

TARGETED LIPID PARTICLES AND COMPOSITIONS AND USES THEREOFCross-Reference to Related Applications

[0001] This application claims priority to U.S. provisional application 63/003,168 entitled “Targeted Lipid Particles and Compositions and Uses Thereof”, filed March 31, 2020, and to U.S. provisional application 63/154,341, entitled “Targeted Lipid Particles and Compositions and Uses Thereof”, filed February 26, 2021, the contents of each of which are incorporated by reference in their entirety for all purposes.

Incorporation by Reference of Sequence Listing

[0002] The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing is provided as a file entitled 186152003640SeqList.TXT, created March 29, 2021, which is 2,071,683 bytes in size. The information in the electronic format of the Sequence Listing is incorporated by reference in its entirety

Field

[0003] The present disclosure relates to lipid particles containing a lipid bilayer enclosing a lumen or cavity, a henipavirus F protein molecule or biologically active portion thereof, and a targeted envelope protein containing a henipavirus envelope attachment glycoprotein G (G protein) or biologically active portion thereof and a binding domain, such as a single domain antibody (sdAb) variable domain. The present disclosure also provides a targeted envelope protein containing a G protein fused or linked to a binding domain, such as a sdAb variable domain, and polynucleotides encoding such proteins. Also disclosed are producer cells and compositions containing such targeted lipid particles and methods of making and using the targeted lipid particles.

Background

[0004] Lipid particles, including virus-like particles and viral vectors, are commonly used for delivery of exogenous agents to cells. However, delivery of the lipid particles to certain target cells can be challenging. For lentiviral vectors, the host range can be altered by pseudotyping with a heterologous envelope protein. Certain retargeted envelope proteins may not be sufficiently stable or expressed on the surface of the lipid particle. Improved lipid

particles, including virus-like particles and viral vectors, for targeting desired cells are needed. The provided disclosure addresses this need.

Summary

[0005] Provided herein is a targeted lipid particle which includes (a) a lipid bilayer enclosing a lumen, (b) a henipavirus F protein molecule or biologically active portion thereof; and (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) single domain antibody (sdAb) variable domain, wherein the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof, wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer. In some embodiments, the single domain antibody is attached to the G protein via a linker. In some embodiments, the linker is a peptide linker.

[0006] Provided herein is a targeted lipid particle which includes (a) a lipid bilayer enclosing a lumen, (b) a henipavirus F protein molecule or biologically active portion thereof; and (c) a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or biologically active portion thereof attached to a single domain antibody (sdAb) variable domain via a peptide linker, wherein the single domain antibody binds to a cell surface molecule of a target cell, wherein the F protein molecule or biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer. In some embodiments, N-terminus of the F protein molecule or biologically active portion thereof is exposed on the outside of lipid bilayer. In some embodiments, the C-terminus of the G protein is exposed on the outside of the lipid bilayer.

[0007] In some embodiments, the single domain antibody binds a cell surface molecule present on a target cell. In some embodiments, the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule. In some of any embodiments, the single domain antibody binds an antigen or portion thereof present on a target cell. In some embodiments, the antigen is the cell surface molecule or a portion of the cell surface molecule that contains an epitope recognized by the single domain antibody. In some of any embodiments, the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells. In some

embodiments, the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell. In some of any embodiments, the target cell is a hepatocyte. In some of any embodiments, the cell surface molecule or antigen is selected from the group consisting of ASGR1, ASGR2 and TM4SF5.

[0008] In some of any embodiments, the target cell is a T cell. In some of any embodiments, the cell surface molecule or antigen is CD8 or CD4.

[0009] In some of any embodiments, the cell surface molecule or antigen is LDL-R.

[0010] Provided herein are targeted lipid particles comprising (a) a lipid bilayer enclosing a lumen, (b) a henipavirus F protein molecule or biologically active portion thereof; and (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2,

wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.

[0011] Provided herein are targeted lipid particles comprising (a) a lipid bilayer enclosing a lumen, (b) a henipavirus F protein molecule or biologically active portion thereof; and (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds a cell surface molecule selected from the group consisting of CD8 and CD4, , optionally human CD8 or human CD4, wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.

[0012] Provided herein are targeted lipid particles comprising (a) a lipid bilayer enclosing a lumen, (b) a henipavirus F protein molecule or biologically active portion thereof; and (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R, wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.

[0013] In some of any embodiments, the lipid particle is a lentiviral vector. In some of any embodiments, the binding domain is attached to the G protein via a linker. In some of any embodiments, the linker is a peptide linker.

[0014] Provided herein is a lentiviral vector, comprising a binding domain that targets a cell surface molecule selected from the group consisting of ASGR1, ASGR2 and TM4SF5, optionally human ASGR1, human ASGR2 and human TM4SF5, wherein the lentiviral vector is pseudotyped with a retargeted viral fusion protein, said retargeted viral fusion protein comprising: (a) a henipavirus F protein molecule or biologically active portion thereof; and (b) a targeted envelope protein comprising the binding domain attached to a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof.

[0015] Provided herein is a lentiviral vector, comprising a binding domain that targets a cell surface molecule selected from the group consisting of CD8 and CD4, optionally human CD8 and human CD4, wherein the lentiviral vector is pseudotyped with a retargeted viral fusion protein, said retargeted viral fusion protein comprising: (a) a henipavirus F protein molecule or biologically active portion thereof; and (b) a targeted envelope protein comprising the binding domain attached to a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof.

[0016] Provided herein is a lentiviral vector, comprising a binding domain that targets low density lipoprotein receptor (LDL-R), optionally wherein the LDL-R is human LDL-R, wherein the lentiviral vector is pseudotyped with a retargeted viral fusion protein comprising (a) a henipavirus F protein molecule or biologically active portion thereof; and (b) a targeted envelope protein comprising the binding domain attached to a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof.

[0017] In some of any embodiments, the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof.

[0018] Provided herein is a lentiviral vector, comprising (a) a henipavirus F protein molecule or biologically active portion thereof; and (b) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds CD4; and (c) a cargo comprising nucleic acid encoding a chimeric antigen receptor (CAR), wherein the CAR comprises (i) an extracellular antigen binding domain that binds an extracellular antigen (e.g., CD19 or BCMA) and (ii) an intracellular signaling region a CD3zeta signaling domain and, optionally a 4-1BB or CD28 co-stimulatory signaling domain. In some embodiments, the extracellular antigen binding domain of the CAR is an scFv.

[0019] In some of any embodiments, the lentiviral vector is capable of delivering the nucleic acid encoding the CAR to T cells. In some embodiments the T cells are in vivo in a subject.

[0020] Provided herein is a lentiviral vector, comprising: (a) a henipavirus F protein molecule or biologically active portion thereof; and (b) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds ASGR1; wherein the lentiviral vector is capable of targeting to hepatocytes. In some of any embodiments, the lentiviral vector further comprises an exogenous agent for delivery to hepatocytes.

[0021] In some of any embodiments, the lentiviral vector is capable of delivering the exogenous agent to hepatocytes, optionally wherein the hepatocytes are in vivo in a subject.

[0022] In some of any embodiments, the binding domain is attached to the G protein via a linker. In some of any embodiments, the linker is a peptide linker. In some of any embodiments, the binding domain is a single domain antibody. In some of any embodiments, the binding domain is a single chain variable fragment (scFv).

[0023] In some of any embodiments, the peptide linker comprises up to 65 amino acids in length. In some of any embodiments, the peptide linker comprises up to 50 amino acids

in length. In some of any embodiments, the peptide linker comprises from or from about 2 to 65 amino acids, 2 to 60 amino acids, 2 to 56 amino acids, 2 to 52 amino acids, 2 to 48 amino acids, 2 to 44 amino acids, 2 to 40 amino acids, 2 to 36 amino acids, 2 to 32 amino acids, 2 to 28 amino acids, 2 to 24 amino acids, 2 to 20 amino acids, 2 to 18 amino acids, 2 to 14 amino acids, 2 to 12 amino acids, 2 to 10 amino acids, 2 to 8 amino acids, 2 to 6 amino acids, 6 to 65 amino acids, 6 to 60 amino acids, 6 to 56 amino acids, 6 to 52 amino acids, 6 to 48 amino acids, 6 to 44 amino acids, 6 to 40 amino acids, 6 to 36 amino acids, 6 to 32 amino acids, 6 to 28 amino acids, 6 to 24 amino acids, 6 to 20 amino acids, 6 to 18 amino acids, 6 to 14 amino acids, 6 to 12 amino acids, 6 to 10 amino acids, 6 to 8 amino acids, 8 to 65 amino acids, 8 to 60 amino acids, 8 to 56 amino acids, 8 to 52 amino acids, 8 to 48 amino acids, 8 to 44 amino acids, 8 to 40 amino acids, 8 to 36 amino acids, 8 to 32 amino acids, 8 to 28 amino acids, 8 to 24 amino acids, 8 to 20 amino acids, 8 to 18 amino acids, 8 to 14 amino acids, 8 to 12 amino acids, 8 to 10 amino acids, 10 to 65 amino acids, 10 to 60 amino acids, 10 to 56 amino acids, 10 to 52 amino acids, 10 to 48 amino acids, 10 to 44 amino acids, 10 to 40 amino acids, 10 to 36 amino acids, 10 to 32 amino acids, 10 to 28 amino acids, 10 to 24 amino acids, 10 to 20 amino acids, 10 to 18 amino acids, 10 to 14 amino acids, 10 to 12 amino acids, 12 to 65 amino acids, 12 to 60 amino acids, 12 to 56 amino acids, 12 to 52 amino acids, 12 to 48 amino acids, 12 to 44 amino acids, 12 to 40 amino acids, 12 to 36 amino acids, 12 to 32 amino acids, 12 to 28 amino acids, 12 to 24 amino acids, 12 to 20 amino acids, 12 to 18 amino acids, 12 to 14 amino acids, 14 to 65 amino acids, 14 to 60 amino acids, 14 to 56 amino acids, 14 to 52 amino acids, 14 to 48 amino acids, 14 to 44 amino acids, 14 to 40 amino acids, 14 to 36 amino acids, 14 to 32 amino acids, 14 to 28 amino acids, 14 to 24 amino acids, 14 to 20 amino acids, 14 to 18 amino acids, 18 to 65 amino acids, 18 to 60 amino acids, 18 to 56 amino acids, 18 to 52 amino acids, 18 to 48 amino acids, 18 to 44 amino acids, 18 to 40 amino acids, 18 to 36 amino acids, 18 to 32 amino acids, 18 to 28 amino acids, 18 to 24 amino acids, 18 to 20 amino acids, 20 to 65 amino acids, 20 to 60 amino acids, 20 to 56 amino acids, 20 to 52 amino acids, 20 to 48 amino acids, 20 to 44 amino acids, 20 to 40 amino acids, 20 to 36 amino acids, 20 to 32 amino acids, 20 to 28 amino acids, 20 to 26 amino acids, 20 to 24 amino acids, 24 to 65 amino acids, 24 to 60 amino acids, 24 to 56 amino acids, 24 to 52 amino acids, 24 to 48 amino acids, 24 to 44 amino acids, 24 to 40 amino acids, 24 to 36 amino acids, 24 to 32 amino acids, 24 to 30 amino acids, 24 to 28 amino acids, 28 to 65 amino acids, 28 to 60 amino acids, 28 to 56 amino acids, 28 to 52 amino acids, 28 to 48 amino acids, 28 to 44 amino acids, 28 to 40 amino acids, 28 to 36 amino acids, 28 to 34 amino acids,

28 to 32 amino acids, 32 to 65 amino acids, 32 to 60 amino acids, 32 to 56 amino acids, 32 to 52 amino acids, 32 to 48 amino acids, 32 to 44 amino acids, 32 to 40 amino acids, 32 to 38 amino acids, 32 to 36 amino acids, 36 to 65 amino acids, 36 to 60 amino acids, 36 to 56 amino acids, 36 to 52 amino acids, 36 to 48 amino acids, 36 to 44 amino acids, 36 to 40 amino acids, 40 to 65 amino acids, 40 to 60 amino acids, 40 to 56 amino acids, 40 to 52 amino acids, 40 to 48 amino acids, 40 to 44 amino acids, 44 to 65 amino acids, 44 to 60 amino acids, 44 to 56 amino acids, 44 to 52 amino acids, 44 to 48 amino acids, 48 to 65 amino acids, 48 to 60 amino acids, 48 to 56 amino acids, 48 to 52 amino acids, 50 to 65 amino acids, 50 to 60 amino acids, 50 to 56 amino acids, 50 to 52 amino acids, 54 to 65 amino acids, 54 to 60 amino acids, 54 to 56 amino acids, 58 to 65 amino acids, 58 to 60 amino acids, or 60 to 65 amino acids. In some of any embodiments, peptide linker comprises a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64 or 65 amino acids in length. In some of any embodiments, wherein the peptide linker is a flexible linker that comprises GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) or combinations thereof. In some of any embodiments, the peptide linker comprises (GGS)_n, wherein n is 1 to 10. In some of any embodiments, the peptide linker comprises (GGGGS)_n (SEQ ID NO: 42), wherein n is 1 to 10. In some of any embodiments, the peptide linker comprises (GGGGGS)_n (SEQ ID NO:27), wherein n is 1 to 6.

[0024] In some of any embodiments, the G protein or the biologically active portion thereof is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein. In some of any embodiments, the G protein or the biologically active portion thereof is a wild-type NiV-G protein or a functionally active variant or biologically active portion thereof. In some of any embodiments, the mutant NiV-G protein or functionally active variant or biologically active portion thereof comprises an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44.

[0025] In some of any embodiments, the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

[0026] In some of any embodiments, the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and has the sequence set forth in any of SEQ ID NOS: 10-15, 35-40 or 45-50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40 or 45-50.

[0027] In some of any embodiments, the NiV-G protein is a biologically active portion that has a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 10 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 35 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 45 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about

82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45.

[0028] In some of any embodiments, the NiV-G protein is a biologically active portion that has a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 36 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 11 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 46 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

[0029] In some of any embodiments, the NiV-G protein or the biologically active portion has a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G

protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 12 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 37 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 47 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47.

[0030] In some of any embodiments, the NiV-G protein is a biologically active portion that has a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 13 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or

about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 38 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 48 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48.

[0031] In some of any embodiments, the NiV-G protein is a biologically active portion has a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 14 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 39 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or

about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 49 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49.

[0032] In some of any embodiments, the NiV-G protein is a biologically active portion has a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 15 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 40 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40.

[0033] In some of any embodiments, the NiV-G protein is a biologically active portion that has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 22 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at

least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:22. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 53 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:53.

[0034] In some of any embodiments, the G-protein, the biologically active portion thereof is a functionally active variant that is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3.

[0035] In some of any embodiments, the mutant NiV-G protein includes one or more amino acid substitutions corresponding to amino acid substitutions selected from the group consisting of E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28. In some of any embodiments, the mutant NiV-G protein includes the amino acid substitutions E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

[0036] In some of any embodiments, the mutant NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 16 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16. In some of any embodiments, the mutant NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 51 or an amino acid sequence having at least at or

about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:51.

[0037] In some of any embodiments, the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a Hendra virus F protein or is a functionally active variant or biologically active portion thereof. In some of any embodiments, the F protein or the biologically active portion thereof is a wild-type NiV-F protein or a functionally active variant or a biologically active portion thereof. In some of any embodiments, the NiV-F-protein or the functionally active variant or biologically active portion thereof comprises the amino acid sequences set forth in SEQ ID NO: 2, or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 2.

[0038] In some of any embodiments, the NiV-F protein is a biologically active portion thereof that has a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2).

[0039] In some of any embodiments, the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:5 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 5.

[0040] In some of any embodiments, the NiV-F protein is a biologically active portion thereof that includes i) a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2); and ii) a point mutation on an N-linked glycosylation site.

[0041] In some of any embodiments, the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:7 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 7.

[0042] In some of any embodiments, the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2).

[0043] In some of any embodiments, NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:8 or an amino acid sequence that is encoded by a sequence of nucleotides encoding a sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 8.

[0044] In some of any embodiments, the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:23 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 23. In some of any embodiments, the F-protein or the biologically active portion thereof comprises an F1 subunit or a fusogenic portion thereof.

[0045] In some of any embodiments, the F protein comprises the sequence set forth in SEQ ID NO:23 and the G protein comprises the sequence set forth in SEQ ID NO:16.

[0046] In some of any embodiments, the F protein consists or consists essentially of the sequence set forth in SEQ ID NO:23 and/or the G protein consists or consists essentially of the sequence set forth in SEQ ID NO:16.

[0047] In some of any embodiments, the F1 subunit is a proteolytically cleaved portion of the F0 precursor. In some of any embodiments, the F1 subunit comprises the sequence set forth in SEQ ID NO: 4, or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:4.

[0048] In some of any embodiments, the lipid bilayer is derived from a membrane of a host cell used for producing a retrovirus or retrovirus-like particle. In some of any embodiments, the host cell is selected from the group consisting of 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 some of any embodiments, the host cell comprises 293T cells. In some of any embodiments, the lipid bilayer is or comprises a viral envelope. In some of any embodiments, the retrovirus-like particle is replication defective.

[0049] In some of any embodiments, the targeted lipid particle comprises one or more viral components other than the F protein molecule and the G protein. In some of any embodiments, the one or more viral components are from a retrovirus. In some of any embodiments, the retrovirus is a lentivirus. In some of any embodiments, the one or more viral components comprise a viral packaging protein selected from one or more of Gag, Pol, Rev and Tat. In some of any embodiments, the one or more viral components 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)/central termination

sequence (CTS) (e.g. DNA flap), Poly A tail sequence, a posttranscriptional regulatory element (e.g. WPRE), a Rev response element (RRE), and 3' LTR (e.g., comprising U5 and lacking a functional U3).

[0050] In some of any embodiments, the targeted lipid particle is a lentiviral vector.

[0051] In some of any embodiments, the targeted lipid particle or the lentiviral vector is replication defective.

[0052] In some of any embodiments, the targeted lipid particle or the lentiviral vector further comprises an exogenous agent. In some of any embodiments, the targeted lipid particle further comprises an exogenous agent. In some embodiments, the lentiviral vector further comprises an exogenous agent.

[0053] In some of any embodiments, the exogenous agent is present in the lumen. In some of any embodiments, the exogenous agent is a protein or a nucleic acid. In some embodiments, the nucleic acid is a DNA or RNA.

[0054] In some of any embodiments, the exogenous agent is a nucleic acid encoding a cargo for delivery to the target cell. In some of any embodiments, the exogenous agent encodes a therapeutic agent or a diagnostic agent.

[0055] In some of any embodiments, the exogenous agent encodes a membrane protein. In some embodiments, the membrane protein is an antigen receptor for targeting cells expressed by or associated with a disease or condition. In some embodiments, the membrane protein is a chimeric antigen receptor (CAR). In some embodiments, the CAR comprises (i) an extracellular antigen binding domain that binds an extracellular antigen (e.g., CD19 or BCMA), optionally wherein the extracellular antigen binding domain is an scFv, (ii) a transmembrane domain and (iii) an intracellular signaling region comprising a CD3zeta signaling domain and, optionally a co-stimulatory signaling domain, e.g., a 4-1BB or CD28 co-stimulatory signaling domain. In some embodiments, the target cell is a T cell. In some embodiments, the cell surface molecule on the target cell is CD4 or CD8. In some embodiments, the binding domain is an scFv that binds CD4 (e.g. human CD4). In some embodiments, the binding domain is a single domain antibody that binds CD4 (e.g. human CD4). In some embodiments, the binding domain is an scFv that binds CD8 (e.g. human CD8). In some embodiments, the binding domain is a single domain antibody that binds CD8 (e.g. human CD8).

[0056] In some of any embodiments, the exogenous agent is a nucleic acid comprising a payload gene for correcting a genetic deficiency, optionally a genetic deficiency in the target cell. In some embodiments, the genetic deficiency is associated with a liver cell or a hepatocyte. In some embodiments, the target cell is a hepatocyte. In some embodiments, the cell surface molecule is a molecule selected from the group consisting of ASGR1, ASGR2 and TM4SF5. In some embodiments, the binding domain is an scFv that binds ASGR1 (e.g. human ASGR1). In some embodiments, the binding domain is a single domain antibody that binds ASGR1 (e.g. human ASGR1). In some embodiments, the binding domain is an scFv that binds ASGR2 (e.g. human ASGR2). In some embodiments, the binding domain is a single domain antibody that binds ASGR2 (e.g. human ASGR2). In some embodiment, the binding domain is a scFv that binds TM4SF5 (e.g. human TM4SF5). In some embodiments, the binding domain is a single domain antibody that binds TM4SF5 (e.g. human TM4SF5).

[0057] In some of any embodiments, the single domain antibody binds a cell surface molecule present on a target cell. In some of any embodiments, the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule. In some of any embodiments, the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells. In some of any embodiments, the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

[0058] In some of any embodiments, the single domain antibody binds an antigen or portion thereof present on a target cell. In some of any embodiments, the cell surface molecule or antigen is selected from the group consisting of ASGR1, ASGR2 and TM4SF5. In some embodiments, the antigen or portion thereof is human ASGR1. In some embodiments, the antigen or portion thereof is human ASGR2. In some embodiments, the antigen or portion thereof is human TM4SF5.

[0059] Provided herein is a polynucleotide comprising a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain that binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5. In some embodiments, the cell surface molecule is human ASGR1. In some embodiments, the cell surface molecule is human ASGR2. In some embodiments, the cell surface molecule is human TM4SF5. In some of any embodiments, the cell surface molecule or antigen is CD8 or CD4.

[0060] Provided herein is a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain that binds a cell surface molecule selected from the group consisting of CD4 and CD8. In some embodiments, the cell surface molecule is human CD4. In some embodiments, the cell surface molecule is human CD8. In some embodiments, the cell surface molecule or antigen is low density lipoprotein receptor (LDL-R). In some embodiments, the cell surface molecule or antigen is human LDL-R.

[0061] Provided herein is a polynucleotide comprising a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain that binds low density lipoprotein receptor (LDL-R). In some embodiments, the binding domain binds human LDL-R. In some of any embodiments, the binding domain is a single domain antibody (sdAb). In some of any embodiments, the binding domain is a single chain variable fragment (scFv).

[0062] Provided herein is a polynucleotide comprising a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a single domain antibody (sdAb) variable domain, wherein the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof. In some of any embodiments, the polynucleotide further comprises (iii) a nucleic acid sequence encoding a henipavirus F protein molecule or a biologically active portion thereof.

[0063] In some embodiments, the nucleic acid sequence is a first nucleic acid sequence and the polynucleotide further comprise a second nucleic acid sequence encoding a henipavirus F protein molecule or a biologically active portion thereof. In some embodiments, the polynucleotide comprise an IRES or a sequence encoding a linking peptide between the first and second nucleic acid sequence. In some embodiments, the

linking peptide is a self-cleaving peptide or a peptide that causes ribosome skipping, optionally a T2A peptide.

[0064] In some of any embodiments, the polynucleotide includes at least one promoter that is operatively linked to control expression of the nucleic acid. In some of any embodiments, the promoter is operatively linked to control expression of the first nucleic acid sequence and the second nucleic acid sequence. In some of any embodiments, the promoter is a constitutive promoter. In some of any embodiments, the promoter is an inducible promoter.

