp His Glu  Asp Tyr Asn Ser Glu  Leu Gln Glu
    3920                 3925                 3930

Val Glu  Lys Trp Leu Leu Gln  Met Ser Gly Arg Leu  Val Ala Pro
    3935                 3940                 3945

Asp Leu  Leu Glu Thr Ser Ser  Leu Glu Thr Ile Thr  Gln Gln Leu
    3950                 3955                 3960

Ala His  His Lys Ala Met Met  Glu Glu Ile Ala Gly  Phe Glu Asp
    3965                 3970                 3975

Arg Leu  Asn Asn Leu Gln Met  Lys Gly Asp Thr Leu  Ile Gly Gln
    3980                 3985                 3990

Cys Ala  Asp His Leu Gln Ala  Lys Leu Lys Gln Asn  Val His Ala
    3995                 4000                 4005

His Leu  Gln Gly Thr Lys Asp  Ser Tyr Ser Ala Ile  Cys Ser Thr
    4010                 4015                 4020

Ala Gln  Arg Met Tyr Gln Ser  Leu Glu His Glu Leu  Gln Lys His
    4025                 4030                 4035

Val Ser  Arg Gln Asp Thr Leu  Gln Gln Cys Gln Ala  Trp Leu Ser
    4040                 4045                 4050
```

```
Ala Val  Gln Pro Asp Leu Glu  Pro Ser Pro Gln Pro  Pro Leu Ser
    4055                 4060                 4065

Arg Ala  Glu Ala Ile Lys Gln  Val Lys His Phe Arg  Ala Leu Gln
    4070                 4075                 4080

Glu Gln  Ala Arg Thr Tyr Leu  Asp Leu Leu Cys Ser  Met Cys Asp
    4085                 4090                 4095

Leu Ser  Asn Ala Ser Val Lys  Thr Thr Ala Lys Asp  Ile Gln Gln
    4100                 4105                 4110

Thr Glu  Gln Thr Ile Glu Gln  Lys Leu Val Gln Ala  Gln Asn Leu
    4115                 4120                 4125

Thr Gln  Gly Trp Glu Glu Ile  Lys His Leu Lys Ser  Glu Leu Trp
    4130                 4135                 4140

Ile Tyr  Leu Gln Asp Ala Asp  Gln Gln Leu Gln Asn  Met Lys Arg
    4145                 4150                 4155

Arg His  Ser Glu Leu Glu Leu  Asn Ile Ala Gln Asn  Met Val Ser
    4160                 4165                 4170

Gln Val  Lys Asp Phe Val Lys  Lys Leu Gln Ser Lys  Gln Ala Ser
    4175                 4180                 4185

Val Asn  Thr Ile Ile Glu Lys  Val Asn Lys Leu Thr  Lys Lys Glu
    4190                 4195                 4200

Glu Ser  Pro Glu His Lys Glu  Ile Asn His Leu Asn  Asp Gln Trp
    4205                 4210                 4215

Leu Asp  Leu Cys Arg Gln Ser  Asn Asn Leu Cys Leu  Gln Arg Glu
    4220                 4225                 4230

Glu Asp  Leu Gln Arg Thr Arg  Asp Tyr His Asp Cys  Met Asn Val
    4235                 4240                 4245

Val Glu  Val Phe Leu Glu Lys  Phe Thr Thr Glu Trp  Asp Asn Leu
    4250                 4255                 4260

Ala Arg  Ser Asp Ala Glu Ser  Thr Ala Val His Leu  Glu Ala Leu
    4265                 4270                 4275

Lys Lys  Leu Ala Leu Ala Leu  Gln Glu Arg Lys Tyr  Ala Ile Glu
    4280                 4285                 4290

Asp Leu  Lys Asp Gln Lys Gln  Lys Met Ile Glu His  Leu Asn Leu
    4295                 4300                 4305
```

```
Asp Asp  Lys Glu Leu Val Lys  Glu Gln Thr Ser His  Leu Glu Gln
    4310                 4315                 4320


Arg Trp  Phe Gln Leu Glu Asp  Leu Ile Lys Arg Lys  Ile Gln Val
    4325                 4330                 4335


Ser Val  Thr Asn Leu Glu Glu  Leu Asn Val Val Gln  Ser Arg Phe
    4340                 4345                 4350


Gln Glu  Leu Met Glu Trp Ala  Glu Glu Gln Gln Pro  Asn Ile Ala
    4355                 4360                 4365


Glu Ala  Leu Lys Gln Ser Pro  Pro Pro Asp Met Ala  Gln Asn Leu
    4370                 4375                 4380


Leu Met  Asp His Leu Ala Ile  Cys Ser Glu Leu Glu  Ala Lys Gln
    4385                 4390                 4395


Met Leu  Leu Lys Ser Leu Ile  Lys Asp Ala Asp Arg  Val Met Ala
    4400                 4405                 4410


Asp Leu  Gly Leu Asn Glu Arg  Gln Val Ile Gln Lys  Ala Leu Ser
    4415                 4420                 4425


Asp Ala  Gln Ser His Val Asn  Cys Leu Ser Asp Leu  Val Gly Gln
    4430                 4435                 4440


Arg Arg  Lys Tyr Leu Asn Lys  Ala Leu Ser Glu Lys  Thr Gln Phe
    4445                 4450                 4455


Leu Met  Ala Val Phe Gln Ala  Thr Ser Gln Ile Gln  Gln His Glu
    4460                 4465                 4470


Arg Lys  Ile Met Phe Arg Glu  His Ile Cys Leu Leu  Pro Asp Asp
    4475                 4480                 4485


Val Ser  Lys Gln Val Lys Thr  Cys Lys Ser Ala Gln  Ala Ser Leu
    4490                 4495                 4500


Lys Thr  Tyr Gln Asn Glu Val  Thr Gly Leu Trp Ala  Gln Gly Arg
    4505                 4510                 4515


Glu Leu  Met Lys Glu Val Thr  Glu Gln Glu Lys Ser  Glu Val Leu
    4520                 4525                 4530


Gly Lys  Leu Gln Glu Leu Gln  Ser Val Tyr Asp Ser  Val Leu Gln
    4535                 4540                 4545


Lys Cys  Ser His Arg Leu Gln  Glu Leu Glu Lys Asn  Leu Val Ser
    4550                 4555                 4560


Arg Lys  His Phe Lys Glu Asp  Phe Asp Lys Ala Cys  His Trp Leu
```

```
     4565                4570                4575

Lys Gln  Ala Asp Ile Val Thr  Phe Pro Glu Ile Asn  Leu Met Asn
     4580                4585                4590

Glu Ser  Ser Glu Leu His Thr  Gln Leu Ala Lys Tyr  Gln Asn Ile
     4595                4600                4605

Leu Glu  Gln Ser Pro Glu Tyr  Glu Asn Leu Leu Leu  Thr Leu Gln
     4610                4615                4620

Arg Thr  Gly Gln Thr Ile Leu  Pro Ser Leu Asn Glu  Val Asp His
     4625                4630                4635

Ser Tyr  Leu Ser Glu Lys Leu  Asn Ala Leu Pro Arg  Gln Phe Asn
     4640                4645                4650

Val Ile  Val Ala Leu Ala Lys  Asp Lys Phe Tyr Lys  Val Gln Glu
     4655                4660                4665

Ala Ile  Leu Ala Arg Lys Glu  Tyr Ala Ser Leu Ile  Glu Leu Thr
     4670                4675                4680

Thr Gln  Ser Leu Ser Glu Leu  Glu Ala Gln Phe Leu  Arg Met Ser
     4685                4690                4695

Lys Val  Pro Thr Asp Leu Ala  Val Glu Glu Ala Leu  Ser Leu Gln
     4700                4705                4710

Asp Gly  Cys Arg Ala Ile Leu  Asp Glu Val Ala Gly  Leu Gly Glu
     4715                4720                4725

Ala Val  Asp Glu Leu Asn Gln  Lys Lys Glu Gly Phe  Arg Ser Thr
     4730                4735                4740

Gly Gln  Pro Trp Gln Pro Asp  Lys Met Leu His Leu  Val Thr Leu
     4745                4750                4755

Tyr His  Arg Leu Lys Arg Gln  Thr Glu Gln Arg Val  Ser Leu Leu
     4760                4765                4770

Glu Asp  Thr Thr Ser Ala Tyr  Gln Glu His Glu Lys  Met Cys Gln
     4775                4780                4785

Gln Leu  Glu Arg Gln Leu Lys  Ser Val Lys Glu Glu  Gln Ser Lys
     4790                4795                4800

Val Asn  Glu Glu Thr Leu Pro  Ala Glu Glu Lys Leu  Lys Met Tyr
     4805                4810                4815

His Ser  Leu Ala Gly Ser Leu  Gln Asp Ser Gly Ile  Val Leu Lys
     4820                4825                4830
```

```
Arg Val  Thr Ile His Leu Glu  Asp Leu Ala Pro His  Leu Asp Pro
    4835                4840                4845

Leu Ala  Tyr Glu Lys Ala Arg  His Gln Ile Gln Ser  Trp Gln Gly
    4850                4855                4860

Glu Leu  Lys Leu Leu Thr Ser  Ala Ile Gly Glu Thr  Val Thr Glu
    4865                4870                4875

Cys Glu  Ser Arg Met Val Gln  Ser Ile Asp Phe Gln  Thr Glu Met
    4880                4885                4890

Ser Arg  Ser Leu Asp Trp Leu  Arg Arg Val Lys Ala  Glu Leu Ser
    4895                4900                4905

Gly Pro  Val Tyr Leu Asp Leu  Asn Leu Gln Asp Ile  Gln Glu Glu
    4910                4915                4920

Ile Arg  Lys Ile Gln Ile His  Gln Glu Glu Val Gln  Ser Ser Leu
    4925                4930                4935

Arg Ile  Met Asn Ala Leu Ser  His Lys Glu Lys Glu  Lys Phe Thr
    4940                4945                4950

Lys Ala  Lys Glu Leu Ile Ser  Ala Asp Leu Glu His  Ser Leu Ala
    4955                4960                4965

Glu Leu  Ser Glu Leu Asp Gly  Asp Ile Gln Glu Ala  Leu Arg Thr
    4970                4975                4980

Arg Gln  Ala Thr Leu Thr Glu  Ile Tyr Ser Gln Cys  Gln Arg Tyr
    4985                4990                4995

Tyr Gln  Val Phe Gln Ala Ala  Asn Asp Trp Leu Glu  Asp Ala Gln
    5000                5005                5010

Glu Leu  Leu Gln Leu Ala Gly  Asn Gly Leu Asp Val  Glu Ser Ala
    5015                5020                5025

Glu Glu  Asn Leu Lys Ser His  Met Glu Phe Phe Ser  Thr Glu Asp
    5030                5035                5040

Gln Phe  His Ser Asn Leu Glu  Glu Leu His Ser Leu  Val Ala Thr
    5045                5050                5055

Leu Asp  Pro Leu Ile Lys Pro  Thr Gly Lys Glu Asp  Leu Glu Gln
    5060                5065                5070

Lys Val  Ala Ser Leu Glu Leu  Arg Ser Gln Arg Met  Ser Arg Asp
    5075                5080                5085
```

```
Ser Gly  Ala Gln Val Asp Leu  Leu Gln Arg Cys Thr  Ala Gln Trp
    5090                5095                5100

His Asp  Tyr Gln Lys Ala Arg  Glu Glu Val Ile Glu  Leu Met Asn
    5105                5110                5115

Asp Thr  Glu Lys Lys Leu Ser  Glu Phe Ser Leu Leu  Lys Thr Ser
    5120                5125                5130

Ser Ser  His Glu Ala Glu Glu  Lys Leu Ser Glu His  Lys Ala Leu
    5135                5140                5145

Val Ser  Val Val Asn Ser Phe  His Glu Lys Ile Val  Ala Leu Glu
    5150                5155                5160

Glu Lys  Ala Ser Gln Leu Glu  Lys Thr Gly Asn Asp  Ala Ser Lys
    5165                5170                5175

Ala Thr  Leu Ser Arg Ser Met  Thr Thr Val Trp Gln  Arg Trp Thr
    5180                5185                5190

Arg Leu  Arg Ala Val Ala Gln  Asp Gln Glu Lys Ile  Leu Glu Asp
    5195                5200                5205

Ala Val  Asp Glu Trp Thr Gly  Phe Asn Asn Lys Val  Lys Lys Ala
    5210                5215                5220

Thr Glu  Met Ile Asp Gln Leu  Gln Asp Lys Leu Pro  Gly Ser Ser
    5225                5230                5235

Ala Glu  Lys Ala Ser Lys Ala  Glu Leu Leu Thr Leu  Leu Glu Tyr
    5240                5245                5250

His Asp  Thr Phe Val Leu Glu  Leu Glu Gln Gln Gln  Ser Ala Leu
    5255                5260                5265

Gly Met  Leu Arg Gln Gln Thr  Leu Ser Met Leu Gln  Asp Gly Ala
    5270                5275                5280

Ala Pro  Thr Pro Gly Glu Glu  Pro Pro Leu Met Gln  Glu Ile Thr
    5285                5290                5295

Ala Met  Gln Asp Arg Cys Leu  Asn Met Gln Glu Lys  Val Lys Thr
    5300                5305                5310

Asn Gly  Lys Leu Val Lys Gln  Glu Leu Lys Asp Arg  Glu Met Val
    5315                5320                5325

Glu Thr  Gln Ile Asn Ser Val  Lys Cys Trp Val Gln  Glu Thr Lys
    5330                5335                5340
```

```
Glu Tyr  Leu Gly Asn Pro Thr  Ile Glu Ile Asp Ala  Gln Leu Glu
    5345                 5350                 5355


Glu Leu  Gln Ile Leu Leu Thr  Glu Ala Thr Asn His  Arg Gln Asn
    5360                 5365                 5370


Ile Glu  Lys Met Ala Glu Glu  Gln Lys Glu Lys Tyr  Leu Gly Leu
    5375                 5380                 5385


Tyr Thr  Ile Leu Pro Ser Glu  Leu Ser Leu Gln Leu  Ala Glu Val
    5390                 5395                 5400


Ala Leu  Asp Leu Lys Ile Arg  Asp Gln Ile Gln Asp  Lys Ile Lys
    5405                 5410                 5415


Glu Val  Glu Gln Ser Lys Ala  Thr Ser Gln Glu Leu  Ser Arg Gln
    5420                 5425                 5430


Ile Gln  Lys Leu Ala Lys Asp  Leu Thr Thr Ile Leu  Thr Lys Leu
    5435                 5440                 5445


Lys Ala  Lys Thr Asp Asn Val  Val Gln Ala Lys Thr  Asp Gln Lys
    5450                 5455                 5460


Val Leu  Gly Glu Glu Leu Asp  Gly Cys Asn Ser Lys  Leu Met Glu
    5465                 5470                 5475


Leu Asp  Ala Ala Val Gln Lys  Phe Leu Glu Gln Asn  Gly Gln Leu
    5480                 5485                 5490


Gly Lys  Pro Leu Ala Lys Lys  Ile Gly Lys Leu Thr  Glu Leu His
    5495                 5500                 5505


Gln Gln  Thr Ile Arg Gln Ala  Glu Asn Arg Leu Ser  Lys Leu Asn
    5510                 5515                 5520


Gln Ala  Ala Ser His Leu Glu  Glu Tyr Asn Glu Met  Leu Glu Leu
    5525                 5530                 5535


Ile Leu  Lys Trp Ile Glu Lys  Ala Lys Val Leu Ala  His Gly Thr
    5540                 5545                 5550


Ile Ala  Trp Asn Ser Ala Ser  Gln Leu Arg Glu Gln  Tyr Ile Leu
    5555                 5560                 5565


His Gln  Thr Leu Leu Glu Glu  Ser Lys Glu Ile Asp  Ser Glu Leu
    5570                 5575                 5580


Glu Ala  Met Thr Glu Lys Leu  Gln Tyr Leu Thr Ser  Val Tyr Cys
    5585                 5590                 5595


Thr Glu  Lys Met Ser Gln Gln  Val Ala Glu Leu Gly  Arg Glu Thr
```

```
     5600                5605                5610

Glu Glu  Leu Arg Gln Met Ile  Lys Ile Arg Leu Gln  Asn Leu Gln
     5615                5620                5625

Asp Ala  Ala Lys Asp Met Lys  Lys Phe Glu Ala Glu  Leu Lys Lys
     5630                5635                5640

Leu Gln  Ala Ala Leu Glu Gln  Ala Gln Ala Thr Leu  Thr Ser Pro
     5645                5650                5655

Glu Val  Gly Arg Leu Ser Leu  Lys Glu Gln Leu Ser  His Arg Gln
     5660                5665                5670

His Leu  Leu Ser Glu Met Glu  Ser Leu Lys Pro Lys  Val Gln Ala
     5675                5680                5685

Val Gln  Leu Cys Gln Ser Ala  Leu Arg Ile Pro Glu  Asp Val Val
     5690                5695                5700

Ala Ser  Leu Pro Leu Cys His  Ala Ala Leu Arg Leu  Gln Glu Glu
     5705                5710                5715

Ala Ser  Arg Leu Gln His Thr  Ala Ile Gln Gln Cys  Asn Ile Met
     5720                5725                5730

Gln Glu  Ala Val Val Gln Tyr  Glu Gln Tyr Glu Gln  Glu Met Lys
     5735                5740                5745

His Leu  Gln Gln Leu Ile Glu  Gly Ala His Arg Glu  Ile Glu Asp
     5750                5755                5760

Lys Pro  Val Ala Thr Ser Asn  Ile Gln Glu Leu Gln  Ala Gln Ile
     5765                5770                5775

Ser Arg  His Glu Glu Leu Ala  Gln Lys Ile Lys Gly  Tyr Gln Glu
     5780                5785                5790

Gln Ile  Ala Ser Leu Asn Ser  Lys Cys Lys Met Leu  Thr Met Lys
     5795                5800                5805

Ala Lys  His Ala Thr Met Leu  Leu Thr Val Thr Glu  Val Glu Gly
     5810                5815                5820

Leu Ala  Glu Gly Thr Glu Asp  Leu Asp Gly Glu Leu  Leu Pro Thr
     5825                5830                5835

Pro Ser  Ala His Pro Ser Val  Val Met Met Thr Ala  Gly Arg Cys
     5840                5845                5850

His Thr  Leu Leu Ser Pro Val  Thr Glu Glu Ser Gly  Glu Glu Gly
     5855                5860                5865
```

```
Thr Asn  Ser Glu Ile Ser Ser  Pro Pro Ala Cys Arg  Ser Pro Ser
    5870                5875                5880

Pro Val  Ala Asn Thr Asp Ala  Ser Val Asn Gln Asp  Ile Ala Tyr
    5885                5890                5895

Tyr Gln  Ala Leu Ser Ala Glu  Arg Leu Gln Thr Asp  Ala Ala Lys
    5900                5905                5910

Ile His  Pro Ser Thr Ser Ala  Ser Gln Glu Phe Tyr  Glu Pro Gly
    5915                5920                5925

Leu Glu  Pro Ser Ala Thr Ala  Lys Leu Gly Asp Leu  Gln Arg Ser
    5930                5935                5940

Trp Glu  Thr Leu Lys Asn Val  Ile Ser Glu Lys Gln  Arg Thr Leu
    5945                5950                5955

Tyr Glu  Ala Leu Glu Arg Gln  Gln Lys Tyr Gln Asp  Ser Leu Gln
    5960                5965                5970

Ser Ile  Ser Thr Lys Met Glu  Ala Ile Glu Leu Lys  Leu Ser Glu
    5975                5980                5985

Ser Pro  Glu Pro Gly Arg Ser  Pro Glu Ser Gln Met  Ala Glu His
    5990                5995                6000

Gln Ala  Leu Met Asp Glu Ile  Leu Met Leu Gln Asp  Glu Ile Asn
    6005                6010                6015

Glu Leu  Gln Ser Ser Leu Ala  Glu Glu Leu Val Ser  Glu Ser Cys
    6020                6025                6030

Glu Ala  Asp Pro Ala Glu Gln  Leu Ala Leu Gln Ser  Thr Leu Thr
    6035                6040                6045

Val Leu  Ala Glu Arg Met Ser  Thr Ile Arg Met Lys  Ala Ser Gly
    6050                6055                6060