[0065] In some of any embodiments, the sdAb variable domain is attached to the G protein via an encoded peptide linker. In some embodiments, the binding domain is attached to the G protein via an encoded peptide linker. In some of any embodiments, the encoded peptide linker comprises up to 25 amino acids in length. In some of any embodiments, the encoded peptide linker comprises up to 65 amino acids in length. In some of any embodiments, the encoded peptide linker comprises from or from about 2 to 65 amino acids, 2 to 60 amino acids, 2 to 56 amino acids, 2 to 52 amino acids, 2 to 48 amino acids, 2 to 44 amino acids, 2 to 40 amino acids, 2 to 36 amino acids, 2 to 32 amino acids, 2 to 28 amino acids, 2 to 24 amino acids, 2 to 20 amino acids, 2 to 18 amino acids, 2 to 14 amino acids, 2 to 12 amino acids, 2 to 10 amino acids, 2 to 8 amino acids, 2 to 6 amino acids, 6 to 65 amino acids, 6 to 60 amino acids, 6 to 56 amino acids, 6 to 52 amino acids, 6 to 48 amino acids, 6 to 44 amino acids, 6 to 40 amino acids, 6 to 36 amino acids, 6 to 32 amino acids, 6 to 28 amino acids, 6 to 24 amino acids, 6 to 20 amino acids, 6 to 18 amino acids, 6 to 14 amino acids, 6 to 12 amino acids, 6 to 10 amino acids, 6 to 8 amino acids, 8 to 65 amino acids, 8 to 60 amino acids, 8 to 56 amino acids, 8 to 52 amino acids, 8 to 48 amino acids, 8 to 44 amino acids, 8 to 40 amino acids, 8 to 36 amino acids, 8 to 32 amino acids, 8 to 28 amino acids, 8 to 24 amino acids, 8 to 20 amino acids, 8 to 18 amino acids, 8 to 14 amino acids, 8 to 12 amino acids, 8 to 10 amino acids, 10 to 65 amino acids, 10 to 60 amino acids, 10 to 56 amino acids, 10 to 52 amino acids, 10 to 48 amino acids, 10 to 44 amino acids, 10 to 40 amino acids, 10 to 36 amino acids, 10 to 32 amino acids, 10 to 28 amino acids, 10 to 24 amino acids, 10 to 20 amino acids, 10 to 18 amino acids, 10 to 14 amino acids, 10 to 12 amino acids, 12 to 65 amino acids, 12 to 60 amino acids, 12 to 56 amino acids, 12 to 52 amino acids, 12 to 48 amino acids, 12 to 44 amino acids, 12 to 40 amino acids, 12 to 36 amino acids, 12 to 32 amino acids, 12 to 28 amino acids, 12 to 24 amino acids, 12 to 20 amino acids, 12 to 18 amino acids, 12 to 14 amino acids, 14 to 65 amino acids, 14 to 60 amino acids, 14 to 56 amino acids, 14 to 52 amino acids, 14 to 48 amino acids, 14 to 44 amino acids,

14 to 40 amino acids, 14 to 36 amino acids, 14 to 32 amino acids, 14 to 28 amino acids, 14 to 24 amino acids, 14 to 20 amino acids, 14 to 18 amino acids, 18 to 65 amino acids, 18 to 60 amino acids, 18 to 56 amino acids, 18 to 52 amino acids, 18 to 48 amino acids, 18 to 44 amino acids, 18 to 40 amino acids, 18 to 36 amino acids, 18 to 32 amino acids, 18 to 28 amino acids, 18 to 24 amino acids, 18 to 20 amino acids, 20 to 65 amino acids, 20 to 60 amino acids, 20 to 56 amino acids, 20 to 52 amino acids, 20 to 48 amino acids, 20 to 44 amino acids, 20 to 40 amino acids, 20 to 36 amino acids, 20 to 32 amino acids, 20 to 28 amino acids, 20 to 26 amino acids, 20 to 24 amino acids, 24 to 65 amino acids, 24 to 60 amino acids, 24 to 56 amino acids, 24 to 52 amino acids, 24 to 48 amino acids, 24 to 44 amino acids, 24 to 40 amino acids, 24 to 36 amino acids, 24 to 32 amino acids, 24 to 30 amino acids, 24 to 28 amino acids, 28 to 65 amino acids, 28 to 60 amino acids, 28 to 56 amino acids, 28 to 52 amino acids, 28 to 48 amino acids, 28 to 44 amino acids, 28 to 40 amino acids, 28 to 36 amino acids, 28 to 34 amino acids, 28 to 32 amino acids, 32 to 65 amino acids, 32 to 60 amino acids, 32 to 56 amino acids, 32 to 52 amino acids, 32 to 48 amino acids, 32 to 44 amino acids, 32 to 40 amino acids, 32 to 38 amino acids, 32 to 36 amino acids, 36 to 65 amino acids, 36 to 60 amino acids, 36 to 56 amino acids, 36 to 52 amino acids, 36 to 48 amino acids, 36 to 44 amino acids, 36 to 40 amino acids, 40 to 65 amino acids, 40 to 60 amino acids, 40 to 56 amino acids, 40 to 52 amino acids, 40 to 48 amino acids, 40 to 44 amino acids, 44 to 65 amino acids, 44 to 60 amino acids, 44 to 56 amino acids, 44 to 52 amino acids, 44 to 48 amino acids, 48 to 65 amino acids, 48 to 60 amino acids, 48 to 56 amino acids, 48 to 52 amino acids, 50 to 65 amino acids, 50 to 60 amino acids, 50 to 56 amino acids, 50 to 52 amino acids, 54 to 65 amino acids, 54 to 60 amino acids, 54 to 56 amino acids, 58 to 65 amino acids, 58 to 60 amino acids, or 60 to 65 amino acids.

[0066] In some of any embodiments, the encoded peptide linker comprises a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64 or 65 amino acids in length. In some of any embodiments, the encoded peptide linker comprises GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) and combinations thereof. In some of any embodiments, the encoded peptide linker comprises (GGS)*n*, wherein *n* is 1 to 10. In some of any embodiments, the encoded peptide linker comprises (GGGGS)*n* (SEQ ID NO:42), wherein *n* is 1 to 10. In some of any embodiments, the encoded peptide linker comprises (GGGGGS)*n* (SEQ ID NO:27), wherein *n* is 1 to 4. . In some of any embodiments, the sequence encoding the G protein is a wild-type

Nipah virus G (NiV-G) protein or a Hendra virus G protein or is a functionally active variant or a biologically active portion thereof. In some embodiments, the variant is a variant thereof that exhibits reduced binding for the native binding partner. In some of any embodiments, the nucleic acid sequence encoding the G protein is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein or is a variant thereof that exhibits reduced binding for the native binding partner. In some embodiments, the encoded G protein is a wild-type NiV-G protein or a functionally active variant or a biologically active portion thereof. In some of any embodiments, the nucleic acid sequence encoding the G protein is a wild-type NiV-G protein. In some of any embodiments, the nucleic acid sequence encoding the G-protein is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3.

[0067] In some of any embodiments, the NiV-G protein or functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO:9, SEQ ID NO: 28 or SEQ ID NO: 44 or comprises an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44. In some of any embodiments, the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and comprises the sequence set forth in any of SEQ ID NOS: 10-15, 35-40 or 45-50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40 or 45-50.

[0068] In some of any embodiments, the NiV-G protein is a biologically active portion that comprises a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 10 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10. In some of any embodiments, NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 35 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 45 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45.

[0069] In some of any embodiments, NiV-G protein is a biologically active portion that comprises a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the mutant NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 11 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or

about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 36 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 46 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

[0070] In some of any embodiments, there is a biologically active portion that NiV-G protein comprises a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 12 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 37

or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 47 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47.

[0071] In some of any embodiments, the NiV-G protein is a biologically active portion that comprises a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 13 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13. In some of any embodiments, NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 38 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38. In some of any

embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 48 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48.

[0072] In some of any embodiments, the NiV-G protein is a biologically active portion that comprises a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 14 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 39 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 49 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or

about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49.

[0073] In some of any embodiments, the NiV-G protein is a biologically active portion that comprises a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44). In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 15 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 40 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40. In some of any embodiments, the NiV-G protein or the biologically active portion comprises the amino acid sequence set forth in SEQ ID NO: 50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 50.

[0074] In some of any embodiments, the NiV-G protein is a biologically active portion that has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9,SEQ ID NO:28 or SEQ ID NO:44).In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in

SEQ ID NO: 22 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:22. In some of any embodiments, the NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 53 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:53.

[0075] In some of any embodiments, the G-protein is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3. In some of any embodiments, the mutant NiV-G protein comprises: one or more amino acid substitutions corresponding to amino acid substitutions selected from the group consisting of E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28. In some of any embodiments, the mutant NiV-G protein comprises amino acid substitutions E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

[0076] In some of any embodiments, the mutant NiV-G protein comprises: i) a truncation at or near the N-terminus; and ii) point mutations selected from the group consisting of E501A, W504A, Q530A and E533A. In some of any embodiments, the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 16 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16. In some of any embodiments, the mutant NiV-G protein comprises the amino acid sequence set forth in

SEQ ID NO: 51 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:51.

[0077] In some of any embodiments, the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a Hendra virus F protein or is a functionally active variant or biologically active portion thereof. In some of any embodiments, the F protein or the biologically active portion thereof is a wild-type NiV-F protein or a functionally active variant or a biologically active portion thereof. In some of any embodiments, the NiV-F-protein or the functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO: 2, or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 2.

[0078] In some of any embodiments, the NiV-F protein is a biologically active portion thereof that has a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2). In some of any embodiments, the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:5 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 5. In some of any embodiments, the NiV-F protein is a biologically active portion thereof that comprises i) a 20 amino acid truncation at or near the

C-terminus of the wild-type NiV-F protein (SEQ ID NO:2); and ii) a point mutation on an N-linked glycosylation site.

[0079] In some of any embodiments, the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:7 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 7.

[0080] In some of any embodiments, the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2). In some of any embodiments, the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:8 or an amino acid sequence that is encoded by a sequence of nucleotides encoding a sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 8.

[0081] In some of any embodiments, the NiV-F protein has the sequence set forth in SEQ ID NO:23 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 23. In some of any embodiments, the F protein comprises the sequence set forth in SEQ ID NO:23 and the G protein comprises the sequence set forth in SEQ ID NO:16. In some of any embodiments, the F protein consists or consists essentially of the sequence set forth in SEQ

ID NO:23 and the G protein consists or consists essentially of the sequence set forth in SEQ ID NO:16.

[0082] Provided herein is a vector, comprising the polynucleotide of any of the embodiments described herein. In some of any embodiments, the vector is a mammalian vector, viral vector or artificial chromosome, optionally wherein the artificial chromosome is a bacterial artificial chromosome (BAC).

[0083] Provided herein is a plasmid, comprising the polynucleotide of any of the embodiments described herein. In some of any embodiments, the plasmid further comprises one or more nucleic acids encoding proteins for lentivirus production.

[0084] Provided herein is a cell comprising the polynucleotide of any of embodiments described herein or the vector of any of the embodiments described herein, or the plasmid of any of the embodiments described herein.

[0085] Provided herein is a method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain, the method comprising a) providing a cell that comprises a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and a nucleic acid encoding a targeted envelope protein, the targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain; b) culturing the cell under conditions that allow for production of a targeted lipid particle, and c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

[0086] Provided herein is a method of making a pseudotyped lentiviral vector, the method comprising a) providing a producer cell that comprises a lentiviral viral nucleic acid(s), a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof, and a nucleic acid encoding a targeted envelope protein, said targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody; b) culturing the cell under conditions that allow for production of the lentiviral vector, and c) separating, enriching, or purifying the lentiviral vector from the cell, thereby making the pseudotyped lentiviral vector.

[0087] In some of any embodiments, the single domain antibody binds a cell surface molecule present on a target cell. In some of any embodiments, the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule. In some of any embodiments, the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells. In some of any embodiments, the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell. In some of any embodiments, the single domain antibody binds an antigen or portion thereof present on a target cell.

[0088] Provided herein is a method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a binding domain, the method comprising a) providing a cell that comprises a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and a nucleic acid encoding a targeted envelope protein, the targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and binding domain, wherein the binding domain (i) binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2; (ii) binds a cell surface molecule selected from the group consisting of CD4 or CD8, optionally human CD4 or human CD8; or (iii) binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R; b) culturing the cell under conditions that allow for production of a targeted lipid particle, and c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

[0089] Provided herein is a method of making a pseudotyped lentiviral vector, the method comprising a) providing a producer cell that comprises a lentiviral viral nucleic acid(s), a

nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof, and a nucleic acid encoding a targeted envelope protein, said targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and binding domain, wherein the binding domain: (i) binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2; (ii) binds a cell surface molecule selected from the group consisting of CD4 or CD8, optionally human CD4 or human CD8; or (iii) binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R; b) culturing the producer cell under conditions that allow for production of a lentiviral vector, and c) separating, enriching, or purifying the lentiviral vector from the cell, thereby making the pseudotyped lentiviral vector.

[0090] In some of any embodiments, the binding domain is a single domain antibody. In some of any embodiments, the binding domain is a single chain variable fragment (scFv). In some of any embodiments, the cell surface molecule is selected from the group consisting of ASGR1, ASGR2 and TM4SF5. In some of any embodiments, the cell surface molecule is CD8 or CD4, In some of any embodiments, the cell surface molecule is LDL-R.

[0091] Provided herein is a method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a) providing a cell that comprises the polynucleotide of any of the embodiments provided herein the vector of any of the embodiments described herein, or the plasmid of any of the embodiments described herein; b) culturing the cell under conditions that allow for production of a targeted lipid particle, and c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

[0092] Provided herein is a method of making a pseudotyped lentiviral vector, comprising: a) providing a producer cell that comprises a lentiviral viral nucleic acid(s), and the polynucleotide of any of the embodiments listed herein or the vector of any of the embodiments listed herein b) culturing the cell under conditions that allow for production of the lentiviral vector, and c) separating, enriching, or purifying the lentiviral vector from the cell, thereby making the pseudotyped lentiviral vector. In some of any embodiments, prior to step (b) the method further comprises providing the cell a polynucleotide encoding a henipavirus F protein molecule or biologically active portion thereof.

[0093] In some of any embodiments, the cell is a mammalian cell.

[0094] In some of any embodiments, the cell is a producer cell comprising viral nucleic acid. In some of any embodiments, the viral nucleic acid is a retroviral nucleic acid or lentiviral nucleic acid and the targeted lipid particle is a viral particle or a viral-like particle. In some of any embodiments, the viral particle or a viral-like particle is a retroviral particle or a retroviral-like particle. In some embodiments, the viral particle or a viral-like particle is a lentiviral particle or lentiviral-like particle.

[0095] In some of any embodiments, the viral nucleic acid(s) lacks one or more genes involved in viral replication. In some of any embodiments, the viral nucleic acid comprises a nucleic acid encoding a viral packaging protein selected from one or more of Gag, Pol, Rev and Tat. In some of any embodiments, the viral 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)/central termination sequence (CTS) (e.g. DNA flap), Poly A tail sequence, a posttranscriptional regulatory element (e.g. WPRE), a Rev response element (RRE), and 3' LTR (e.g., comprising U5 and lacking a functional U3).

[0096] Provided herein is a producer cell comprising the polynucleotide of any of the embodiments listed herein or the vector of any of the embodiments listed herein, or the plasmid of any of the embodiments described herein.

[0097] In some of any embodiments, the producer cell further comprises a nucleic acid encoding a henipavirus F protein or a biologically active portion thereof.

[0098] In some of any embodiments, the cell further comprises a viral nucleic acid. In some of any embodiments, the viral nucleic acid is a lentiviral nucleic acid. Provided herein is a producer cell comprising (i) a viral nucleic acid(s) and (ii) nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and (iii) a nucleic acid encoding a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain, optionally wherein the viral nucleic acid(s) are lentiviral nucleic acids. In some of any embodiments the single domain antibody binds a cell surface molecule present on a target cell. In some of any embodiments the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule.

[0099] In some of any embodiments the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem

cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells. In some of any embodiments the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell. In some of any embodiments the single domain antibody binds an antigen or portion thereof present on a target cell.

[0100] Provided herein is a producer cell comprising (i) a viral nucleic acid(s) and (ii) nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and (iii) a nucleic acid encoding a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and binding domain, wherein the binding domain (i) binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2; (ii) binds a cell surface molecule selected from the group consisting of CD4 or CD8, optionally human CD4 or human CD8; or (iii) binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R. In some of any embodiments the viral nucleic acid(s) are lentiviral nucleic acid.

[0101] In some of any embodiments the cell surface molecule or antigen is selected from the group consisting of ASGR1, ASGR2 and TM4SF5. In some of any embodiments, the cell surface molecule or antigen is CD8 or CD4. In some of any embodiments, the cell surface molecule or antigen is LDL-R.

[0102] In some of any embodiments, the viral nucleic acid(s) lacks one or more genes involved in viral replication. In some of any embodiments, the viral nucleic acid comprises a nucleic acid encoding a viral packaging protein selected from one or more of Gag, Pol, Rev and Tat.

[0103] In some of any embodiments, the viral 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)/central termination sequence (CTS) (e.g. DNA flap), Poly A tail sequence, a posttranscriptional

regulatory element (e.g. WPRE), a Rev response element (RRE), and 3' LTR (e.g., comprising U5 and lacking a functional U3).

[0104] In some of any embodiments, the henipavirus F protein molecule or biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 2; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:2. In some of any embodiments, the henipavirus F protein molecule or biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 5; (ii) an amino acid sequence having at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:5.

[0105] In some of any embodiments, the henipavirus F protein molecule or biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 7; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:7. In some of any embodiments, the henipavirus F protein molecule or biologically active portion thereof comprises (i) a sequence encoding by a nucleotide sequence encoding the sequence set forth in SEQ ID NO: 8; (ii) a amino acid sequence encoded by a nucleotide sequence encoding a sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:8.

[0106] In some of any embodiments, the henipavirus F protein molecule or biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 23; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:23.

[0107] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44.

[0108] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 10; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10.

[0109] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 35; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or

about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35.

[0110] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 45; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45.

[0111] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 11; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11.

[0112] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 36; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36.

[0113] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ

ID NO: 46; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

[0114] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 12; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12.

[0115] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 37; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37.

[0116] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 47; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at

or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47.

[0117] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 13; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13.

[0118] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 38; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38.

[0119] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 48; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48.

[0120] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 14; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%,

at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14.

[0121] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 39; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39.

[0122] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 49; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49.

[0123] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 15; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at

or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15.

[0124] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 40; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40.

[0125] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises: (i) the sequence set forth in SEQ ID NO: 50; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:50.

[0126] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 16; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16.

[0127] In some of any embodiments, the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises (i) the sequence set forth in SEQ ID NO: 51; (ii) an amino acid sequence having at least at or about 80%, at least at or about 81%,

at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:51.

[0128] In some aspects of the provided embodiments, the targeted lipid particle has greater expression of the targeted envelope protein compared to a reference lipid particle that has incorporated into a similar lipid bilayer the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting moiety is a single chain variable fragment (scFv). In some of any embodiments, the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more. In some embodiments, the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, preferably at or about or greater than 10-fold or more. In some of any embodiments, the titer in target cells following transduction is at or greater than 1×10^6 transduction units (TU)/mL, at or greater than 2×10^6 TU/mL, at or greater than 3×10^6 TU/mL, at or greater than 4×10^6 TU/mL, at or greater than 5×10^6 TU/mL, at or greater than 6×10^6 TU/mL, at or greater than 7×10^6 TU/mL, at or greater than 8×10^6 TU/mL, at or greater than 9×10^6 TU/mL, or at or greater than 1×10^7 TU/mL. Also provided herein is a composition wherein among the population of lipid particles, greater than at or about 50%, greater than at or about 55%, greater than at or about 60%, greater than at or about 65%, greater than at or about 70%, or greater than at or about 75% are surface positive for the targeted envelope protein. In some of any embodiments, the targeted envelope protein is present on the surface of the targeted lipid particle at a density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

[0129] Provided herein is a viral vector particle or viral-like particle produced from the producer cell of any of the embodiments provided herein.

[0130] Provided herein is a composition comprising a plurality of targeted lipid particles of any of the embodiments provided herein. In some embodiments, the composition further includes a pharmaceutically acceptable carrier. In some of any embodiments, the targeted lipid particles comprise an average diameter of less than 1 μ m. In some of any embodiments, the

composition further includes a targeted envelope protein present on the surface of the targeted lipid particles at an average density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

[0131] Provided herein is a producer cell containing greater membrane (e.g., plasma membrane) expression of the targeted envelope protein compared to a reference producer cell that has incorporated into its membrane (e.g. plasma membrane) the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting moiety is a single chain variable fragment (scFv). In some embodiments, the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more. In some embodiments, the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, preferably at or about or greater than 10-fold or more. In some embodiments, the producer cell has the expression of the targeted envelope protein on a membrane (e.g., plasma membrane) of the producer cell is at least 20 proteins (e.g., at least 50, 100, 200, 500, 1000, 2000, 5000, or 10,000 proteins) per square micron. In some of any embodiments, the targeted envelope protein comprises at least 0.1% (e.g., at least 0.2%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10%) of the total membrane (e.g., plasma membrane) proteins of the producer cell (e.g., by total protein weight).

[0132] Provided herein is a method of transducing a cell comprising transducing a cell with any of the viral vectors described herein or with any of the compositions described herein. In some of any embodiments, the targeted envelope protein of the lentiviral vector or targeted lipid particle targets CD4 and the cell is a CD4⁺ cell. In some of any embodiments, the targeted envelope protein of the lentiviral vector targets CD8 and the cell is a CD8⁺ cell. In some of any embodiments, the targeted envelope protein of the lentiviral vector targets ASGR1, ASGR2 or TM4SF5 and the cell is a hepatocyte.

[0133] Provided herein is a method of delivering an exogenous agent to a subject (e.g., a human subject), the method comprising administering to the subject the targeted lipid particle of any of the embodiments provided herein or the composition of any of the embodiments provided herein, wherein the targeted lipid particle or lentiviral vector comprise the exogenous agent..

[0134] Provided herein is a method of delivering an exogenous agent to a subject (e.g., a human subject), the method comprising administering to the subject any of the compositions

described herein, wherein targeted lipid particle or lentiviral vectors of the plurality comprise the exogenous agent.

[0135] Provided herein is a method of delivering a chimeric antigen receptor (CAR) to a cell, comprising contacting a cell with any of the lentiviral vectors described herein or a targeted lipid particle of any of the embodiments described herein, wherein the lentiviral vector or targeted lipid particle comprise nucleic acid encoding the CAR.

[0136] Provided herein is a method of delivering a chimeric antigen receptor (CAR) to a cell, comprising contacting a cell with any of the compositions described herein, wherein lentiviral vectors or targeted lipid particles of the plurality comprise nucleic acid encoding the CAR.

[0137] Provided herein is a method of delivering an exogenous agent to a hepatocyte, comprising contacting a cell with any of the lentiviral vectors described herein, or a targeted lipid particle or lentiviral vector of any of the embodiments described herein.

[0138] Provided herein is a method of delivering an exogenous agent to a hepatocyte, comprising contacting a cell with any of the compositions described herein, wherein lentiviral vectors or targeted lipid particles of the plurality comprise an exogenous agent for delivery to the hepatocyte. In some of any embodiments, the contacting transduces the cell with lentiviral vector or the targeted lipid particle.

[0139] Provided herein is a method of treating a disease or disorder in a subject (e.g., a human subject), the method comprising administering to the subject the targeted lipid particle of any of the embodiments provided herein or the composition of any of the embodiments provided herein.

[0140] Provided herein is a method of fusing a mammalian cell to a targeted lipid particle, the method comprising administering to the subject the targeted lipid particle of any of the embodiments provided herein or the composition of any of the embodiments provided herein. In some of any embodiments, the fusing of the mammalian cell to the targeted lipid particle delivers an exogenous agent to a subject (e.g., a human subject). In some of any embodiments, the fusing of the mammalian cell to the targeted lipid particle treats a disease or disorder in a subject (e.g., a human subject). In some of any embodiments, the targeted envelope protein of the lentiviral vector or targeted lipid particle targets CD4 and the cell is a CD4+ cell. In some of any embodiments, the targeted envelope protein of the lentiviral vector targets CD8

and the cell is a CD8+ cell. In some of any embodiments, the targeted envelope protein of the lentiviral vector targets ASGR1, ASGR2 or TM4SF5 and the cell is a hepatocyte.

[0141] In some of any embodiments, the targeted lipid particle has greater expression of the targeted envelope protein compared to a reference lipid particle that has incorporated into a similar lipid bilayer the same envelope protein but that is fused to an alternative targeting moiety. In some embodiments, the alternative targeting moiety is a single chain variable fragment (scFv). In some of any embodiments, the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more. In some of any embodiments, the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, preferably at or about or greater than 10-fold or more.

[0142] In some of any embodiments, the titer in target cells following transduction is at or greater than 1×10^6 transduction units (TU)/mL, at or greater than 2×10^6 TU/mL, at or greater than 3×10^6 TU/mL, at or greater than 4×10^6 TU/mL, at or greater than 5×10^6 TU/mL, at or greater than 6×10^6 TU/mL, at or greater than 7×10^6 TU/mL, at or greater than 8×10^6 TU/mL, at or greater than 9×10^6 TU/mL, or at or greater than 1×10^7 TU/mL.

[0143] In some of any embodiments, among the population of lipid particles or lentiviral vectors in the composition, greater than at or about 50%, greater than at or about 55%, greater than at or about 60%, greater than at or about 65%, greater than at or about 70%, or greater than at or about 75% are surface positive for the targeted envelope protein. In some of any embodiments, the targeted envelope protein is present on the surface of the targeted lipid particle at a density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

[0144] Provided herein is a composition comprising a plurality of the targeted lipid particles of any of the embodiments described herein or a plurality of lentiviral vectors of any of the embodiments described herein, wherein the targeted envelope protein is present on the surface of the targeted lipid particles at an average density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

[0145] In some of any embodiments, the producer cell has greater membrane (e.g., plasma membrane) expression of the targeted envelope protein compared to a reference producer cell that has incorporated into its membrane (e.g. plasma membrane) the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting

moiety is a single chain variable fragment (scFv). In some of any embodiments, the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more. In some of any embodiments, the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, preferably at or about or greater than 10-fold or more. In some of any embodiments, the producer cell has the expression of the targeted envelope protein on a membrane (e.g., plasma membrane) of the producer cell is at least 20 proteins (e.g., at least 50, 100, 200, 500, 1000, 2000, 5000, or 10,000 proteins) per square micron. In some of any embodiments, the targeted envelope protein comprises at least 0.1% (e.g., at least 0.2%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10%) of the total membrane (e.g., plasma membrane) proteins of the producer cell (e.g., by total protein weight).

Detailed Description

[0146] Provided herein are targeted lipid particles containing a lipid bilayer enclosing a lumen or cavity and a targeted envelope protein containing (1) a henipavirus envelope attachment glycoprotein G (G protein) or biologically active portion thereof and (2) a binding domain, such as a single domain antibody (sdAb) variable domain, in which the targeted envelope protein is embedded in the lipid bilayer of the lipid particles. In particular embodiments, the binding domain, such as a single domain antibody, is an antibody with the ability to bind, such as specifically bind, to a desired target molecule. Exemplary binding domains are described in Section II.A.2. In some embodiments, the targeted lipid particles also contains a henipavirus fusion (F) protein molecule or a biologically active portion thereof embedded in the lipid bilayer. In particular embodiments, the lipid particles can be a virus-like particle, a virus, or a viral vector, such as a lentiviral vector.

[0147] In some embodiments, one or both of the G protein and the F protein is from a Hendra (HeV) or a Nipah (NiV) virus, or is a biologically active portion thereof or is a variant or mutant thereof. In particular embodiments, both the G protein and the F protein is from a Hendra (HeV) or a Nipah (NiV) virus. In some embodiments, the fusion and attachment glycoproteins mediate cellular entry of Nipah virus.

[0148] The F protein, such as NiV-F, is a class I fusion protein that has structural and functional features in common with fusion proteins of many families (e.g., HIV-1 gp41 or influenza virus hemagglutinin [HA]), such as an ectodomain with a hydrophobic fusion peptide

and two heptad repeat regions (White JM et al. 2008. Crit Rev Biochem Mol Biol 43:189–219). F proteins are synthesized as inactive precursors F₀ and are activated by proteolytic cleavage into the two disulfide-linked subunits F₁ and F₂ (Moll M. et al. 2004. J. Virol. 78(18): 9705-9712).

[0149] G proteins are attachment proteins of henipavirus (e.g. Nipah virus or Hendra virus) that are type II transmembrane glycoproteins containing an N-terminal cytoplasmic tail, a transmembrane domain, an extracellular stalk, and a globular head (Liu, Q. et al. 2015. Journal of Virology, 89(3):1838-1850). The attachment protein, NiV-G, recognizes the receptors EphrinB2 and EphrinB3. Binding of the receptor to NiV-G triggers a series of conformational changes that eventually lead to the triggering of NiV-F, which exposes the fusion peptide of NiV-F, allowing another series of conformational changes that lead to virus-cell membrane fusion (Stone J.A. et al. 2016. J Virol. 90(23): 10762-10773). EphrinB2 was previously identified as the primary NiV receptor (Negrete et al., 2005), as well as EphrinB3 as an alternate receptor (Negrete et al., 2006). In fact, NiV-G has a high affinity for EphrinB2 and B3, with affinity binding constants (K_d) in the picomolar range (Negrete et al., 2006) (K_d=0.06 nM and 0.58 nM for cell surface expressed ephrinB2 and B3, respectively).

[0150] The efficiency of transduction of targeted lipid particles can be improved by engineering hyperfusogenic mutations in one or both of NiV-F and NiV-G. Several such mutations have been previously described (see, e.g., Lee et al, 2011, Trends in Microbiology). This could be useful, for example, for maintaining the specificity and picomolar affinity of NiV-G for EphrinB2 and/or B3. Additionally, mutations in NiV-G that completely abrogate EphrinB2 and B3 binding, but that do not impact the association of this NiV-G with NiV-F, have been identified. Methods to improve targeting of lipid particles can be achieved by fusion of a binding molecule with a G protein (e.g. Niv-G, including a Niv-G with mutations to abrogate ephrin B2 and ephrin B3 binding). This could allow for altered G protein tropism allowing for targeting of other desired cell types that are not EphrinB2+ through the addition of the binding molecule directed against a different cell surface molecule.

[0151] While retargeted lipid particles incorporating such binding molecules fused to a G protein have been generated, it is found herein that some binding molecules when fused with a G protein (e.g. NiV-G) express better on the surface of lipid particles than others. For example, it is found that single domain antibodies (sdAbs), such as VHH, may express 10-fold better than a single chain variable fragment (scFv). Without wishing to be bound by theory, the

increase in expression may be due to an increased stability of the retargeted G protein on the surface of the lipid particle. This greater expression can improve the ability of the lipid particle to target the target molecule (e.g. a cell surface molecule) compared to a similar lipid particle but containing an alternative binding domain, e.g. scFv, against the same target molecule.

[0152] Thus, provided herein are targeted lipid particles containing a G protein of a henipavirus (e.g. Hendra or Nipah, e.g. NiV-G) attached to a sdAb variable domain directed against or that is able to bind to a cell surface molecule on a target cell. sdAb variable domains can include those of a VL or VH only sdAb, nanobodies, camelid VHH domains, shark IgNAR or fragments thereof. In some embodiments, the sdAb is a VHH.

[0153] In aspects of the provided embodiments, a targeted lipid particle can be engineered to express a henipavirus F protein molecule or biologically active portion thereof; and a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) single domain antibody (sdAb) variable domain, wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer. In some embodiments, the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof. In some embodiments, the sdAb variable domain is attached to the G protein via a linker.

[0154] Also provided are targeted lipid particles additionally containing one or more exogenous agents, such as for delivery of a diagnostic or therapeutic agent to cells, including following in vivo administration to a subject. Also provided herein are methods and uses of the targeted lipid particles, such in diagnostic and therapeutic methods. Also provided are polynucleotides, methods for engineering, preparing, and producing the targeted lipid non-cell particles, compositions containing the particles, and kits and devices containing and for using, producing and administering the particles.

[0155] All publications, including patent documents, scientific articles and databases, referred to in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication were individually incorporated by reference. If a definition set forth herein is contrary to or otherwise inconsistent with a definition set forth in the patents, applications, published applications and other publications that are herein incorporated by reference, the definition set forth herein prevails over the definition that is incorporated herein by reference.

[0156] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

BRIEF DESCRIPTION OF THE DRAWINGS

[0157] **FIGS. 1A-1C** depict characterization of cells transfected with constructs containing scFv or VHH binding modalities. **FIG. 1A** depicts surface expression of cells transfected with constructs containing scFv or VHH binding modalities, analyzed by flow cytometry, and depicted as median fluorescence intensity (MFI), quantified by % of His⁺ cells. **FIG. 1B** depicts binding to soluble hCD4-Fc protein of cells transfected with constructs containing scFv or VHH binding modalities analyzed by flow cytometry, and depicted as median fluorescence intensity (MFI), quantified by %Fc⁺ cell. **FIG. 1C** depicts surface expression of targeted binding sequences on 293 cells for cells transfected with constructs containing VHH binding modalities, compared to the scFv binding modalities, analyzed by flow cytometry, and depicted as median fluorescence intensity (MFI), as quantified by % of His⁺ cells. Empty vector and the expression vector without the binder domain were used as negative controls.

[0158] **FIG. 2** depicts transduction efficacy of four exemplary constructs containing scFv or VHH binding modalities on PanT cells from peripheral blood that were negatively selected to enrich for T cells were thawed and activated with anti CD3/ anti-CD28. Cells were analyzed by flow cytometry, and titer determined by % of CD4-positive cells that were GFP⁺.

[0159] **FIGS. 3A-3B** depict transduction efficiency of CD8 retargeted pseudotyped lentiviruses in an in vivo model using activated PBMCs injected intraperitoneally into NOD-scid-IL2r^{null} mice, as analyzed by flow cytometry. Transduction efficiency of CD8 retargeted pseudotyped lentiviruses is depicted on CD8⁺ (**FIG. 3A**) or CD8⁻ (**FIG. 3B**) T cells, and titer was determined by % of CD8 positive or negative cells that were GFP⁺.

[0160] **FIGS. 4A-4B** depict the ability of CD8 retargeted pseudotyped lentiviruses containing chimeric antigen receptors (CARs) to effect killing of leukemic cells in vitro. **FIG. 4A** shows the ability to detect CD19⁺ CAR expression on CD8⁺ cells at 4 days post transduction. **FIG. 4B** shows the elimination of Nalm6 cells evaluated at 18 hours post incubation, analyzed by flow cytometry

I. DEFINITIONS

[0161] Unless defined otherwise, all terms of art, notations and other technical and scientific terms or terminology used herein are intended to have the same meaning as is commonly understood by one of ordinary skill in the art to which the claimed subject matter pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art.

[0162] Unless defined otherwise, all technical and scientific terms, acronyms, and abbreviations used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Unless indicated otherwise, abbreviations and symbols for chemical and biochemical names is per IUPAC-IUB nomenclature. Unless indicated otherwise, all numerical ranges are inclusive of the values defining the range as well as all integer values in-between.

[0163] As used herein, the articles “a” and “an” refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element.

[0164] As used herein, the term “about” will be understood by persons of ordinary skill in the art and will vary to some extent on the context in which it is used. As used herein, “about” when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of $\pm 20\%$ or $\pm 10\%$, more preferably $\pm 5\%$, even more preferably $\pm 1\%$, and still more preferably $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods.

[0165] As used herein, “lipid particle” refers to any biological or synthetic particle that contains a bilayer of amphipathic lipids enclosing a lumen or cavity. Typically a lipid particle does not contain a nucleus. Examples of lipid particles include solid particles such as nanoparticles, viral-derived particles or cell-derived particles. Such lipid particles include, but are not limited to, viral particles (e.g. lentiviral particles), virus-like particles, viral vectors (e.g., lentiviral vectors) exosomes, enucleated cells, various vesicles, such as a microvesicle, a membrane vesicle, an extracellular membrane vesicle, a plasma membrane vesicle, a giant plasma membrane vesicle, an apoptotic body, a mitoparticle, a pyrenocyte, or a lysosome. In

some embodiments, a lipid particle can be a fusosome. In some embodiments, the lipid particle is not a platelet.

[0166] As used herein a “biologically active portion,” such as with reference to a protein such as a G protein or an F protein, refers to a portion of the protein that exhibits or retains an activity or property of the full-length of the protein. For example, a biologically active portion of an F protein retains fusogenic activity in conjunction with the G protein when each are embedded in a lipid bilayer. A biologically active portion of the G protein retains fusogenic activity in conjunction with an F protein when each is embedded in a lipid bilayer. The retained activity and include 10%-150% or more of the activity of a full-length or wild-type F protein or G protein. Examples of biologically active portions of F and G proteins include truncations of the cytoplasmic domain, e.g. truncations of up to 1, 2, 3, 4, 5, 6, 7, 8 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35 or more contiguous amino acids, see e.g. Khetawat and Broder 2010 Virology Journal 7:312; Witting et al. 2013 Gene Therapy 20:997-1005; published international; patent application No. WO/2013/148327.

[0167] As used herein, “fusosome” refers to a particle containing 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.

[0168] As used herein, “fusosome composition” refers to a composition comprising one or more fusosomes.

[0169] 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.

[0170] 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.

[0171] As used herein, a “targeted envelope protein” refers to a polypeptide that contains a henipavirus G protein attached to a single domain antibody (sdAb) variable domain, such as a VL or VH only sdAb, nanobodies, camelid VHH domains, shark IgNAR or fragments thereof, that targets a molecule on a desired cell type. In some such embodiments, the attachment may be directly or indirectly via a linker, such as a peptide linker.

[0172] As used herein, a “targeted lipid particle” refers to a lipid particle that contains a targeted envelope protein embedded in the lipid bilayer.

[0173] 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.

[0174] As used herein, a “target cell” refers to a cell of a type to which it is desired that a targeted lipid particle delivers an exogenous agent. In embodiments, a target cell is a cell of a specific tissue type or class, e.g., an immune effector cell, e.g., a T 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 leads to preferential delivery of the exogenous agent to a target cell compared to a non-target cell.

[0175] As used herein a “non-target cell” refers to a cell of a type to which it is not desired that a targeted lipid particle 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 leads to lower delivery of the exogenous agent to a non-target cell compared to a target cell.

[0176] As used herein, a “single domain antibody” or “sdAb” refers to an antibody having a single monomeric domain antigen binding/recognition domain. Such antibodies include nanobodies, camelid antibodies (e.g. VHH), or shark antibodies (e.g. IgNAR). In some embodiments, a variable domain of a sdAb comprises three CDRs and four framework regions, designated FR1, CDR1, FR2, CDR2, FR3, CDR3, and FR4. In some embodiments, a sdAb variable domain may be truncated at the N-terminus or C-terminus such that it comprise only a partial FR1 and/or FR4, or lacks one or both of those framework regions, so long as the sdAb variable domain substantially maintains antigen binding and specificity.

[0177] The term “CDR” denotes a complementarity determining region as defined by at least one manner of identification to one of skill in the art. The precise amino acid sequence boundaries of a given CDR or FR can be readily determined using any of a number of well-known schemes, including those described by Kabat *et al.* (1991), “Sequences of Proteins of Immunological Interest,” 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD (“Kabat” numbering scheme); Al-Lazikani *et al.*, (1997) JMB 273,927-948 (“Chothia” numbering scheme); MacCallum *et al.*, J. Mol. Biol. 262:732-745 (1996), “Antibody-antigen interactions: Contact analysis and binding site topography,” J. Mol. Biol. 262, 732-745.” (“Contact” numbering scheme); Lefranc MP *et al.*, “IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains,” Dev Comp Immunol, 2003 Jan;27(1):55-77 (“IMGT” numbering scheme); Honegger A and Plückthun A, “Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool,” J Mol Biol, 2001 Jun 8;309(3):657-70, (“Aho” numbering scheme); and Martin *et al.*, “Modeling antibody hypervariable loops: a combined algorithm,” PNAS, 1989, 86(23):9268-9272, (“AbM” numbering scheme).

[0178] The boundaries of a given CDR or FR may vary depending on the scheme used for identification. For example, the Kabat scheme is based on structural alignments, while the Chothia scheme is based on structural information. Numbering for both the Kabat and Chothia

schemes is based upon the most common antibody region sequence lengths, with insertions accommodated by insertion letters, for example, “30a,” and deletions appearing in some antibodies. The two schemes place certain insertions and deletions (“indels”) at different positions, resulting in differential numbering. The Contact scheme is based on analysis of complex crystal structures and is similar in many respects to the Chothia numbering scheme. The AbM scheme is a compromise between Kabat and Chothia definitions based on that used by Oxford Molecular’s AbM antibody modeling software.

[0179] In some embodiments, CDRs can be defined in accordance with any of the Chothia numbering schemes, the Kabat numbering scheme, a combination of Kabat and Chothia, the AbM definition, and/or the contact definition. A sdAb variable domain comprises three CDRs, designated CDR1, CDR2, and CDR3. Table 1, below, lists exemplary position boundaries of CDR-H1, CDR-H2, CDR-H3 as identified by Kabat, Chothia, AbM, and Contact schemes, respectively. For CDR-H1, residue numbering is listed using both the Kabat and Chothia numbering schemes. FRs are located between CDRs, for example, with FR-H1 located before CDR-H1, FR-H2 located between CDR-H1 and CDR-H2, FR-H3 located between CDR-H2 and CDR-H3 and so forth. It is noted that because the shown Kabat numbering scheme places insertions at H35A and H35B, the end of the Chothia CDR-H1 loop when numbered using the shown Kabat numbering convention varies between H32 and H34, depending on the length of the loop.

Table 1. Boundaries of CDRs according to various numbering schemes.				
CDR	Kabat	Chothia	AbM	Contact
CDR-H1 (Kabat Numbering ¹)	H31--H35B	H26--H32...34	H26--H35B	H30--H35B
CDR-H1 (Chothia Numbering ²)	H31--H35	H26--H32	H26--H35	H30--H35
CDR-H2	H50--H65	H52--H56	H50--H58	H47--H58
CDR-H3	H95--H102	H95--H102	H95--H102	H93--H101

1 - Kabat *et al.* (1991), “Sequences of Proteins of Immunological Interest,” 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD

2 - Al-Lazikani *et al.*, (1997) JMB 273,927-948

[0180] Thus, unless otherwise specified, a “CDR” or “complementary determining region,” or individual specified CDRs (*e.g.*, CDR-H1, CDR-H2, CDR-H3), of a given antibody or region thereof, such as a variable region thereof, should be understood to encompass a (or the specific) complementary determining region as defined by any of the aforementioned schemes. For

example, where it is stated that a particular CDR (*e.g.*, a CDR-H3) contains the amino acid sequence of a corresponding CDR in a given sdAb amino acid sequence, it is understood that such a CDR has a sequence of the corresponding CDR (*e.g.*, CDR-H3) within the sdAb, as defined by any of the aforementioned schemes. It is understood that any antibody, such as a sdAb, includes CDRs and such can be identified according to any of the other aforementioned numbering schemes or other numbering schemes known to a skilled artisan.

[0181] As used herein, the term “specifically binds” to a target molecule, such as an antigen, means that a binding molecule, such as a single domain antibody, reacts or associates more frequently, more rapidly, with greater duration and/or with greater affinity with a particular target molecule than it does with alternative molecules. A binding molecule, such as a sdAb variable domain, “specifically binds” to a target molecule if it binds with greater affinity, avidity, more readily, and/or with greater duration than it binds to other molecules. It is understood that a binding molecule, such as a sdAb, that specifically binds to a first target may or may not specifically bind to a second target. As such, “specific binding” does not necessarily require (although it can include) exclusive binding.

[0182] As used herein, “percent (%) amino acid sequence identity” and “homology” with respect to a peptide, polypeptide or antibody sequence are defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific peptide or polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or MEGALIGN™ (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

[0183] An amino acid substitution may include but are not limited to the replacement of one amino acid in a polypeptide with another amino acid. Exemplary substitutions are shown in **Table 2**. Amino acid substitutions may be introduced into an antibody of interest and the products screened for a desired activity, for example, retained/improved binding.

Table 2

Original Residue	Exemplary Substitutions
Ala (A)	Val; Leu; Ile
Arg (R)	Lys; Gln; Asn
Asn (N)	Gln; His; Asp, Lys; Arg
Asp (D)	Glu; Asn
Cys (C)	Ser; Ala
Gln (Q)	Asn; Glu
Glu (E)	Asp; Gln
Gly (G)	Ala
His (H)	Asn; Gln; Lys; Arg
Ile (I)	Leu; Val; Met; Ala; Phe; Norleucine
Leu (L)	Norleucine; Ile; Val; Met; Ala; Phe
Lys (K)	Arg; Gln; Asn
Met (M)	Leu; Phe; Ile
Phe (F)	Trp; Leu; Val; Ile; Ala; Tyr
Pro (P)	Ala
Ser (S)	Thr
Thr (T)	Val; Ser
Trp (W)	Tyr; Phe
Tyr (Y)	Trp; Phe; Thr; Ser
Val (V)	Ile; Leu; Met; Phe; Ala; Norleucine

[0184] Amino acids may be grouped according to common side-chain properties:

- (1) hydrophobic: Norleucine, Met, Ala, Val, Leu, Ile;
- (2) neutral hydrophilic: Cys, Ser, Thr, Asn, Gln;
- (3) acidic: Asp, Glu;
- (4) basic: His, Lys, Arg;
- (5) residues that influence chain orientation: Gly, Pro;
- (6) aromatic: Trp, Tyr, Phe.

[0185] Non-conservative substitutions will entail exchanging a member of one of these classes for another class.

[0186] The term, “corresponding to” with reference to positions of a protein, such as recitation that nucleotides or amino acid positions “correspond to” nucleotides or amino acid positions in a disclosed sequence, such as set forth in the Sequence listing, refers to nucleotides or amino acid positions identified upon alignment with the disclosed sequence based on structural sequence alignment or using a standard alignment algorithm, such as the GAP algorithm. For example, corresponding residues of a similar sequence (e.g. fragment or species variant) can be determined by alignment to a reference sequence by structural alignment

methods. By aligning the sequences, one skilled in the art can identify corresponding residues, for example, using conserved and identical amino acid residues as guides.

[0187] The term “isolated” as used herein refers to a molecule that has been separated from at least some of the components with which it is typically found in nature or produced. For example, a polypeptide is referred to as “isolated” when it is separated from at least some of the components of the cell in which it was produced. Where a polypeptide is secreted by a cell after expression, physically separating the supernatant containing the polypeptide from the cell that produced it is considered to be “isolating” the polypeptide. Similarly, a polynucleotide is referred to as “isolated” when it is not part of the larger polynucleotide (such as, for example, genomic DNA or mitochondrial DNA, in the case of a DNA polynucleotide) in which it is typically found in nature, or is separated from at least some of the components of the cell in which it was produced, for example, in the case of an RNA polynucleotide. Thus, a DNA polynucleotide that is contained in a vector inside a host cell may be referred to as “isolated”.

[0188] 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.

[0189] An “exogenous agent” as used herein with reference to a targeted lipid particle, 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.

[0190] 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.

[0191] As used herein, a composition refers to any mixture of two or more products, substances, or compounds, including cells. It may be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous or any combination thereof.

[0192] As used herein, the term “pharmaceutically acceptable” refers to a material, such as a carrier or diluent, which does not abrogate the biological activity or properties of the compound, and is relatively nontoxic, i.e., the material may be administered to an individual without causing undesirable biological effects or interacting in a deleterious manner with any of the components of the composition in which it is contained.

[0193] As used herein, the term “pharmaceutical composition” refers to a mixture of at least one compound of the invention with other chemical components, such as carriers, stabilizers, diluents, dispersing agents, suspending agents, thickening agents, and/or excipients. The pharmaceutical composition facilitates administration of the compound to an organism. Multiple techniques of administering a compound exist in the art including, but not limited to, intravenous, oral, aerosol, parenteral, ophthalmic, pulmonary and topical administration.

[0194] A “disease” or “disorder” as used herein refers to a condition where treatment is needed and/or desired.

[0195] 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 or reducing at least one of the clinical symptoms thereof. For purposes of this disclosure, ameliorating a disease or disorder can include obtaining a beneficial or desired clinical result that includes, but is not limited to, any one or more of: alleviation of one or more symptoms, diminishment of extent of disease, preventing or delaying spread (for example, metastasis, for example metastasis to the lung or to the lymph node) of disease, preventing or delaying recurrence of disease, delay or slowing of disease progression, amelioration of the disease state, inhibiting the disease or progression of the disease, inhibiting or slowing the disease or its progression, arresting its development, and remission (whether partial or total).

[0196] The terms “individual” and “subject” are used interchangeably herein to refer to an animal; for example a mammal. The term patient includes human and veterinary subjects. In

some embodiments, methods of treating mammals, including, but not limited to, humans, rodents, simians, felines, canines, equines, bovines, porcines, ovines, caprines, mammalian laboratory animals, mammalian farm animals, mammalian sport animals, and mammalian pets, are provided. The subject can be male or female and can be any suitable age, including infant, juvenile, adolescent, adult, and geriatric subjects. In some examples, an “individual” or “subject” refers to an individual or subject in need of treatment for a disease or disorder. In some embodiments, the subject to receive the treatment can be a patient, designating the fact that the subject has been identified as having a disorder of relevance to the treatment, or being at adequate risk of contracting the disorder. In particular embodiments, the subject is a human, such as a human patient.