Lys Arg  Gln Leu Leu Glu Glu  Lys Leu Asn Asp Gln  Leu Glu Glu
    6065                6070                6075

Gln Arg  Gln Glu Gln Ala Leu  Gln Arg Tyr Arg Cys  Glu Ala Asp
    6080                6085                6090

Glu Leu  Asp Ser Trp Leu Leu  Ser Thr Lys Ala Thr  Leu Asp Thr
    6095                6100                6105

Ala Leu  Ser Pro Pro Lys Glu  Pro Met Asp Met Glu  Ala Gln Leu
    6110                6115                6120
```

```
Met Asp  Cys Gln Asn Met Leu  Val Glu Ile Glu Gln  Lys Val Val
    6125                 6130                 6135

Ala Leu  Ser Glu Leu Ser Val  His Asn Glu Asn Leu  Leu Leu Glu
    6140                 6145                 6150

Gly Lys  Ala His Thr Lys Asp  Glu Ala Glu Gln Leu  Ala Gly Lys
    6155                 6160                 6165

Leu Arg  Arg Leu Lys Gly Ser  Leu Leu Glu Leu Gln  Arg Ala Leu
    6170                 6175                 6180

His Asp  Lys Gln Leu Asn Met  Gln Gly Thr Ala Gln  Glu Lys Glu
    6185                 6190                 6195

Glu Ser  Asp Val Asp Leu Thr  Ala Thr Gln Ser Pro  Gly Val Gln
    6200                 6205                 6210

Glu Trp  Leu Ala Gln Ala Arg  Thr Thr Trp Thr Gln  Gln Arg Gln
    6215                 6220                 6225

Ser Ser  Leu Gln Gln Gln Lys  Glu Leu Glu Gln Glu  Leu Ala Glu
    6230                 6235                 6240

Gln Lys  Ser Leu Leu Arg Ser  Val Ala Ser Arg Gly  Glu Glu Ile
    6245                 6250                 6255

Leu Ile  Gln His Ser Ala Ala  Glu Thr Ser Gly Asp  Ala Gly Glu
    6260                 6265                 6270

Lys Pro  Asp Val Leu Ser Gln  Glu Leu Gly Met Glu  Gly Glu Lys
    6275                 6280                 6285

Ser Ser  Ala Glu Asp Gln Met  Arg Met Lys Trp Glu  Ser Leu His
    6290                 6295                 6300

Gln Glu  Phe Ser Thr Lys Gln  Lys Leu Leu Gln Asn  Val Leu Glu
    6305                 6310                 6315

Gln Glu  Gln Glu Gln Val Leu  Tyr Ser Arg Pro Asn  Arg Leu Leu
    6320                 6325                 6330

Ser Gly  Val Pro Leu Tyr Lys  Gly Asp Val Pro Thr  Gln Asp Lys
    6335                 6340                 6345

Ser Ala  Val Thr Ser Leu Leu  Asp Gly Leu Asn Gln  Ala Phe Glu
    6350                 6355                 6360

Glu Val  Ser Ser Gln Ser Gly  Gly Ala Lys Arg Gln  Ser Ile His
    6365                 6370                 6375
```

```
Leu Glu  Gln Lys Leu Tyr Asp  Gly Val Ser Ala Thr  Ser Thr Trp
    6380                 6385                 6390


Leu Asp  Asp Val Glu Glu Arg  Leu Phe Val Ala Thr  Ala Leu Leu
    6395                 6400                 6405


Pro Glu  Glu Thr Glu Thr Cys  Leu Phe Asn Gln Glu  Ile Leu Ala
    6410                 6415                 6420


Lys Asp  Ile Lys Glu Met Ser  Glu Glu Met Asp Lys  Asn Lys Asn
    6425                 6430                 6435


Leu Phe  Ser Gln Ala Phe Pro  Glu Asn Gly Asp Asn  Arg Asp Val
    6440                 6445                 6450


Ile Glu  Asp Thr Leu Gly Cys  Leu Leu Gly Arg Leu  Ser Leu Leu
    6455                 6460                 6465


Asp Ser  Val Val Asn Gln Arg  Cys His Gln Met Lys  Glu Arg Leu
    6470                 6475                 6480


Gln Gln  Ile Leu Asn Phe Gln  Asn Asp Leu Lys Val  Leu Phe Thr
    6485                 6490                 6495


Ser Leu  Ala Asp Asn Lys Tyr  Ile Ile Leu Gln Lys  Leu Ala Asn
    6500                 6505                 6510


Val Phe  Glu Gln Pro Val Ala  Glu Gln Ile Glu Ala  Ile Gln Gln
    6515                 6520                 6525


Ala Glu  Asp Gly Leu Lys Glu  Phe Asp Ala Gly Ile  Ile Glu Leu
    6530                 6535                 6540


Lys Arg  Arg Gly Asp Lys Leu  Gln Val Glu Gln Pro  Ser Met Gln
    6545                 6550                 6555


Glu Leu  Ser Lys Leu Gln Asp  Met Tyr Asp Glu Leu  Met Met Ile
    6560                 6565                 6570


Ile Gly  Ser Arg Arg Ser Gly  Leu Asn Gln Asn Leu  Thr Leu Lys
    6575                 6580                 6585


Ser Gln  Tyr Glu Arg Ala Leu  Gln Asp Leu Ala Asp  Leu Leu Glu
    6590                 6595                 6600


Thr Gly  Gln Glu Lys Met Ala  Gly Asp Gln Lys Ile  Ile Val Ser
    6605                 6610                 6615


Ser Lys  Glu Glu Ile Gln Gln  Leu Leu Asp Lys His  Lys Glu Tyr
    6620                 6625                 6630


Phe Gln  Gly Leu Glu Ser His  Met Ile Leu Thr Glu  Thr Leu Phe
```

```
     6635                6640                6645

Arg Lys  Ile Ile Ser Phe Ala  Val Gln Lys Glu Thr  Gln Phe His
     6650                6655                6660

Thr Glu  Leu Met Ala Gln Ala  Ser Ala Val Leu Lys  Arg Ala His
     6665                6670                6675

Lys Arg  Gly Val Glu Leu Glu  Tyr Ile Leu Glu Thr  Trp Ser His
     6680                6685                6690

Leu Asp  Glu Asp Gln Gln Glu  Leu Ser Arg Gln Leu  Glu Val Val
     6695                6700                6705

Glu Ser  Ser Ile Pro Ser Val  Gly Leu Val Glu Glu  Asn Glu Asp
     6710                6715                6720

Arg Leu  Ile Asp Arg Ile Thr  Leu Tyr Gln His Leu  Lys Ser Ser
     6725                6730                6735

Leu Asn  Glu Tyr Gln Pro Lys  Leu Tyr Gln Val Leu  Asp Asp Gly
     6740                6745                6750

Lys Arg  Leu Leu Ile Ser Ile  Ser Cys Ser Asp Leu  Glu Ser Gln
     6755                6760                6765

Leu Asn  Gln Leu Gly Glu Cys  Trp Leu Ser Asn Thr  Asn Lys Met
     6770                6775                6780

Ser Lys  Glu Leu His Arg Leu  Glu Thr Ile Leu Lys  His Trp Thr
     6785                6790                6795

Arg Tyr  Gln Ser Glu Ser Ala  Asp Leu Ile His Trp  Leu Gln Ser
     6800                6805                6810

Ala Lys  Asp Arg Leu Glu Phe  Trp Thr Gln Gln Ser  Val Thr Val
     6815                6820                6825

Pro Gln  Glu Leu Glu Met Val  Arg Asp His Leu Asn  Ala Phe Leu
     6830                6835                6840

Glu Phe  Ser Lys Glu Val Asp  Ala Gln Ser Ser Leu  Lys Ser Ser
     6845                6850                6855

Val Leu  Ser Thr Gly Asn Gln  Leu Leu Arg Leu Lys  Lys Val Asp
     6860                6865                6870

Thr Ala  Thr Leu Arg Ser Glu  Leu Ser Arg Ile Asp  Ser Gln Trp
     6875                6880                6885

Thr Asp  Leu Leu Thr Asn Ile  Pro Ala Val Gln Glu  Lys Leu His
     6890                6895                6900
```

```
Gln Leu  Gln Met Asp Lys Leu  Pro Ser Arg His Ala  Ile Ser Glu
    6905                6910                6915

Val Met  Ser Trp Ile Ser Leu  Met Glu Asn Val Ile  Gln Lys Asp
    6920                6925                6930

Glu Asp  Asn Ile Lys Asn Ser  Ile Gly Tyr Lys Ala  Ile His Glu
    6935                6940                6945

Tyr Leu  Gln Lys Tyr Lys Gly  Phe Lys Ile Asp Ile  Asn Cys Lys
    6950                6955                6960

Gln Leu  Thr Val Asp Phe Val  Asn Gln Ser Val Leu  Gln Ile Ser
    6965                6970                6975

Ser Gln  Asp Val Glu Ser Lys  Arg Ser Asp Lys Thr  Asp Phe Ala
    6980                6985                6990

Glu Gln  Leu Gly Ala Met Asn  Lys Ser Trp Gln Ile  Leu Gln Gly
    6995                7000                7005

Leu Val  Thr Glu Lys Ile Gln  Leu Leu Glu Gly Leu  Leu Glu Ser
    7010                7015                7020

Trp Ser  Glu Tyr Glu Asn Asn  Val Gln Cys Leu Lys  Thr Trp Phe
    7025                7030                7035

Glu Thr  Gln Glu Lys Arg Leu  Lys Gln Gln His Arg  Ile Gly Asp
    7040                7045                7050

Gln Ala  Ser Val Gln Asn Ala  Leu Lys Asp Cys Gln  Asp Leu Glu
    7055                7060                7065

Asp Leu  Ile Lys Ala Lys Glu  Lys Glu Val Glu Lys  Ile Glu Gln
    7070                7075                7080

Asn Gly  Leu Ala Leu Ile Gln  Asn Lys Lys Glu Asp  Val Ser Ser
    7085                7090                7095

Ile Val  Met Ser Thr Leu Arg  Glu Leu Gly Gln Thr  Trp Ala Asn
    7100                7105                7110

Leu Asp  His Met Val Gly Gln  Leu Lys Ile Leu Leu  Lys Ser Val
    7115                7120                7125

Leu Asp  Gln Trp Ser Ser His  Lys Val Ala Phe Asp  Lys Ile Asn
    7130                7135                7140

Ser Tyr  Leu Met Glu Ala Arg  Tyr Ser Leu Ser Arg  Phe Arg Leu
    7145                7150                7155
```

```
Leu Thr  Gly Ser Leu Glu Ala  Val Gln Val Gln Val  Asp Asn Leu
    7160                 7165                 7170

Gln Asn  Leu Gln Asp Asp Leu  Glu Lys Gln Glu Arg  Ser Leu Gln
    7175                 7180                 7185

Lys Phe  Gly Ser Ile Thr Asn  Gln Leu Leu Lys Glu  Cys His Pro
    7190                 7195                 7200

Pro Val  Thr Glu Thr Leu Thr  Asn Thr Leu Lys Glu  Val Asn Met
    7205                 7210                 7215

Arg Trp  Asn Asn Leu Leu Glu  Glu Ile Ala Glu Gln  Leu Gln Ser
    7220                 7225                 7230

Ser Lys  Ala Leu Leu Gln Leu  Trp Gln Arg Tyr Lys  Asp Tyr Ser
    7235                 7240                 7245

Lys Gln  Cys Ala Ser Thr Val  Gln Gln Gln Glu Asp  Arg Thr Asn
    7250                 7255                 7260

Glu Leu  Leu Lys Ala Ala Thr  Asn Lys Asp Ile Ala  Asp Asp Glu
    7265                 7270                 7275

Val Ala  Thr Trp Ile Gln Asp  Cys Asn Asp Leu Leu  Lys Gly Leu
    7280                 7285                 7290

Gly Thr  Val Lys Asp Ser Leu  Phe Phe Leu His Glu  Leu Gly Glu
    7295                 7300                 7305

Gln Leu  Lys Gln Gln Val Asp  Ala Ser Ala Ala Ser  Ala Ile Gln
    7310                 7315                 7320

Ser Asp  Gln Leu Ser Leu Ser  Gln His Leu Cys Ala  Leu Glu Gln
    7325                 7330                 7335

Ala Leu  Cys Lys Gln Gln Thr  Ser Leu Gln Ala Gly  Val Leu Asp
    7340                 7345                 7350

Tyr Glu  Thr Phe Ala Lys Ser  Leu Glu Ala Leu Glu  Ala Trp Ile
    7355                 7360                 7365

Val Glu  Ala Glu Glu Ile Leu  Gln Gly Gln Asp Pro  Ser His Ser
    7370                 7375                 7380

Ser Asp  Leu Ser Thr Ile Gln  Glu Arg Met Glu Glu  Leu Lys Gly
    7385                 7390                 7395

Gln Met  Leu Lys Phe Ser Ser  Met Ala Pro Asp Leu  Asp Arg Leu
    7400                 7405                 7410
```

```
Asn Glu  Leu Gly Tyr Arg Leu  Pro Leu Asn Asp Lys  Glu Ile Lys
    7415                 7420                 7425


Arg Met  Gln Asn Leu Asn Arg  His Trp Ser Leu Ile  Ser Ser Gln
    7430                 7435                 7440


Thr Thr  Glu Arg Phe Ser Lys  Leu Gln Ser Phe Leu  Leu Gln His
    7445                 7450                 7455


Gln Thr  Phe Leu Glu Lys Cys  Glu Thr Trp Met Glu  Phe Leu Val
    7460                 7465                 7470


Gln Thr  Glu Gln Lys Leu Ala  Val Glu Ile Ser Gly  Asn Tyr Gln
    7475                 7480                 7485


His Leu  Leu Glu Gln Gln Arg  Ala His Glu Leu Phe  Gln Ala Glu
    7490                 7495                 7500


Met Phe  Ser Arg Gln Gln Ile  Leu His Ser Ile Ile  Ile Asp Gly
    7505                 7510                 7515


Gln Arg  Leu Leu Glu Gln Gly  Gln Val Asp Asp Arg  Asp Glu Phe
    7520                 7525                 7530


Asn Leu  Lys Leu Thr Leu Leu  Ser Asn Gln Trp Gln  Gly Val Ile
    7535                 7540                 7545


Arg Arg  Ala Gln Gln Arg Arg  Gly Ile Ile Asp Ser  Gln Ile Arg
    7550                 7555                 7560


Gln Trp  Gln Arg Tyr Arg Glu  Met Ala Glu Lys Leu  Arg Lys Trp
    7565                 7570                 7575


Leu Val  Glu Val Ser Tyr Leu  Pro Met Ser Gly Leu  Gly Ser Val
    7580                 7585                 7590


Pro Ile  Pro Leu Gln Gln Ala  Arg Thr Leu Phe Asp  Glu Val Gln
    7595                 7600                 7605


Phe Lys  Glu Lys Val Phe Leu  Arg Gln Gln Gly Ser  Tyr Ile Leu
    7610                 7615                 7620


Thr Val  Glu Ala Gly Lys Gln  Leu Leu Leu Ser Ala  Asp Ser Gly
    7625                 7630                 7635


Ala Glu  Ala Ala Leu Gln Ala  Glu Leu Ala Glu Ile  Gln Glu Lys
    7640                 7645                 7650


Trp Lys  Ser Ala Ser Met Arg  Leu Glu Glu Gln Lys  Lys Lys Leu
    7655                 7660                 7665


Ala Phe  Leu Leu Lys Asp Trp  Glu Lys Cys Glu Lys  Gly Ile Ala
```

7670 7675 7680

Asp Ser Leu Glu Lys Leu Arg Thr Phe Lys Lys Lys Leu Ser Gln
7685 7690 7695

Ser Leu Pro Asp His His Glu Glu Leu His Ala Glu Gln Met Arg
7700 7705 7710

Cys Lys Glu Leu Glu Asn Ala Val Gly Ser Trp Thr Asp Asp Leu
7715 7720 7725

Thr Gln Leu Ser Leu Leu Lys Asp Thr Leu Ser Ala Tyr Ile Ser
7730 7735 7740

Ala Asp Asp Ile Ser Ile Leu Asn Glu Arg Val Glu Leu Leu Gln
7745 7750 7755

Arg Gln Trp Glu Glu Leu Cys His Gln Leu Ser Leu Arg Arg Gln
7760 7765 7770

Gln Ile Gly Glu Arg Leu Asn Glu Trp Ala Val Phe Ser Glu Lys
7775 7780 7785

Asn Lys Glu Leu Cys Glu Trp Leu Thr Gln Met Glu Ser Lys Val
7790 7795 7800

Ser Gln Asn Gly Asp Ile Leu Ile Glu Glu Met Ile Glu Lys Leu
7805 7810 7815

Lys Lys Asp Tyr Gln Glu Glu Ile Ala Ile Ala Gln Glu Asn Lys
7820 7825 7830

Ile Gln Leu Gln Gln Met Gly Glu Arg Leu Ala Lys Ala Ser His
7835 7840 7845

Glu Ser Lys Ala Ser Glu Ile Glu Tyr Lys Leu Gly Lys Val Asn
7850 7855 7860

Asp Arg Trp Gln His Leu Leu Asp Leu Ile Ala Ala Arg Val Lys
7865 7870 7875

Lys Leu Lys Glu Thr Leu Val Ala Val Gln Gln Leu Asp Lys Asn
7880 7885 7890

Met Ser Ser Leu Arg Thr Trp Leu Ala His Ile Glu Ser Glu Leu
7895 7900 7905

Ala Lys Pro Ile Val Tyr Asp Ser Cys Asn Ser Glu Glu Ile Gln
7910 7915 7920

Arg Lys Leu Asn Glu Gln Gln Glu Leu Gln Arg Asp Ile Glu Lys
7925 7930 7935

```
His Ser  Thr Gly Val Ala Ser  Val Leu Asn Leu Cys  Glu Val Leu
    7940                7945                7950

Leu His  Asp Cys Asp Ala Cys  Ala Thr Asp Ala Glu  Cys Asp Ser
    7955                7960                7965

Ile Gln  Gln Ala Thr Arg Asn  Leu Asp Arg Arg Trp  Arg Asn Ile
    7970                7975                7980

Cys Ala  Met Ser Met Glu Arg  Arg Leu Lys Ile Glu  Glu Thr Trp
    7985                7990                7995

Arg Leu  Trp Gln Lys Phe Leu  Asp Asp Tyr Ser Arg  Phe Glu Asp
    8000                8005                8010

Trp Leu  Lys Ser Ser Glu Arg  Thr Ala Ala Phe Pro  Ser Ser Ser
    8015                8020                8025

Gly Val  Ile Tyr Thr Val Ala  Lys Glu Glu Leu Lys  Lys Phe Glu
    8030                8035                8040

Ala Phe  Gln Arg Gln Val His  Glu Cys Leu Thr Gln  Leu Glu Leu
    8045                8050                8055

Ile Asn  Lys Gln Tyr Arg Arg  Leu Ala Arg Glu Asn  Arg Thr Asp
    8060                8065                8070

Ser Ala  Cys Ser Leu Lys Gln  Met Val His Glu Gly  Asn Gln Arg
    8075                8080                8085

Trp Asp  Asn Leu Gln Lys Arg  Val Thr Ser Ile Leu  Arg Arg Leu
    8090                8095                8100

Lys His  Phe Ile Gly Gln Arg  Glu Glu Phe Glu Thr  Ala Arg Asp
    8105                8110                8115

Ser Ile  Leu Val Trp Leu Thr  Glu Met Asp Leu Gln  Leu Thr Asn
    8120                8125                8130

Ile Glu  His Phe Ser Glu Cys  Asp Val Gln Ala Lys  Ile Lys Gln
    8135                8140                8145

Leu Lys  Ala Phe Gln Gln Glu  Ile Ser Leu Asn His  Asn Lys Ile
    8150                8155                8160

Glu Gln  Ile Ile Ala Gln Gly  Glu Gln Leu Ile Glu  Lys Ser Glu
    8165                8170                8175

Pro Leu  Asp Ala Ala Ile Ile  Glu Glu Glu Leu Asp  Glu Leu Arg
    8180                8185                8190
```

```
Arg Tyr  Cys Gln Glu Val Phe  Gly Arg Val Glu Arg  Tyr His Lys
    8195                 8200                 8205