II. TARGETED LIPID PARTICLES (E.G. LENTIVIRAL VECTORS)

[0197] Provided herein are targeted lipid particles that comprise a henipavirus F protein molecule or biologically active portion thereof, and a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion, wherein each of (i) and (ii) is exposed on the outer surface of the targeted lipid particle. In some embodiments, the binding domain is a single domain antibody. In some embodiments, the binding domain is a single chain variable fragment. In particular embodiments, the provided lipid particles exhibit fusogenic activity, which is mediated by the targeted envelope protein that facilitates binding to a target cell and contains the G protein or biologically active portion thereof, and the F glycoprotein that is involved in facilitating the merger or fusion of the two lumens of the lipid particle and the target cell membranes.

[0198] Provided herein are targeted lipid particles that comprise a henipavirus F protein molecule or biologically active portion thereof, and a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a single domain antibody (sdAb) variable domain, wherein the single domain antibody is attached to the C-terminus of the G protein or the biologically active portion, wherein each of (i) and (ii) is exposed on the outer surface of the targeted lipid particle. In particular embodiments, the provided lipid particles exhibit fusogenic activity, which is mediated by the targeted envelope protein that facilitates binding to a target cell and contains the G protein or biologically active portion thereof, and the F glycoprotein that is involved in

facilitating the merger or fusion of the two lumens of the lipid particle and the target cell membranes.

[0199] In some of any embodiment, the targeted lipid particles are viral particles or viral-like particles. In some aspects, such targeted lipid particles contain viral nucleic acid, such as retroviral nucleic acid, for example lentiviral nucleic acid. In particular embodiments, any provided targeted lipid particles, such as a viral particle or viral-like particle, is replication defective. In some embodiments, the targeted lipid particle is a lentiviral vector, in which the lentiviral vector is pseudotyped with the henipavirus F protein and the targeted envelope protein.

[0200] For instance, provided herein is a pseudotyped lentiviral vector that comprises a henipavirus F protein molecule or biologically active portion thereof, and a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion, wherein each of (i) and (ii) is exposed on the outer surface of the targeted lipid particle. In some embodiments, the binding domain is a single domain antibody. In some embodiments, the binding domain is a single chain variable fragment.

[0201] In some embodiments, the targeted lipid particle provided herein (e.g. targeted lentiviral vector) has increased or greater expression of the targeted envelope protein compared to a reference lipid particle (e.g. reference lentiviral vector) that incorporates a similar envelope protein but that is fused to an alternative targeting moiety other than a sdAb variable domain, such as a single chain variable fragment (scFv). In some embodiments, such targeted lipid particles are produced by pseudotyping of lipid particles (e.g. lentiviral particles) following co-transfection of the packaging cells with the transfer, envelope, and gag-pol plasmids.

[0202] In some embodiments, the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more, compared to a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some examples, the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, compared to a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some embodiments, expression can be assayed in vitro using flow cytometry, e.g. FACs. In some embodiments, expression can be

depicted as the number or density of targeted envelope protein on the surface of a targeted lipid particle (e.g. targeted lentiviral vector). In some embodiments, expression can be depicted as the mean fluorescent intensity (MFI) of surface expression of the targeted envelope protein on the surface of a targeted lipid particle (e.g. targeted lentiviral vector). In some embodiments, expression can be depicted as the percent of lipid particle (e.g. lentiviral vectors) in a population that are surface positive for the targeted envelope protein.

[0203] In some embodiments, in a population of targeted lipid particles (e.g. targeted lentiviral vectors) greater than at or about 50% of the lipid particles are surface positive for the targeted envelope protein. For example, in a population of provided targeted lipid particles (e.g. targeted lentiviral vectors) greater than at or about 55%, greater than at or about 60%, greater than at or about 65%, greater than at or about 70%, greater than at or about 75% of the cells in the population are surface positive for the targeted envelope protein.

[0204] In some embodiments, titer of the targeted lipid particles following introduction into target cells, such as by transduction (e.g. transduced cells), is increased compared to titer into the same target cells of reference lipid particles (e.g. reference lentiviral vector) that incorporate a similar envelope protein but fused to an alternative targeting moiety other than a sdAb variable domain, such as a single chain variable fragment (scFv). Typically, the alternative targeting moiety recognizes or binds the same target molecule as the sdAb variable domain of the targeted envelope protein of the targeted lipid particles. In some embodiments, the titer is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more, compared to titer of a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some examples, the titer is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, compared to the titer of a reference lipid particle (e.g. reference lentiviral vector), e.g. a reference lipid particle containing a similar envelope protein but that is fused to an scFv. In some embodiments, the titer of the targeted lipid particles in target cells (e.g. transduced cells) is greater than at or about 1×10^6 transduction units (TU)/mL. For example, the titer of the targeted lipid particles in target cells (e.g. transduced cells) is greater than at or about 2×10^6 TU/mL, greater than at or about 3×10^6 TU/mL, greater than at or about 4×10^6 TU/mL, greater than at or about 5×10^6 TU/mL, greater than at or about 6×10^6 TU/mL, greater than at or about

7×10^6 TU/mL, greater than at or about 8×10^6 TU/mL, greater than at or about 9×10^6 TU/mL, or greater than at or about 1×10^7 TU/mL.

A. Targeted Envelope Protein (e.g. henipavirus plus binding domain)

[0205] In some embodiments, the targeted lipid particle (e.g. lentiviral vector) includes a targeted envelope protein exposed on the surface of the targeted lipid particle (e.g. lentiviral vector).

[0206] In some embodiments, the targeted envelope protein contains a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a binding domain that binds to a cell surface molecule on a target cell. In some embodiments, the binding domain is a single domain antibody (sdAb). In some embodiments, the binding domain is a single chain variable fragment (scFv). The binding domain can be linked directly or indirectly to the G protein. In particular embodiments, the binding domain is linked to the C-terminus (C-terminal amino acid) of the G protein or the biologically active portion thereof. The linkage can be via a peptide linker, such as a flexible peptide linker.

I. G Protein

[0207] In some embodiments, the targeted envelope protein contains a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain or biologically active portion thereof. In some embodiments, the sdAb binds to a cell surface molecule on a target cell. The sdAb variable domain can be linked directly or indirectly to the G protein. In particular embodiments, the sdAb variable domain is linked to the C-terminus (C-terminal amino acid) of the G protein or the biologically active portion thereof. The linkage can be via a peptide linker, such as a flexible peptide linker.

[0208] In some embodiments, an binding domain (e.g. sdAb) binds to a cell surface antigen of a cell. In some embodiments, a cell surface antigen is characteristic of one type of cell. In some embodiments, a cell surface antigen is characteristic of more than one type of cell.

[0209] In some embodiments, the binding domain (e.g. sdAb) variable domain binds a cell surface molecule or antigen. In some embodiments, the cell surface molecule is ASGR1, ASGR2, TM4SF5, CD8, CD4, or low density lipoprotein receptor (LDL-R). In some embodiments, the cell surface molecule is ASGR1. In some embodiments, the cell surface molecule is ASGR2. In some embodiments, the cell surface molecule is TM4SF5. In some

embodiments, the cell surface molecule is CD8. In some embodiments, the cell surface molecule is CD4. In some embodiments, the cell surface molecule is LDL-R.

[0210] In some embodiments the G protein is a Henipavirus G protein or a biologically active portion thereof. In some embodiments, the Henipavirus G protein is a Hendra (HeV) virus G protein, a Nipah (NiV) virus G-protein (NiV-G), a Cedar (CedPV) virus G-protein, a Mojiang virus G-protein, a bat Paramyxovirus G-protein or a biologically active portion thereof. **Table 3** provides non-limiting examples of G proteins.

[0211] The attachment G proteins are type II transmembrane glycoproteins containing an N-terminal cytoplasmic tail (e.g. corresponding to amino acids 1-49 of SEQ ID NO:9), a transmembrane domain (e.g. corresponding to amino acids 50-70 of SEQ ID NO:9), and an extracellular domain containing an extracellular stalk (e.g. corresponding to amino acids 71-187 of SEQ ID NO:9), and a globular head (corresponding to amino acids 188-602 of SEQ ID NO:9). The N-terminal cytoplasmic domain is within the inner lumen of the lipid bilayer and the C-terminal portion is the extracellular domain that is exposed on the outside of the lipid bilayer. Regions of the stalk in the C-terminal region (e.g. corresponding to amino acids 159-167 of NiV-G) have been shown to be involved in interactions with F protein and triggering of F protein fusion (Liu et al. 2015 J of Virology 89:1838). In wild-type G protein, the globular head mediates receptor binding to henipavirus entry receptors ephrin B2 and ephrin B3, but is dispensable for membrane fusion (Brandel-Tretheway et al. Journal of Virology. 2019. 93(13)e00577-19). In particular embodiments herein, tropism of the G protein is altered by linkage of the G protein or biologically active fragment thereof (e.g. cytoplasmic truncation) to a sdAb variable domain. Binding of the G protein to a binding partner can trigger fusion mediated by a compatible F protein or biologically active portion thereof. G protein sequences disclosed herein are predominantly disclosed as expressed sequences including an N-terminal methionine required for start of translation. As such N-terminal methionines are commonly cleaved co- or post-translationally, the mature protein sequences for all G protein sequences disclosed herein are also contemplated as lacking the N-terminal methionine.

[0212] G glycoproteins are highly conserved between henipavirus species. For example, the G protein of NiV and HeV viruses share 79% amino acids identity. Studies have shown a high degree of compatibility among G proteins with F proteins of different species as demonstrated by heterotypic fusion activation (Brandel-Tretheway et al. Journal of

Virology. 2019). As described further below, a re-targeted lipid particle can contain heterologous G and F proteins from different species.

Table 3. Henipavirus protein G 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. Column 6 provides the SEQ ID numbers for the described sequences.

Genbank ID	Nucleotides of CDS	Full sequence ID	Sequence	#Sequences/Cluster	SEQ ID NO	SEQ ID NO (without N-terminal methionine)
AF017149	8913-10727	gb:AF017149 Organism:Hen dra virus Strain Name:UNKN OWN- AF017149 Pro tein Name:glycop rotein Gene Symbol:G	MMADSKLVSLNNLSGKIKDQGKVIKN YYGTMDIKKINDGLLDISKILGAFNTVIA LLGSIIIVMNIMIIQNYTRTTDNQALIKES LQSVQQQIKALTDKIGTEIGPKVSLIDTSS TTIPANIGLLGSKISQSTSSINENVNDKC KFTLPPLKIHCECNISCPNPLPFREYRPI SQGVSDLVGLPNQICLQKTTSTILKPR LISYTLPI NTREGVCITDPLLAVDNGFF AYSHLEKIGSCTRGIAKQRIIGVGEVLD RGDKVPSMFMTNVWTPPNPSTIIHCSST YHEDFYTLCAVSHVGDPI LNSTSWTES LSLIRLAVRPKSDSGDYNQKYIAITKVER GKYDKVMPYGPSGIKQGDLYFPAVGFL PRTEFYQYNDNCPIIHCYKSKAENCRLS MGVNSKSHYILRSGLLKYNLSLGGDIIL QFIEIADNRLTIGSPSKIYNSLGQPVFY QASYSWDTMIKLGDVDTVDPLRVQWRN NSVISRPGQSQCPRFNVCPVCWEGTYN DAFLIDRLNWVSAGVYLNSENQTAENPV FAVFKDNEILYQVPLAEDDTNAQKTITD CFLENNVWCISLVEIYDTGDSVIRPKL FAVKIPAQCSSES	14	18	52
AF212302	8943-10751	gb:AF212302 Organism:Nip ah virus Strain Name:UNKN OWN-	MPAENKKVRFENTTSKDGKIPSKVIKSY YGTMDIKKINEGLLDISKILSAFNTVIA LLGSIIIVMNIMIIQNYTRSTDNQA VIKDALQGIQQQIKGLADKIGTEIGPKV SLIDTSSTTIPANIGLLGSKISQSTASIN ENVNEKC	14	28	44

		AF212302 Protein Name:attachm ent glycoprotein Gene Symbol:G	KFTLPPLKIHECNISCPNLPFREYRPQTE GVSNLVGLPNNICLQKTSNQLKPKLISY TLPVVGQSGTCITDPLLAMDEGYFAYSH LERIGSCSRGVSKQRIIGVGEVLDGRDEV PSLFMTNVWTPPNPNTVYHCSAVYNNE FYYVLCVSTVGDPILNSTYWSGSLMM TRLAVKPKSNGGGYNQHLALRSIEKG RYDKVMPYGPSGIKQGDLYFPAVGFL VRTEFKYNDSPITKQCYSKPENCRLS MGIRPNSHYILRSGLLKYNLSDGENPKV VFIEISDQRLSIGSPSKIYDSLGPVFYQA SFSWDTMIKFGDVLTVNPLVVNWRNNT VISRPGQSQCPRFNTCPEICWEGVYNDA FLIDRINWISAGVFLDSNQTAENPVFTVF KDNEILYRAQLASEDTNAQKTITNCFL KNKIWCISLVEIYDTGDNVIRPKLFAVKI PEQCT			
JQ001 776	8170- 10275	gb:JQ001776: 8170- 10275 Organis m:Cedar virus Strain Name:CG1a Pro tein Name:attachm ent glycoprotein Gene Symbol:G	MLSQLQKNYLDNSNQGGDKMNNPDKK LSVNFNPLELDKGQKDLNKSYYVKNKN YNVSNLLNESLHDIKFCIYCIFSLLIITIN IITISIVITRLKVHEENNGMESPNLQSIQD SLSSLTNMINTETPRIGILVTATSVTLSSS INYVGTKTNQLVNELKDYITKSCGFKVP ELKLHECNISCADPKISKSAMYSTNAY A ELAGPPKIFCKSVSKDPDFRLKQIDYVIP VQQDRSICMNNPLLDISDGFFTYIHYE NSCKKSDSFKVLLSHGEIVDRGDYRPSL YLLSSHYHPYSMQVINCVPVTCNQSSFV FCHISNNTKTLDNSDYSSDEYYITYFNGI DRPKTKKIPINMTADNRYIHFTFSGGG GVCLGEEFIIPVTTVINTDVFTHDYCESF NCSVQTGKSLKEICSESLRSPTNSSRYNL NGIMIISQNNMTDFKIQLNGITYNKL SPGRLSKTLGQVLYYQSSMSWDYTLKA GFVEKWKPFPTNWMNNTVISRPNQGN PRYHKCPEICYGGTYNDIAPLDLGKDMY VSVILDSQLAENPEITVFNSTTILYKER VSKDELNTRSTTTSCFLFLDEPWCISVLE TNRFNKGSIRPEIYSYKIPKYC	3	29	54
NC_02 5256	9117- 11015	gb:NC_02525 6:9117- 11015 Organis m:Bat Paramyxovirus Eid_hel/GH- M74a/GHA/20 09 Strain Name:BatPV/ Eid_hel/GH- M74a/GHA/20 09 Protein Name:glycopr otein Gene Symbol:G	MPQKTVEFINMNSPLERGVSTLSDKKTL NQSKITKQGYFGLGSHSERNWKKQKNQ NDHYMTVSTMLEILVVLGIMFNLIVLT MVYYQNDNINQRMALTSNITVLNLNL NQLTNKIQREIIPRITLIDTATTITPSAIT ILATLTTRISELLPSINQKCEFKTPTLV DCRINCTPPLNPDSGVKMSSLATNLVAH GPSPCRNFSSVPTIYYRIPGLYNRTALD ERCILNPRLTISSTKFAYVHSEYDKNCTR GFKYYELMTFGEILEGPEKEPRMFSRSF YSPTNAVNYHSCTPIVTVNEGFLCLEC TSSDPLYKANLSNSTFHLVILRHNDKEKI VSMPSFNLSTDQEYVQIIPAEAGGTAESG NLYFPCIGRLLHKRVTHPLCKKSNCSRT DDESLKSYYNQGSPOHQVNVNCLIRIRN AQRDNPTWDVITVDLTNTYPGSRSRIFG SFSKPMYLYQSSVSWHTLLQVAEITDL YQLDWLDTPIYISRPGGSECPFGNYCPTV CWEGTYNDVYSLTPNNDLFVTYVLKSE	2	30	55

			QVAENPYFAIFSRDQILKEFPLDAWISSA RTTISCFMFNNEIWCIAALEITRLNDDII RPIYYSFWLPTDCRTPYPHTGKMTRVPL RSTYNY			
NC_02 5352	8716- 11257	gb:NC_02535 2:8716- 11257 Organis m:Mojiang virus Strain Name:Tonggu an1 Protein Name:attachm ent glycoprotein G ene Symbol:G	MATNRDNTITSAEVSQEDKVKKYYGVE TAEKVADSGNKNVFLMNTLLILTGAIT ITLNTNLTAAKSQQNMLKIIQDDVNAK LEMFVNLDQLVKGEIKPKVSLINTAVSV SIPGQISNLQTKFLQKYVYLEESITKQCT CNPLSGIFPTSGPTYPPDTKPDDDTTDDD KVDTTIKPIEYKPDGCNRTGDHFTMEP GANFYTVPNLGPASSNSDECYTNPSFSIG SSIYMFSQEIRKTDCTAGEILSIQIVLGRI VDKGQQGPQASPLLVAVPNPKIINSCA VAAGDEMGWVLCSTLTAAASGEPIPHM FDGFWLYKLEPDTEVVSRYITGYAYLLD KQYDSVFIGKGGGIQKGNLDLYFQMYGL SRNRQSFKALCEHGSCLGTGGGGYQVL CDRAVMSFGSEESLITNAYLKVNDLASG KPVIIQGTFFPSDSYKGSNGRMYTIGDKY GLYLAPSSWNRYLRFGITPDISVRSTTWL KSQDPIMKILSTCTNTDRDMCPEICNTRG YQDIFPLSESEYYTYIGITPNNGGTKNF VAVRSDSGHIASIDILQNYYSITSATISCF MYKDEIWCIATEGKKQKQDNPPRIYAH YKIRQMCYNMKSATVTVGNAKNITIRR Y	2	31	56

[0213] In some embodiments, the G protein has a sequence set forth in any of SEQ ID NOS: 9, 18, 28, 29, 30, 31, 44, 52, or 54-56 or is a functionally active variant or biologically active portion thereof that has a sequence that is at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% identical to any one of SEQ ID NOS: 9, 18, 28, 29, 30, 31, 44, 52, or 54-56. In particular embodiments, the G protein or functionally active variant or biologically active portion is a protein that retains fusogenic activity in conjunction with a Henipavirus F protein, such as an F protein set forth in Section I.B (e.g. NiV-F or HeV-F). Fusogenic activity includes the activity of the G protein in conjunction with a Henipavirus F protein to promote or facilitate fusion of two membrane lumens, such as the lumen of the targeted lipid particle having embedded in its lipid bilayer a henipavirus F and G protein, and a cytoplasm of a target cell, e.g. a cell that contains a surface receptor or molecule that is recognized or bound by the

targeted envelope protein. In some embodiments, the F protein and G protein are from the same Henipavirus species (e.g. NiV-G and NiV-F). In some embodiments, the F protein and G protein are from different Henipavirus species (e.g. NiV-G and HeV-F).

[0214] In particular embodiments, the G protein has the sequence of amino acids set forth in SEQ ID NO: 9, SEQ ID NO: 28, SEQ ID NO: 18, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 44, SEQ ID NO: 52 or SEQ ID NO: 54-56 or is a functionally active variant thereof or a biologically active portion thereof that retains fusogenic activity. In some embodiments, the functionally active variant comprises an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28, SEQ ID NO: 18, SEQ ID NO:30, SEQ ID NO: 31, SEQ ID NO: 44, SEQ ID NO: 52 or SEQ ID NO: 54-56 and retains fusogenic activity in conjunction with a Henipavirus F protein (e.g., NiV-F or HeV-F). In some embodiments, the biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28, SEQ ID NO: 18, SEQ ID NO:30, SEQ ID NO: 31, SEQ ID NO: 44, SEQ ID NO: 52 or SEQ ID NO: 54-56 and retains fusogenic activity in conjunction with a Henipavirus F protein (e.g., NiV-F or HeV-F).

[0215] Reference to retaining fusogenic activity includes activity (in conjunction with a Henipavirus F protein) that is between at or about 10% and at or about 150% or more of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:28, SEQ ID NO: 18, SEQ ID NO:30, SEQ ID NO: 31, SEQ ID NO: 44, SEQ ID NO: 52 or SEQ ID NO: 54-56 such as at least or at least about 10% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 15% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 20% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 25% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 30% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such

as at least or at least about 35% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 40% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 45% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 50% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 55% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 60% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 65% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 70% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 75% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 80% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 85% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 90% of the level or degree of fusogenic activity of the corresponding wild-type G protein, such as at least or at least about 95% of the level or degree of fusogenic activity of the corresponding wild-type G protein, or such as at least or at least about 100% of the level or degree of fusogenic activity of the corresponding wild-type G protein, or such as at least or at least about 120% of the level or degree of fusogenic activity of the corresponding wild-type G protein.

[0216] In some embodiments the G protein is a mutant G protein that is a functionally active variant or biologically active portion containing one or more amino acid mutations, such as one or more amino acid insertions, deletions, substitutions or truncations. In some embodiments, the mutations described herein relate to amino acid insertions, deletions, substitutions or truncations of amino acids compared to a reference G protein sequence. In some embodiments, the reference G protein sequence is the wild-type sequence of a G protein or a biologically active portion thereof. In some embodiments, the functionally active variant or the biologically active portion thereof is a mutant of a wild-type Hendra (HeV) virus G protein, a wild-type Nipah (NiV) virus G-protein (NiV-G), a wild-type Cedar (CedPV) virus G-protein, a wild-type Mojiang virus G-protein, a wild-type bat Paramyxovirus G-protein or biologically active portion thereof. In some embodiments, the wild-type G

protein has the sequence set forth in any one of SEQ ID NOS: 9, 18, 28, 29, 30, 31 SEQ ID NO: 44, SEQ ID NO: 52 or SEQ ID NO: 54-56.

[0217] In some embodiments, the G protein is a mutant G protein that is a biologically active portion that is an N-terminally and/or C-terminally truncated fragment of a wild-type Hendra (HeV) virus G protein, a wild-type Nipah (NiV) virus G-protein (NiV-G), a wild-type Cedar (CedPV) virus G-protein, a wild-type Mojiang virus G-protein, a wild-type bat Paramyxovirus G-protein. In particular embodiments, the truncation is an N-terminal truncation of all or a portion of the cytoplasmic domain. In some embodiments, the mutant G protein is a biologically active portion that is truncated and lacks up to 49 contiguous amino acid residues at or near the N-terminus of the wild-type G protein, such as a wild-type G protein set forth in any one of SEQ ID NOS: 9, 18, 28, 29, 30, 31, SEQ ID NO: 44, SEQ ID NO: 52 or SEQ ID NO: 54-56. In some embodiments, the mutant F protein is truncated and lacks up to 49 contiguous amino acids, such as up to 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 contiguous amino acids at the N-terminus of the wild-type G protein.

[0218] In some embodiments, the G protein is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein, or is a functionally active variant or biologically active portion thereof. In some embodiments, the G protein is a NiV-G protein that has the sequence set forth in SEQ ID NO: 9, SEQ ID NO: 28 or SEQ ID NO: 44, or is a functional variant or a biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, at least at or about 99% sequence identity to SEQ ID NO: 9, SEQ ID NO: 28 or SEQ ID NO: 44.

[0219] In some embodiments, the G protein is a mutant NiV-G protein that is a biologically active portion of a wild-type NiV-G. In some embodiments, the biologically active portion is an N-terminally truncated fragment. In some embodiments, the mutant NiV-G protein is truncated and lacks up to 5 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO: 28 or SEQ ID

[illegible]

[illegible]

terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), or up to 45 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

[0220] In some embodiments, the NiV-G protein is a biologically active portion that does not contain a cytoplasmic domain. In some embodiments, the NiV-G protein without the cytoplasmic domain is encoded by SEQ ID NO: 32.

[0221] In some embodiments, the mutant NiV-G protein comprises a sequence set forth in any of SEQ ID NOS: 10-15, 35-40, 45-50, 22, 53 or SEQ ID NO: 32, or is a functional variant thereof that has an amino acid sequence having at least at or 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40, 45-50, 22, 53 or SEQ ID NO:32.

[0222] In some embodiments, the mutant NiV-G protein has a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), such as set forth in SEQ ID NO: 10 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10 or such as set forth in SEQ ID NO: 35 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35 or such as set forth in SEQ ID NO: 45 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least

at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45. In some embodiments, the mutant NiV-G protein has a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), such as set forth in SEQ ID NO: 11 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11, or such as set forth in SEQ ID NO: 36 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36 or such as set forth in SEQ ID NO: 46 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

[0223] In some embodiments, the mutant NiV-G protein has a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), such as set forth in SEQ ID NO: 12 or a functional variant thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at

least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12 or such as set forth in SEQ ID NO: 37 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37 or such as set forth in SEQ ID NO: 47 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47. In some embodiments, the mutant NiV-G protein has a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44) such as set forth in SEQ ID NO: 13, or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13 or such as set forth in SEQ ID NO: 38 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38 or such as set forth in SEQ ID NO: 48 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about

82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48. In some embodiments, the mutant NiV-G protein has a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), such as set forth in SEQ ID NO: 14 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14 or such as set forth in SEQ ID NO: 39 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39 or such as set forth in SEQ ID NO: 49 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49. In some embodiments, the mutant NiV-G protein has a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), such as set forth in SEQ ID NO: 15 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least

at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15 or such as set forth in SEQ ID NO: 40 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40, or such as set forth in SEQ ID NO: 50 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:50. In some embodiments, the mutant NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44), such as set forth in SEQ ID NO: 22 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:22 or such as set forth in SEQ ID NO: 53 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:53. In some embodiments, the mutant NiV-G protein lacks the N-terminal cytoplasmic domain of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID

NO:44), such as set forth in SEQ ID NO:32 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:32.

[0224] In some embodiments, the mutant G protein is a mutant HeV-G protein that has the sequence set forth in SEQ ID NO:18 or 52, or is a functional variant or biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at or about 85%, at least at or about 86%, at least at or about 87%, at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:18 or 52.

[0225] In some embodiments, the G protein is a mutant HeV-G protein that is a biologically active portion of a wild-type HeV-G. In some embodiments, the biologically active portion is an N-terminally truncated fragment. In some embodiments, the mutant HeV-G protein is truncated and lacks up to 5 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 6 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 7 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 8 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 9 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 10 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 11 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 12 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 13 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 14 contiguous

[illegible]

terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 38 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 39 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 41 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 42 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 43 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), up to 44 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52), or up to 45 contiguous amino acid residues at or near the N-terminus of the wild-type HeV-G protein (SEQ ID NO:18 or 52). In some embodiments, the HeV-G protein is a biologically active portion that does not contain a cytoplasmic domain. In some embodiments, the mutant HeV-G protein lacks the N-terminal cytoplasmic domain of the wild-type HeV-G protein (SEQ ID NO:18 or 52), such as set forth in SEQ ID NO:33 or a functional variant thereof having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:33.

[0226] In some embodiments, the G protein or the functionally active variant or biologically active portion thereof binds to Ephrin B2 or Ephrin B3. In some aspects, the G protein has the sequence of amino acids set forth in any one of SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31 , or is a functionally active variant thereof or a biologically active portion thereof that is able to bind to Ephrin B2 or Ephrin B3. In some embodiments, the functionally active variant or biologically active portion has an amino acid sequence having at least about 80%, at least about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID

NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, and retains binding to Ephrin B2 or B3. Reference to retaining binding to Ephrin B2 or B3 includes binding that is at least or at least about 5% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 10% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 15% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 20% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 25% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 30% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 35% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 40% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 45% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 50% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ

ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 55% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 60% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 65% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, 70% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, such as at least or at least about 75% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, such as at least or at least about 80% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, such as at least or at least about 85% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, such as at least or at least about 90% of the level or degree of binding of the corresponding wild-type G protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof, or such as at least or at least about 95% of the level or degree of binding of the corresponding wild-type protein, such as set forth in SEQ ID NO:9, SEQ ID NO:18 or SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO: 44, SEQ ID NO:30 or SEQ ID NO:31, or a functionally active variant or biologically active portion thereof. In some embodiments, the G protein is NiV-G or a functionally active variant or biologically active portion thereof and binds to

Ephrin B2 or Ephrin B3. In some aspects, the NiV-G has the sequence of amino acids set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, or is a functionally active variant thereof or a biologically active portion thereof that is able to bind to Ephrin B2 or Ephrin B3. In some embodiments, the functionally active variant or biologically active portion has an amino acid sequence having at least about 80%, at least about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44 and retains binding to Ephrin B2 or B3. Exemplary biologically active portions include N-terminally truncated variants lacking all or a portion of the cytoplasmic domain, e.g. 1 or more, such as 1 to 49 contiguous N-terminal amino acid residues, e.g. set forth in any one of SEQ ID NOS: 10-15, 35-40, 45-50 and 32. Reference to retaining binding to Ephrin B2 or B3 includes binding that is at least or at least about 5% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 10% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 15% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 20% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 25% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 30% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 35% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 40% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 45% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 50% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 55% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 60% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 65% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ

ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, 70% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, such as at least or at least about 75% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, such as at least or at least about 80% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, such as at least or at least about 85% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, such as at least or at least about 90% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44, or such as at least or at least about 95% of the level or degree of binding of the corresponding wild-type NiV-G, such as set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44.

[0227] In some embodiments, the G protein is HeV-G or a functionally active variant or biologically active portion thereof and binds to Ephrin B2 or Ephrin B3. In some aspects, the HeV-G has the sequence of amino acids set forth in SEQ ID NO:18 or 52, or is a functionally active variant thereof or a biologically active portion thereof that is able to bind to Ephrin B2 or Ephrin B3. In some embodiments, the functionally active variant or biologically active portion has an amino acid sequence having at least about 80%, at least about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:18 or 52 and retains binding to Ephrin B2 or B3. Exemplary biologically active portions include N-terminally truncated variants lacking all or a portion of the cytoplasmic domain, e.g. 1 or more, such as 1 to 49 contiguous N-terminal amino acid residues, e.g. set forth in any one of SEQ ID NO:33. Reference to retaining binding to Ephrin B2 or B3 includes binding that is at least or at least about 5% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 10% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 15% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 20% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 25% of the level or degree of binding of the corresponding wild-type

HeV-G, such as set forth in SEQ ID NO:18 or 52, 30% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 35% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 40% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 45% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 50% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 55% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 60% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 65% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, 70% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, such as at least or at least about 75% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, such as at least or at least about 80% of the level or degree of binding of the corresponding wild-type NIV-G, such as set forth in SEQ ID NO:18 or 52, such as at least or at least about 85% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, such as at least or at least about 90% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52, or such as at least or at least about 95% of the level or degree of binding of the corresponding wild-type HeV-G, such as set forth in SEQ ID NO:18 or 52.

[0228] In some embodiments, the G protein or the biologically thereof is a mutant G protein that exhibits reduced binding for the native binding partner of a wild-type G protein. In some embodiments, the mutant G protein or the biologically active portion thereof is a mutant of wild-type Niv-G and exhibits reduced binding to one or both of the native binding partners Ephrin B2 or Ephrin B3. In some embodiments, the mutant G-protein or the biologically active portion, such as a mutant NiV-G protein, exhibits reduced binding to the native binding partner. In some embodiments, the reduced binding to Ephrin B2 or Ephrin B3 is reduced by greater than at or about 5%, at or about 10%, at or about 15%, at or about 20%, at or about 25%, at or about 30%, at or about 40%, at or about 50%, at or about 60%, at or about 70%, at or about 80%, at or about 90%, or at or about 100%.

[0229] In some embodiments, the mutations described herein can improve transduction efficiency. In some embodiments, the mutations described herein allow for specific

targeting of other desired cell types that are not Ephrin B2 or Ephrin B3. In some embodiments, the mutations described herein result in at least the partial inability to bind at least one natural receptor, such as reduce the binding to at least one of Ephrin B2 or Ephrin B3. In some embodiments, the mutations described herein interfere with natural receptor recognition.

[0230] In some embodiments, the G protein contains one or more amino acid substitutions in a residue that is involved in the interaction with one or both of Ephrin B2 and Ephrin B3. In some embodiments, the amino acid substitutions correspond to mutations E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

[0231] In some embodiments, the G protein is a mutant G protein containing one or more amino acid substitutions selected from the group consisting of E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28. In some embodiments, the G protein is a mutant G protein that contains one or more amino acid substitutions elected from the group consisting of E501A, W504A, Q530A and E533A with reference to SEQ ID NO:28 and is a biologically active portion thereof containing an N-terminal truncation. In some embodiments, the mutant NiV-G protein or the biologically active portion thereof is truncated and lacks up to 5 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 6 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 7 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 8 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 9 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), up to 10 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 11 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 12 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 13 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 14 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), up to 15 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 16 contiguous amino acid residues at or near the N-terminus of the wild-

type NiV-G protein (SEQ ID NO:28), 17 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 18 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 19 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), up to 20 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 21 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 22 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 23 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 24 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), up to 25 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 26 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 27 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 28 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 29 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), up to 30 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (EQ ID NO:28), up to 31 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 32 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 33 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 34 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), 35 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:28), up to 36 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (EQ ID NO:28), up to 37 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (EQ ID NO:28), up to 38 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (EQ ID NO:28), up to 39 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (EQ ID NO:28), or up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (EQ ID NO:28).

[0232] In some embodiments, the mutant NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 16 or 51 or an amino acid sequence having at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16 or 51. In particular embodiments, the G protein has the sequence of amino acids set forth in SEQ ID NO: 16 or 51.

[0233] In some embodiments, the targeted envelope protein contains a G protein or a functionally active variant or biologically active portion and an sdAb variable domain, in which the targeted envelope protein exhibits increased binding for another molecule that is different from the native binding partner of a wild-type G protein. In some embodiments, the molecule can be a protein expressed on the surface of desired target cell. In some embodiments, the increased binding to the other molecule is increased by greater than at or about 25%, at or about 30%, at or about 40%, at or about 50%, at or about 60%, at or about 70%, at or about 80%, at or about 90%, or at or about 100%. In particular embodiments, the binding confers re-targeted binding compared to the binding of a wild-type G protein in which a new or different binding activity is conferred.

2. Binding domain

[0234] In some embodiments, the binding domain can be any agent that binds to a cell surface molecule on a target cells. In some embodiments, the binding domain can be an antibody or an antibody portion or fragment.

[0235] The binding domain may be modulated to have different binding strengths. For example, scFvs and antibodies with various binding strengths may be used to alter the fusion activity of the chimeric attachment proteins towards cells that display high or low amounts of the target antigen. For example DARPin with different affinities may be used to alter the fusion activity towards cells that display high or low amounts of the target antigen. Binding domains may also be modulated to target different regions on the target ligand, which will affect the fusion rate with cells displaying the target..

[0236] The binding domain may comprise a humanized antibody molecule, intact IgA, IgG, IgE or IgM antibody; bi- or multi- specific antibody (e.g., Zybodies®, etc); antibody fragments such as Fab fragments, Fab' fragments, F(ab')₂ fragments, Fd' fragments, Fd fragments, and isolated CDRs or sets thereof; single chain Fvs; polypeptide-Fc fusions; single domain antibodies (e.g., shark single domain antibodies such as IgNAR or fragments

thereof); cameloid antibodies; masked antibodies (e.g., Probodies®); Small Modular ImmunoPharmaceuticals (“SMIPs™”); single chain or Tandem diabodies (TandAb®); VHHs; Anticalins®; Nanobodies®; minibodies; BiTE®s; ankyrin repeat proteins or DARPINs®; Avimers®; DARTs; TCR-like antibodies; Adnectins®; Affilins®; Transbodies®; Affibodies®; TrimerX®; MicroProteins; Fynomers®, Centyrins®; and KALBITOR®s. A targeting moiety can also include an antibody or an antigen-binding fragment thereof (e.g., Fab, Fab', F(ab')₂, Fv fragments, scFv antibody fragments, disulfide-linked Fvs (sdFv), a Fd fragment consisting of the VH and CH1 domains, linear antibodies, single domain antibodies such as sdAb (either VL or VH), nanobodies, or camelid VHH domains), an antigen-binding fibronectin type III (Fn3) scaffold such as a fibronectin polypeptide minibody, a ligand, a cytokine, a chemokine, or a T cell receptor (TCRs).

[0237] In some embodiments, the binding domain is a single chain molecule. In some embodiments, the binding domain is a single domain antibody. In some embodiments, the binding domain is a single chain variable fragment. In particular embodiments, the binding domain contains an antibody variable sequence (s) that is human or humanized.

[0238] In some embodiments, the binding domain is a single domain antibody. In some embodiments, the single domain antibody can be human or humanized. In some embodiments, the single domain antibody or portion thereof is naturally occurring. In some embodiments, the single domain antibody or portion thereof is synthetic.

[0239] In some embodiments, the single domain antibodies are antibodies whose complementary determining regions are part of a single domain polypeptide. In some embodiments, the single domain antibody is a heavy chain only antibody variable domain. In some embodiments, the single domain antibody does not include light chains.

[0240] In some embodiments, the heavy chain antibody devoid of light chains is referred to as VHH. In some embodiments, the single domain antibody antibodies have a molecular weight of 12-15 kDa. In some embodiments, the single domain antibody antibodies include camelid antibodies or shark antibodies. In some embodiments, the single domain antibody molecule is derived from antibodies raised in Camelidae species, for example in camel, llama, dromedary, alpaca, vicuna and guanaco. In some embodiments, the single domain antibody is referred to as immunoglobulin new antigen receptors (IgNARs) and is derived from cartilaginous fishes. In some embodiments, the single domain antibody is generated by splitting dimeric variable domains of human or mouse IgG into monomers and camelizing critical residues.

[0241] In some embodiments, the single domain antibody can be generated from phage display libraries. In some embodiments, the phage display libraries are generated from a VHH repertoire of camelids immunized with various antigens, as described in Arbabi et al., *FEBS Letters*, 414, 521-526 (1997); Lauwereys et al., *EMBO J.*, 17, 3512-3520 (1998); Decanniere et al., *Structure*, 7, 361-370 (1999). In some embodiments, the phage display library is generated comprising antibody fragments of a non-immunized camelid. In some embodiments, single domain antibodies a library of human single domain antibodies is synthetically generated by introducing diversity into one or more scaffolds.

[0242] In some embodiments, the C-terminus of the single domain antibody is attached to the C-terminus of the G protein or biologically active portion thereof. In some embodiments, the N-terminus of the single domain antibody is exposed on the exterior surface of the lipid bilayer. In some embodiments, the N-terminus of the single domain antibody binds to a cell surface molecule of a target cell. In some embodiments, the single domain antibody specifically binds to a cell surface molecule present on a target cell. In some embodiments, the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule.

[0243] In some embodiments, the cell surface molecule of a target cell is an antigen or portion thereof. In some embodiments, the single domain antibody or portion thereof is an antibody having a single monomeric domain antigen binding/recognition domain that is able to bind selectively to a specific antigen. In some embodiments, the single domain antibody binds an antigen present on a target cell.

[0244] Exemplary cells include polymorphonuclear cells (also known as PMN, PML, PMNL, or granulocytes), stem cells, embryonic stem cells, neural stem cells, mesenchymal stem cells (MSCs), hematopoietic stem cells (HSCs), human myogenic stem cells, muscle-derived stem cells (MuStem), embryonic stem cells (ES or ESCs), limbal epithelial stem cells, cardio-myogenic stem cells, cardiomyocytes, progenitor cells, immune effector cells, lymphocytes, macrophages, dendritic cells, natural killer cells, T cells, cytotoxic T lymphocytes, allogenic cells, resident cardiac cells, induced pluripotent stem cells (iPS), adipose-derived or phenotypic modified stem or progenitor cells, CD133+ cells, aldehyde dehydrogenase-positive cells (ALDH+), umbilical cord blood (UCB) cells, peripheral blood stem cells (PBSCs), neurons, neural progenitor cells, pancreatic beta cells, glial cells, or hepatocytes,

[0245] In some embodiments, the target cell is a cell of a target tissue. The target tissue can include 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.

[0246] In some embodiments, the target cell is a muscle cell (e.g., skeletal muscle cell), kidney cell, liver cell (e.g. hepatocyte), or a cardiac cell (e.g. cardiomyocyte). In some 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 T cell (e.g. a naive T cell), a macrophage (e.g., a tumor infiltrating macrophage), or a fibroblast (e.g., a cardiac fibroblast).

[0247] In some embodiments, the target cell is a tumor-infiltrating lymphocyte, a T cell, a neoplastic or tumor cell, a virus-infected cell, a stem cell, a central nervous system (CNS) cell, a hematopoietic stem cell (HSC), a liver cell or a fully differentiated cell. In some embodiments, the target cell is a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

[0248] In some embodiments, the target cell is 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).

[0249] In some embodiments, the cell surface molecule is any one of CD8, CD4, asialoglycoprotein receptor 2 (ASGR2), transmembrane 4 L6 family member 5 (TM4SF5), low density lipoprotein receptor (LDLR) or asialoglycoprotein 1 (ASGR1).

[0250] In some embodiments, the G protein or functionally active variant or biologically active portion thereof is linked directly to the sdAb variable domain. In some embodiments, the targeted envelope protein is a fusion protein that has the following structure: (N'-single domain antibody-C')-(C'-G protein-N').

[0251] In some embodiments, the G protein or functionally active variant or biologically active portion thereof is linked indirectly via a linker to the the sdAb variable domain. In some

embodiments, the linker is a peptide linker. In some embodiments, the linker is a chemical linker.

[0252] In some embodiments, the linker is a peptide linker and the targeted envelope protein is a fusion protein containing the G protein or functionally active variant or biologically active portion thereof linked via a peptide linker to the sdAb variable domain. In some embodiments, the targeted envelope protein is a fusion protein that has the following structure: (N'-single domain antibody-C')-Linker-(C'-G protein-N').

[0253] In some embodiments, the peptide linker is up to 65 amino acids in length. In some embodiments, the peptide linker comprises from or from about 2 to 65 amino acids, 2 to 60 amino acids, 2 to 56 amino acids, 2 to 52 amino acids, 2 to 48 amino acids, 2 to 44 amino acids, 2 to 40 amino acids, 2 to 36 amino acids, 2 to 32 amino acids, 2 to 28 amino acids, 2 to 24 amino acids, 2 to 20 amino acids, 2 to 18 amino acids, 2 to 14 amino acids, 2 to 12 amino acids, 2 to 10 amino acids, 2 to 8 amino acids, 2 to 6 amino acids, 6 to 65 amino acids, 6 to 60 amino acids, 6 to 56 amino acids, 6 to 52 amino acids, 6 to 48 amino acids, 6 to 44 amino acids, 6 to 40 amino acids, 6 to 36 amino acids, 6 to 32 amino acids, 6 to 28 amino acids, 6 to 24 amino acids, 6 to 20 amino acids, 6 to 18 amino acids, 6 to 14 amino acids, 6 to 12 amino acids, 6 to 10 amino acids, 6 to 8 amino acids, 8 to 65 amino acids, 8 to 60 amino acids, 8 to 56 amino acids, 8 to 52 amino acids, 8 to 48 amino acids, 8 to 44 amino acids, 8 to 40 amino acids, 8 to 36 amino acids, 8 to 32 amino acids, 8 to 28 amino acids, 8 to 24 amino acids, 8 to 20 amino acids, 8 to 18 amino acids, 8 to 14 amino acids, 8 to 12 amino acids, 8 to 10 amino acids, 10 to 65 amino acids, 10 to 60 amino acids, 10 to 56 amino acids, 10 to 52 amino acids, 10 to 48 amino acids, 10 to 44 amino acids, 10 to 40 amino acids, 10 to 36 amino acids, 10 to 32 amino acids, 10 to 28 amino acids, 10 to 24 amino acids, 10 to 20 amino acids, 10 to 18 amino acids, 10 to 14 amino acids, 10 to 12 amino acids, 12 to 65 amino acids, 12 to 60 amino acids, 12 to 56 amino acids, 12 to 52 amino acids, 12 to 48 amino acids, 12 to 44 amino acids, 12 to 40 amino acids, 12 to 36 amino acids, 12 to 32 amino acids, 12 to 28 amino acids, 12 to 24 amino acids, 12 to 20 amino acids, 12 to 18 amino acids, 12 to 14 amino acids, 14 to 65 amino acids, 14 to 60 amino acids, 14 to 56 amino acids, 14 to 52 amino acids, 14 to 48 amino acids, 14 to 44 amino acids, 14 to 40 amino acids, 14 to 36 amino acids, 14 to 32 amino acids, 14 to 28 amino acids, 14 to 24 amino acids, 14 to 20 amino acids, 14 to 18 amino acids, 18 to 65 amino acids, 18 to 60 amino acids, 18 to 56 amino acids, 18 to 52 amino acids, 18 to 48 amino acids, 18 to 44 amino acids, 18 to 40 amino acids, 18 to 36 amino acids, 18 to 32 amino acids, 18 to 28 amino acids, 18 to 24

amino acids, 18 to 20 amino acids, 20 to 65 amino acids, 20 to 60 amino acids, 20 to 56 amino acids, 20 to 52 amino acids, 20 to 48 amino acids, 20 to 44 amino acids, 20 to 40 amino acids, 20 to 36 amino acids, 20 to 32 amino acids, 20 to 28 amino acids, 20 to 26 amino acids, 20 to 24 amino acids, 24 to 65 amino acids, 24 to 60 amino acids, 24 to 56 amino acids, 24 to 52 amino acids, 24 to 48 amino acids, 24 to 44 amino acids, 24 to 40 amino acids, 24 to 36 amino acids, 24 to 32 amino acids, 24 to 30 amino acids, 24 to 28 amino acids, 28 to 65 amino acids, 28 to 60 amino acids, 28 to 56 amino acids, 28 to 52 amino acids, 28 to 48 amino acids, 28 to 44 amino acids, 28 to 40 amino acids, 28 to 36 amino acids, 28 to 34 amino acids, 28 to 32 amino acids, 32 to 65 amino acids, 32 to 60 amino acids, 32 to 56 amino acids, 32 to 52 amino acids, 32 to 48 amino acids, 32 to 44 amino acids, 32 to 40 amino acids, 32 to 38 amino acids, 32 to 36 amino acids, 36 to 65 amino acids, 36 to 60 amino acids, 36 to 56 amino acids, 36 to 52 amino acids, 36 to 48 amino acids, 36 to 44 amino acids, 36 to 40 amino acids, 40 to 65 amino acids, 40 to 60 amino acids, 40 to 56 amino acids, 40 to 52 amino acids, 40 to 48 amino acids, 40 to 44 amino acids, 44 to 65 amino acids, 44 to 60 amino acids, 44 to 56 amino acids, 44 to 52 amino acids, 44 to 48 amino acids, 48 to 65 amino acids, 48 to 60 amino acids, 48 to 56 amino acids, 48 to 52 amino acids, 50 to 65 amino acids, 50 to 60 amino acids, 50 to 56 amino acids, 50 to 52 amino acids, 54 to 65 amino acids, 54 to 60 amino acids, 54 to 56 amino acids, 58 to 65 amino acids, 58 to 60 amino acids, or 60 to 65 amino acids. In some embodiments, the peptide linker is a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, or 65 amino acids in length.

[0254] In particular embodiments, the linker is a flexible peptide linker. In some such embodiments, the linker is 1-20 amino acids, such as 1-20 amino acids predominantly composed of glycine. In some embodiments, the linker is 1-20 amino acids, such as 1-20 amino acids predominantly composed of glycine and serine. In some embodiments, the linker is a flexible peptide linker containing amino acids Glycine and Serine, referred to as GS-linkers. In some embodiments, the peptide linker includes the sequences GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) or combinations thereof. In some embodiments, the polypeptide linker has the sequence (GGS)*n*, wherein *n* is 1 to 10. In some embodiments, the polypeptide linker has the sequence (GGGGS)*n*, (SEQ ID NO:42) wherein *n* is 1 to 10. In some embodiments, the polypeptide linker has the sequence (GGGGGS)*n* (SEQ ID NO:27), wherein *n* is 1 to 6.

3. Polynucleotides

[0255] Provided herein are polynucleotides comprising a nucleic acid sequence encoding a targeted envelope protein. In some embodiments, the polynucleotides comprise a nucleic acid sequence encoding a G protein or biologically active portion thereof. In some embodiments, the polynucleotides further comprise a nucleic acid sequence encoding a single domain antibody (sdAb) variable domain or biologically active portion thereof. The polynucleotides may include a sequence of nucleotides encoding any of the targeted envelope proteins described above. The polynucleotide can be a synthetic nucleic acid. Also provided are expression vector containing any of the provided polynucleotides.

[0256] In some of any embodiments, 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. In some embodiments, vectors can be suitable for replication and integration in eukaryotes. In some embodiments, cloning vectors contain transcription and translation terminators, initiation sequences, and promoters useful for expression of the desired nucleic acid sequence. In some of any embodiments, a plasmid comprises a promoter suitable for expression in a cell.