Lys Leu  Ile Arg Leu Pro Leu  Pro Asp Asp Glu His  Asp Leu Ser
    8210                 8215                 8220

Asp Arg  Glu Leu Glu Leu Glu  Asp Ser Ala Ala Leu  Ser Asp Leu
    8225                 8230                 8235

His Trp  His Asp Arg Ser Ala  Asp Ser Leu Leu Ser  Pro Gln Pro
    8240                 8245                 8250

Ser Ser  Asn Leu Ser Leu Ser  Leu Ala Gln Pro Leu  Arg Ser Glu
    8255                 8260                 8265

Arg Ser  Gly Arg Asp Thr Pro  Ala Ser Val Asp Ser  Ile Pro Leu
    8270                 8275                 8280

Glu Trp  Asp His Asp Tyr Asp  Leu Ser Arg Asp Leu  Glu Ser Ala
    8285                 8290                 8295

Met Ser  Arg Ala Leu Pro Ser  Glu Asp Glu Glu Gly  Gln Asp Asp
    8300                 8305                 8310

Lys Asp  Phe Tyr Leu Arg Gly  Ala Val Gly Leu Ser  Gly Asp His
    8315                 8320                 8325

Ser Ala  Leu Glu Ser Gln Ile  Arg Gln Leu Gly Lys  Ala Leu Asp
    8330                 8335                 8340

Asp Ser  Arg Phe Gln Ile Gln  Gln Thr Glu Asn Ile  Ile Arg Ser
    8345                 8350                 8355

Lys Thr  Pro Thr Gly Pro Glu  Leu Asp Thr Ser Tyr  Lys Gly Tyr
    8360                 8365                 8370

Met Lys  Leu Leu Gly Glu Cys  Ser Ser Ser Ile Asp  Ser Val Lys
    8375                 8380                 8385

Arg Leu  Glu His Lys Leu Lys  Glu Glu Glu Glu Ser  Leu Pro Gly
    8390                 8395                 8400

Phe Val  Asn Leu His Ser Thr  Glu Thr Gln Thr Ala  Gly Val Ile
    8405                 8410                 8415

Asp Arg  Trp Glu Leu Leu Gln  Ala Gln Ala Leu Ser  Lys Glu Leu
    8420                 8425                 8430

Arg Met  Lys Gln Asn Leu Gln  Lys Trp Gln Gln Phe  Asn Ser Asp
    8435                 8440                 8445
```

```
Leu Asn  Ser Ile Trp Ala Trp  Leu Gly Asp Thr Glu  Glu Glu Leu
    8450                 8455                 8460

Glu Gln  Leu Gln Arg Leu Glu  Leu Ser Thr Asp Ile  Gln Thr Ile
    8465                 8470                 8475

Glu Leu  Gln Ile Lys Lys Leu  Lys Glu Leu Gln Lys  Ala Val Asp
    8480                 8485                 8490

His Arg  Lys Ala Ile Ile Leu  Ser Ile Asn Leu Cys  Ser Pro Glu
    8495                 8500                 8505

Phe Thr  Gln Ala Asp Ser Lys  Glu Ser Arg Asp Leu  Gln Asp Arg
    8510                 8515                 8520

Leu Ser  Gln Met Asn Gly Arg  Trp Asp Arg Val Cys  Ser Leu Leu
    8525                 8530                 8535

Glu Glu  Trp Arg Gly Leu Leu  Gln Asp Ala Leu Met  Gln Cys Gln
    8540                 8545                 8550

Gly Phe  His Glu Met Ser His  Gly Leu Leu Leu Met  Leu Glu Asn
    8555                 8560                 8565

Ile Asp  Arg Arg Lys Asn Glu  Ile Val Pro Ile Asp  Ser Asn Leu
    8570                 8575                 8580

Asp Ala  Glu Ile Leu Gln Asp  His His Lys Gln Leu  Met Gln Ile
    8585                 8590                 8595

Lys His  Glu Leu Leu Glu Ser  Gln Leu Arg Val Ala  Ser Leu Gln
    8600                 8605                 8610

Asp Met  Ser Cys Gln Leu Leu  Val Asn Ala Glu Gly  Thr Asp Cys
    8615                 8620                 8625

Leu Glu  Ala Lys Glu Lys Val  His Val Ile Gly Asn  Arg Leu Lys
    8630                 8635                 8640

Leu Leu  Leu Lys Glu Val Ser  Arg His Ile Lys Glu  Leu Glu Lys
    8645                 8650                 8655

Leu Leu  Asp Val Ser Ser Ser  Gln Gln Asp Leu Ser  Ser Trp Ser
    8660                 8665                 8670

Ser Ala  Asp Glu Leu Asp Thr  Ser Gly Ser Val Ser  Pro Thr Ser
    8675                 8680                 8685

Gly Arg  Ser Thr Pro Asn Arg  Gln Lys Thr Pro Arg  Gly Lys Cys
    8690                 8695                 8700

Ser Leu  Ser Gln Pro Gly Pro  Ser Val Ser Ser Pro  His Ser Arg
```

8705 8710 8715

Ser Thr Lys Gly Gly Ser Asp Ser Ser Leu Ser Glu Pro Gly Pro
8720 8725 8730

Gly Arg Ser Gly Arg Gly Phe Leu Phe Arg Val Leu Arg Ala Ala
8735 8740 8745

Leu Pro Leu Gln Leu Leu Leu Leu Leu Leu Ile Gly Leu Ala Cys
8750 8755 8760

Leu Val Pro Met Ser Glu Glu Asp Tyr Ser Cys Ala Leu Ser Asn
8765 8770 8775

Asn Phe Ala Arg Ser Phe His Pro Met Leu Arg Tyr Thr Asn Gly
8780 8785 8790

Pro Pro Pro Leu
8795

<210> 135
<211> 2677
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> SETX

<400> 135

Met Ser Thr Cys Cys Trp Cys Thr Pro Gly Gly Ala Ser Thr Ile Asp
1 5 10 15

Phe Leu Lys Arg Tyr Ala Ser Asn Thr Pro Ser Gly Glu Phe Gln Thr
20 25 30

Ala Asp Glu Asp Leu Cys Tyr Cys Leu Glu Cys Val Ala Glu Tyr His
35 40 45

Lys Ala Arg Asp Glu Leu Pro Phe Leu His Glu Val Leu Trp Glu Leu
50 55 60

Glu Thr Leu Arg Leu Ile Asn His Phe Glu Lys Ser Met Lys Ala Glu
65 70 75 80

Ile Gly Asp Asp Asp Glu Leu Tyr Ile Val Asp Asn Asn Gly Glu Met
85 90 95

Pro Leu Phe Asp Ile Thr Gly Gln Asp Phe Glu Asn Lys Leu Arg Val
100 105 110

Pro Leu Leu Glu Ile Leu Lys Tyr Pro Tyr Leu Leu Leu His Glu Arg
115 120 125

```
Val Asn Glu Leu Cys Val Glu Ala Leu Cys Arg Met Glu Gln Ala Asn
    130                 135                 140

Cys Ser Phe Gln Val Phe Asp Lys His Pro Gly Ile Tyr Leu Phe Leu
145                 150                 155                 160

Val His Pro Asn Glu Met Val Arg Arg Trp Ala Ile Leu Thr Ala Arg
                165                 170                 175

Asn Leu Gly Lys Val Asp Arg Asp Asp Tyr Tyr Asp Leu Gln Glu Val
            180                 185                 190

Leu Leu Cys Leu Phe Lys Val Ile Glu Leu Gly Leu Leu Glu Ser Pro
        195                 200                 205

Asp Ile Tyr Thr Ser Ser Val Leu Glu Lys Gly Lys Leu Ile Leu Leu
    210                 215                 220

Pro Ser His Met Tyr Asp Thr Thr Asn Tyr Lys Ser Tyr Trp Leu Gly
225                 230                 235                 240

Ile Cys Met Leu Leu Thr Ile Leu Glu Glu Gln Ala Met Asp Ser Leu
                245                 250                 255

Leu Leu Gly Ser Asp Lys Gln Asn Asp Phe Met Gln Ser Ile Leu His
            260                 265                 270

Thr Met Glu Arg Glu Ala Asp Asp Asp Ser Val Asp Pro Phe Trp Pro
        275                 280                 285

Ala Leu His Cys Phe Met Val Ile Leu Asp Arg Leu Gly Ser Lys Val
    290                 295                 300

Trp Gly Gln Leu Met Asp Pro Ile Val Ala Phe Gln Thr Ile Ile Asn
305                 310                 315                 320

Asn Ala Ser Tyr Asn Arg Glu Ile Arg His Ile Arg Asn Ser Ser Val
                325                 330                 335

Arg Thr Lys Leu Glu Pro Glu Ser Tyr Leu Asp Asp Met Val Thr Cys
            340                 345                 350

Ser Gln Ile Val Tyr Asn Tyr Asn Pro Glu Lys Thr Lys Lys Asp Ser
        355                 360                 365

Gly Trp Arg Thr Ala Ile Cys Pro Asp Tyr Cys Pro Asn Met Tyr Glu
    370                 375                 380

Glu Met Glu Thr Leu Ala Ser Val Leu Gln Ser Asp Ile Gly Gln Asp
385                 390                 395                 400
```

```
Met Arg Val His Asn Ser Thr Phe Leu Trp Phe Ile Pro Phe Val Gln
                405                 410                 415

Ser Leu Met Asp Leu Lys Asp Leu Gly Val Ala Tyr Ile Ala Gln Val
            420                 425                 430

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

Ala Val Cys Asp Lys Val Thr Glu Phe Phe Leu Leu Ile Leu Val Ser
    450                 455                 460

Val Ile Glu Leu His Arg Asn Lys Lys Cys Leu His Leu Leu Trp Val
465                 470                 475                 480

Ser Ser Gln Gln Trp Val Glu Ala Val Val Lys Cys Ala Lys Leu Pro
                485                 490                 495

Thr Thr Ala Phe Thr Arg Ser Ser Glu Lys Ser Ser Gly Asn Cys Ser
            500                 505                 510

Lys Gly Thr Ala Met Ile Ser Ser Leu Ser Leu His Ser Met Pro Ser
        515                 520                 525

Asn Ser Val Gln Leu Ala Tyr Val Gln Leu Ile Arg Ser Leu Leu Lys
    530                 535                 540

Glu Gly Tyr Gln Leu Gly Gln Gln Ser Leu Cys Lys Arg Phe Trp Asp
545                 550                 555                 560

Lys Leu Asn Leu Phe Leu Arg Gly Asn Leu Ser Leu Gly Trp Gln Leu
                565                 570                 575

Thr Ser Gln Glu Thr His Glu Leu Gln Ser Cys Leu Lys Gln Ile Ile
            580                 585                 590

Arg Asn Ile Lys Phe Lys Ala Pro Pro Cys Asn Thr Phe Val Asp Leu
        595                 600                 605

Thr Ser Ala Cys Lys Ile Ser Pro Ala Ser Tyr Asn Lys Glu Glu Ser
    610                 615                 620

Glu Gln Met Gly Lys Thr Ser Arg Lys Asp Met His Cys Leu Glu Ala
625                 630                 635                 640

Ser Ser Pro Thr Phe Ser Lys Glu Pro Met Lys Val Gln Asp Ser Val
                645                 650                 655

Leu Ile Lys Ala Asp Asn Thr Ile Glu Gly Asp Asn Asn Glu Gln Asn
            660                 665                 670
```

Tyr Ile Lys Asp Val Lys Leu Glu Asp His Leu Leu Ala Gly Ser Cys
675 680 685

Leu Lys Gln Ser Ser Lys Asn Ile Phe Thr Glu Arg Ala Glu Asp Gln
690 695 700

Ile Lys Ile Ser Thr Arg Lys Gln Lys Ser Val Lys Glu Ile Ser Ser
705 710 715 720

Tyr Thr Pro Lys Asp Cys Thr Ser Arg Asn Gly Pro Glu Arg Gly Cys
725 730 735

Asp Arg Gly Ile Ile Val Ser Thr Arg Leu Leu Thr Asp Ser Ser Thr
740 745 750

Asp Ala Leu Glu Lys Val Ser Thr Ser Asn Glu Asp Phe Ser Leu Lys
755 760 765

Asp Asp Ala Leu Ala Lys Thr Ser Lys Arg Lys Thr Lys Val Gln Lys
770 775 780

Asp Glu Ile Cys Ala Lys Leu Ser His Val Ile Lys Lys Gln His Arg
785 790 795 800

Lys Ser Thr Leu Val Asp Asn Thr Ile Asn Leu Asp Glu Asn Leu Thr
805 810 815

Val Ser Asn Ile Glu Ser Phe Tyr Ser Arg Lys Asp Thr Gly Val Gln
820 825 830

Lys Gly Asp Gly Phe Ile His Asn Leu Ser Leu Asp Pro Ser Gly Val
835 840 845

Leu Asp Asp Lys Asn Gly Glu Gln Lys Ser Gln Asn Asn Val Leu Pro
850 855 860

Lys Glu Lys Gln Leu Lys Asn Glu Glu Leu Val Ile Phe Ser Phe His
865 870 875 880

Glu Asn Asn Cys Lys Ile Gln Glu Phe His Val Asp Gly Lys Glu Leu
885 890 895

Ile Pro Phe Thr Glu Met Thr Asn Ala Ser Glu Lys Lys Ser Ser Pro
900 905 910

Phe Lys Asp Leu Met Thr Val Pro Glu Ser Arg Asp Glu Glu Met Ser
915 920 925

Asn Ser Thr Ser Val Ile Tyr Ser Asn Leu Thr Arg Glu Gln Ala Pro
930 935 940

Asp Ile Ser Pro Lys Ser Asp Thr Leu Thr Asp Ser Gln Ile Asp Arg

945 950 955 960

Asp Leu His Lys Leu Ser Leu Leu Ala Gln Ala Ser Val Ile Thr Phe
965 970 975

Pro Ser Asp Ser Pro Gln Asn Ser Ser Gln Leu Gln Arg Lys Val Lys
980 985 990

Glu Asp Lys Arg Cys Phe Thr Ala Asn Gln Asn Asn Val Gly Asp Thr
995 1000 1005

Ser Arg Gly Gln Val Ile Ile Ile Ser Asp Ser Asp Asp Asp Asp
1010 1015 1020

Asp Glu Arg Ile Leu Ser Leu Glu Lys Leu Thr Lys Gln Asp Lys
1025 1030 1035

Ile Cys Leu Glu Arg Glu His Pro Glu Gln His Val Ser Thr Val
1040 1045 1050

Asn Ser Lys Glu Glu Lys Asn Pro Val Lys Glu Glu Lys Thr Glu
1055 1060 1065

Thr Leu Phe Gln Phe Glu Glu Ser Asp Ser Gln Cys Phe Glu Phe
1070 1075 1080

Glu Ser Ser Ser Glu Val Phe Ser Val Trp Gln Asp His Pro Asp
1085 1090 1095

Asp Asn Asn Ser Val Gln Asp Gly Glu Lys Lys Cys Leu Ala Pro
1100 1105 1110

Ile Ala Asn Thr Thr Asn Gly Gln Gly Cys Thr Asp Tyr Val Ser
1115 1120 1125

Glu Val Val Lys Lys Gly Ala Glu Gly Ile Glu Glu His Thr Arg
1130 1135 1140

Pro Arg Ser Ile Ser Val Glu Glu Phe Cys Glu Ile Glu Val Lys
1145 1150 1155

Lys Pro Lys Arg Lys Arg Ser Glu Lys Pro Met Ala Glu Asp Pro
1160 1165 1170

Val Arg Pro Ser Ser Ser Val Arg Asn Glu Gly Gln Ser Asp Thr
1175 1180 1185

Asn Lys Arg Asp Leu Val Gly Asn Asp Phe Lys Ser Ile Asp Arg
1190 1195 1200

Arg Thr Ser Thr Pro Asn Ser Arg Ile Gln Arg Ala Thr Thr Val
1205 1210 1215

```
Ser Gln  Lys Lys Ser Ser Lys  Leu Cys Thr Cys Thr  Glu Pro Ile
    1220                 1225                 1230

Arg Lys  Val Pro Val Ser Lys  Thr Pro Lys Lys Thr  His Ser Asp
    1235                 1240                 1245

Ala Lys  Lys Gly Gln Asn Arg  Ser Ser Asn Tyr Leu  Ser Cys Arg
    1250                 1255                 1260

Thr Thr  Pro Ala Ile Val Pro  Pro Lys Lys Phe Arg  Gln Cys Pro
    1265                 1270                 1275

Glu Pro  Thr Ser Thr Ala Glu  Lys Leu Gly Leu Lys  Lys Gly Pro
    1280                 1285                 1290

Arg Lys  Ala Tyr Glu Leu Ser  Gln Arg Ser Leu Asp  Tyr Val Ala
    1295                 1300                 1305

Gln Leu  Arg Asp His Gly Lys  Thr Val Gly Val Val  Asp Thr Arg
    1310                 1315                 1320

Lys Lys  Thr Lys Leu Ile Ser  Pro Gln Asn Leu Ser  Val Arg Asn
    1325                 1330                 1335

Asn Lys  Lys Leu Leu Thr Ser  Gln Glu Leu Gln Met  Gln Arg Gln
    1340                 1345                 1350

Ile Arg  Pro Lys Ser Gln Lys  Asn Arg Arg Arg Leu  Ser Asp Cys
    1355                 1360                 1365

Glu Ser  Thr Asp Val Lys Arg  Ala Gly Ser His Thr  Ala Gln Asn
    1370                 1375                 1380

Ser Asp  Ile Phe Val Pro Glu  Ser Asp Arg Ser Asp  Tyr Asn Cys
    1385                 1390                 1395

Thr Gly  Gly Thr Glu Val Leu  Ala Asn Ser Asn Arg  Lys Gln Leu
    1400                 1405                 1410

Ile Lys  Cys Met Pro Ser Glu  Pro Glu Thr Ile Lys  Ala Lys His
    1415                 1420                 1425

Gly Ser  Pro Ala Thr Asp Asp  Ala Cys Pro Leu Asn  Gln Cys Asp
    1430                 1435                 1440

Ser Val  Val Leu Asn Gly Thr  Val Pro Thr Asn Glu  Val Ile Val
    1445                 1450                 1455

Ser Thr  Ser Glu Asp Pro Leu  Gly Gly Gly Asp Pro  Thr Ala Arg
    1460                 1465                 1470
```

```
His Ile  Glu Met Ala Ala Leu  Lys Glu Gly Glu Pro  Asp Ser Ser
    1475                1480                1485

Ser Asp  Ala Glu Glu Asp Asn  Leu Phe Leu Thr Gln  Asn Asp Pro
    1490                1495                1500

Glu Asp  Met Asp Leu Cys Ser  Gln Met Glu Asn Asp  Asn Tyr Lys
    1505                1510                1515

Leu Ile  Glu Leu Ile His Gly  Lys Asp Thr Val Glu  Val Glu Glu
    1520                1525                1530

Asp Ser  Val Ser Arg Pro Gln  Leu Glu Ser Leu Ser  Gly Thr Lys
    1535                1540                1545

Cys Lys  Tyr Lys Asp Cys Leu  Glu Thr Thr Lys Asn  Gln Gly Glu
    1550                1555                1560

Tyr Cys  Pro Lys His Ser Glu  Val Lys Ala Ala Asp  Glu Asp Val
    1565                1570                1575

Phe Arg  Lys Pro Gly Leu Pro  Pro Pro Ala Ser Lys  Pro Leu Arg
    1580                1585                1590

Pro Thr  Thr Lys Ile Phe Ser  Ser Lys Ser Thr Ser  Arg Ile Ala
    1595                1600                1605

Gly Leu  Ser Lys Ser Leu Glu  Thr Ser Ser Ala Leu  Ser Pro Ser
    1610                1615                1620

Leu Lys  Asn Lys Ser Lys Gly  Ile Gln Ser Ile Leu  Lys Val Pro
    1625                1630                1635

Gln Pro  Val Pro Leu Ile Ala  Gln Lys Pro Val Gly  Glu Met Lys
    1640                1645                1650

Asn Ser  Cys Asn Val Leu His  Pro Gln Ser Pro Asn  Asn Ser Asn
    1655                1660                1665

Arg Gln  Gly Cys Lys Val Pro  Phe Gly Glu Ser Lys  Tyr Phe Pro
    1670                1675                1680

Ser Ser  Ser Pro Val Asn Ile  Leu Leu Ser Ser Gln  Ser Val Ser
    1685                1690                1695

Asp Thr  Phe Val Lys Glu Val  Leu Lys Trp Lys Tyr  Glu Met Phe
    1700                1705                1710

Leu Asn  Phe Gly Gln Cys Gly  Pro Pro Ala Ser Leu  Cys Gln Ser
    1715                1720                1725
```

```
Ile Ser  Arg Pro Val Pro Val  Arg Phe His Asn Tyr  Gly Asp Tyr
    1730                1735                1740

Phe Asn  Val Phe Phe Pro Leu  Met Val Leu Asn Thr  Phe Glu Thr
    1745                1750                1755

Val Ala  Gln Glu Trp Leu Asn  Ser Pro Asn Arg Glu  Asn Phe Tyr
    1760                1765                1770

Gln Leu  Gln Val Arg Lys Phe  Pro Ala Asp Tyr Ile  Lys Tyr Trp
    1775                1780                1785

Glu Phe  Ala Val Tyr Leu Glu  Glu Cys Glu Leu Ala  Lys Gln Leu
    1790                1795                1800

Tyr Pro  Lys Glu Asn Asp Leu  Val Phe Leu Ala Pro  Glu Arg Ile
    1805                1810                1815

Asn Glu  Glu Lys Lys Asp Thr  Glu Arg Asn Asp Ile  Gln Asp Leu
    1820                1825                1830

His Glu  Tyr His Ser Gly Tyr  Val His Lys Phe Arg  Arg Thr Ser
    1835                1840                1845

Val Met  Arg Asn Gly Lys Thr  Glu Cys Tyr Leu Ser  Ile Gln Thr
    1850                1855                1860

Gln Glu  Asn Phe Pro Ala Asn  Leu Asn Glu Leu Val  Asn Cys Ile
    1865                1870                1875

Val Ile  Ser Ser Leu Val Thr  Thr Gln Arg Lys Leu  Lys Ala Met
    1880                1885                1890

Ser Leu  Leu Gly Ser Arg Asn  Gln Leu Ala Arg Ala  Val Leu Asn
    1895                1900                1905

Pro Asn  Pro Met Asp Phe Cys  Thr Lys Asp Leu Leu  Thr Thr Thr
    1910                1915                1920

Ser Glu  Arg Ile Ile Ala Tyr  Leu Arg Asp Phe Asn  Glu Asp Gln
    1925                1930                1935

Lys Lys  Ala Ile Glu Thr Ala  Tyr Ala Met Val Lys  His Ser Pro
    1940                1945                1950

Ser Val  Ala Lys Ile Cys Leu  Ile His Gly Pro Pro  Gly Thr Gly
    1955                1960                1965

Lys Ser  Lys Thr Ile Val Gly  Leu Leu Tyr Arg Leu  Leu Thr Glu
    1970                1975                1980

Asn Gln  Arg Lys Gly His Ser  Asp Glu Asn Ser Asn  Ala Lys Ile
```

```
     1985                1990                1995

 Lys Gln  Asn Arg Val Leu Val  Cys Ala Pro Ser Asn  Ala Ala Val
     2000                2005                2010

 Asp Glu  Leu Met Lys Lys Ile  Ile Leu Glu Phe Lys  Glu Lys Cys
     2015                2020                2025

 Lys Asp  Lys Lys Asn Pro Leu  Gly Asn Cys Gly Asp  Ile Asn Leu
     2030                2035                2040

 Val Arg  Leu Gly Pro Glu Lys  Ser Ile Asn Ser Glu  Val Leu Lys
     2045                2050                2055

 Phe Ser  Leu Asp Ser Gln Val  Asn His Arg Met Lys  Lys Glu Leu
     2060                2065                2070

 Pro Ser  His Val Gln Ala Met  His Lys Arg Lys Glu  Phe Leu Asp
     2075                2080                2085

 Tyr Gln  Leu Asp Glu Leu Ser  Arg Gln Arg Ala Leu  Cys Arg Gly
     2090                2095                2100

 Gly Arg  Glu Ile Gln Arg Gln  Glu Leu Asp Glu Asn  Ile Ser Lys
     2105                2110                2115

 Val Ser  Lys Glu Arg Gln Glu  Leu Ala Ser Lys Ile  Lys Glu Val
     2120                2125                2130

 Gln Gly  Arg Pro Gln Lys Thr  Gln Ser Ile Ile Ile  Leu Glu Ser
     2135                2140                2145

 His Ile  Ile Cys Cys Thr Leu  Ser Thr Ser Gly Gly  Leu Leu Leu
     2150                2155                2160

 Glu Ser  Ala Phe Arg Gly Gln  Gly Gly Val Pro Phe  Ser Cys Val
     2165                2170                2175

 Ile Val  Asp Glu Ala Gly Gln  Ser Cys Glu Ile Glu  Thr Leu Thr
     2180                2185                2190

 Pro Leu  Ile His Arg Cys Asn  Lys Leu Ile Leu Val  Gly Asp Pro
     2195                2200                2205

 Lys Gln  Leu Pro Pro Thr Val  Ile Ser Met Lys Ala  Gln Glu Tyr
     2210                2215                2220

 Gly Tyr  Asp Gln Ser Met Met  Ala Arg Phe Cys Arg  Leu Leu Glu
     2225                2230                2235

 Glu Asn  Val Glu His Asn Met  Ile Ser Arg Leu Pro  Ile Leu Gln
     2240                2245                2250
```

```
Leu Thr  Val Gln Tyr Arg Met  His Pro Asp Ile Cys  Leu Phe Pro
    2255                 2260                 2265

Ser Asn  Tyr Val Tyr Asn Arg  Asn Leu Lys Thr Asn  Arg Gln Thr
    2270                 2275                 2280

Glu Ala  Ile Arg Cys Ser Ser  Asp Trp Pro Phe Gln  Pro Tyr Leu
    2285                 2290                 2295

Val Phe  Asp Val Gly Asp Gly  Ser Glu Arg Arg Asp  Asn Asp Ser
    2300                 2305                 2310

Tyr Ile  Asn Val Gln Glu Ile  Lys Leu Val Met Glu  Ile Ile Lys
    2315                 2320                 2325

Leu Ile  Lys Asp Lys Arg Lys  Asp Val Ser Phe Arg  Asn Ile Gly
    2330                 2335                 2340

Ile Ile  Thr His Tyr Lys Ala  Gln Lys Thr Met Ile  Gln Lys Asp
    2345                 2350                 2355

Leu Asp  Lys Glu Phe Asp Arg  Lys Gly Pro Ala Glu  Val Asp Thr
    2360                 2365                 2370

Val Asp  Ala Phe Gln Gly Arg  Gln Lys Asp Cys Val  Ile Val Thr
    2375                 2380                 2385

Cys Val  Arg Ala Asn Ser Ile  Gln Gly Ser Ile Gly  Phe Leu Ala
    2390                 2395                 2400

Ser Leu  Gln Arg Leu Asn Val  Thr Ile Thr Arg Ala  Lys Tyr Ser
    2405                 2410                 2415

Leu Phe  Ile Leu Gly His Leu  Arg Thr Leu Met Glu  Asn Gln His
    2420                 2425                 2430

Trp Asn  Gln Leu Ile Gln Asp  Ala Gln Lys Arg Gly  Ala Ile Ile
    2435                 2440                 2445

Lys Thr  Cys Asp Lys Asn Tyr  Arg His Asp Ala Val  Lys Ile Leu
    2450                 2455                 2460

Lys Leu  Lys Pro Val Leu Gln  Arg Ser Leu Thr His  Pro Pro Thr
    2465                 2470                 2475

Ile Ala  Pro Glu Gly Ser Arg  Pro Gln Gly Gly Leu  Pro Ser Ser
    2480                 2485                 2490

Lys Leu  Asp Ser Gly Phe Ala  Lys Thr Ser Val Ala  Ala Ser Leu
    2495                 2500                 2505
```

```
Tyr His  Thr Pro Ser Asp Ser  Lys Glu Ile Thr Leu  Thr Val Thr
    2510                2515                2520


Ser Lys  Asp Pro Glu Arg Pro  Pro Val His Asp Gln  Leu Gln Asp
    2525                2530                2535


Pro Arg  Leu Leu Lys Arg Met  Gly Ile Glu Val Lys  Gly Gly Ile
    2540                2545                2550


Phe Leu  Trp Asp Pro Gln Pro  Ser Ser Pro Gln His  Pro Gly Ala
    2555                2560                2565


Thr Pro  Pro Thr Gly Glu Pro  Gly Phe Pro Val Val  His Gln Asp
    2570                2575                2580


Leu Ser  His Ile Gln Gln Pro  Ala Ala Val Val Ala  Ala Leu Ser
    2585                2590                2595


Ser His  Lys Pro Pro Val Arg  Gly Glu Pro Pro Ala  Ala Ser Pro
    2600                2605                2610


Glu Ala  Ser Thr Cys Gln Ser  Lys Cys Asp Asp Pro  Glu Glu Glu
    2615                2620                2625


Leu Cys  His Arg Arg Glu Ala  Arg Ala Phe Ser Glu  Gly Glu Gln
    2630                2635                2640


Glu Lys  Cys Gly Ser Glu Thr  His His Thr Arg Arg  Asn Ser Arg
    2645                2650                2655


Trp Asp  Lys Arg Thr Leu Glu  Gln Glu Asp Ser Ser  Ser Lys Lys
    2660                2665                2670


Arg Lys  Leu Leu
    2675


<210>  136
<211>  632
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  FMR1

<400>  136

Met Glu Glu Leu Val Val Glu Val Arg Gly Ser Asn Gly Ala Phe Tyr
1               5                   10                  15


Lys Ala Phe Val Lys Asp Val His Glu Asp Ser Ile Thr Val Ala Phe
            20                  25                  30
```

```
Glu Asn Asn Trp Gln Pro Asp Arg Gln Ile Pro Phe His Asp Val Arg
        35                  40                  45

Phe Pro Pro Pro Val Gly Tyr Asn Lys Asp Ile Asn Glu Ser Asp Glu
    50                  55                  60

Val Glu Val Tyr Ser Arg Ala Asn Glu Lys Glu Pro Cys Cys Trp Trp
65                  70                  75                  80

Leu Ala Lys Val Arg Met Ile Lys Gly Glu Phe Tyr Val Ile Glu Tyr
                85                  90                  95

Ala Ala Cys Asp Ala Thr Tyr Asn Glu Ile Val Thr Ile Glu Arg Leu
            100                 105                 110

Arg Ser Val Asn Pro Asn Lys Pro Ala Thr Lys Asp Thr Phe His Lys
        115                 120                 125

Ile Lys Leu Asp Val Pro Glu Asp Leu Arg Gln Met Cys Ala Lys Glu
    130                 135                 140

Ala Ala His Lys Asp Phe Lys Lys Ala Val Gly Ala Phe Ser Val Thr
145                 150                 155                 160

Tyr Asp Pro Glu Asn Tyr Gln Leu Val Ile Leu Ser Ile Asn Glu Val
                165                 170                 175

Thr Ser Lys Arg Ala His Met Leu Ile Asp Met His Phe Arg Ser Leu
            180                 185                 190

Arg Thr Lys Leu Ser Leu Ile Met Arg Asn Glu Glu Ala Ser Lys Gln
        195                 200                 205

Leu Glu Ser Ser Arg Gln Leu Ala Ser Arg Phe His Glu Gln Phe Ile
    210                 215                 220

Val Arg Glu Asp Leu Met Gly Leu Ala Ile Gly Thr His Gly Ala Asn
225                 230                 235                 240

Ile Gln Gln Ala Arg Lys Val Pro Gly Val Thr Ala Ile Asp Leu Asp
                245                 250                 255

Glu Asp Thr Cys Thr Phe His Ile Tyr Gly Glu Asp Gln Asp Ala Val
            260                 265                 270

Lys Lys Ala Arg Ser Phe Leu Glu Phe Ala Glu Asp Val Ile Gln Val
        275                 280                 285

Pro Arg Asn Leu Val Gly Lys Val Ile Gly Lys Asn Gly Lys Leu Ile
    290                 295                 300

Gln Glu Ile Val Asp Lys Ser Gly Val Val Arg Val Arg Ile Glu Ala
```

305 310 315 320

Glu Asn Glu Lys Asn Val Pro Gln Glu Glu Glu Ile Met Pro Pro Asn
325 330 335

Ser Leu Pro Ser Asn Asn Ser Arg Val Gly Pro Asn Ala Pro Glu Glu
340 345 350

Lys Lys His Leu Asp Ile Lys Glu Asn Ser Thr His Phe Ser Gln Pro
355 360 365

Asn Ser Thr Lys Val Gln Arg Val Leu Val Ala Ser Ser Val Val Ala
370 375 380

Gly Glu Ser Gln Lys Pro Glu Leu Lys Ala Trp Gln Gly Met Val Pro
385 390 395 400

Phe Val Phe Val Gly Thr Lys Asp Ser Ile Ala Asn Ala Thr Val Leu
405 410 415

Leu Asp Tyr His Leu Asn Tyr Leu Lys Glu Val Asp Gln Leu Arg Leu
420 425 430

Glu Arg Leu Gln Ile Asp Glu Gln Leu Arg Gln Ile Gly Ala Ser Ser
435 440 445

Arg Pro Pro Pro Asn Arg Thr Asp Lys Glu Lys Ser Tyr Val Thr Asp
450 455 460

Asp Gly Gln Gly Met Gly Arg Gly Ser Arg Pro Tyr Arg Asn Arg Gly
465 470 475 480

His Gly Arg Arg Gly Pro Gly Tyr Thr Ser Gly Thr Asn Ser Glu Ala
485 490 495

Ser Asn Ala Ser Glu Thr Glu Ser Asp His Arg Asp Glu Leu Ser Asp
500 505 510

Trp Ser Leu Ala Pro Thr Glu Glu Glu Arg Glu Ser Phe Leu Arg Arg
515 520 525

Gly Asp Gly Arg Arg Arg Gly Gly Gly Gly Arg Gly Gln Gly Gly Arg
530 535 540

Gly Arg Gly Gly Gly Phe Lys Gly Asn Asp Asp His Ser Arg Thr Asp
545 550 555 560

Asn Arg Pro Arg Asn Pro Arg Glu Ala Lys Gly Arg Thr Thr Asp Gly
565 570 575

Ser Leu Gln Ile Arg Val Asp Cys Asn Asn Glu Arg Ser Val His Thr
580 585 590

```
Lys Thr Leu Gln Asn Thr Ser Ser Glu Gly Ser Arg Leu Arg Thr Gly
        595                 600                 605


Lys Asp Arg Asn Gln Lys Lys Glu Lys Pro Asp Ser Val Asp Gly Gln
    610                 615                 620


Gln Pro Leu Val Asn Gly Val Pro
625                 630


<210>  137
<211>  635
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  SLC6A8

<400>  137

Met Ala Lys Lys Ser Ala Glu Asn Gly Ile Tyr Ser Val Ser Gly Asp
1               5                   10                  15


Glu Lys Lys Gly Pro Leu Ile Ala Pro Gly Pro Asp Gly Ala Pro Ala
            20                  25                  30


Lys Gly Asp Gly Pro Val Gly Leu Gly Thr Pro Gly Gly Arg Leu Ala
        35                  40                  45


Val Pro Pro Arg Glu Thr Trp Thr Arg Gln Met Asp Phe Ile Met Ser
    50                  55                  60


Cys Val Gly Phe Ala Val Gly Leu Gly Asn Val Trp Arg Phe Pro Tyr
65                  70                  75                  80


Leu Cys Tyr Lys Asn Gly Gly Gly Val Phe Leu Ile Pro Tyr Val Leu
                85                  90                  95


Ile Ala Leu Val Gly Gly Ile Pro Ile Phe Phe Leu Glu Ile Ser Leu
            100                 105                 110


Gly Gln Phe Met Lys Ala Gly Ser Ile Asn Val Trp Asn Ile Cys Pro
        115                 120                 125


Leu Phe Lys Gly Leu Gly Tyr Ala Ser Met Val Ile Val Phe Tyr Cys
    130                 135                 140


Asn Thr Tyr Tyr Ile Met Val Leu Ala Trp Gly Phe Tyr Tyr Leu Val
145                 150                 155                 160


Lys Ser Phe Thr Thr Thr Leu Pro Trp Ala Thr Cys Gly His Thr Trp
                165                 170                 175
```

```
Asn Thr Pro Asp Cys Val Glu Ile Phe Arg His Glu Asp Cys Ala Asn
            180                 185                 190

Ala Ser Leu Ala Asn Leu Thr Cys Asp Gln Leu Ala Asp Arg Arg Ser
        195                 200                 205

Pro Val Ile Glu Phe Trp Glu Asn Lys Val Leu Arg Leu Ser Gly Gly
    210                 215                 220

Leu Glu Val Pro Gly Ala Leu Asn Trp Glu Val Thr Leu Cys Leu Leu
225                 230                 235                 240

Ala Cys Trp Val Leu Val Tyr Phe Cys Val Trp Lys Gly Val Lys Ser
                245                 250                 255

Thr Gly Lys Ile Val Tyr Phe Thr Ala Thr Phe Pro Tyr Val Val Leu
            260                 265                 270

Val Val Leu Leu Val Arg Gly Val Leu Leu Pro Gly Ala Leu Asp Gly
        275                 280                 285

Ile Ile Tyr Tyr Leu Lys Pro Asp Trp Ser Lys Leu Gly Ser Pro Gln
    290                 295                 300

Val Trp Ile Asp Ala Gly Thr Gln Ile Phe Phe Ser Tyr Ala Ile Gly
305                 310                 315                 320

Leu Gly Ala Leu Thr Ala Leu Gly Ser Tyr Asn Arg Phe Asn Asn Asn
                325                 330                 335

Cys Tyr Lys Asp Ala Ile Ile Leu Ala Leu Ile Asn Ser Gly Thr Ser
            340                 345                 350

Phe Phe Ala Gly Phe Val Val Phe Ser Ile Leu Gly Phe Met Ala Ala
        355                 360                 365

Glu Gln Gly Val His Ile Ser Lys Val Ala Glu Ser Gly Pro Gly Leu
    370                 375                 380

Ala Phe Ile Ala Tyr Pro Arg Ala Val Thr Leu Met Pro Val Ala Pro
385                 390                 395                 400

Leu Trp Ala Ala Leu Phe Phe Phe Met Leu Leu Leu Leu Gly Leu Asp
                405                 410                 415

Ser Gln Phe Val Gly Val Glu Gly Phe Ile Thr Gly Leu Leu Asp Leu
            420                 425                 430