[0257] In some embodiments, the polynucleotides contain at least one promoter that is operatively linked to control expression of the targeted envelope protein containing the G protein and the single domain antibody (sdAb) variable domain. For expression of the targeted envelope protein, at least one module in each promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some promoters lacking a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 genes, a discrete element overlying the start site itself helps to fix the place of initiation.

[0258] In some embodiments, additional promoter elements, e.g., enhancers, regulate the frequency of transcriptional initiation. In some embodiments, additional promoter elements 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. In some embodiments, spacing between promoter elements frequently is flexible, so that promoter function is preserved when elements are inverted or moved relative to one another. In some embodiments, the thymidine kinase (tk) promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. In some embodiments, depending

on the promoter, individual elements can function either cooperatively or independently to activate transcription.

[0259] A promoter may be one naturally associated with a gene or polynucleotide sequence, as may be obtained by isolating the 5' non-coding sequences located upstream of the coding segment and/or exon. Such a promoter can be referred to as "endogenous." Similarly, an enhancer may be one naturally associated with a polynucleotide sequence, located either downstream or upstream of that sequence. Alternatively, certain advantages will be gained by positioning the coding polynucleotide segment under the control of a recombinant or heterologous promoter, which refers to a promoter that is not normally associated with a polynucleotide sequence in its natural environment. A recombinant or heterologous enhancer refers also to an enhancer not normally associated with a polynucleotide sequence in its natural environment. Such promoters or enhancers may include promoters or enhancers of other genes, and promoters or enhancers isolated from any other prokaryotic, viral, or eukaryotic cell, and promoters or enhancers not "naturally occurring," i.e., containing different elements of different transcriptional regulatory regions, and/or mutations that alter expression. In addition to producing nucleic acid sequences of promoters and enhancers synthetically, sequences may be produced using recombinant cloning and/or nucleic acid amplification technology, including PCR, in connection with the compositions disclosed herein (U.S. Pat. Nos. 4,683,202 and 5,928,906).

[0260] In some embodiments, a suitable promoter is the immediate early cytomegalovirus (CMV) promoter sequence. In some embodiments, the promoter sequence is a strong constitutive promoter sequence capable of driving high levels of expression of any polynucleotide sequence operatively linked thereto. In some embodiments, a suitable promoter is Elongation Growth Factor-1 α (EF-1 α). In some embodiments, 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.

[0261] In some embodiments, the promoter is an inducible promoter. In some embodiments, the 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. In some embodiments, inducible promoters comprise metallothionine promoter, a glucocorticoid promoter, a progesterone promoter, and a tetracycline promoter.

[0262] In some embodiments, exogenously controlled inducible promoters can be used to regulate expression of the G protein and single domain antibody (sdAb) variable domain. For example, radiation-inducible promoters, heat-inducible promoters, and/or drug-inducible promoters can be used to selectively drive transgene expression in, for example, targeted regions. In such embodiments, the location, duration, and level of transgene expression can be regulated by the administration of the exogenous source of induction.

[0263] In some embodiments, expression of the targeted envelope protein containing a G protein and single domain antibody (sdAb) variable domain is regulated using a drug-inducible promoter. For example, in some cases, the promoter, enhancer, or transactivator comprises a Lac operator sequence, a tetracycline operator sequence, a galactose operator sequence, a doxycycline operator sequence, a rapamycin operator sequence, a tamoxifen operator sequence, or a hormone-responsive operator sequence, or an analog thereof. In some instances, the inducible promoter comprises a tetracycline response element (TRE). In some embodiments, the inducible promoter comprises an estrogen response element (ERE), which can activate gene expression in the presence of tamoxifen. In some instances, a drug-inducible element, such as a TRE, can be combined with a selected promoter to enhance transcription in the presence of drug, such as doxycycline. In some embodiments, the drug-inducible promoter is a small molecule-inducible promoter.

[0264] Any of the provided polynucleotides can be modified to remove CpG motifs and/or to optimize codons for translation in a particular species, such as human, canine, feline, equine, ovine, bovine, etc. species. In some embodiments, the polynucleotides are optimized for human codon usage (i.e., human codon-optimized). In some embodiments, the polynucleotides are modified to remove CpG motifs. In other embodiments, the provided polynucleotides are modified to remove CpG motifs and are codon-optimized, such as human codon-optimized. Methods of codon optimization and CpG motif detection and modification are well-known. Typically, polynucleotide optimization enhances transgene expression, increases transgene stability and preserves the amino acid sequence of the encoded polypeptide.

[0265] In order to assess the expression of the targeted envelope protein, the expression vector to be introduced into a cell can also contain either a selectable marker gene or a reporter gene or both to facilitate identification and selection of expressing particles, e.g. viral particles. In other embodiments, 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 are known in the art and include, for example, antibiotic-resistance genes, such as neo and the like.

[0266] Reporter genes are used for identifying potentially transfected cells and for evaluating the functionality of regulatory sequences. Reporter genes that encode for easily assayable proteins are well known in the art. In general, a reporter gene is a gene that is not present in or expressed by the recipient organism or tissue and that encodes a protein 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.

[0267] Suitable reporter genes may include genes encoding luciferase, beta-galactosidase, chloramphenicol acetyl transferase, secreted alkaline phosphatase, or the green fluorescent protein gene (see, e.g., Ui-Tei et al., 2000, FEBS Lett. 479:79-82). Suitable expression systems are well known and may be prepared using well known techniques or obtained commercially. Internal deletion constructs may be generated using unique internal restriction sites or by partial digestion of non-unique restriction sites. Constructs may then be transfected into cells that display high levels of the desired polynucleotide and/or polypeptide expression. 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.

B. Fusogen (e.g. henipavirus F protein)

[0268] In some embodiments, the targeted lipid particle comprises one or more fusogens. In some embodiments, the targeted lipid particle contains an exogenous or overexpressed fusogen. In some embodiments, the fusogen is disposed in the lipid bilayer. In some embodiments, the fusogen facilitates the fusion of the targeted lipid particle to a membrane. In some embodiments, the membrane is a plasma cell membrane.

[0269] In some embodiments, fusogens comprise protein based, lipid based, and chemical based fusogens. In some embodiments, the targeted lipid particle comprises a first fusogen comprising a protein fusogen and a second fusogen comprising a lipid fusogen or chemical fusogen. In some embodiments, the fusogen binds fusogen binding partner on a target cell surface.

[0270] In some embodiments, the fusogen comprises a protein with a hydrophobic fusion peptide domain. In some embodiments, the fusogen comprises a henipavirus F protein molecule or biologically active portion thereof. In some embodiments, the Henipavirus F protein is a Hendra (Hev) virus F protein, a Nipah (NiV) virus F-protein, a Cedar (CedPV) virus F protein, a Mojang virus F protein or a bat Paramyxovirus F protein or a biologically active portion thereof.

[0271] **Table 4** provides non-limiting examples of F proteins. In some embodiments, the N-terminal hydrophobic fusion peptide domain of the F protein molecule or biologically active portion thereof is exposed on the outside of lipid bilayer.

[0272] F proteins of henipaviruses are encoded as F₀ precursors containing a signal peptide (e.g. corresponding to amino acid residues 1-26 of SEQ ID NO:1). Following cleavage of the signal peptide, the mature F₀ (e.g. SEQ ID NO:2) is transported to the cell surface, then endocytosed and cleaved by cathepsin L (e.g. between amino acids 109-110 of SEQ ID NO:1) into the mature fusogenic subunits F1 (e.g. corresponding to amino acids 110-546 of SEQ ID NO:1; set forth in SEQ ID NO:4) and F2 (e.g. corresponding to amino acid residues 27-109 of SEQ ID NO:1; set forth in SEQ ID NO:3). The F1 and F2 subunits are associated by a disulfide bond and recycled back to the cell surface. The F1 subunit contains the fusion peptide domain located at the N terminus of the F1 subunit (e.g. .g. corresponding to amino acids 110-129 of SEQ ID NO:1) where it is able to insert into a cell membrane to drive fusion. In particular cases, fusion activity is blocked by association of the F protein with G protein, until G engages with a target molecule resulting in its disassociation from F and exposure of the fusion peptide to mediate membrane fusion.

[0273] Among different henipavirus species, the sequence and activity of the F protein is highly conserved. For examples, the F protein of NiV and HeV viruses share 89% amino acid sequence identity. Further, in some cases, the henipavirus F proteins exhibit compatibility with G proteins from other species to trigger fusion (Brandel-Tretheway et al. Journal of Virology. 2019. 93(13):e00577-19). In some aspects or the provided re-targeted

lipid particles, the F protein is heterologous to the G protein, i.e. the F and G protein or biologically active portions are from different henipavirus species. For example, the F protein is from Hendra virus and the G protein is from Nipah virus. In other aspects, the F protein can be a chimeric F protein containing regions of F proteins from different species of Henipavirus. In some embodiments, switching a region of amino acid residues of the F protein from one species of Henipavirus to another can result in fusion to the G protein of the species comprising the amino acid insertion. (Brandel-Tretheway et al. 2019). In some cases, the chimeric F protein contains an extracellular domain from one henipavirus species and a transmembrane and/or cytoplasmic domain from a different henipavirus species. For example, the F protein contains an extracellular domain of Hendra virus and a transmembrane/cytoplasmic domain of Nipah virus. F protein sequences disclosed herein are predominantly disclosed as expressed sequences including an N-terminal signal sequence. As such N-terminal signal sequences are commonly cleaved co- or post-translationally, the mature protein sequences for all F protein sequences disclosed herein are also contemplated as lacking the N-terminal signal sequence.

Table 4. Henipavirus 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. Nipah virus F protein is >80% identical to that of Hendra virus and is found within the same sequence cluster. 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. Column 6 provides the SEQ ID numbers for the described sequences.

Genbank ID	Nucleotides of CDS	Full Gene Name	Sequence	#Sequences/Cluster	SEQ ID NO	SEQ ID NO (with out signal sequence)
AF017149	6618 - 8258	gb:AF017149 Organism:Hendra virus Strain Name:UNKNOWN-	MATQEVRLKCLLCGIIVLVLSLEGLGILHYEKL LSKIGLVKGITRKYKIKSNPLTKDIVIKMIPNVS NVSKCTGTVMENYKSRLTGILSPIKGAIELYN NNTHDLVGDVKLAVVMMAGIAIGIATAAQIT AGVALYEAMKNADNINKLKSSIESTNEAVVK	29	17	59

		AF017149 Protein Name:fusion Gene Symbol:F	LQETAECTVYVLTALQDYINTNLVPTIDQISC KQTEALDLALSKYLSDDLFFVFGPNLQDPVSN SMTIQAISQAFGGNYETLLRTLGYATEDFDDL LESDSIAGQIVYVDLSSYYIIVRVYFIPILTEIQQ AYVQELLPVSFNNDNSEWISIVPNFVLRNTLI SNIEVKYCLITKKSVCINQDYATPMTASVREC LTGSTDKCPRELVVSSHVPRFALSGGVLFANC ISVTCQCQTGRAISQSGETLLMIDNTTCTTV VLGNIIISLGKYLGSINYNSEIAVGPPVYTDK VDISSQISSMNQSLQQSKDYIKEAQKILDTVNP SLISMLSMIILYVLSIAALCIGLITFISFVIVEKK RGNYSRLDDRQVRPVSNGDLYYIGT			
Q9I H6 3		Additional in cluster: sp Q9I H63 FUS_NIPAV Fusion glycoprotein F0 OS=Nipah virus	MVVILDKRCYCNLLILMISECSVGILHYEKL SKIGLVKGVTRKYKIKSNPLTKDIVIKMIPNVS NMSQCTGSMENYKTRLNGILTPIKGALEIYK NNTHDLVGDVRLAGVIMAGVAIGIATAAQIT AGVALYEAMKNADNINKLKSSIESTNEAVVK LQETAECTVYVLTALQDYINTNLVPTIDKISC KQTELSLDLALSKYLSDDLFFVFGPNLQDPVSN SMTIQAISQAFGGNYETLLRTLGYATEDFDDL LESDSITGQIYVDLSSYYIIVRVYFIPILTEIQQA YIQELLPVSFNNDNSEWISIVPNFILVRNTLISN IEIGFCLITKRSVCINQDYATPMTNNMRECLTG STEKCPRELVVSSHVPRFALSNGVLFANCISVT CQCQTGRAISQSGETLLMIDNTTCPTAVLG NVIISLGKYLGSVNYNSEGIAIGPPVFTDKVDI SSQISSMNQSLQQSKDYIKEAQRLDVTNPSLI SMLSMIILYVLSIASLCIGLITFISFIIVEKKRNT YSRLEDRRVRPTSSGDLYYIGT		1	2
JQ 001 776	6129 - 8166	gb:JQ001776:6129-8166 Organism: Cedar virus Strain Name:CG1a Protein Name:fusion glycoprotein Gene Symbol:F	MSNKRRTVLIHSYTLFYLNNAIVGFDFDKLN KIGVVQGRVLNYKIKGDPMTKDLVLKFPNIV NITECVREPLSRYNETVRRLLLPIHNMGLLYL NNTNAKMTGLMIAGVIMGGIAIGIATAAQITA GFALYEAKKNTENIQKLTDSIMKTQDSIDKLT DSVGTSLILNKLQTYINNQLVPLELLSCRQN KIEFDLMLTKYLVDMTVIGPNINNPVNKDM TIQSLSLFDGNYDIMMSELGYTPQDFDLIES KSITGQIYVDMENLYVVIRTYLPTLIEVPDAQI YEFNKITMSSNGGEYLSITPNFILIRGNYSNI DVATCYMTKASVICNQDYSPLMSQNLRSYQ GETEYCPVEAVIASHSRPFALTNGVIFANCINT ICRCQDNGKTITQINQFVSMIDNSTCNDVMV DKFTIKVGKYMGRKDINNINIQIGPQIIDKVD LSNEINKMNQSLKDSIFYLREAKRILDSVNISLI SPSVQLFLIISVLSFIILLIIVLYCKSKHSYKY NKFIDDPDYNDYKRERINGKASKSNNIYYV GD	3	24	57
NC _02 535 2	5950 - 8712	gb:NC_025352:5950-8712 Organism: Mojiang virus Strain Name:Tongguan Protein	MALNKNMFSSFLGYLLVYATTVQSSIHYS LSKVGVIKGLTYNYKIKGSPSTKLMVVKLIPNI DSVKNCTQKQYDEYKNLVRKALEPVMMAID TMLNNVKSNGNKKYRFAGAIMAGVALGVATA ATVTAGIALHRSNENAAIANMKSIAIQTNE AVKQLQLANKQTLAVIDTIRGEINNNIIPVINQ LSCDTIGLSVGIRLTQYYSEITAFGPALQNPV	2	25	60

		Name:fusion protein Gene Symbol:F	NTRITQAISSVFNGNFDELLKIMGYTSGDLYE ILHSELIRGNIIDVDVDAGYIALEIEFPNLTLP NAVQELMPISYNIDGDEWVTLVPRFVLTRTT LLSNIDTSRCTITDSSVICDNDYALPMSHELIG CLQGDTSKCAREKVVSSYVPKFALSDGLVYA NCLNTICRCMDTDTPIQSLSGATVSLLDNKRC SVYQVGDLISVGSYLGDEYNADNVELGPPI VIDKIDIGNQLAGINQTLQEAEDYIEKSEEFK GVNPSIITLGSMMVLYIFMILIAIVSVIALVLSIK LTVKGNVVRQQFTYTQHVPSMENINYVSH			
NC _02 525 6	6865 - 8853	gb:NC_025256 :6865- 8853 Organism :Bat Paramyxovirus Eid_hel/GH- M74a/GHA/20 09 Strain Name:BatPV/E id_hel/GH- M74a/GHA/20 09 Protein Name:fusion protein Gene Symbol:F	MKKKTDNPTISKRGHNHSRGIKSRAALLRETDN YSNGLIVENLVRNCHHPKNNLNKYTKTKQRD STIPYRVEERKGGHYPKIKHLIDKSYKHIKRGKR RNGHNGNIITILLILILKTQMSEGAHYETLS KIGLIKGITREYKVKGTPSSKDIVIKLIPNVTGL NKCTNISMENYKEQLDKILIPINNIELYANSTK SAPGNARFAGVIIAGVALGVAAAAQITAGIAL HEARQNAERINLLKDSISATNNAVAELQEATG GIVNVITGMQDYINTNLVPQIDKLQCSQIKTA LDISLSQYYSEILTVFGPNLQNPVTTSMISQAI QSFGGNIDLLNLLGYTANDLLDLLESKSITG QITYINLEHYFMVIRVYYPIMTTISNAYVQELI KISFNVDSGEWVSLVPSYILIRNSYLSNIDISEC LITKNSVICRHFAMPMSYTLKECLTGDTEKC PREAVVTSYVPRFAISGGVIYANCLSTTCQCY QTGKVIAQDGSQTLMMIDNQTCSIVRIEILIS TGKYLGSQEYNTMHVSVGNPVFTDKLDITSQI SNINQSIEQSKFYLDKSKAILDKINLNIGSVPI SILFIIAILSLISITTFVIVMIIVRRYKNKYTPLINS DPSSRRSTIQDVYIIPNPGHSIRSAARSIDRDR D	2	26	58

[0274] In some embodiments, the F protein is encoded by a nucleotide sequence that encodes the sequence set forth by any one of SEQ ID NOs: 1, 2, 17, 24, 25, 26 or 57-60 or is a functionally active variant or a biologically active portion thereof that has a sequence that is at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% identical to any one of SEQ ID NOS: 1, 2, 17, 24, 25, 26 or 57-60. In particular embodiments, the F protein or the functionally active variant or biologically active portion thereof retains fusogenic activity in conjunction with a Henipavirus G protein, such as a G protein set forth in Section I.A (e.g. NiV-G or HeV-G). Fusogenic activity includes the activity of the F protein in conjunction with a Henipavirus G protein to promote or facilitate fusion of two membrane lumens, such as the lumen of the targeted lipid particle having embedded in its lipid bilayer a henipavirus F and G protein, and a cytoplasm of a target cell, e.g. a cell that contains a surface receptor or molecule that

is recognized or bound by the targeted envelope protein. In some embodiments, the F protein and G protein are from the same Henipavirus species (e.g. NiV-G and NiV-F). In some embodiments, the F protein and G protein are from different Henipavirus species (e.g. NiV-G and HeV-F). In particular embodiments, the F protein of the functionally active variant or biologically active portion retains the cleavage site cleaved by cathepsin L (e.g. corresponding to the cleavage site between amino acids 109-110 of SEQ ID NO:1).

[0275] In particular embodiments, the F protein has the sequence of amino acids set forth in SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:17, SEQ ID NO: 24, SEQ ID NO:25, SEQ ID NO: 26, SEQ ID NO: 57, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, or SEQ ID NO: 60 or is a functionally active variant thereof or a biologically active portion thereof that retains fusogenic activity. In some embodiments, the functionally active variant comprises an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:17, SEQ ID NO: 24, SEQ ID NO:25, SEQ ID NO: 26, SEQ ID NO: 57, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, or SEQ ID NO: 60 and retains fusogenic activity in conjunction with a Henipavirus G protein (e.g., NiV-G or HeV-G). In some embodiments, the biologically active portion has an amino acid sequence having at least at or about 80%, at least at or about 85%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:17, SEQ ID NO: 24, SEQ ID NO:25, SEQ ID NO: 26, SEQ ID NO: 57, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, or SEQ ID NO: 60 and retains fusogenic activity in conjunction with a Henipavirus G protein (e.g., NiV-G or HeV-G).

[0276] Reference to retaining fusogenic activity includes activity (in conjunction with a Henipavirus G protein) that between at or about 10% and at or about 150% or more of the level or degree of binding of the corresponding wild-type F protein, such as set forth in SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:17, SEQ ID NO: 24, SEQ ID NO:25, SEQ ID NO: 26, SEQ ID NO: 57, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, or SEQ ID NO: 60, such as at least or at least about 10% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 15% of the level or degree of fusogenic activity of the

corresponding wild-type F protein, such as at least or at least about 20% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 25% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 30% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 35% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 40% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 45% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 50% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 55% of the level or degree of fusogenic activity of the corresponding wild-type f protein, such as at least or at least about 60% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 65% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 70% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 75% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 80% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 85% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 90% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 95% of the level or degree of fusogenic activity of the corresponding wild-type F protein, such as at least or at least about 100% of the level or degree of fusogenic activity of the corresponding wild-type F protein, or such as at least or at least about 120% of the level or degree of fusogenic activity of the corresponding wild-type F protein.

[0277] In some embodiments, the F protein is a mutant F protein that is a functionally active fragment or a biologically active portion containing one or more amino acid mutations, such as one or more amino acid insertions, deletions, substitutions or truncations. In some embodiments, the mutations described herein relate to amino acid insertions, deletions, substitutions or truncations of amino acids compared to a reference F protein sequence. In some embodiments, the reference F protein sequence is the wild-type sequence of an F protein or a biologically active portion thereof. In some embodiments, the mutant F protein or the biologically active portion thereof is a mutant of a wild-type Hendra (Hev)

virus F protein, a Nipah (NiV) virus F-protein, a Cedar (CedPV) virus F protein, a Mojiang virus F protein or a bat Paramyxovirus F protein. In some embodiments, the wild-type F protein is encoded by a sequence of nucleotides that encodes any one of SEQ ID NO: 1, 2, 17, 24, 25, 26, or 57-60.

[0278] In some embodiments, the mutant F protein is a biologically active portion of a wild-type F protein that is an N-terminally and/or C-terminally truncated fragment. In some embodiments, the mutant F protein or the biologically active portion of a wild-type F protein thereof comprises one or more amino acid substitutions. In some embodiments, the mutations described herein can improve transduction efficiency. In some embodiments, the mutations described herein can increase fusogenic capacity. Exemplary mutations include any as described, see e.g. Khetawat and Broder 2010 Virology Journal 7:312; Witting et al. 2013 Gene Therapy 20:997-1005; published international; patent application No. WO/2013/148327.

[0279] In some embodiments, the mutant F protein is a biologically active portion that is truncated and lacks up to 20 contiguous amino acid residues at or near the C-terminus of the wild-type F protein, such as a wild-type F protein encoded by a sequence of nucleotides encoding the F protein set forth in any one of SEQ ID NOS: 1, 17, 24, 25 or 26. In some embodiments, the mutant F protein is truncated and lacks up to 19 contiguous amino acids, such as up to 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 contiguous amino acids at the C-terminus of the wild-type F protein.

[0280] In some embodiments, the F protein or the functionally active variant or biologically active portion thereof comprises an F1 subunit or a fusogenic portion thereof. In some embodiments, the F1 subunit is a proteolytically cleaved portion of the F0 precursor. In some embodiments, the F0 precursor is inactive. In some embodiments, the cleavage of the F0 precursor forms a disulfide-linked F1+F2 heterodimer. In some embodiments, the cleavage exposes the fusion peptide and produces a mature F protein. In some embodiments, the cleavage occurs at or around a single basic residue. In some embodiments, the cleavage occurs at Arginine 109 of NiV-F protein. In some embodiments, cleavage occurs at Lysine 109 of the Hendra virus F protein.

[0281] In some embodiments, the F protein is a wild-type Nipah virus F (NiV-F) protein or is a functionally active variant or biologically active portion thereof. In some embodiments, the F₀ precursor is encoded by a sequence of nucleotides encoding the

sequence set forth in SEQ ID NO: 1. The encoding nucleic acid can encode a signal peptide sequence that has the sequence MVVILDKRCY CNLLILILMI SECSVG (SEQ ID NO: 34). In some embodiments, the F protein has the sequence set forth in SEQ ID NO:2. In some examples, the F protein is cleaved into an F1 subunit comprising the sequence set forth in SEQ ID NO:4 and an F2 subunit comprising the sequence set forth in SEQ ID NO: 3.

[0282] In some embodiments, the F protein is a NiV-F protein that is encoded by a sequence of nucleotides encoding the sequence set forth in SEQ ID NO:1, or is a functionally active variant or biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at or about 86%, at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 1. In some embodiments, the NiV-F-protein has the sequence of set forth in SEQ ID NO: 2, or is a functionally active variant or a biologically active portion thereof that has an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at or about 86%, at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 2. In particular embodiments, the F protein or the functionally active variant or biologically active portion thereof retains the cleavage site cleaved by cathepsin L (e.g. corresponding to the cleavage site between amino acids 109-110 of SEQ ID NO:1).

[0283] In some embodiments, the F protein or the functionally active variant or the biologically active portion thereof includes an F1 subunit that has the sequence set forth in SEQ ID NO: 4, or an amino acid sequence having, at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at or about 86%, at least at or about 87%, at least at or about 88%, or at least at or about 89% at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or

about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:4.

[0284] In some embodiments, the F protein or the functionally active variant or biologically active portion thereof includes an F2 subunit that has the sequence set forth in SEQ ID NO: 3, or an amino acid sequence having, at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at or about 86%, at least at or about 87%, at least at or about 88%, or at least at or about 89% at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:3.

[0285] In some embodiments, the F protein is a mutant NiV-F protein that is a biologically active portion thereof that is truncated and lacks up to 20 contiguous amino acid residues at or near the C-terminus of the wild-type NiV-F protein (e.g. set forth SEQ ID NO:2). In some embodiments, the mutant NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO:5. In some embodiments, the mutant NiV-F protein has a sequence that has at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 5. In some embodiments, the mutant F protein contains an F1 protein that has the sequence set forth in SEQ ID NO:6. In some embodiments, the mutant F protein has a sequence that has at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 6.

[0286] In some embodiments, the F protein is a mutant NiV-F protein that is a biologically active portion thereof that comprises a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2); and a point mutation on an N-linked glycosylation site. In some embodiments, the mutant NiV-F protein comprises an amino acid sequence set forth in SEQ ID NO: 7. In some embodiments, the mutant NiV-F protein has a sequence that has at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least

at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 7.

[0287] In some embodiments, the F protein is a mutant NiV-F protein that is a biologically active portion thereof that comprises a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2). In some embodiments, the NiV-F protein is encoded by a nucleotide sequence that encodes the sequence set forth in SEQ ID NO: 8. In some embodiments, the NiV-F proteins is encoded by a nucleotide sequence that encodes sequence having at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 8. In particular embodiments, the variant F protein is a mutant Niv-F protein that has the sequence of amino acids set forth in SEQ ID NO:23. In some embodiments, the NiV-F proteins is encoded by a a sequence having at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 23.

C. Lipid Bilayer

[0288] In some embodiments, the targeted lipid particle includes a naturally derived bilayer of amphipathic lipids that encloses lumen or cavity. In some embodiments, the targeted lipid particle comprises a lipid bilayer as the outermost surface. In some embodiments, the lipid bilayer encloses a lumen. In some embodiments, the lumen is aqueous. In some embodiments, the lumen is in contact with the hydrophilic head groups on the interior of the lipid bilayer. In some embodiments, the lumen is a cytosol. In some embodiments, the cytosol contains cellular components present in a source cell. In some embodiments, the cytosol does not contain components present in a source cell. In some embodiments, the lumen is a cavity. In some embodiments, the cavity contains an aqueous environment. In some embodiments, the cavity does not contain an aqueous environment.

[0289] In some aspects, the lipid bilayer is derived from a source cell during a process to produce a lipid-containing particle. Exemplary methods for producing lipid-containing particles are provided in Section I.E. In some embodiments, the lipid bilayer includes membrane components of the cell from which the lipid bilayer is produced, e.g., phospholipids, membrane proteins, etc. In some embodiments, the lipid bilayer includes a cytosol that includes

components found in the cell from which the micro-vesicle is produced, e.g., solutes, proteins, nucleic acids, etc., but not all of the components of a cell, e.g., they lack a nucleus. In some embodiments, the lipid bilayer is considered to be exosome-like. The lipid bilayer may vary in size, and in some instances have a diameter ranging from 30 and 300 nm, such as from 30 and 150 nm, and including from 40 to 100 nm.

[0290] In some embodiments, the lipid bilayer is a viral envelope. In some embodiments, the viral envelope is obtained from a source cell. In some embodiments, the viral envelope is obtained by the viral capsid from the source cell plasma membrane. In some embodiments, the lipid bilayer is obtained from a membrane other than the plasma membrane of a host cell. In some embodiments, the viral envelope lipid bilayer is embedded with viral proteins, including viral glycoproteins.

[0291] In other aspects, the lipid bilayer includes synthetic lipid complex. In some embodiments, the synthetic lipid complex is a liposome. In some embodiments, the lipid bilayer is a vesicular structure characterized by a phospholipid bilayer membrane and an inner aqueous medium. In some embodiments, the lipid bilayer has multiple lipid layers separated by aqueous medium. In some embodiments, the lipid bilayer forms spontaneously when phospholipids are suspended in an excess of aqueous solution. In some examples, the lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers.

[0292] In some embodiments, a targeted envelope protein and fusogen, such as any described above including any that are exogenous or overexpressed relative to the source cell, is disposed in the lipid bilayer.

[0293] In some embodiments, the targeted lipid particle comprises several different types of lipids. In some embodiments, the lipids are amphipathic lipids. In some embodiments, the amphipathic lipids are phospholipids. In some embodiments, the phospholipids comprise phosphatidylcholine, phosphatidylethanolamine, phosphatidylinositol, and phosphatidylserine. In some embodiments, the lipids comprise phospholipids such as phosphocholines and phosphoinositols. In some embodiments, the lipids comprise DMPC, DOPC, and DSPC.

[0294] In some embodiments, the bilayer may be comprised of one or more lipids of the same or different type. In some embodiments, the source cell comprises a cell selected from 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.

D. Exogenous Agent

[0295] In embodiments, the targeted lipid particle, such as a lentiviral vector, further comprises an agent that is exogenous relative to the source cell (hereinafter also called “cargo” or “payload”). In some embodiments, the exogenous agent is 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 some embodiments, the exogenous agent is a nucleic acid that encodes a protein. The protein can be any protein as is desired for targeted delivery to a target cell. In some embodiments, the protein is a therapeutic agent or a diagnostic agent. In some embodiments, the protein is an antigen receptor for targeting cells expressed by or associated with a disease or condition, for instance a chimeric antigen receptor (CAR) or a T cell receptor (TCR). Reference to the coding sequence of a nucleic acid encoding the protein also is referred to herein as a payload gene. In some embodiments, the exogenous agent or the nucleic acid encoding the exogenous agent are present in the lumen of the non-cell particle.

[0296] In some embodiments, the exogenous agent or cargo comprises or encodes a cytosolic protein. In some embodiments the exogenous agent or cargo comprises or encodes a membrane protein. In some embodiments, the exogenous agent or cargo comprises or encodes a 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.

[0297] In embodiments, the exogenous agent is present 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. In embodiments, the targeted lipid particle has an altered, e.g., increased or decreased level of one or more endogenous molecule, e.g., protein or nucleic acid (e.g., in some embodiments, endogenous relative to the source cell, and in some embodiments, endogenous relative to the target cell), e.g., due to treatment of the source cell, e.g., mammalian source cell with a siRNA or gene editing enzyme. In embodiments, the endogenous molecule is present 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. 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 , 5.0×10^3 , 10^4 , 5.0×10^4 , 10^5 , 5.0×10^5 , 10^6 , 5.0×10^6 , 1.0×10^7 , 5.0×10^7 , or 1.0×10^8 , greater than its concentration in the source cell. 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 , 5.0×10^3 , 10^4 , 5.0×10^4 , 10^5 , 5.0×10^5 , 10^6 , 5.0×10^6 , 1.0×10^7 , 5.0×10^7 , or 1.0×10^8 less than its concentration in the source cell.

[0298] In some embodiments, the targeted lipid particle 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 exogenous therapeutic agent) comprised by the fusosome. In some embodiments, the targeted lipid particle that fuses with the target cell(s) delivers 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 exogenous therapeutic agent) comprised by the lipid particles that fuse with the target cell(s). In some embodiments, the targeted lipid particle 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 exogenous therapeutic agent) comprised by the targeted lipid particle compositions.

[0299] In some embodiments, the exogenous agent or cargo is not expressed naturally in the cell from which the targeted lipid particle is derived. In some embodiments, the exogenous agent or cargo is expressed naturally in the cell from which the targeted lipid particle is derived. In some embodiments, the exogenous agent or cargo is loaded into the targeted lipid particle via expression in the cell from which the lipid particle is derived (e.g. expression from DNA or mRNA introduced via transfection, transduction, or electroporation). In some embodiments, the exogenous agent or cargo is expressed from DNA integrated into the genome or maintained episomally. In some embodiments, expression of the exogenous agent or cargo is constitutive. In some embodiments, expression of the exogenous agent or cargo is induced. In some embodiments, expression of the exogenous agent or cargo is induced immediately prior to generating the targeted lipid particle. In some embodiments, expression of the exogenous agent or cargo is induced at the same time as expression of the fusogen.

[0300] In some embodiments, the exogenous agent or cargo is loaded into the lipid particle via electroporation into the lipid particle itself or into the cell from which the fusosome is

derived. In some embodiments, the exogenous agent or cargo is loaded into the lipid particle via transfection (e.g., of a DNA or mRNA encoding the cargo) into the lipid particle itself or into the cell from which the lipid particle is derived.

[0301] In some embodiments, the exogenous agent or cargo may include one or more nucleic acid sequences, one or more polypeptides, a combination of nucleic acid sequences and/or polypeptides, one or more organelles, and any combination thereof. In some embodiments, the exogenous agent or cargo may include one or more cellular components. In some embodiments, the exogenous agent or cargo includes one or more cytosolic and/or nuclear components.

[0302] In some embodiments, the exogenous agent or cargo includes a nucleic acid, e.g., DNA, nDNA (nuclear DNA), mtDNA (mitochondrial DNA), protein coding DNA, gene, operon, chromosome, genome, transposon, retrotransposon, viral genome, intron, exon, modified DNA, mRNA (messenger RNA), tRNA (transfer RNA), modified RNA, microRNA, siRNA (small interfering RNA), tmRNA (transfer messenger RNA), rRNA (ribosomal RNA), mtRNA (mitochondrial RNA), snRNA (small nuclear RNA), small nucleolar RNA (snoRNA), SmY RNA (mRNA trans-splicing RNA), gRNA (guide RNA), TERC (telomerase RNA component), aRNA (antisense RNA), cis-NAT (Cis-natural antisense transcript), CRISPR RNA (crRNA), lncRNA (long noncoding RNA), piRNA (piwi-interacting RNA), shRNA (short hairpin RNA), tasiRNA (trans-acting siRNA), eRNA (enhancer RNA), satellite RNA, pcRNA (protein coding RNA), dsRNA (double stranded RNA), RNAi (interfering RNA), circRNA (circular RNA), reprogramming RNAs, aptamers, and any combination thereof. In some embodiments, the nucleic acid is a wild-type nucleic acid. In some embodiments, the protein is a mutant nucleic acid. In some embodiments the nucleic acid is a fusion or chimera of multiple nucleic acid sequences.

[0303] In some embodiments, the exogenous agent or cargo may include a nucleic acid. For example, the exogenous agent or cargo may comprise RNA to enhance expression of an endogenous protein, or a siRNA or miRNA that inhibits protein expression of an endogenous protein. For example, the endogenous protein may modulate structure or function in the target cells. In some embodiments, the cargo may include a nucleic acid encoding an engineered protein that modulates structure or function in the target cells. In some embodiments, the exogenous agent or cargo is a nucleic acid that targets a transcriptional activator that modulate structure or function in the target cells.

[0304] In some embodiments, the exogenous agent or cargo is or encodes a polypeptide, e.g., enzymes, structural polypeptides, signaling polypeptides, regulatory polypeptides, transport polypeptides, sensory polypeptides, motor polypeptides, defense polypeptides, storage polypeptides, transcription factors, antibodies, cytokines, hormones, catabolic polypeptides, anabolic polypeptides, proteolytic polypeptides, metabolic polypeptides, kinases, transferases, hydrolases, lyases, isomerases, ligases, enzyme modulator polypeptides, protein binding polypeptides, lipid binding polypeptides, membrane fusion polypeptides, cell differentiation polypeptides, epigenetic polypeptides, cell death polypeptides, nuclear transport polypeptides, nucleic acid binding polypeptides, reprogramming polypeptides, DNA editing polypeptides, DNA repair polypeptides, DNA recombination polypeptides, transposase polypeptides, DNA integration polypeptides, targeted endonucleases (e.g. Zinc -finger nucleases, transcription-activator-like nucleases (TALENs), cas9 and homologs thereof), recombinases, and any combination thereof. In some embodiments the protein targets a protein in the cell for degradation. In some embodiments the protein targets a protein in the cell for degradation by localizing the protein to the proteasome. In some embodiments, the protein is a wild-type protein. In some embodiments, the protein is a mutant protein. In some embodiments the protein is a fusion or chimeric protein.

[0305] In some embodiments, the exogenous agent or cargo is a small molecule, e.g., ions (e.g. Ca^{2+} , Cl^- , Fe^{2+}), carbohydrates, lipids, reactive oxygen species, reactive nitrogen species, isoprenoids, signaling molecules, heme, polypeptide cofactors, electron accepting compounds, electron donating compounds, metabolites, ligands, and any combination thereof. In some embodiments the small molecule is a pharmaceutical that interacts with a target in the cell. In some embodiments the small molecule targets a protein in the cell for degradation. In some embodiments the small molecule targets a protein in the cell for degradation by localizing the protein to the proteasome. In some embodiments that small molecule is a proteolysis targeting chimera molecule (PROTAC).

[0306] In some embodiments, the exogenous agent or cargo includes a mixture of proteins, nucleic acids, or metabolites, e.g., multiple polypeptides, multiple nucleic acids, multiple small molecules; combinations of nucleic acids, polypeptides, and small molecules; ribonucleoprotein complexes (e.g. Cas9-gRNA complex); multiple transcription factors, multiple epigenetic factors, reprogramming factors (e.g. Oct4, Sox2, cMyc, and Klf4); multiple regulatory RNAs; and any combination thereof.

[0307] In some embodiments, the exogenous agent or cargo includes one or more organelles, e.g., chondrisomes, mitochondria, lysosomes, nucleus, cell membrane, cytoplasm, endoplasmic reticulum, ribosomes, vacuoles, endosomes, spliceosomes, polymerases, capsids, acrosome, autophagosome, centriole, glycosome, glyoxysome, hydrogenosome, melanosome, mitosome, myofibril, cnidocyst, peroxisome, proteasome, vesicle, stress granule, networks of organelles, and any combination thereof.

[0308] In some embodiments, the exogenous agent is or encodes 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 is or encodes a secreted protein, e.g., a protein that is produced and secreted by the recipient cell. In some embodiments, the exogenous agent is or encodes 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 is or encodes 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.

[0309] In some embodiments, the exogenous agent is capable of being delivered to a hepatocyte or liver cell. In some embodiments, the exogenous agents or cargo can be delivered to treat a disease or disorder in a hepatocyte or liver cell.

[0310] In some embodiments, the exogenous agent is encoded by a gene from among OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMACHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAH, PAL, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LNBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTB, HLCS, PC, SLC7A7, CPT2, ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL, GRHPR, HOGA1, LDLR, ACAD8, ACADSB, ACAT1, ACSF3, ASPA, AUH, DNAJC19, ETHE1, FBP1, FTCD, GSS, HIBCH, IDH2, L2HGDH, MLYCD, OPA3, OPLAH, OXCT1, POLG, PPM1K, SERAC1, SLC25A1, SUCLA2, SUCLG1, TAZ, AGK, CLPB, TMEM70, ALDH18A1, OAT, CA5A, GLUD1, GLUL, UMPS, SLC22A5, CPT1A, HADHA, HADH, SLC52A1, SLC52A2,

SLC52A3, HADHB, GYS2, PYGL, SLC2A2, ALG1, ALG2, ALG3, ALG6, ALG8, ALG9, ALG11, ALG12, ALG13, ATP6V0A2, B3GLCT, CHST14, COG1, COG2, COG4, COG5, COG6, COG7, COG8, DOLK, DHDDS, DPAGT1, DPM1, DPM2, DPM3, G6PC3, GFPT1, GMPPA, GMPPB, MAGT1, MAN1B1, MGAT2, MOGS, MPDU1, MPI, NGLY1, PGM1, PGM3, RFT1, SEC23B, SLC35A1, SLC35A2, SLC35C1, SSR4, SRD5A3, TMEM165, TRIP11, TUSC3, ALG14, B4GALT1, DDOST, NUS1, RPN2, SEC23A, SLC35A3, ST3GAL3, STT3A, STT3B, AGA, ARSA, ARSB, ASAH1, ATP13A2, CLN3, CLN5, CLN6, CLN8, CTNS, CTSA, CTSD, CTSF, CTSK, DNAJC5, FUCA1, GAA, GALC, GALNS, GLA, GLB1, GM2A, GNPTAB, GNPTG, GNS, GRN, GUSB, HEXA, HEXB, HGSNAT, HYAL1, IDS, IDUA, KCTD7, LAMP2, MAN2B1, MANBA, MCOLN1, MFSD8, NAGA, NAGLU, NEU1, NPC1, NPC2, SGSH, PPT1, PSAP, SLC17A5, SMPD1, SUMF1, TPP1, AHCY, GNMT, MAT1A, GCH1, PCBD1, PTS, QDPR, SPR, DNAJC12, ALDH4A1, PRODH, HPD, GBA, HGD, AMN, CD320, CUBN, GIF, TCN1, TCN2, PREPL, PHGDH, PSAT1, PSPH, AMT, GCSH, GLDC, LIAS, NFU1, SLC6A9, SLC2A1, ATP7A, AP1S1, CP, SLC33A1, PEX7, PHYH, AGPS, GNPAT, ABCD1, ACOX1, PEX1, PEX2, PEX3, PEX5, PEX6, PEX10, PEX12, PEX13, PEX14, PEX16, PEX19, PEX26, AMACR, ADA, ADSL, AMPD1, GPHN, MOCOS, MOCS1, PNP, XDH, SUOX, OGDH, SLC25A19, DHTKD1, SLC13A5, FH, DLAT, MPC1, PDHA1, PDHB, PDHX, PDP1, ABCC2, SLC01B1, SLC01B3, HFE2, ADAMTS13, PYGM, COL1A2, TNFRSF11B, TSC1, TSC2, DHCR7, PGK1, VLDLR, KYNU, F5, C3, COL4A1, CFH, SLC12A2, GK, SFTPC, CRTAP, P3H1, COL7A1, PKLR, TALDO1, TF, EPCAM, VHL, GC, SERPINA1, ABCC6, F8, F9, ApoB, PCSK9, LDLRAP1, ABCG5, ABCG8, LCAT, SPINK5, or GNE.

[0311] In some embodiments, the exogenous agent is encoded by a gene from among OTC, CPS1, NAGS, BCKDHA, BCKDHB, DBT, DLD, MUT, MMAA, MMAB, MMACHC, MMADHC, MCEE, PCCA, PCCB, UGT1A1, ASS1, PAL, PAH, ATP8B1, ABCB11, ABCB4, TJP2, IVD, GCDH, ETFA, ETFB, ETFDH, ASL, D2HGDH, HMGCL, MCCC1, MCCC2, ABCD4, HCFC1, LMBRD1, ARG1, SLC25A15, SLC25A13, ALAD, CPOX, HMBS, PPOX, BTB, HLCS, PC, SLC7A7, CPT2, ACADM, ACADS, ACADVL, AGL, G6PC, GBE1, PHKA1, PHKA2, PHKB, PHKG2, SLC37A4, PMM2, CBS, FAH, TAT, GALT, GALK1, GALE, G6PD, SLC3A1, SLC7A9, MTHFR, MTR, MTRR, ATP7B, HPRT1, HJV, HAMP, JAG1, TTR, AGXT, LIPA, SERPING1, HSD17B4, UROD, HFE, LPL, GRHPR, HOGA1, or

LDLR. In some embodiments, the exogenous agent is the enzyme phenylalanine ammonia lyase (PAL).

[0312] In some embodiments, the exogenous agents or cargo can be delivered to treat and disease or indication listed in **Table 5**. In some embodiments, the indications are specific for a liver cell or hepatocyte.

[0313] 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 sequence 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 4 of **Table 5** or an amino acid sequence of column 5 of **Table 5**. In some embodiments, the payload gene encoding an exogenous agent 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 5**. In some embodiments, the payload gene encoding an exogenous agent has a nucleic acid sequence having at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%, identity to a nucleic acid sequence of **Table 5**, e.g., an Ensemble Gene Accession Number of column 3 of **Table 5**.

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

Gene	Entrez Accession Number	Ensembl Gene(s) Accession Number (ENSG0000+ number shown)	Uniprot Protein(s) Accession Number	Amino Acid Sequence (first Uniprot Accession Number) SEQ ID NO	Disease / Disorder	Category
OTC	5009	0036473	P00480	61	ornithine transcarbamylase (OTC) deficiency	Urea cycle disorder
CPS1	1373	0021826	P31327, Q6PEK7, B7ZAW0, A0A024R454	62	carbamoyl phosphate synthetase I (CPSI) deficiency	Urea cycle disorder

NAGS	162417	0161653	Q8N159, Q2NKP2	63	N-acetylglutamate synthase (NAGS) deficiency	Urea cycle disorder
BCKD HA	593	0248098	A0A024R0K 3, P12694, Q59EI3	64	maple syrup urine disease (MSUD); Classic Maple Syrup Urine Disease (CMSUD)	Organic acidemia
BCKD HB	594	0083123	A0A140VKB 3, P21953, B4E2N3, B7ZB80	65	maple syrup urine disease (MSUD); Classic Maple Syrup Urine Disease (CMSUD)	Organic acidemia
DBT	1629	0137992	P11182	66	maple syrup urine disease (MSUD); Classic Maple Syrup Urine Disease (CMSUD)	Organic acidemia
DLD	1738	0091140	A0A024R713 , P09622, E9PEX6	67	maple syrup urine disease (MSUD) Dihydrolipoamide dehydrogenase deficiency	Urea cycle disorder
MUT	4594	0146085	A0A024RD8 2, B2R6K1, P22033	68	methylmalonic acidemia due to methylmalonyl- CoA mutase deficiency	Organic acidemia
MMA A	166785	0151611	Q8IVH4	69	cobalamin A deficiency (methylmalonic acidemia)	Organic acidemia
MMAB	326625	0139428	Q96EY8	70	cobalamin B deficiency (methylmalonic acidemia)	Organic acidemia
MMAC HC	25974	0132763	A0A0C4DG U2, Q9Y4U1	71	cobalamin C deficiency (methylmalonic acidemia); Methylmalonic Acidemia with Homocystinuria	Organic acidemia
MMA DHC	27249	0168288	Q9H3L0	72	cobalamin D deficiency (methylmalonic acidemia); Methylmalonic Acidemia with Homocystinuria; Homocystinuria; Cobalamin C Deficiency	Organic acidemia
MCEE	84693	0124370	Q96PE7	73	methylmalonic acidemia; Cobalamin D Deficiency	Organic acidemia

PCCA	5095	0175198	P05165	74	propionic acidemia	Organic acidemia
PCCB	5096	0114054	P05166	75	propionic acidemia	Organic acidemia
UGT1A1	54658	0241635	P22309, Q5DT03	76	Crigler-Najjar syndrome type 1 Crigler-Najjar syndrome type 2, Gilbert syndrome	
ASS1	445	0130707	P00966, Q5T6L4	77	citrullinemia type I	Urea cycle disorder
PAH	5053	0171759	A0A024RBG4, P00439	78	Phenylalanine hydroxylase deficiency	Aminoacidopathy
PAL				79	Phenylalanine hydroxylase deficiency	Aminoacidopathy
ATP8B1	5205	0081923	O43520	80	Progressive familial intrahepatic cholestasis Type 1	
ABCB11	8647	0073734,0276582	O95342	81	Progressive familial intrahepatic cholestasis Type 2; Progressive Familial Intrahepatic Cholestasis Type 3	
ABCB4	5244	0005471	P21439	82	Progressive familial intrahepatic cholestasis Type 3; Progressive Familial Intrahepatic Cholestasis Type 2	
TJP2	9414	0119139	B7Z2R3, Q9UDY2, B7Z954	83	Progressive familial intrahepatic cholestasis Type 4	
IVD	3712	0128928	P26440, A0A0A0MT83	84	isovaleric acidemia (IVD)	Organic acidemia
GCDH	2639	0105607	A0A024R7F9, Q92947	85	glutaric acidemia type I	Organic acidemia
ETFA	2108	0140374	A0A0S2Z3L0, P13804	86	multiple acyl-CoA dehydrogenase deficiency (a.k.a. glutaric aciduria type II)	Organic acidemia
ETFB	2109	0105379	P38117	87	multiple acyl-CoA dehydrogenase deficiency (a.k.a. glutaric aciduria type II)	Organic acidemia