Leu Pro Ala Ser Tyr Tyr Phe Arg Phe Gln Arg Glu Ile Ser Val Ala
        435                 440                 445
```

Leu Cys Cys Ala Leu Cys Phe Val Ile Asp Leu Ser Met Val Thr Asp
450 455 460

Gly Gly Met Tyr Val Phe Gln Leu Phe Asp Tyr Tyr Ser Ala Ser Gly
465 470 475 480

Thr Thr Leu Leu Trp Gln Ala Phe Trp Glu Cys Val Val Val Ala Trp
485 490 495

Val Tyr Gly Ala Asp Arg Phe Met Asp Asp Ile Ala Cys Met Ile Gly
500 505 510

Tyr Arg Pro Cys Pro Trp Met Lys Trp Cys Trp Ser Phe Phe Thr Pro
515 520 525

Leu Val Cys Met Gly Ile Phe Ile Phe Asn Val Val Tyr Tyr Glu Pro
530 535 540

Leu Val Tyr Asn Asn Thr Tyr Val Tyr Pro Trp Trp Gly Glu Ala Met
545 550 555 560

Gly Trp Ala Phe Ala Leu Ser Ser Met Leu Cys Val Pro Leu His Leu
565 570 575

Leu Gly Cys Leu Leu Arg Ala Lys Gly Thr Met Ala Glu Arg Trp Gln
580 585 590

His Leu Thr Gln Pro Ile Trp Gly Leu His His Leu Glu Tyr Arg Ala
595 600 605

Gln Asp Ala Asp Val Arg Gly Leu Thr Thr Leu Thr Pro Val Ser Glu
610 615 620

Ser Ser Lys Val Val Val Val Glu Ser Val Met
625 630 635

<210> 138
<211> 875
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> UBE3A

<400> 138

Met Glu Lys Leu His Gln Cys Tyr Trp Lys Ser Gly Glu Pro Gln Ser
1 5 10 15

Asp Asp Ile Glu Ala Ser Arg Met Lys Arg Ala Ala Ala Lys His Leu
20 25 30

Ile Glu Arg Tyr Tyr His Gln Leu Thr Glu Gly Cys Gly Asn Glu Ala

35 40 45

Cys Thr Asn Glu Phe Cys Ala Ser Cys Pro Thr Phe Leu Arg Met Asp
50 55 60

Asn Asn Ala Ala Ala Ile Lys Ala Leu Glu Leu Tyr Lys Ile Asn Ala
65 70 75 80

Lys Leu Cys Asp Pro His Pro Ser Lys Lys Gly Ala Ser Ser Ala Tyr
85 90 95

Leu Glu Asn Ser Lys Gly Ala Pro Asn Asn Ser Cys Ser Glu Ile Lys
100 105 110

Met Asn Lys Lys Gly Ala Arg Ile Asp Phe Lys Asp Val Thr Tyr Leu
115 120 125

Thr Glu Glu Lys Val Tyr Glu Ile Leu Glu Leu Cys Arg Glu Arg Glu
130 135 140

Asp Tyr Ser Pro Leu Ile Arg Val Ile Gly Arg Val Phe Ser Ser Ala
145 150 155 160

Glu Ala Leu Val Gln Ser Phe Arg Lys Val Lys Gln His Thr Lys Glu
165 170 175

Glu Leu Lys Ser Leu Gln Ala Lys Asp Glu Asp Lys Asp Glu Asp Glu
180 185 190

Lys Glu Lys Ala Ala Cys Ser Ala Ala Ala Met Glu Glu Asp Ser Glu
195 200 205

Ala Ser Ser Ser Arg Ile Gly Asp Ser Ser Gln Gly Asp Asn Asn Leu
210 215 220

Gln Lys Leu Gly Pro Asp Asp Val Ser Val Asp Ile Asp Ala Ile Arg
225 230 235 240

Arg Val Tyr Thr Arg Leu Leu Ser Asn Glu Lys Ile Glu Thr Ala Phe
245 250 255

Leu Asn Ala Leu Val Tyr Leu Ser Pro Asn Val Glu Cys Asp Leu Thr
260 265 270

Tyr His Asn Val Tyr Ser Arg Asp Pro Asn Tyr Leu Asn Leu Phe Ile
275 280 285

Ile Val Met Glu Asn Arg Asn Leu His Ser Pro Glu Tyr Leu Glu Met
290 295 300

Ala Leu Pro Leu Phe Cys Lys Ala Met Ser Lys Leu Pro Leu Ala Ala
305 310 315 320

```
Gln Gly Lys Leu Ile Arg Leu Trp Ser Lys Tyr Asn Ala Asp Gln Ile
                325                 330                 335

Arg Arg Met Met Glu Thr Phe Gln Gln Leu Ile Thr Tyr Lys Val Ile
            340                 345                 350

Ser Asn Glu Phe Asn Ser Arg Asn Leu Val Asn Asp Asp Asp Ala Ile
        355                 360                 365

Val Ala Ala Ser Lys Cys Leu Lys Met Val Tyr Tyr Ala Asn Val Val
    370                 375                 380

Gly Gly Glu Val Asp Thr Asn His Asn Glu Glu Asp Asp Glu Glu Pro
385                 390                 395                 400

Ile Pro Glu Ser Ser Glu Leu Thr Leu Gln Glu Leu Leu Gly Glu Glu
                405                 410                 415

Arg Arg Asn Lys Lys Gly Pro Arg Val Asp Pro Leu Glu Thr Glu Leu
            420                 425                 430

Gly Val Lys Thr Leu Asp Cys Arg Lys Pro Leu Ile Pro Phe Glu Glu
        435                 440                 445

Phe Ile Asn Glu Pro Leu Asn Glu Val Leu Glu Met Asp Lys Asp Tyr
    450                 455                 460

Thr Phe Phe Lys Val Glu Thr Glu Asn Lys Phe Ser Phe Met Thr Cys
465                 470                 475                 480

Pro Phe Ile Leu Asn Ala Val Thr Lys Asn Leu Gly Leu Tyr Tyr Asp
                485                 490                 495

Asn Arg Ile Arg Met Tyr Ser Glu Arg Arg Ile Thr Val Leu Tyr Ser
            500                 505                 510

Leu Val Gln Gly Gln Gln Leu Asn Pro Tyr Leu Arg Leu Lys Val Arg
        515                 520                 525

Arg Asp His Ile Ile Asp Asp Ala Leu Val Arg Leu Glu Met Ile Ala
    530                 535                 540

Met Glu Asn Pro Ala Asp Leu Lys Lys Gln Leu Tyr Val Glu Phe Glu
545                 550                 555                 560

Gly Glu Gln Gly Val Asp Glu Gly Gly Val Ser Lys Glu Phe Phe Gln
                565                 570                 575

Leu Val Val Glu Glu Ile Phe Asn Pro Asp Ile Gly Met Phe Thr Tyr
            580                 585                 590
```

```
Asp Glu Ser Thr Lys Leu Phe Trp Phe Asn Pro Ser Ser Phe Glu Thr
        595                 600                 605

Glu Gly Gln Phe Thr Leu Ile Gly Ile Val Leu Gly Leu Ala Ile Tyr
    610                 615                 620

Asn Asn Cys Ile Leu Asp Val His Phe Pro Met Val Val Tyr Arg Lys
625                 630                 635                 640

Leu Met Gly Lys Lys Gly Thr Phe Arg Asp Leu Gly Asp Ser His Pro
                645                 650                 655

Val Leu Tyr Gln Ser Leu Lys Asp Leu Leu Glu Tyr Glu Gly Asn Val
            660                 665                 670

Glu Asp Asp Met Met Ile Thr Phe Gln Ile Ser Gln Thr Asp Leu Phe
        675                 680                 685

Gly Asn Pro Met Met Tyr Asp Leu Lys Glu Asn Gly Asp Lys Ile Pro
    690                 695                 700

Ile Thr Asn Glu Asn Arg Lys Glu Phe Val Asn Leu Tyr Ser Asp Tyr
705                 710                 715                 720

Ile Leu Asn Lys Ser Val Glu Lys Gln Phe Lys Ala Phe Arg Arg Gly
                725                 730                 735

Phe His Met Val Thr Asn Glu Ser Pro Leu Lys Tyr Leu Phe Arg Pro
            740                 745                 750

Glu Glu Ile Glu Leu Leu Ile Cys Gly Ser Arg Asn Leu Asp Phe Gln
        755                 760                 765

Ala Leu Glu Glu Thr Thr Glu Tyr Asp Gly Gly Tyr Thr Arg Asp Ser
    770                 775                 780

Val Leu Ile Arg Glu Phe Trp Glu Ile Val His Ser Phe Thr Asp Glu
785                 790                 795                 800

Gln Lys Arg Leu Phe Leu Gln Phe Thr Thr Gly Thr Asp Arg Ala Pro
                805                 810                 815

Val Gly Gly Leu Gly Lys Leu Lys Met Ile Ile Ala Lys Asn Gly Pro
            820                 825                 830

Asp Thr Glu Arg Leu Pro Thr Ser His Thr Cys Phe Asn Val Leu Leu
        835                 840                 845

Leu Pro Glu Tyr Ser Ser Lys Glu Lys Leu Lys Glu Arg Leu Leu Lys
    850                 855                 860
```

Ala Ile Thr Tyr Ala Lys Gly Phe Gly Met Leu
865 870 875

<210> 139
<211> 154
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> SOD1

<400> 139

Met Ala Thr Lys Ala Val Cys Val Leu Lys Gly Asp Gly Pro Val Gln
1 5 10 15

Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
20 25 30

Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
35 40 45

His Glu Phe Gly Asp Asn Thr Ala Gly Cys Thr Ser Ala Gly Pro His
50 55 60

Phe Asn Pro Leu Ser Arg Lys His Gly Gly Pro Lys Asp Glu Glu Arg
65 70 75 80

His Val Gly Asp Leu Gly Asn Val Thr Ala Asp Lys Asp Gly Val Ala
85 90 95

Asp Val Ser Ile Glu Asp Ser Val Ile Ser Leu Ser Gly Asp His Cys
100 105 110

Ile Ile Gly Arg Thr Leu Val Val His Glu Lys Ala Asp Asp Leu Gly
115 120 125

Lys Gly Gly Asn Glu Glu Ser Thr Lys Thr Gly Asn Ala Gly Ser Arg
130 135 140

Leu Ala Cys Gly Val Ile Gly Ile Ala Gln
145 150

<210> 140
<211> 414
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> TDP43

<400> 140

Met Ser Glu Tyr Ile Arg Val Thr Glu Asp Glu Asn Asp Glu Pro Ile

1 5 10 15

Glu Ile Pro Ser Glu Asp Asp Gly Thr Val Leu Leu Ser Thr Val Thr
20 25 30

Ala Gln Phe Pro Gly Ala Cys Gly Leu Arg Tyr Arg Asn Pro Val Ser
35 40 45

Gln Cys Met Arg Gly Val Arg Leu Val Glu Gly Ile Leu His Ala Pro
50 55 60

Asp Ala Gly Trp Gly Asn Leu Val Tyr Val Val Asn Tyr Pro Lys Asp
65 70 75 80

Asn Lys Arg Lys Met Asp Glu Thr Asp Ala Ser Ser Ala Val Lys Val
85 90 95

Lys Arg Ala Val Gln Lys Thr Ser Asp Leu Ile Val Leu Gly Leu Pro
100 105 110

Trp Lys Thr Thr Glu Gln Asp Leu Lys Glu Tyr Phe Ser Thr Phe Gly
115 120 125

Glu Val Leu Met Val Gln Val Lys Lys Asp Leu Lys Thr Gly His Ser
130 135 140

Lys Gly Phe Gly Phe Val Arg Phe Thr Glu Tyr Glu Thr Gln Val Lys
145 150 155 160

Val Met Ser Gln Arg His Met Ile Asp Gly Arg Trp Cys Asp Cys Lys
165 170 175

Leu Pro Asn Ser Lys Gln Ser Gln Asp Glu Pro Leu Arg Ser Arg Lys
180 185 190

Val Phe Val Gly Arg Cys Thr Glu Asp Met Thr Glu Asp Glu Leu Arg
195 200 205

Glu Phe Phe Ser Gln Tyr Gly Asp Val Met Asp Val Phe Ile Pro Lys
210 215 220

Pro Phe Arg Ala Phe Ala Phe Val Thr Phe Ala Asp Asp Gln Ile Ala
225 230 235 240

Gln Ser Leu Cys Gly Glu Asp Leu Ile Ile Lys Gly Ile Ser Val His
245 250 255

Ile Ser Asn Ala Glu Pro Lys His Asn Ser Asn Arg Gln Leu Glu Arg
260 265 270

Ser Gly Arg Phe Gly Gly Asn Pro Gly Gly Phe Gly Asn Gln Gly Gly
275 280 285

```
Phe Gly Asn Ser Arg Gly Gly Gly Ala Gly Leu Gly Asn Asn Gln Gly
    290                 295                 300

Ser Asn Met Gly Gly Gly Met Asn Phe Gly Ala Phe Ser Ile Asn Pro
305                 310                 315                 320

Ala Met Met Ala Ala Ala Gln Ala Ala Leu Gln Ser Ser Trp Gly Met
                325                 330                 335

Met Gly Met Leu Ala Ser Gln Gln Asn Gln Ser Gly Pro Ser Gly Asn
            340                 345                 350

Asn Gln Asn Gln Gly Asn Met Gln Arg Glu Pro Asn Gln Ala Phe Gly
        355                 360                 365

Ser Gly Asn Asn Ser Tyr Ser Gly Ser Asn Ser Gly Ala Ala Ile Gly
    370                 375                 380

Trp Gly Ser Ala Ser Asn Ala Gly Ser Gly Ser Gly Phe Asn Gly Gly
385                 390                 395                 400

Phe Gly Ser Ser Met Asp Ser Lys Ser Ser Gly Trp Gly Met
                405                 410


<210>  141
<211>  481
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  C9orf72

<400>  141

Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala Val Ala Lys Thr Glu
1               5                   10                  15

Ile Ala Leu Ser Gly Lys Ser Pro Leu Leu Ala Ala Thr Phe Ala Tyr
            20                  25                  30

Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro Lys
        35                  40                  45

Thr Glu Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala Asn
    50                  55                  60

His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala Ile
65                  70                  75                  80

Asp Val Lys Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile Val Ser
                85                  90                  95
```

```
Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly Leu
            100                 105                 110

Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu His
        115                 120                 125

Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile
    130                 135                 140

Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Ile Leu Glu
145                 150                 155                 160

Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu
                165                 170                 175

Thr Gly Glu Val Ile Pro Val Met Glu Leu Leu Ser Ser Met Lys Ser
            180                 185                 190

His Ser Val Pro Glu Glu Ile Asp Ile Ala Asp Thr Val Leu Asn Asp
        195                 200                 205

Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile
    210                 215                 220

Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser
225                 230                 235                 240

Ala Glu Lys Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe Leu Thr
                245                 250                 255

Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser Ser Phe
            260                 265                 270

Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp Ser Thr
        275                 280                 285

Gly Ser Phe Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro Tyr Pro
    290                 295                 300

Thr Thr His Ile Asp Val Asp Val Asn Thr Val Lys Gln Met Pro Pro
305                 310                 315                 320

Cys His Glu His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser Glu Leu
                325                 330                 335

Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln Asp Thr
            340                 345                 350

Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile Phe Gln
        355                 360                 365
```

Asp Val Leu His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp Gln Val
370 375 380

Phe Gln Leu Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu Ala Gln
385 390 395 400

Phe Leu Leu Val Leu His Arg Lys Ala Leu Thr Leu Ile Lys Tyr Ile
405 410 415

Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu Arg Asn
420 425 430

Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile
435 440 445

Met Ala Leu Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe Ile Phe
450 455 460

Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr
465 470 475 480

Phe

<210> 142
<211> 210
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> FXN

<400> 142

Met Trp Thr Leu Gly Arg Arg Ala Val Ala Gly Leu Leu Ala Ser Pro
1 5 10 15

Ser Pro Ala Gln Ala Gln Thr Leu Thr Arg Val Pro Arg Pro Ala Glu
20 25 30

Leu Ala Pro Leu Cys Gly Arg Arg Gly Leu Arg Thr Asp Ile Asp Ala
35 40 45

Thr Cys Thr Pro Arg Arg Ala Ser Ser Asn Gln Arg Gly Leu Asn Gln
50 55 60

Ile Trp Asn Val Lys Lys Gln Ser Val Tyr Leu Met Asn Leu Arg Lys
65 70 75 80

Ser Gly Thr Leu Gly His Pro Gly Ser Leu Asp Glu Thr Thr Tyr Glu
85 90 95

Arg Leu Ala Glu Glu Thr Leu Asp Ser Leu Ala Glu Phe Phe Glu Asp

100 105 110

Leu Ala Asp Lys Pro Tyr Thr Phe Glu Asp Tyr Asp Val Ser Phe Gly
115 120 125

Ser Gly Val Leu Thr Val Lys Leu Gly Gly Asp Leu Gly Thr Tyr Val
130 135 140

Ile Asn Lys Gln Thr Pro Asn Lys Gln Ile Trp Leu Ser Ser Pro Ser
145 150 155 160

Ser Gly Pro Lys Arg Tyr Asp Trp Thr Gly Lys Asn Trp Val Tyr Ser
165 170 175

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

Ala Leu Lys Thr Lys Leu Asp Leu Ser Ser Leu Ala Tyr Ser Gly Lys
195 200 205

Asp Ala
210

<210> 143
<211> 486
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> MECP2

<400> 143

Met Val Ala Gly Met Leu Gly Leu Arg Glu Glu Lys Ser Glu Asp Gln
1 5 10 15

Asp Leu Gln Gly Leu Lys Asp Lys Pro Leu Lys Phe Lys Lys Val Lys
20 25 30

Lys Asp Lys Lys Glu Glu Lys Glu Gly Lys His Glu Pro Val Gln Pro
35 40 45

Ser Ala His His Ser Ala Glu Pro Ala Glu Ala Gly Lys Ala Glu Thr
50 55 60

Ser Glu Gly Ser Gly Ser Ala Pro Ala Val Pro Glu Ala Ser Ala Ser
65 70 75 80

Pro Lys Gln Arg Arg Ser Ile Ile Arg Asp Arg Gly Pro Met Tyr Asp
85 90 95

Asp Pro Thr Leu Pro Glu Gly Trp Thr Arg Lys Leu Lys Gln Arg Lys
100 105 110

```
Ser Gly Arg Ser Ala Gly Lys Tyr Asp Val Tyr Leu Ile Asn Pro Gln
        115                 120                 125

Gly Lys Ala Phe Arg Ser Lys Val Glu Leu Ile Ala Tyr Phe Glu Lys
    130                 135                 140

Val Gly Asp Thr Ser Leu Asp Pro Asn Asp Phe Asp Phe Thr Val Thr
145                 150                 155                 160

Gly Arg Gly Ser Pro Ser Arg Arg Glu Gln Lys Pro Pro Lys Lys Pro
                165                 170                 175

Lys Ser Pro Lys Ala Pro Gly Thr Gly Arg Gly Arg Gly Arg Pro Lys
            180                 185                 190

Gly Ser Gly Thr Thr Arg Pro Lys Ala Ala Thr Ser Glu Gly Val Gln
        195                 200                 205

Val Lys Arg Val Leu Glu Lys Ser Pro Gly Lys Leu Leu Val Lys Met
    210                 215                 220

Pro Phe Gln Thr Ser Pro Gly Gly Lys Ala Glu Gly Gly Gly Ala Thr
225                 230                 235                 240

Thr Ser Thr Gln Val Met Val Ile Lys Arg Pro Gly Arg Lys Arg Lys
                245                 250                 255

Ala Glu Ala Asp Pro Gln Ala Ile Pro Lys Lys Arg Gly Arg Lys Pro
            260                 265                 270

Gly Ser Val Val Ala Ala Ala Ala Ala Glu Ala Lys Lys Lys Ala Val
        275                 280                 285

Lys Glu Ser Ser Ile Arg Ser Val Gln Glu Thr Val Leu Pro Ile Lys
    290                 295                 300

Lys Arg Lys Thr Arg Glu Thr Val Ser Ile Glu Val Lys Glu Val Val
305                 310                 315                 320

Lys Pro Leu Leu Val Ser Thr Leu Gly Glu Lys Ser Gly Lys Gly Leu
                325                 330                 335

Lys Thr Cys Lys Ser Pro Gly Arg Lys Ser Lys Glu Ser Ser Pro Lys
            340                 345                 350

Gly Arg Ser Ser Ser Ala Ser Ser Pro Pro Lys Lys Glu His His His
        355                 360                 365

His His His His Ser Glu Ser Pro Lys Ala Pro Val Pro Leu Leu Pro
    370                 375                 380
```

```
Pro Leu Pro Pro Pro Pro Pro Glu Pro Glu Ser Ser Glu Asp Pro Thr
385                 390                 395                 400

Ser Pro Pro Glu Pro Gln Asp Leu Ser Ser Ser Val Cys Lys Glu Glu
                405                 410                 415

Lys Met Pro Arg Gly Gly Ser Leu Glu Ser Asp Gly Cys Pro Lys Glu
            420                 425                 430

Pro Ala Lys Thr Gln Pro Ala Val Ala Thr Ala Ala Thr Ala Ala Glu
        435                 440                 445

Lys Tyr Lys His Arg Gly Glu Gly Glu Arg Lys Asp Ile Val Ser Ser
    450                 455                 460

Ser Met Pro Arg Pro Asn Arg Glu Glu Pro Val Asp Ser Arg Thr Pro
465                 470                 475                 480

Val Thr Glu Arg Val Ser
                485
```

<210> 144
<211> 313
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> ASPA

<400> 144

```
Met Thr Ser Cys His Ile Ala Glu Glu His Ile Gln Lys Val Ala Ile
1               5                   10                  15

Phe Gly Gly Thr His Gly Asn Glu Leu Thr Gly Val Phe Leu Val Lys
            20                  25                  30

His Trp Leu Glu Asn Gly Ala Glu Ile Gln Arg Thr Gly Leu Glu Val
        35                  40                  45

Lys Pro Phe Ile Thr Asn Pro Arg Ala Val Lys Lys Cys Thr Arg Tyr
    50                  55                  60

Ile Asp Cys Asp Leu Asn Arg Ile Phe Asp Leu Glu Asn Leu Gly Lys
65                  70                  75                  80

Lys Met Ser Glu Asp Leu Pro Tyr Glu Val Arg Arg Ala Gln Glu Ile
                85                  90                  95

Asn His Leu Phe Gly Pro Lys Asp Ser Glu Asp Ser Tyr Asp Ile Ile
            100                 105                 110
```

Phe Asp Leu His Asn Thr Thr Ser Asn Met Gly Cys Thr Leu Ile Leu
115 120 125

Glu Asp Ser Arg Asn Asn Phe Leu Ile Gln Met Phe His Tyr Ile Lys
130 135 140

Thr Ser Leu Ala Pro Leu Pro Cys Tyr Val Tyr Leu Ile Glu His Pro
145 150 155 160

Ser Leu Lys Tyr Ala Thr Thr Arg Ser Ile Ala Lys Tyr Pro Val Gly
165 170 175

Ile Glu Val Gly Pro Gln Pro Gln Gly Val Leu Arg Ala Asp Ile Leu
180 185 190

Asp Gln Met Arg Lys Met Ile Lys His Ala Leu Asp Phe Ile His His
195 200 205

Phe Asn Glu Gly Lys Glu Phe Pro Pro Cys Ala Ile Glu Val Tyr Lys
210 215 220

Ile Ile Glu Lys Val Asp Tyr Pro Arg Asp Glu Asn Gly Glu Ile Ala
225 230 235 240

Ala Ile Ile His Pro Asn Leu Gln Asp Gln Asp Trp Lys Pro Leu His
245 250 255

Pro Gly Asp Pro Met Phe Leu Thr Leu Asp Gly Lys Thr Ile Pro Leu
260 265 270

Gly Gly Asp Cys Thr Val Tyr Pro Val Phe Val Asn Glu Ala Ala Tyr
275 280 285

Tyr Glu Lys Lys Glu Ala Phe Ala Lys Thr Thr Lys Leu Thr Leu Asn
290 295 300

Ala Lys Ser Ile Arg Cys Cys Leu His
305 310

<210> 145
<211> 539
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> ALDH7A1

<400> 145

Met Trp Arg Leu Pro Arg Ala Leu Cys Val His Ala Ala Lys Thr Ser
1 5 10 15

Lys Leu Ser Gly Pro Trp Ser Arg Pro Ala Ala Phe Met Ser Thr Leu

20 25 30

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

Glu Glu Asn Glu Gly Val Tyr Asn Gly Ser Trp Gly Gly Arg Gly Glu
50 55 60

Val Ile Thr Thr Tyr Cys Pro Ala Asn Asn Glu Pro Ile Ala Arg Val
65 70 75 80

Arg Gln Ala Ser Val Ala Asp Tyr Glu Glu Thr Val Lys Lys Ala Arg
85 90 95

Glu Ala Trp Lys Ile Trp Ala Asp Ile Pro Ala Pro Lys Arg Gly Glu
100 105 110

Ile Val Arg Gln Ile Gly Asp Ala Leu Arg Glu Lys Ile Gln Val Leu
115 120 125

Gly Ser Leu Val Ser Leu Glu Met Gly Lys Ile Leu Val Glu Gly Val
130 135 140

Gly Glu Val Gln Glu Tyr Val Asp Ile Cys Asp Tyr Ala Val Gly Leu
145 150 155 160

Ser Arg Met Ile Gly Gly Pro Ile Leu Pro Ser Glu Arg Ser Gly His
165 170 175

Ala Leu Ile Glu Gln Trp Asn Pro Val Gly Leu Val Gly Ile Ile Thr
180 185 190

Ala Phe Asn Phe Pro Val Ala Val Tyr Gly Trp Asn Asn Ala Ile Ala
195 200 205

Met Ile Cys Gly Asn Val Cys Leu Trp Lys Gly Ala Pro Thr Thr Ser
210 215 220

Leu Ile Ser Val Ala Val Thr Lys Ile Ile Ala Lys Val Leu Glu Asp
225 230 235 240

Asn Lys Leu Pro Gly Ala Ile Cys Ser Leu Thr Cys Gly Gly Ala Asp
245 250 255

Ile Gly Thr Ala Met Ala Lys Asp Glu Arg Val Asn Leu Leu Ser Phe
260 265 270

Thr Gly Ser Thr Gln Val Gly Lys Gln Val Gly Leu Met Val Gln Glu
275 280 285

Arg Phe Gly Arg Ser Leu Leu Glu Leu Gly Gly Asn Asn Ala Ile Ile
290 295 300

```
Ala Phe Glu Asp Ala Asp Leu Ser Leu Val Val Pro Ser Ala Leu Phe
305                 310                 315                 320

Ala Ala Val Gly Thr Ala Gly Gln Arg Cys Thr Thr Ala Arg Arg Leu
                325                 330                 335

Phe Ile His Glu Ser Ile His Asp Glu Val Val Asn Arg Leu Lys Lys
            340                 345                 350

Ala Tyr Ala Gln Ile Arg Val Gly Asn Pro Trp Asp Pro Asn Val Leu
        355                 360                 365

Tyr Gly Pro Leu His Thr Lys Gln Ala Val Ser Met Phe Leu Gly Ala
    370                 375                 380

Val Glu Glu Ala Lys Lys Glu Gly Gly Thr Val Val Tyr Gly Gly Lys
385                 390                 395                 400

Val Met Asp Arg Pro Gly Asn Tyr Val Glu Pro Thr Ile Val Thr Gly
                405                 410                 415

Leu Gly His Asp Ala Ser Ile Ala His Thr Glu Thr Phe Ala Pro Ile
            420                 425                 430

Leu Tyr Val Phe Lys Phe Lys Asn Glu Glu Glu Val Phe Ala Trp Asn
        435                 440                 445

Asn Glu Val Lys Gln Gly Leu Ser Ser Ser Ile Phe Thr Lys Asp Leu
    450                 455                 460

Gly Arg Ile Phe Arg Trp Leu Gly Pro Lys Gly Ser Asp Cys Gly Ile
465                 470                 475                 480

Val Asn Val Asn Ile Pro Thr Ser Gly Ala Glu Ile Gly Gly Ala Phe
                485                 490                 495

Gly Gly Glu Lys His Thr Gly Gly Gly Arg Glu Ser Gly Ser Asp Ala
            500                 505                 510

Trp Lys Gln Tyr Met Arg Arg Ser Thr Cys Thr Ile Asn Tyr Ser Lys
        515                 520                 525

Asp Leu Pro Leu Ala Gln Gly Ile Lys Phe Gln
    530                 535


<210>  146
<211>  563
<212>  PRT
<213>  homo sapiens


<220>
```

<221> MISC_FEATURE
<223> TPP1

<400> 146

```
Met Gly Leu Gln Ala Cys Leu Leu Gly Leu Phe Ala Leu Ile Leu Ser
1               5                  10                  15

Gly Lys Cys Ser Tyr Ser Pro Glu Pro Asp Gln Arg Arg Thr Leu Pro
            20                  25                  30

Pro Gly Trp Val Ser Leu Gly Arg Ala Asp Pro Glu Glu Glu Leu Ser
        35                  40                  45

Leu Thr Phe Ala Leu Arg Gln Gln Asn Val Glu Arg Leu Ser Glu Leu
    50                  55                  60

Val Gln Ala Val Ser Asp Pro Ser Ser Pro Gln Tyr Gly Lys Tyr Leu
65                  70                  75                  80

Thr Leu Glu Asn Val Ala Asp Leu Val Arg Pro Ser Pro Leu Thr Leu
                85                  90                  95

His Thr Val Gln Lys Trp Leu Leu Ala Ala Gly Ala Gln Lys Cys His
            100                 105                 110

Ser Val Ile Thr Gln Asp Phe Leu Thr Cys Trp Leu Ser Ile Arg Gln
        115                 120                 125

Ala Glu Leu Leu Leu Pro Gly Ala Glu Phe His His Tyr Val Gly Gly
    130                 135                 140

Pro Thr Glu Thr His Val Val Arg Ser Pro His Pro Tyr Gln Leu Pro
145                 150                 155                 160

Gln Ala Leu Ala Pro His Val Asp Phe Val Gly Gly Leu His Arg Phe
                165                 170                 175

Pro Pro Thr Ser Ser Leu Arg Gln Arg Pro Glu Pro Gln Val Thr Gly
            180                 185                 190

Thr Val Gly Leu His Leu Gly Val Thr Pro Ser Val Ile Arg Lys Arg
        195                 200                 205

Tyr Asn Leu Thr Ser Gln Asp Val Gly Ser Gly Thr Ser Asn Asn Ser
    210                 215                 220

Gln Ala Cys Ala Gln Phe Leu Glu Gln Tyr Phe His Asp Ser Asp Leu
225                 230                 235                 240

Ala Gln Phe Met Arg Leu Phe Gly Gly Asn Phe Ala His Gln Ala Ser
                245                 250                 255
```

```
Val Ala Arg Val Val Gly Gln Gln Gly Arg Gly Arg Ala Gly Ile Glu
            260                 265                 270

Ala Ser Leu Asp Val Gln Tyr Leu Met Ser Ala Gly Ala Asn Ile Ser
        275                 280                 285

Thr Trp Val Tyr Ser Ser Pro Gly Arg His Glu Gly Gln Glu Pro Phe
    290                 295                 300

Leu Gln Trp Leu Met Leu Leu Ser Asn Glu Ser Ala Leu Pro His Val
305                 310                 315                 320

His Thr Val Ser Tyr Gly Asp Asp Glu Asp Ser Leu Ser Ser Ala Tyr
                325                 330                 335

Ile Gln Arg Val Asn Thr Glu Leu Met Lys Ala Ala Ala Arg Gly Leu
            340                 345                 350

Thr Leu Leu Phe Ala Ser Gly Asp Ser Gly Ala Gly Cys Trp Ser Val
        355                 360                 365

Ser Gly Arg His Gln Phe Arg Pro Thr Phe Pro Ala Ser Ser Pro Tyr
    370                 375                 380

Val Thr Thr Val Gly Gly Thr Ser Phe Gln Glu Pro Phe Leu Ile Thr
385                 390                 395                 400

Asn Glu Ile Val Asp Tyr Ile Ser Gly Gly Gly Phe Ser Asn Val Phe
                405                 410                 415

Pro Arg Pro Ser Tyr Gln Glu Glu Ala Val Thr Lys Phe Leu Ser Ser
            420                 425                 430

Ser Pro His Leu Pro Pro Ser Ser Tyr Phe Asn Ala Ser Gly Arg Ala
        435                 440                 445

Tyr Pro Asp Val Ala Ala Leu Ser Asp Gly Tyr Trp Val Val Ser Asn
    450                 455                 460

Arg Val Pro Ile Pro Trp Val Ser Gly Thr Ser Ala Ser Thr Pro Val
465                 470                 475                 480

Phe Gly Gly Ile Leu Ser Leu Ile Asn Glu His Arg Ile Leu Ser Gly
                485                 490                 495

Arg Pro Pro Leu Gly Phe Leu Asn Pro Arg Leu Tyr Gln Gln His Gly
            500                 505                 510

Ala Gly Leu Phe Asp Val Thr Arg Gly Cys His Glu Ser Cys Leu Asp
        515                 520                 525

Glu Glu Val Glu Gly Gln Gly Phe Cys Ser Gly Pro Gly Trp Asp Pro
```

530 535 540

Val Thr Gly Trp Gly Thr Pro Asn Phe Pro Ala Leu Leu Lys Thr Leu
545 550 555 560

Leu Asn Pro

<210> 147
<211> 466
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> FUCA1

<400> 147

Met Arg Ala Pro Gly Met Arg Ser Arg Pro Ala Gly Pro Ala Leu Leu
1 5 10 15

Leu Leu Leu Leu Phe Leu Gly Ala Ala Glu Ser Val Arg Arg Ala Gln
20 25 30

Pro Pro Arg Arg Tyr Thr Pro Asp Trp Pro Ser Leu Asp Ser Arg Pro
35 40 45

Leu Pro Ala Trp Phe Asp Glu Ala Lys Phe Gly Val Phe Ile His Trp
50 55 60

Gly Val Phe Ser Val Pro Ala Trp Gly Ser Glu Trp Phe Trp Trp His
65 70 75 80

Trp Gln Gly Glu Gly Arg Pro Gln Tyr Gln Arg Phe Met Arg Asp Asn
85 90 95

Tyr Pro Pro Gly Phe Ser Tyr Ala Asp Phe Gly Pro Gln Phe Thr Ala
100 105 110

Arg Phe Phe His Pro Glu Glu Trp Ala Asp Leu Phe Gln Ala Ala Gly
115 120 125

Ala Lys Tyr Val Val Leu Thr Thr Lys His His Glu Gly Phe Thr Asn
130 135 140

Trp Pro Ser Pro Val Ser Trp Asn Trp Asn Ser Lys Asp Val Gly Pro
145 150 155 160

His Arg Asp Leu Val Gly Glu Leu Gly Thr Ala Leu Arg Lys Arg Asn
165 170 175

Ile Arg Tyr Gly Leu Tyr His Ser Leu Leu Glu Trp Phe His Pro Leu
180 185 190

```
Tyr Leu Leu Asp Lys Lys Asn Gly Phe Lys Thr Gln His Phe Val Ser
        195                 200                 205

Ala Lys Thr Met Pro Glu Leu Tyr Asp Leu Val Asn Ser Tyr Lys Pro
    210                 215                 220

Asp Leu Ile Trp Ser Asp Gly Glu Trp Glu Cys Pro Asp Thr Tyr Trp
225                 230                 235                 240

Asn Ser Thr Asn Phe Leu Ser Trp Leu Tyr Asn Asp Ser Pro Val Lys
                245                 250                 255

Asp Glu Val Val Val Asn Asp Arg Trp Gly Gln Asn Cys Ser Cys His
            260                 265                 270

His Gly Gly Tyr Tyr Asn Cys Glu Asp Lys Phe Lys Pro Gln Ser Leu
        275                 280                 285

Pro Asp His Lys Trp Glu Met Cys Thr Ser Ile Asp Lys Phe Ser Trp
    290                 295                 300

Gly Tyr Arg Arg Asp Met Ala Leu Ser Asp Val Thr Glu Glu Ser Glu
305                 310                 315                 320

Ile Ile Ser Glu Leu Val Gln Thr Val Ser Leu Gly Gly Asn Tyr Leu
                325                 330                 335

Leu Asn Ile Gly Pro Thr Lys Asp Gly Leu Ile Val Pro Ile Phe Gln
            340                 345                 350

Glu Arg Leu Leu Ala Val Gly Lys Trp Leu Ser Ile Asn Gly Glu Ala
        355                 360                 365

Ile Tyr Ala Ser Lys Pro Trp Arg Val Gln Trp Glu Lys Asn Thr Thr
    370                 375                 380

Ser Val Trp Tyr Thr Ser Lys Gly Ser Ala Val Tyr Ala Ile Phe Leu
385                 390                 395                 400

His Trp Pro Glu Asn Gly Val Leu Asn Leu Glu Ser Pro Ile Thr Thr
                405                 410                 415

Ser Thr Thr Lys Ile Thr Met Leu Gly Ile Gln Gly Asp Leu Lys Trp
            420                 425                 430

Ser Thr Asp Pro Asp Lys Gly Leu Phe Ile Ser Leu Pro Gln Leu Pro
        435                 440                 445

Pro Ser Ala Val Pro Ala Glu Phe Ala Trp Thr Ile Lys Leu Thr Gly
    450                 455                 460
```

Val Lys
465

<210> 148
<211> 685
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> GALC

<400> 148

Met Ala Glu Trp Leu Leu Ser Ala Ser Trp Gln Arg Arg Ala Lys Ala
1 5 10 15

Met Thr Ala Ala Ala Gly Ser Ala Gly Arg Ala Ala Val Pro Leu Leu
20 25 30

Leu Cys Ala Leu Leu Ala Pro Gly Gly Ala Tyr Val Leu Asp Asp Ser
35 40 45

Asp Gly Leu Gly Arg Glu Phe Asp Gly Ile Gly Ala Val Ser Gly Gly
50 55 60

Gly Ala Thr Ser Arg Leu Leu Val Asn Tyr Pro Glu Pro Tyr Arg Ser
65 70 75 80

Gln Ile Leu Asp Tyr Leu Phe Lys Pro Asn Phe Gly Ala Ser Leu His
85 90 95

Ile Leu Lys Val Glu Ile Gly Gly Asp Gly Gln Thr Thr Asp Gly Thr
100 105 110

Glu Pro Ser His Met His Tyr Ala Leu Asp Glu Asn Tyr Phe Arg Gly
115 120 125

Tyr Glu Trp Trp Leu Met Lys Glu Ala Lys Lys Arg Asn Pro Asn Ile
130 135 140

Thr Leu Ile Gly Leu Pro Trp Ser Phe Pro Gly Trp Leu Gly Lys Gly
145 150 155 160

Phe Asp Trp Pro Tyr Val Asn Leu Gln Leu Thr Ala Tyr Tyr Val Val
165 170 175

Thr Trp Ile Val Gly Ala Lys Arg Tyr His Asp Leu Asp Ile Asp Tyr
180 185 190

Ile Gly Ile Trp Asn Glu Arg Ser Tyr Asn Ala Asn Tyr Ile Lys Ile
195 200 205

```
Leu Arg Lys Met Leu Asn Tyr Gln Gly Leu Gln Arg Val Lys Ile Ile
    210                 215                 220

Ala Ser Asp Asn Leu Trp Glu Ser Ile Ser Ala Ser Met Leu Leu Asp
225                 230                 235                 240

Ala Glu Leu Phe Lys Val Val Asp Val Ile Gly Ala His Tyr Pro Gly
                245                 250                 255

Thr His Ser Ala Lys Asp Ala Lys Leu Thr Gly Lys Lys Leu Trp Ser
            260                 265                 270

Ser Glu Asp Phe Ser Thr Leu Asn Ser Asp Met Gly Ala Gly Cys Trp
        275                 280                 285

Gly Arg Ile Leu Asn Gln Asn Tyr Ile Asn Gly Tyr Met Thr Ser Thr
    290                 295                 300

Ile Ala Trp Asn Leu Val Ala Ser Tyr Tyr Glu Gln Leu Pro Tyr Gly
305                 310                 315                 320

Arg Cys Gly Leu Met Thr Ala Gln Glu Pro Trp Ser Gly His Tyr Val
                325                 330                 335

Val Glu Ser Pro Val Trp Val Ser Ala His Thr Thr Gln Phe Thr Gln
            340                 345                 350

Pro Gly Trp Tyr Tyr Leu Lys Thr Val Gly His Leu Glu Lys Gly Gly
        355                 360                 365

Ser Tyr Val Ala Leu Thr Asp Gly Leu Gly Asn Leu Thr Ile Ile Ile
    370                 375                 380

Glu Thr Met Ser His Lys His Ser Lys Cys Ile Arg Pro Phe Leu Pro
385                 390                 395                 400

Tyr Phe Asn Val Ser Gln Gln Phe Ala Thr Phe Val Leu Lys Gly Ser
                405                 410                 415

Phe Ser Glu Ile Pro Glu Leu Gln Val Trp Tyr Thr Lys Leu Gly Lys
            420                 425                 430

Thr Ser Glu Arg Phe Leu Phe Lys Gln Leu Asp Ser Leu Trp Leu Leu
        435                 440                 445

Asp Ser Asp Gly Ser Phe Thr Leu Ser Leu His Glu Asp Glu Leu Phe
    450                 455                 460

Thr Leu Thr Thr Leu Thr Thr Gly Arg Lys Gly Ser Tyr Pro Leu Pro
465                 470                 475                 480

Pro Lys Ser Gln Pro Phe Pro Ser Thr Tyr Lys Asp Asp Phe Asn Val
```

485 490 495

Asp Tyr Pro Phe Phe Ser Glu Ala Pro Asn Phe Ala Asp Gln Thr Gly
500 505 510

Val Phe Glu Tyr Phe Thr Asn Ile Glu Asp Pro Gly Glu His His Phe
515 520 525

Thr Leu Arg Gln Val Leu Asn Gln Arg Pro Ile Thr Trp Ala Ala Asp
530 535 540

Ala Ser Asn Thr Ile Ser Ile Ile Gly Asp Tyr Asn Trp Thr Asn Leu
545 550 555 560

Thr Ile Lys Cys Asp Val Tyr Ile Glu Thr Pro Asp Thr Gly Gly Val
565 570 575

Phe Ile Ala Gly Arg Val Asn Lys Gly Gly Ile Leu Ile Arg Ser Ala
580 585 590

Arg Gly Ile Phe Phe Trp Ile Phe Ala Asn Gly Ser Tyr Arg Val Thr
595 600 605

Gly Asp Leu Ala Gly Trp Ile Ile Tyr Ala Leu Gly Arg Val Glu Val
610 615 620

Thr Ala Lys Lys Trp Tyr Thr Leu Thr Leu Thr Ile Lys Gly His Phe
625 630 635 640

Thr Ser Gly Met Leu Asn Asp Lys Ser Leu Trp Thr Asp Ile Pro Val
645 650 655

Asn Phe Pro Lys Asn Gly Trp Ala Ala Ile Gly Thr His Ser Phe Glu
660 665 670

Phe Ala Gln Phe Asp Asn Phe Leu Val Glu Ala Thr Arg
675 680 685

<210> 149
<211> 529
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> HEXA

<400> 149

Met Thr Ser Ser Arg Leu Trp Phe Ser Leu Leu Leu Ala Ala Ala Phe
1 5 10 15

Ala Gly Arg Ala Thr Ala Leu Trp Pro Trp Pro Gln Asn Phe Gln Thr
20 25 30

```
Ser Asp Gln Arg Tyr Val Leu Tyr Pro Asn Asn Phe Gln Phe Gln Tyr
        35                  40                  45

Asp Val Ser Ser Ala Ala Gln Pro Gly Cys Ser Val Leu Asp Glu Ala
    50                  55                  60

Phe Gln Arg Tyr Arg Asp Leu Leu Phe Gly Ser Gly Ser Trp Pro Arg
65                  70                  75                  80

Pro Tyr Leu Thr Gly Lys Arg His Thr Leu Glu Lys Asn Val Leu Val
                85                  90                  95

Val Ser Val Val Thr Pro Gly Cys Asn Gln Leu Pro Thr Leu Glu Ser
            100                 105                 110

Val Glu Asn Tyr Thr Leu Thr Ile Asn Asp Asp Gln Cys Leu Leu Leu
        115                 120                 125

Ser Glu Thr Val Trp Gly Ala Leu Arg Gly Leu Glu Thr Phe Ser Gln
    130                 135                 140

Leu Val Trp Lys Ser Ala Glu Gly Thr Phe Phe Ile Asn Lys Thr Glu
145                 150                 155                 160

Ile Glu Asp Phe Pro Arg Phe Pro His Arg Gly Leu Leu Leu Asp Thr
                165                 170                 175

Ser Arg His Tyr Leu Pro Leu Ser Ser Ile Leu Asp Thr Leu Asp Val
            180                 185                 190

Met Ala Tyr Asn Lys Leu Asn Val Phe His Trp His Leu Val Asp Asp
        195                 200                 205

Pro Ser Phe Pro Tyr Glu Ser Phe Thr Phe Pro Glu Leu Met Arg Lys
    210                 215                 220

Gly Ser Tyr Asn Pro Val Thr His Ile Tyr Thr Ala Gln Asp Val Lys
225                 230                 235                 240

Glu Val Ile Glu Tyr Ala Arg Leu Arg Gly Ile Arg Val Leu Ala Glu
                245                 250                 255

Phe Asp Thr Pro Gly His Thr Leu Ser Trp Gly Pro Gly Ile Pro Gly
            260                 265                 270

Leu Leu Thr Pro Cys Tyr Ser Gly Ser Glu Pro Ser Gly Thr Phe Gly
        275                 280                 285

Pro Val Asn Pro Ser Leu Asn Asn Thr Tyr Glu Phe Met Ser Thr Phe
    290                 295                 300
```

Phe Leu Glu Val Ser Ser Val Phe Pro Asp Phe Tyr Leu His Leu Gly
305 310 315 320

Gly Asp Glu Val Asp Phe Thr Cys Trp Lys Ser Asn Pro Glu Ile Gln
325 330 335

Asp Phe Met Arg Lys Lys Gly Phe Gly Glu Asp Phe Lys Gln Leu Glu
340 345 350

Ser Phe Tyr Ile Gln Thr Leu Leu Asp Ile Val Ser Ser Tyr Gly Lys
355 360 365

Gly Tyr Val Val Trp Gln Glu Val Phe Asp Asn Lys Val Lys Ile Gln
370 375 380

Pro Asp Thr Ile Ile Gln Val Trp Arg Glu Asp Ile Pro Val Asn Tyr
385 390 395 400

Met Lys Glu Leu Glu Leu Val Thr Lys Ala Gly Phe Arg Ala Leu Leu
405 410 415

Ser Ala Pro Trp Tyr Leu Asn Arg Ile Ser Tyr Gly Pro Asp Trp Lys
420 425 430

Asp Phe Tyr Ile Val Glu Pro Leu Ala Phe Glu Gly Thr Pro Glu Gln
435 440 445

Lys Ala Leu Val Ile Gly Gly Glu Ala Cys Met Trp Gly Glu Tyr Val
450 455 460

Asp Asn Thr Asn Leu Val Pro Arg Leu Trp Pro Arg Ala Gly Ala Val
465 470 475 480

Ala Glu Arg Leu Trp Ser Asn Lys Leu Thr Ser Asp Leu Thr Phe Ala
485 490 495

Tyr Glu Arg Leu Ser His Phe Arg Cys Glu Leu Leu Arg Arg Gly Val
500 505 510

Gln Ala Gln Pro Leu Asn Val Gly Phe Cys Glu Gln Glu Phe Glu Gln
515 520 525

Thr

<210> 150
<211> 556
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE

<223> HEXB

<400> 150

```
Met Glu Leu Cys Gly Leu Gly Leu Pro Arg Pro Pro Met Leu Leu Ala
1               5                   10                  15

Leu Leu Leu Ala Thr Leu Leu Ala Ala Met Leu Ala Leu Leu Thr Gln
            20                  25                  30

Val Ala Leu Val Val Gln Val Ala Glu Ala Ala Arg Ala Pro Ser Val
        35                  40                  45

Ser Ala Lys Pro Gly Pro Ala Leu Trp Pro Leu Pro Leu Ser Val Lys
    50                  55                  60

Met Thr Pro Asn Leu Leu His Leu Ala Pro Glu Asn Phe Tyr Ile Ser
65                  70                  75                  80

His Ser Pro Asn Ser Thr Ala Gly Pro Ser Cys Thr Leu Leu Glu Glu
                85                  90                  95

Ala Phe Arg Arg Tyr His Gly Tyr Ile Phe Gly Phe Tyr Lys Trp His
            100                 105                 110

His Glu Pro Ala Glu Phe Gln Ala Lys Thr Gln Val Gln Gln Leu Leu
        115                 120                 125

Val Ser Ile Thr Leu Gln Ser Glu Cys Asp Ala Phe Pro Asn Ile Ser
    130                 135                 140

Ser Asp Glu Ser Tyr Thr Leu Leu Val Lys Glu Pro Val Ala Val Leu
145                 150                 155                 160

Lys Ala Asn Arg Val Trp Gly Ala Leu Arg Gly Leu Glu Thr Phe Ser
                165                 170                 175

Gln Leu Val Tyr Gln Asp Ser Tyr Gly Thr Phe Thr Ile Asn Glu Ser
            180                 185                 190

Thr Ile Ile Asp Ser Pro Arg Phe Ser His Arg Gly Ile Leu Ile Asp
        195                 200                 205

Thr Ser Arg His Tyr Leu Pro Val Lys Ile Ile Leu Lys Thr Leu Asp
    210                 215                 220

Ala Met Ala Phe Asn Lys Phe Asn Val Leu His Trp His Ile Val Asp
225                 230                 235                 240

Asp Gln Ser Phe Pro Tyr Gln Ser Ile Thr Phe Pro Glu Leu Ser Asn
                245                 250                 255

Lys Gly Ser Tyr Ser Leu Ser His Val Tyr Thr Pro Asn Asp Val Arg
```

260 265 270

Met Val Ile Glu Tyr Ala Arg Leu Arg Gly Ile Arg Val Leu Pro Glu
275 280 285

Phe Asp Thr Pro Gly His Thr Leu Ser Trp Gly Lys Gly Gln Lys Asp
290 295 300

Leu Leu Thr Pro Cys Tyr Ser Arg Gln Asn Lys Leu Asp Ser Phe Gly
305 310 315 320

Pro Ile Asn Pro Thr Leu Asn Thr Thr Tyr Ser Phe Leu Thr Thr Phe
325 330 335

Phe Lys Glu Ile Ser Glu Val Phe Pro Asp Gln Phe Ile His Leu Gly
340 345 350

Gly Asp Glu Val Glu Phe Lys Cys Trp Glu Ser Asn Pro Lys Ile Gln
355 360 365

Asp Phe Met Arg Gln Lys Gly Phe Gly Thr Asp Phe Lys Lys Leu Glu
370 375 380

Ser Phe Tyr Ile Gln Lys Val Leu Asp Ile Ile Ala Thr Ile Asn Lys
385 390 395 400

Gly Ser Ile Val Trp Gln Glu Val Phe Asp Asp Lys Ala Lys Leu Ala
405 410 415

Pro Gly Thr Ile Val Glu Val Trp Lys Asp Ser Ala Tyr Pro Glu Glu
420 425 430

Leu Ser Arg Val Thr Ala Ser Gly Phe Pro Val Ile Leu Ser Ala Pro
435 440 445

Trp Tyr Leu Asp Leu Ile Ser Tyr Gly Gln Asp Trp Arg Lys Tyr Tyr
450 455 460

Lys Val Glu Pro Leu Asp Phe Gly Gly Thr Gln Lys Gln Lys Gln Leu
465 470 475 480

Phe Ile Gly Gly Glu Ala Cys Leu Trp Gly Glu Tyr Val Asp Ala Thr
485 490 495

Asn Leu Thr Pro Arg Leu Trp Pro Arg Ala Ser Ala Val Gly Glu Arg
500 505 510

Leu Trp Ser Ser Lys Asp Val Arg Asp Met Asp Asp Ala Tyr Asp Arg
515 520 525

Leu Thr Arg His Arg Cys Arg Met Val Glu Arg Gly Ile Ala Ala Gln
530 535 540

```
Pro Leu Tyr Ala Gly Tyr Cys Asn His Glu Asn Met
545                 550                 555


<210>  151
<211>  879
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  MANBA

<400>  151

Met Arg Leu His Leu Leu Leu Leu Leu Ala Leu Cys Gly Ala Gly Thr
1               5                   10                  15


Thr Ala Ala Glu Leu Ser Tyr Ser Leu Arg Gly Asn Trp Ser Ile Cys
            20                  25                  30


Asn Gly Asn Gly Ser Leu Glu Leu Pro Gly Ala Val Pro Gly Cys Val
        35                  40                  45


His Ser Ala Leu Phe Gln Gln Gly Leu Ile Gln Asp Ser Tyr Tyr Arg
    50                  55                  60


Phe Asn Asp Leu Asn Tyr Arg Trp Val Ser Leu Asp Asn Trp Thr Tyr
65                  70                  75                  80


Ser Lys Glu Phe Lys Ile Pro Phe Glu Ile Ser Lys Trp Gln Lys Val
                85                  90                  95


Asn Leu Ile Leu Glu Gly Val Asp Thr Val Ser Lys Ile Leu Phe Asn
            100                 105                 110


Glu Val Thr Ile Gly Glu Thr Asp Asn Met Phe Asn Arg Tyr Ser Phe
        115                 120                 125


Asp Ile Thr Asn Val Val Arg Asp Val Asn Ser Ile Glu Leu Arg Phe
    130                 135                 140


Gln Ser Ala Val Leu Tyr Ala Ala Gln Gln Ser Lys Ala His Thr Arg
145                 150                 155                 160


Tyr Gln Val Pro Pro Asp Cys Pro Pro Leu Val Gln Lys Gly Glu Cys
                165                 170                 175


His Val Asn Phe Val Arg Lys Glu Gln Cys Ser Phe Ser Trp Asp Trp
            180                 185                 190


Gly Pro Ser Phe Pro Thr Gln Gly Ile Trp Lys Asp Val Arg Ile Glu
        195                 200                 205
```

```
Ala Tyr Asn Ile Cys His Leu Asn Tyr Phe Thr Phe Ser Pro Ile Tyr
    210                 215                 220

Asp Lys Ser Ala Gln Glu Trp Asn Leu Glu Ile Glu Ser Thr Phe Asp
225                 230                 235                 240

Val Val Ser Ser Lys Pro Val Gly Gly Gln Val Ile Val Ala Ile Pro
                245                 250                 255

Lys Leu Gln Thr Gln Gln Thr Tyr Ser Ile Glu Leu Gln Pro Gly Lys
            260                 265                 270

Arg Ile Val Glu Leu Phe Val Asn Ile Ser Lys Asn Ile Thr Val Glu
        275                 280                 285

Thr Trp Trp Pro His Gly His Gly Asn Gln Thr Gly Tyr Asn Met Thr
    290                 295                 300

Val Leu Phe Glu Leu Asp Gly Gly Leu Asn Ile Glu Lys Ser Ala Lys
305                 310                 315                 320

Val Tyr Phe Arg Thr Val Glu Leu Ile Glu Glu Pro Ile Lys Gly Ser
                325                 330                 335

Pro Gly Leu Ser Phe Tyr Phe Lys Ile Asn Gly Phe Pro Ile Phe Leu
            340                 345                 350

Lys Gly Ser Asn Trp Ile Pro Ala Asp Ser Phe Gln Asp Arg Val Thr
        355                 360                 365

Ser Glu Leu Leu Arg Leu Leu Leu Gln Ser Val Val Asp Ala Asn Met
    370                 375                 380

Asn Thr Leu Arg Val Trp Gly Gly Gly Ile Tyr Glu Gln Asp Glu Phe
385                 390                 395                 400

Tyr Glu Leu Cys Asp Glu Leu Gly Ile Met Val Trp Gln Asp Phe Met
                405                 410                 415

Phe Ala Cys Ala Leu Tyr Pro Thr Asp Gln Gly Phe Leu Asp Ser Val
            420                 425                 430

Thr Ala Glu Val Ala Tyr Gln Ile Lys Arg Leu Lys Ser His Pro Ser
        435                 440                 445

Ile Ile Ile Trp Ser Gly Asn Asn Glu Asn Glu Glu Ala Leu Met Met
    450                 455                 460

Asn Trp Tyr His Ile Ser Phe Thr Asp Arg Pro Ile Tyr Ile Lys Asp
465                 470                 475                 480
```

```
Tyr Val Thr Leu Tyr Val Lys Asn Ile Arg Glu Leu Val Leu Ala Gly
                485                 490                 495

Asp Lys Ser Arg Pro Phe Ile Thr Ser Ser Pro Thr Asn Gly Ala Glu
            500                 505                 510

Thr Val Ala Glu Ala Trp Val Ser Gln Asn Pro Asn Ser Asn Tyr Phe
        515                 520                 525

Gly Asp Val His Phe Tyr Asp Tyr Ile Ser Asp Cys Trp Asn Trp Lys
    530                 535                 540

Val Phe Pro Lys Ala Arg Phe Ala Ser Glu Tyr Gly Tyr Gln Ser Trp
545                 550                 555                 560

Pro Ser Phe Ser Thr Leu Glu Lys Val Ser Ser Thr Glu Asp Trp Ser
                565                 570                 575

Phe Asn Ser Lys Phe Ser Leu His Arg Gln His His Glu Gly Gly Asn
            580                 585                 590

Lys Gln Met Leu Tyr Gln Ala Gly Leu His Phe Lys Leu Pro Gln Ser
        595                 600                 605

Thr Asp Pro Leu Arg Thr Phe Lys Asp Thr Ile Tyr Leu Thr Gln Val
    610                 615                 620

Met Gln Ala Gln Cys Val Lys Thr Glu Thr Glu Phe Tyr Arg Arg Ser
625                 630                 635                 640

Arg Ser Glu Ile Val Asp Gln Gln Gly His Thr Met Gly Ala Leu Tyr
                645                 650                 655

Trp Gln Leu Asn Asp Ile Trp Gln Ala Pro Ser Trp Ala Ser Leu Glu
            660                 665                 670

Tyr Gly Gly Lys Trp Lys Met Leu His Tyr Phe Ala Gln Asn Phe Phe
        675                 680                 685

Ala Pro Leu Leu Pro Val Gly Phe Glu Asn Glu Asn Thr Phe Tyr Ile
    690                 695                 700

Tyr Gly Val Ser Asp Leu His Ser Asp Tyr Ser Met Thr Leu Ser Val
705                 710                 715                 720

Arg Val His Thr Trp Ser Ser Leu Glu Pro Val Cys Ser Arg Val Thr
                725                 730                 735

Glu Arg Phe Val Met Lys Gly Gly Glu Ala Val Cys Leu Tyr Glu Glu
            740                 745                 750

Pro Val Ser Glu Leu Leu Arg Arg Cys Gly Asn Cys Thr Arg Glu Ser
```

```
        755                 760                 765


Cys Val Val Ser Phe Tyr Leu Ser Ala Asp His Glu Leu Leu Ser Pro
    770                 775                 780


Thr Asn Tyr His Phe Leu Ser Ser Pro Lys Glu Ala Val Gly Leu Cys
785                 790                 795                 800


Lys Ala Gln Ile Thr Ala Ile Ile Ser Gln Gln Gly Asp Ile Phe Val
                805                 810                 815


Phe Asp Leu Glu Thr Ser Ala Val Ala Pro Phe Val Trp Leu Asp Val
            820                 825                 830


Gly Ser Ile Pro Gly Arg Phe Ser Asp Asn Gly Phe Leu Met Thr Glu
        835                 840                 845


Lys Thr Arg Thr Ile Leu Phe Tyr Pro Trp Glu Pro Thr Ser Lys Asn
    850                 855                 860


Glu Leu Glu Gln Ser Phe His Val Thr Ser Leu Thr Asp Ile Tyr
865                 870                 875


<210>  152
<211>  507
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  ARSA

<400>  152

Met Gly Ala Pro Arg Ser Leu Leu Leu Ala Leu Ala Ala Gly Leu Ala
1               5                   10                  15


Val Ala Arg Pro Pro Asn Ile Val Leu Ile Phe Ala Asp Asp Leu Gly
            20                  25                  30


Tyr Gly Asp Leu Gly Cys Tyr Gly His Pro Ser Ser Thr Thr Pro Asn
        35                  40                  45


Leu Asp Gln Leu Ala Ala Gly Gly Leu Arg Phe Thr Asp Phe Tyr Val
    50                  55                  60


Pro Val Ser Leu Cys Thr Pro Ser Arg Ala Ala Leu Leu Thr Gly Arg
65                  70                  75                  80


Leu Pro Val Arg Met Gly Met Tyr Pro Gly Val Leu Val Pro Ser Ser
                85                  90                  95


Arg Gly Gly Leu Pro Leu Glu Glu Val Thr Val Ala Glu Val Leu Ala
            100                 105                 110
```

```
Ala Arg Gly Tyr Leu Thr Gly Met Ala Gly Lys Trp His Leu Gly Val
        115                 120                 125

Gly Pro Glu Gly Ala Phe Leu Pro Pro His Gln Gly Phe His Arg Phe
    130                 135                 140

Leu Gly Ile Pro Tyr Ser His Asp Gln Gly Pro Cys Gln Asn Leu Thr
145                 150                 155                 160

Cys Phe Pro Pro Ala Thr Pro Cys Asp Gly Gly Cys Asp Gln Gly Leu
                165                 170                 175

Val Pro Ile Pro Leu Leu Ala Asn Leu Ser Val Glu Ala Gln Pro Pro
            180                 185                 190

Trp Leu Pro Gly Leu Glu Ala Arg Tyr Met Ala Phe Ala His Asp Leu
        195                 200                 205

Met Ala Asp Ala Gln Arg Gln Asp Arg Pro Phe Phe Leu Tyr Tyr Ala
    210                 215                 220

Ser His His Thr His Tyr Pro Gln Phe Ser Gly Gln Ser Phe Ala Glu
225                 230                 235                 240

Arg Ser Gly Arg Gly Pro Phe Gly Asp Ser Leu Met Glu Leu Asp Ala
                245                 250                 255

Ala Val Gly Thr Leu Met Thr Ala Ile Gly Asp Leu Gly Leu Leu Glu
            260                 265                 270

Glu Thr Leu Val Ile Phe Thr Ala Asp Asn Gly Pro Glu Thr Met Arg
        275                 280                 285

Met Ser Arg Gly Gly Cys Ser Gly Leu Leu Arg Cys Gly Lys Gly Thr
    290                 295                 300

Thr Tyr Glu Gly Gly Val Arg Glu Pro Ala Leu Ala Phe Trp Pro Gly
305                 310                 315                 320

His Ile Ala Pro Gly Val Thr His Glu Leu Ala Ser Ser Leu Asp Leu
                325                 330                 335

Leu Pro Thr Leu Ala Ala Leu Ala Gly Ala Pro Leu Pro Asn Val Thr
            340                 345                 350

Leu Asp Gly Phe Asp Leu Ser Pro Leu Leu Leu Gly Thr Gly Lys Ser
        355                 360                 365

Pro Arg Gln Ser Leu Phe Phe Tyr Pro Ser Tyr Pro Asp Glu Val Arg
    370                 375                 380
```

```
Gly Val Phe Ala Val Arg Thr Gly Lys Tyr Lys Ala His Phe Phe Thr
385                 390                 395                 400


Gln Gly Ser Ala His Ser Asp Thr Thr Ala Asp Pro Ala Cys His Ala
                405                 410                 415


Ser Ser Ser Leu Thr Ala His Glu Pro Pro Leu Leu Tyr Asp Leu Ser
            420                 425                 430


Lys Asp Pro Gly Glu Asn Tyr Asn Leu Leu Gly Gly Val Ala Gly Ala
        435                 440                 445


Thr Pro Glu Val Leu Gln Ala Leu Lys Gln Leu Gln Leu Leu Lys Ala
    450                 455                 460


Gln Leu Asp Ala Ala Val Thr Phe Gly Pro Ser Gln Val Ala Arg Gly
465                 470                 475                 480


Glu Asp Pro Ala Leu Gln Ile Cys Cys His Pro Gly Cys Thr Pro Arg
                485                 490                 495


Pro Ala Cys Cys His Cys Pro Asp Pro His Ala
            500                 505


<210>  153
<211>  1256
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  GNPTAB

<400>  153

Met Leu Phe Lys Leu Leu Gln Arg Gln Thr Tyr Thr Cys Leu Ser His
1               5                   10                  15


Arg Tyr Gly Leu Tyr Val Cys Phe Leu Gly Val Val Val Thr Ile Val
            20                  25                  30


Ser Ala Phe Gln Phe Gly Glu Val Val Leu Glu Trp Ser Arg Asp Gln
        35                  40                  45


Tyr His Val Leu Phe Asp Ser Tyr Arg Asp Asn Ile Ala Gly Lys Ser
    50                  55                  60


Phe Gln Asn Arg Leu Cys Leu Pro Met Pro Ile Asp Val Val Tyr Thr
65                  70                  75                  80


Trp Val Asn Gly Thr Asp Leu Glu Leu Leu Lys Glu Leu Gln Gln Val
                85                  90                  95
```

```
Arg Glu Gln Met Glu Glu Glu Gln Lys Ala Met Arg Glu Ile Leu Gly
            100                 105                 110

Lys Asn Thr Thr Glu Pro Thr Lys Lys Ser Glu Lys Gln Leu Glu Cys
        115                 120                 125

Leu Leu Thr His Cys Ile Lys Val Pro Met Leu Val Leu Asp Pro Ala
    130                 135                 140

Leu Pro Ala Asn Ile Thr Leu Lys Asp Leu Pro Ser Leu Tyr Pro Ser
145                 150                 155                 160

Phe His Ser Ala Ser Asp Ile Phe Asn Val Ala Lys Pro Lys Asn Pro
                165                 170                 175

Ser Thr Asn Val Ser Val Val Val Phe Asp Ser Thr Lys Asp Val Glu
            180                 185                 190

Asp Ala His Ser Gly Leu Leu Lys Gly Asn Ser Arg Gln Thr Val Trp
        195                 200                 205

Arg Gly Tyr Leu Thr Thr Asp Lys Glu Val Pro Gly Leu Val Leu Met
    210                 215                 220

Gln Asp Leu Ala Phe Leu Ser Gly Phe Pro Pro Thr Phe Lys Glu Thr
225                 230                 235                 240

Asn Gln Leu Lys Thr Lys Leu Pro Glu Asn Leu Ser Ser Lys Val Lys
                245                 250                 255

Leu Leu Gln Leu Tyr Ser Glu Ala Ser Val Ala Leu Leu Lys Leu Asn
            260                 265                 270

Asn Pro Lys Asp Phe Gln Glu Leu Asn Lys Gln Thr Lys Lys Asn Met
        275                 280                 285

Thr Ile Asp Gly Lys Glu Leu Thr Ile Ser Pro Ala Tyr Leu Leu Trp
    290                 295                 300

Asp Leu Ser Ala Ile Ser Gln Ser Lys Gln Asp Glu Asp Ile Ser Ala
305                 310                 315                 320

Ser Arg Phe Glu Asp Asn Glu Glu Leu Arg Tyr Ser Leu Arg Ser Ile
                325                 330                 335

Glu Arg His Ala Pro Trp Val Arg Asn Ile Phe Ile Val Thr Asn Gly
            340                 345                 350

Gln Ile Pro Ser Trp Leu Asn Leu Asp Asn Pro Arg Val Thr Ile Val
        355                 360                 365

Thr His Gln Asp Val Phe Arg Asn Leu Ser His Leu Pro Thr Phe Ser
```

370 375 380

Ser Pro Ala Ile Glu Ser His Ile His Arg Ile Glu Gly Leu Ser Gln
385 390 395 400

Lys Phe Ile Tyr Leu Asn Asp Asp Val Met Phe Gly Lys Asp Val Trp
405 410 415

Pro Asp Asp Phe Tyr Ser His Ser Lys Gly Gln Lys Val Tyr Leu Thr
420 425 430

Trp Pro Val Pro Asn Cys Ala Glu Gly Cys Pro Gly Ser Trp Ile Lys
435 440 445

Asp Gly Tyr Cys Asp Lys Ala Cys Asn Asn Ser Ala Cys Asp Trp Asp
450 455 460

Gly Gly Asp Cys Ser Gly Asn Ser Gly Gly Ser Arg Tyr Ile Ala Gly
465 470 475 480

Gly Gly Gly Thr Gly Ser Ile Gly Val Gly Gln Pro Trp Gln Phe Gly
485 490 495

Gly Gly Ile Asn Ser Val Ser Tyr Cys Asn Gln Gly Cys Ala Asn Ser
500 505 510

Trp Leu Ala Asp Lys Phe Cys Asp Gln Ala Cys Asn Val Leu Ser Cys
515 520 525

Gly Phe Asp Ala Gly Asp Cys Gly Gln Asp His Phe His Glu Leu Tyr
530 535 540

Lys Val Ile Leu Leu Pro Asn Gln Thr His Tyr Ile Ile Pro Lys Gly
545 550 555 560

Glu Cys Leu Pro Tyr Phe Ser Phe Ala Glu Val Ala Lys Arg Gly Val
565 570 575

Glu Gly Ala Tyr Ser Asp Asn Pro Ile Ile Arg His Ala Ser Ile Ala
580 585 590

Asn Lys Trp Lys Thr Ile His Leu Ile Met His Ser Gly Met Asn Ala
595 600 605

Thr Thr Ile His Phe Asn Leu Thr Phe Gln Asn Thr Asn Asp Glu Glu
610 615 620

Phe Lys Met Gln Ile Thr Val Glu Val Asp Thr Arg Glu Gly Pro Lys
625 630 635 640

Leu Asn Ser Thr Ala Gln Lys Gly Tyr Glu Asn Leu Val Ser Pro Ile
645 650 655

```
Thr Leu Leu Pro Glu Ala Glu Ile Leu Phe Glu Asp Ile Pro Lys Glu
            660                 665                 670

Lys Arg Phe Pro Lys Phe Lys Arg His Asp Val Asn Ser Thr Arg Arg
        675                 680                 685

Ala Gln Glu Glu Val Lys Ile Pro Leu Val Asn Ile Ser Leu Leu Pro
    690                 695                 700

Lys Asp Ala Gln Leu Ser Leu Asn Thr Leu Asp Leu Gln Leu Glu His
705                 710                 715                 720

Gly Asp Ile Thr Leu Lys Gly Tyr Asn Leu Ser Lys Ser Ala Leu Leu
                725                 730                 735

Arg Ser Phe Leu Met Asn Ser Gln His Ala Lys Ile Lys Asn Gln Ala
            740                 745                 750

Ile Ile Thr Asp Glu Thr Asn Asp Ser Leu Val Ala Pro Gln Glu Lys
        755                 760                 765

Gln Val His Lys Ser Ile Leu Pro Asn Ser Leu Gly Val Ser Glu Arg
    770                 775                 780

Leu Gln Arg Leu Thr Phe Pro Ala Val Ser Val Lys Val Asn Gly His
785                 790                 795                 800

Asp Gln Gly Gln Asn Pro Pro Leu Asp Leu Glu Thr Thr Ala Arg Phe
                805                 810                 815

Arg Val Glu Thr His Thr Gln Lys Thr Ile Gly Gly Asn Val Thr Lys
            820                 825                 830

Glu Lys Pro Pro Ser Leu Ile Val Pro Leu Glu Ser Gln Met Thr Lys
        835                 840                 845

Glu Lys Lys Ile Thr Gly Lys Glu Lys Glu Asn Ser Arg Met Glu Glu
    850                 855                 860

Asn Ala Glu Asn His Ile Gly Val Thr Glu Val Leu Leu Gly Arg Lys
865                 870                 875                 880

Leu Gln His Tyr Thr Asp Ser Tyr Leu Gly Phe Leu Pro Trp Glu Lys
                885                 890                 895

Lys Lys Tyr Phe Gln Asp Leu Leu Asp Glu Glu Glu Ser Leu Lys Thr
            900                 905                 910

Gln Leu Ala Tyr Phe Thr Asp Ser Lys Asn Thr Gly Arg Gln Leu Lys
        915                 920                 925
```

```
Asp Thr Phe Ala Asp Ser Leu Arg Tyr Val Asn Lys Ile Leu Asn Ser
    930                 935                 940

Lys Phe Gly Phe Thr Ser Arg Lys Val Pro Ala His Met Pro His Met
945                 950                 955                 960

Ile Asp Arg Ile Val Met Gln Glu Leu Gln Asp Met Phe Pro Glu Glu
                965                 970                 975

Phe Asp Lys Thr Ser Phe His Lys Val Arg His Ser Glu Asp Met Gln
            980                 985                 990

Phe Ala Phe Ser Tyr Phe Tyr Tyr  Leu Met Ser Ala Val  Gln Pro Leu
        995                 1000                1005

Asn Ile  Ser Gln Val Phe Asp  Glu Val Asp Thr Asp  Gln Ser Gly
    1010                1015                1020

Val Leu  Ser Asp Arg Glu Ile  Arg Thr Leu Ala Thr  Arg Ile His
    1025                1030                1035

Glu Leu  Pro Leu Ser Leu Gln  Asp Leu Thr Gly Leu  Glu His Met
    1040                1045                1050

Leu Ile  Asn Cys Ser Lys Met  Leu Pro Ala Asp Ile  Thr Gln Leu
    1055                1060                1065

Asn Asn  Ile Pro Pro Thr Gln  Glu Ser Tyr Tyr Asp  Pro Asn Leu
    1070                1075                1080

Pro Pro  Val Thr Lys Ser Leu  Val Thr Asn Cys Lys  Pro Val Thr
    1085                1090                1095

Asp Lys  Ile His Lys Ala Tyr  Lys Asp Lys Asn Lys  Tyr Arg Phe
    1100                1105                1110

Glu Ile  Met Gly Glu Glu Glu  Ile Ala Phe Lys Met  Ile Arg Thr
    1115                1120                1125

Asn Val  Ser His Val Val Gly  Gln Leu Asp Asp Ile  Arg Lys Asn
    1130                1135                1140

Pro Arg  Lys Phe Val Cys Leu  Asn Asp Asn Ile Asp  His Asn His
    1145                1150                1155

Lys Asp  Ala Gln Thr Val Lys  Ala Val Leu Arg Asp  Phe Tyr Glu
    1160                1165                1170

Ser Met  Phe Pro Ile Pro Ser  Gln Phe Glu Leu Pro  Arg Glu Tyr
    1175                1180                1185
```

```
Arg Asn  Arg Phe Leu His Met  His Glu Leu Gln Glu  Trp Arg Ala
    1190                 1195                 1200


Tyr Arg  Asp Lys Leu Lys Phe  Trp Thr His Cys Val  Leu Ala Thr
    1205                 1210                 1215


Leu Ile  Met Phe Thr Ile Phe  Ser Phe Phe Ala Glu  Gln Leu Ile
    1220                 1225                 1230


Ala Leu  Lys Arg Lys Ile Phe  Pro Arg Arg Arg Ile  His Lys Glu
    1235                 1240                 1245


Ala Ser  Pro Asn Arg Ile Arg  Val
    1250                 1255


<210>  154
<211>  580
<212>  PRT
<213>  homo sapiens


<220>
<221>  MISC_FEATURE
<223>  MCOLN1

<400>  154

Met Thr Ala Pro Ala Gly Pro Arg Gly Ser Glu Thr Glu Arg Leu Leu
1               5                   10                  15


Thr Pro Asn Pro Gly Tyr Gly Thr Gln Ala Gly Pro Ser Pro Ala Pro
            20                  25                  30


Pro Thr Pro Pro Glu Glu Glu Asp Leu Arg Arg Arg Leu Lys Tyr Phe
        35                  40                  45


Phe Met Ser Pro Cys Asp Lys Phe Arg Ala Lys Gly Arg Lys Pro Cys
    50                  55                  60


Lys Leu Met Leu Gln Val Val Lys Ile Leu Val Val Thr Val Gln Leu
65                  70                  75                  80


Ile Leu Phe Gly Leu Ser Asn Gln Leu Ala Val Thr Phe Arg Glu Glu
                85                  90                  95


Asn Thr Ile Ala Phe Arg His Leu Phe Leu Leu Gly Tyr Ser Asp Gly
            100                 105                 110


Ala Asp Asp Thr Phe Ala Ala Tyr Thr Arg Glu Gln Leu Tyr Gln Ala
        115                 120                 125


Ile Phe His Ala Val Asp Gln Tyr Leu Ala Leu Pro Asp Val Ser Leu
    130                 135                 140


Gly Arg Tyr Ala Tyr Val Arg Gly Gly Gly Asp Pro Trp Thr Asn Gly
```

145 150 155 160

Ser Gly Leu Ala Leu Cys Gln Arg Tyr Tyr His Arg Gly His Val Asp
165 170 175

Pro Ala Asn Asp Thr Phe Asp Ile Asp Pro Met Val Val Thr Asp Cys
180 185 190

Ile Gln Val Asp Pro Pro Glu Arg Pro Pro Pro Pro Pro Ser Asp Asp
195 200 205

Leu Thr Leu Leu Glu Ser Ser Ser Ser Tyr Lys Asn Leu Thr Leu Lys
210 215 220

Phe His Lys Leu Val Asn Val Thr Ile His Phe Arg Leu Lys Thr Ile
225 230 235 240

Asn Leu Gln Ser Leu Ile Asn Asn Glu Ile Pro Asp Cys Tyr Thr Phe
245 250 255

Ser Val Leu Ile Thr Phe Asp Asn Lys Ala His Ser Gly Arg Ile Pro
260 265 270

Ile Ser Leu Glu Thr Gln Ala His Ile Gln Glu Cys Lys His Pro Ser
275 280 285

Val Phe Gln His Gly Asp Asn Ser Phe Arg Leu Leu Phe Asp Val Val
290 295 300

Val Ile Leu Thr Cys Ser Leu Ser Phe Leu Leu Cys Ala Arg Ser Leu
305 310 315 320

Leu Arg Gly Phe Leu Leu Gln Asn Glu Phe Val Gly Phe Met Trp Arg
325 330 335

Gln Arg Gly Arg Val Ile Ser Leu Trp Glu Arg Leu Glu Phe Val Asn
340 345 350

Gly Trp Tyr Ile Leu Leu Val Thr Ser Asp Val Leu Thr Ile Ser Gly
355 360 365

Thr Ile Met Lys Ile Gly Ile Glu Ala Lys Asn Leu Ala Ser Tyr Asp
370 375 380

Val Cys Ser Ile Leu Leu Gly Thr Ser Thr Leu Leu Val Trp Val Gly
385 390 395 400

Val Ile Arg Tyr Leu Thr Phe Phe His Asn Tyr Asn Ile Leu Ile Ala
405 410 415

Thr Leu Arg Val Ala Leu Pro Ser Val Met Arg Phe Cys Cys Cys Val
420 425 430

Ala Val Ile Tyr Leu Gly Tyr Cys Phe Cys Gly Trp Ile Val Leu Gly
435 440 445

Pro Tyr His Val Lys Phe Arg Ser Leu Ser Met Val Ser Glu Cys Leu
450 455 460

Phe Ser Leu Ile Asn Gly Asp Asp Met Phe Val Thr Phe Ala Ala Met
465 470 475 480

Gln Ala Gln Gln Gly Arg Ser Ser Leu Val Trp Leu Phe Ser Gln Leu
485 490 495

Tyr Leu Tyr Ser Phe Ile Ser Leu Phe Ile Tyr Met Val Leu Ser Leu
500 505 510

Phe Ile Ala Leu Ile Thr Gly Ala Tyr Asp Thr Ile Lys His Pro Gly
515 520 525

Gly Ala Gly Ala Glu Glu Ser Glu Leu Gln Ala Tyr Ile Ala Gln Cys
530 535 540

Gln Asp Ser Pro Thr Ser Gly Lys Phe Arg Arg Gly Ser Gly Ser Ala
545 550 555 560

Cys Ser Leu Leu Cys Cys Cys Gly Arg Asp Pro Ser Glu Glu His Ser
565 570 575

Leu Leu Val Asn
580

<210> 155
<211> 5
<212> PRT
<213> artificial sequence

<220>
<223> Caspase 2 protein sequence

<400> 155

Val Asp Val Ala Asp
1 5

<210> 156
<211> 22
<212> RNA
<213> artificial sequence

<220>
<223> miR-338-3p

<400> 156
uccagcauca gugauuuugu ug 22

<210> 157

<211> 23
<212> RNA
<213> artificial sequence

<220>
<223> miR-9-5p

<400> 157
ucuuugguua ucuagcugua uga 23

<210> 158
<211> 22
<212> RNA
<213> artificial sequence

<220>
<223> miR-9-3p

<400> 158
auaaagcuag auaaccgaaa gu 22

<210> 159
<211> 22
<212> RNA
<213> artificial sequence

<220>
<223> miR-125b-5p

<400> 159
ucccugagac ccuaacuugu ga 22

<210> 160
<211> 23
<212> RNA
<213> artificial sequence

<220>
<223> miR-342-3p

<400> 160
ucucacacag aaaucgcacc cgu 23

<210> 161
<211> 22
<212> RNA
<213> artificial sequence

<220>
<223> miR-124-3p

<400> 161
uaaggcacgc ggugaaugcc aa 22

<210> 162
<211> 22
<212> RNA
<213> artificial sequence

<220>
<223> miR-124-5p

<400> 162
cguguucaca gcggaccuug au 22