ETFD H	2110	0171503	B4DEQ0, Q16134	88	multiple acyl-CoA dehydrogenase deficiency (a.k.a. glutaric aciduria type II)	Organic acidemia
ASL	435	0126522	A0A024RDL 8, P04424, A0A0S2Z316	89	argininosuccinate lyase (ASL) deficiency	Urea cycle disorder
D2HG DH	728294	0180902	B3KSR6, B4E3K7, B5MCV2, Q8N465	90	D-2- hydroxyglutaric aciduria type I	Organic acidemia
HMGC L	3155	0117305	P35914	91	3-hydroxy-3- methylglutaryl- CoA lyase (3HMG) deficiency	Organic acidemia Urea cycle disorder
MCCC 1	56922	0078070	Q68D27, Q96RQ3, A0A0S2Z693 , E9PHF7	92	3-methylcrotonyl- CoA carboxylase (3MCC) deficiency	Organic acidemia
MCCC 2	64087	0131844,0281 742,0275300	A0A140VK2 9, Q9HCC0	93	3-methylcrotonyl- CoA carboxylase (3MCC) deficiency	Organic acidemia
ABCD 4	5826	0119688	A0A024R6B 9, O14678, A0A024R6C 8	94	methylmalonic acidemia with homocystinuria	Organic acidemia
HCFC1	3054	0172534	P51610, A6NEM2	95	methylmalonic acidemia with homocystinuria	Organic acidemia
LMBR D1	55788	0168216	Q9NUN5	96	methylmalonic acidemia with homocystinuria	Organic acidemia
ARG1	383	0118520	P05089	97	arginase (ARG1) deficiency	Urea cycle disorder
SLC25 A15	10166	0102743	Q9Y619	98	hyperammonemia- hyperornithinemia- homocitrullinuria (HHH) syndrome	Urea cycle disorder
SLC25 A13	10165	0004864	Q9UJS0	99	citrin deficiency citrullinemia type II	Urea cycle disorder
ALAD	210	0148218	P13716	100	Acute Hepatic porphyria	Porphyria
CPOX	1371	0080819	P36551	101	Acute Hepatic porphyria	Porphyria
HMBS	3145	0256269,0281 702	P08397	102	Acute Hepatic porphyria; Acute Intermittent Porphyria	Porphyria
PPOX	5498	0143224	P50336, B4DY76	103	Acute Hepatic porphyria	Porphyria
BTD	686	0169814	P43251	104	Biotinidase Deficiency	Organic acidemia

HLCS	3141	0159267	P50747	105	Holocarboxylase Synthetase Deficiency	Organic acidemia
PC	5091	0173599	P11498 A0A024R5C5	106	Pyruvate Carboxylase Deficiency	Urea cycle disorder
SLC7A7	9056	0155465	Q9UM01 A0A0S2Z502	107	Lysinuric Protein Intolerance	Urea cycle disorder
CPT2	1376	0157184	P23786 A0A140VK13 A0A1B0GTB8	108	Carnitine Palmitoyltransferase Type II (CPT II) Deficiency	Fatty Acid Oxidation
ACADM	34	0117054	P11310 A0A0S2Z366 , B7Z911, Q5HYG7, Q5T4U5, B4DJE7	109	Medium Chain Acyl-CoA Dehydrogenase (MCAD) Deficiency	Fatty Acid Oxidation
ACADS	35	0122971	P16219 E5KSD5, B4DUH1, E9PE82	110	Short Chain Acyl-CoA (SCAD) Dehydrogenase Deficiency	Fatty acid oxidation
ACADVL	37	0072778	P49748 B3KPA6	111	Very Long Chain Acyl-CoA Dehydrogenase (VLCAD) Deficiency	Fatty acid oxidation
AGL	178	0162688	P35573 A0A0S2A4E4	112	GSD III (Cori / Forbe Disease or Debrancher)	Liver glycogen storage disorder
G6PC	2538	0131482	P35575	113	GSD Ia (Von Gierke Disease)	Liver glycogen storage disorder
GBE1	2632	0114480	Q04446 Q59ET0	114	GSD IV (Andersen Disease, Brancher Enzyme)	Liver glycogen storage disorder
PHKA1	5255	0067177	P46020	115	GSD IXa	
PHKA2	0044446 5256	5256 0044446	P46019	116	GSD IXa	Liver glycogen storage disorder
PHKB	5257	0102893	Q93100	117	GSD IXb	Liver glycogen storage disorder
PHKG2	5261	0156873	P15735	118	GSD IXc	Liver glycogen storage disorder
SLC37A4	2542	0281500 0137700	O43826	119	GSD Ib. c, d	Liver glycogen storage disorder

			A0A024R3H9, A8K0S7, A0A024R3L1, B4DUH2			
PMM2	5373	0140650	O15305, A0A0S2Z4J6, Q59F02	120	PMM2-CDG	Glycosylation disorder
CBS	102724560, 875	0160200	P35520, P0DN79, Q9NTF0, B7Z2D6	121	Cystathionine Beta-Synthase Deficiency (Classic Homocystinuria); Homocystinuria	Aminoacidopathy
FAH	2184	0103876	P16930	122	Tyrosinemia Type I	Aminoacidopathy
TAT	6898	0198650	P17735, A0A140VKB7	123	Tyrosinemia Type II Tyrosinemia Type III	Aminoacidopathy
GALT	2592	0213930	P07902, A0A0S2Z3Y7, B2RAT6	124	Galactosemia due to galactose-1-phosphate uridylyltransferase (GALT) deficiency	Carbohydrate disorder
GALK1	2584	0108479	P51570	125	Galactosemia	Carbohydrate disorder
GALE	2582	0117308	Q14376	126	Galactosemia	Carbohydrate disorder
G6PD	2539	0160211	P11413	127	Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency	Carbohydrate disorder
SLC3A1	6519	0138079	Q07837, A0A0S2Z4E1, B8ZZK1	128	Cystinuria	Aminoacidopathy
SLC7A9	11136	0021488	P82251	129	Cystinuria	Aminoacidopathy
MTHFR	4524	0177000	P42898, Q59GJ6, Q81U67	130	Homocystinuria	Aminoacidopathy
MTR	4548	0116984	Q99707	131	Homocystinuria	Aminoacidopathy
MTRR	4552	0124275	Q9UBK8	132	Homocystinuria	Aminoacidopathy
ATP7B	540	0123191	P35670, A0A024RDX3, B7ZLR4, B7ZLR3, E7ET55	133	Wilson Disease Copper Metabolism Disorder	Metal transport disorder
HPRT1	3251	0165704	P00492, A0A140VJL3	134	Lesch-Nyhan Syndrome Purine Metabolism Disorder	Purine Metabolism Disorder

HJV	148738	0168509	Q6ZVN8	135	Hemochromatosis, Type 2A	
HAMP	57817	0105697	P81172	136	Hemochromatosis Type 2B: Primary Hemochromatosis	
JAG1	182	0101384	P78504, Q99740	137	Alagille Syndrome 1	
TTR	7276	0118271	P02766, E9KL36	138	Familial TTR Amyloidosis; Familial amyloid polyneuropathy	
AGXT	189	0172482	P21549	139	Primary Hyperoxaluria Type I	
LIPA	3988	0107798	P38571 A0A0A0MT32	140	Lysosomal Acid Lipase Deficiency	Lysosomal storage disorder
SERPING1	710	0149131	P05155, A0A0S2Z4J1, B2R659, E7EWE5, B3KSP2, G5E9S2	141	Hereditary Angioedema	
HSD17B4	3295	0133835	P51659	142	D-Bifunctional Protein Deficiency X-linked Adrenoleukodystrophy	Peroxisomal disorders
UROD	7389	0126088	P06132	143	Porphyria Cutanea Tarda	
HFE	3077	0010704	Q30201	144	Porphyria Cutanea Tarda	
LPL	4023	0175445	P06858, A0A1B1RV A9	145	Lipoprotein Lipase Deficiency ("hyperlipoproteinemia type Ia; Buerger-Grutz syndrome, or Familial hyperchylomicronemia)	
GRHP R	9380	0137106	Q9UBQ7	146	Primary Hyperoxaluria Type II	
HOGA1	112817	0241935	Q86XE5	147	Primary Hyperoxaluria Type III	
LDLR	3949	0130164	P01130, A0A024R7D5	148	Homozygous Familial Hypercholesterolemia	
ACAD8	27034	0151498	Q9UKU7	149	isobutyryl-CoA dehydrogenase (IBD) deficiency	Organic acidemia

ACAD SB	36	0196177	P45954, A0A0S2Z3P9	150	short-branched chain acyl-CoA dehydrogenase (SBCAD) deficiency	Organic acidemia
ACAT 1	38	0075239	A0A140VJX 1, P24752	151	beta-ketothiolase deficiency	Organic acidemia
ACSF3	197322	0176715	Q4G176, F5H5A1	152	combined malonic and methylmalonic aciduria	Organic acidemia
ASPA	443	0108381	P45381, Q6FH48	153	Canavan disease	Organic acidemia
AUH	549	0148090	Q13825, B4DYI6	154	3- methylglutaconic acidemia type I	Organic acidemia
DNAJ C19	131118	0205981	Q96DA6, A0A0S2Z5X 1	155	dilated cardiomyopathy with ataxia syndrome (causes 3- methylglutaconic aciduria)	Organic acidemia
ETHE1	23474	0105755	A0A0S2Z580 , O95571, A0A0S2Z5N 8, A0A0S2Z5B 3, B2RCZ7	156	ethylmalonic encephalopathy	Organic acidemia
FBP1	2203	0165140	P09467, Q2TU34	157	fructose 1,6- Bisphosphatase deficiency	Organic acidemia
FTCD	10841	0160282,0281 775	O95954	158	glutamate formiminotransfer ase deficiency (FIGLU)	Organic acidemia
GSS	2937	0100983	P48637, V9HWJ1	159	glutathione synthetase deficiency	Organic acidemia
HIBCH	26275	0198130	A0A140VJL0 , Q6NVY1	160	3- hydroxyisobutyryl- CoA hydrolase deficiency	Organic acidemia
IDH2	3418	0182054	P48735, B4DSZ6	161	D-2- hydroxyglutaric aciduria type II	Organic acidemia
L2HG DH	79944	0087299	Q9H9P8	162	L-2- hydroxyglutaric aciduria	Organic acidemia
MLYC D	23417	0103150	O95822	163	malonic acidemia	Organic acidemia
OPA3	80207	0125741	Q9H6K4, B4DK77	164	Costeff syndrome / 3- methylglutaconic aciduria type III	Organic acidemia

OPLA H	26873	0178814	O14841	165	5-oxoprolinase deficiency	Organic acidemia
OXCT I	5019	0083720	A0A024R040 , P55809	166	SCOT deficiency	Organic acidemia
POLG	5428	0140521	E5KNU5, P54098	167	3- methylglutaconic aciduria	Organic acidemia
PPM1 K	152926	0163644	Q8N3J5	168	maple syrup urine disease (MSUD), variant type	Organic acidemia
SERA C1	84947	0122335	Q96JX3	169	Megdel Syndrome	Organic acidemia
SLC25 A1	6576	0100075	D9HTE9, B4DP62, P53007	170	D,L-2- hydroxyglutaric aciduria	Organic acidemia
SUCL A2	8803	0136143	E5KS60, Q9P2R7, Q9Y4T0	171	succinate-CoA ligase deficiency, methylmalonic aciduria	Organic acidemia
SUCL G1	8802	0163541	P53597	172	succinate-CoA ligase deficiency, methylmalonic aciduria	Organic acidemia
TAZ	6901	0102125	A0A0S2Z4K 0, Q16635, A6XNE1, A0A0S2Z4E6 , A0A0S2Z4K 9, A0A0S2Z4F4	173	Barth syndrome	Organic acidemia
AGK	55750	0006530,0262 327	A4D1U5, Q53H12	174	3- methylglutaconic aciduria	Organic acidemia
CLPB	81570	0162129	Q9H078, A0A140VK1 1	175	3- methylglutaconic aciduria	Organic acidemia
TMEM 70	54968	0175606	Q9BUB7	176	3- methylglutaconic aciduria	Organic acidemia
ALDH 18A1	5832	0059573	P54886	177	ALDH18A1- related cutis laxa	Urea cycle disorder
OAT	4942	0065154	A0A140VJQ 4, P04181	178	gyrate atrophy (OAT)	Urea cycle disorder
CA5A	763	0174990	P35218	179	carbonic anhydrase deficiency	Urea cycle disorder
GLUD 1	2746	0148672	P00367, E9KL48	180	glutamate dehydrogenase deficiency	Urea cycle disorder
GLUL	2752	0135821	A8YXX4, P15104	181	glutamine synthetase deficiency	Urea cycle disorder
UMPS	7372	0114491	A8K5J1, P11172	182	Orotic Aciduria	Urea cycle disorder

SLC22A5	6584	0197375	O76082	183	carnitine-acylcarnitine translocase (CACT) deficiency	Fatty acid oxidation
CPT1A	1374	0110090	P50416, A0A024R5F4, B2RAQ8, Q8WZ48	184	carnitine palmitoyltransferase type I (CPT I) deficiency	Fatty acid oxidation
HADHA	3030	0084754	E9KL44, P40939	185	long chain 3-hydroxyacyl-CoA dehydrogenase (LCHAD) deficiency	Fatty acid oxidation
HADH	3033	0138796	Q16836, B3KTT6	186	medium/short chain acyl-CoA dehydrogenase (M/SCHAD) deficiency	Fatty acid oxidation
SLC52A1	55065	0132517	Q9NWF4	187	Riboflavin transporter deficiency	Fatty acid oxidation
SLC52A2	79581	0185803	Q9HAB3	188	Riboflavin transporter deficiency	Fatty acid oxidation
SLC52A3	113278	0101276	K0A6P4, Q9NQ40	189	Riboflavin transporter deficiency	Fatty acid oxidation
HADHB	3032	0138029	P55084, F5GZQ3	190	Trifunctional protein deficiency	Fatty acid oxidation
GYS2	2998	0111713	P54840	191	GSD 0 (Glycogen synthase, liver isoform)	Liver glycogen storage disorder
PYGL	5836	0100504	P06737	192	GSD VI (Hers disease)	Liver glycogen storage disorder
SLC2A2	6514	0163581	P11168, Q6PAU8	193	Fanconi-Bickel syndrome	Liver glycogen storage disorder
ALG1	56052	0033011	Q9BT22	194	ALG1-CDG	Glycosylation disorder
ALG2	85365	0119523	A0A024R184, Q9H553	195	ALG2-associated myasthenic syndrome	Glycosylation disorder
ALG3	10195	0214160	Q92685, C9J7S5	196	ALG3-CDG	Glycosylation disorder
ALG6	29929	0088035	Q9Y672	197	ALG6-CDG	Glycosylation disorder
ALG8	79053	0159063	Q9BVK2, A0A024R5K5	198	ALG8-CDG	Glycosylation disorder
ALG9	79796	0086848	Q9H6U8	199	ALG9-CDG	Glycosylation disorder
ALG11	440138	0253710	Q2TAA5	200	ALG11-CDG	Glycosylation disorder
ALG12	79087	0182858	A0A024R4V6, Q9BV10	201	ALG12-CDG	Glycosylation disorder
ALG13	79868	0101901	Q9NP73, A0A087WX4	202	ALG13-CDG	Glycosylation disorder

			3, A0A087WT1 5			
ATP6V 0A2	23545	0185344	Q9Y487	203	ATP6V0A2- associated cutis laxa	Glycosylation disorder
B3GLC T	145173	0187676	Q6Y288	204	B3GLCT-CDG	Glycosylation disorder
CHST1 4	113189	0169105	Q8NCH0	205	CHST14-CDG	Glycosylation disorder
COG1	9382	0166685	Q8WTW3	206	COG1-CDG	Glycosylation disorder
COG2	22796	0135775	Q14746, B1ALW7	207	COG2-CDG	Glycosylation disorder
COG4	25839	0103051	A0A0A0MS4 5, Q8N8L9, Q9H9E3, J3KNII	208	COG4-CDG	Glycosylation disorder
COG5	10466	0164597,0284 369	Q9UP83	209	COG5-CDG	Glycosylation disorder
COG6	57511	0133103	A0A140VJG 7, Q9Y2V7, A0A024RD W5	210	COG6-CDG	Glycosylation disorder
COG7	91949	0168434	A0A0S2Z652 , P83436	211	COG7-CDG	Glycosylation disorder
COG8	84342	0272617	A0A024R6Z6 , Q96MW5	212	COG8-CDG	Glycosylation disorder
DOLK	22845	0175283	A0A0S2Z597 , Q9UPQ8	213	DOLK-CDG	Glycosylation disorder
DHDD S	79947	0117682	Q86SQ9	214	DHDDS-CDG	Glycosylation disorder
DPAG T1	1798	0172269	A0A024R3H 8, Q9H3H5	215	DPAGT1-CDG	Glycosylation disorder
DPM1	8813	0000419	Q60762, Q5QPK2, A0A0S2Z4Y 5	216	DPM1-CDG	Glycosylation disorder
DPM2	8818	0136908	Q94777	217	DPM2-CDG	Glycosylation disorder
DPM3	54344	0179085	A0A140VJI4, Q9P2X0, Q86TM7	218	DPM3-CDG	Glycosylation disorder
G6PC3	92579	0141349	Q9BUM1	219	Congenital neutropenia	Glycosylation disorder
GFPT1	2673	0198380	Q06210	220	Congenital myasthenic syndrome	Glycosylation disorder
GMPP A	29926	0144591	A0A024R482 , Q96IJ6	221	GMPPA-CDG	Glycosylation disorder
GMPP B	29925	0173540	Q9Y5P6	222	Congenital muscular dystrophy, congenital myasthenic syndrome, and	Glycosylation disorder

					dystroglycanopathy	
MAGT1	84061	0102158	A0A087WU53, Q9H0U3	223	MAGT1-CDG; X-linked immunodeficiency with magnesium defect, Epstein-Barr virus infection and neoplasia (XMEN) syndrome	Glycosylation disorder
MAN1B1	11253	0177239	Q9UKM7	224	MAN1B1-CDG	Glycosylation disorder
MGAT2	4247	0168282	Q10469	225	MGAT2-CDG	Glycosylation disorder
MOGS	7841	0115275	Q13724, Q58F09	226	MOGS-CDG	Glycosylation disorder
MPDU1	9526	0129255	J3QW43, O75352, A0A0S2Z4W8, B4DLH7	227	MPDU1-CDG	Glycosylation disorder
MPI	4351	0178802	H3BPP3, Q8NHZ6, B4DW50, F5GX71, P34949, H3BPP8	228	MPI-CDG	Glycosylation disorder
NGLY1	55768	0151092	Q96IV0	229	NGLY1-CDG	Glycosylation disorder
PGM1	5236	0079739	B7Z6C2, P36871, B4DDQ8	230	PGM1-CDG	Glycosylation disorder
PGM3	5238	0013375	O95394, A0A087WT27	231	PGM3-CDG	Glycosylation disorder
RFT1	91869	0163933	Q96AA3	232	RFT1-CDG	Glycosylation disorder
SEC23B	10483	0101310	Q15437, B4DJW8	233	SEC23B-CDG	Glycosylation disorder
SLC35A1	10559	0164414	P78382	234	SLC35A1-CDG	Glycosylation disorder
SLC35A2	7355	0102100	P78381, A6NFI1, A6NKM8, B4DE15	235	SLC35A2-CDG	Glycosylation disorder
SLC35C1	55343	0181830	Q96A29, B3KQH0	236	SLC35C1-CDG	Glycosylation disorder
SSR4	6748	0180879	P51571	237	SSR4-CDG	Glycosylation disorder
SRD5A3	79644	0128039	Q9H8P0	238	SRD5A3-CDG	Glycosylation disorder
TMEM165	55858	0134851	Q9HC07	239	TMEM165-CDG	Glycosylation disorder
TRIP11	9321	0100815	Q15643	240	TRIP11-CDG	Glycosylation disorder
TUSC3	7991	0104723	Q13454	241	TUSC3-CDG	Glycosylation disorder

ALG14	199857	0172339	Q96F25	242	ALG14-CDG	Glycosylation disorder
B4GALT1	2683	0086062	P15291, W6MEN3	243	B4GALT1-CDG	Glycosylation disorder
DDOST	1650	0244038	A0A024RAD5, P39656	244	DDOST-CDG	Glycosylation disorder
NUS1	116150	0153989	Q96E22	245	NUS1-CDG	Glycosylation disorder
RPN2	6185	0118705	P04844	246	RPN2-CDG	Glycosylation disorder
SEC23A	10484	0100934	Q15436	247	SEC23A-CDG	Glycosylation disorder
SLC35A3	23443	0117620	Q9Y2D2, A0A1W2PRT7, A0A1W2PSD1, A0A1W2PQL8	248	SLC35A3-CDG	Glycosylation disorder
ST3GAL3	6487	0126091	Q11203	249	ST3GAL3-CDG	Glycosylation disorder
STT3A	3703	0134910	P46977	250	STT3A-CDG	Glycosylation disorder
STT3B	201595	0163527	Q8TCJ2	251	STT3B-CDG	Glycosylation disorder
AGA	175	0038002	P20933	252	Aspartylglucosaminuria	Lyosomal storage disorder
ARSA	410	0100299	A0A0C4DFZ2, B4DVI5, P15289	253	Metachromatic leukodystrophy	Lyosomal storage disorder
ARSB	411	0113273	A0A024RAJ9, P15848, A8K4A0	254	Mucopolysaccharidosis type VI	Lyosomal storage disorder
ASAH1	427	0104763	A8K0B6, Q13510, Q53H01	255	Farber disease	Lyosomal storage disorder
ATP13A2	23400	0159363	Q8N4D4, Q9NQ11, Q8NBS1	256	Neuronal ceroid lipofuscinosis 12 (CLN12), Kufor-Rakeb syndrome (KRS)	Lyosomal storage disorder
CLN3	1201	0188603,0261832	A0A024QZB8, Q13286, B4DMY6, Q2TA70, B4DFF3	257	Neuronal ceroid lipofuscinosis 3 (CLN3)	Lyosomal storage disorder
CLN5	1203	0102805	A0A024R644, O75503	258	Neuronal ceroid lipofuscinosis 5 (CLN5)	Lyosomal storage disorder
CLN6	54982	0128973	A0A024R601, Q9NWW5	259	Neuronal ceroid lipofuscinosis 6 (CLN6)	Lyosomal storage disorder
CLN8	2055	0182372,0278220	A0A024QZ57, Q9UBY8	260	Neuronal ceroid lipofuscinosis 8 (CLN8)	Lyosomal storage disorder
CTNS	1497	0040531	A0A0S2Z3I9, O60931,	261	cystinosis	Lyosomal storage disorder

			A0A0S2Z3K3			
CTSA	5476	0064601	P10619, X6R8A1, B4E324, X6R5C5	262	Galactosialidosis	Lyososomal storage disorder
CTSD	1509	0117984	P07339, V9HWI3	263	Neuronal ceroid lipofuscinosis 10 (CLN10)	Lyososomal storage disorder
CTSF	8722	0174080	Q9UBX1	264	Neuronal ceroid lipofuscinosis 13 (CLN13)	Lyososomal storage disorder
CTSK	1513	0143387	P43235	265	Pycnodysostosis	Lyososomal storage disorder
DNAJC5	80331	0101152	Q6AHX3, Q9H3Z4	266	Neuronal ceroid lipofuscinosis 4 (CLN4)	Lyososomal storage disorder
FUCA1	2517	0179163	P04066, B5MDC5	267	Fucosidosis	Lyososomal storage disorder
GAA	2548	0171298	P10253	268	Pompe disease	Lyososomal storage disorder
GALC	2581	0054983	A0A0A0MQV0, P54803	269	Krabbe disease	Lyososomal storage disorder
GALNS	2588	0141012	P34059, Q96I49, Q6YL38	270	Mucopolysaccharidosis type IVa	Lyososomal storage disorder
GLA	2717	0102393	P06280, Q53Y83	271	Fabry disease	Lyososomal storage disorder
GLB1	2720	0170266	P16278, B7Z6Q5	272	GM1 gangliosidosis, Mucopolysaccharidosis IVb	Lyososomal storage disorder
GM2A	2760	0196743	P17900	273	GM2-gangliosidosis, AB variant	Lyososomal storage disorder
GNPTAB	79158	0111670	Q3T906	274	Mucopolipidosis type II alpha/beta, Mucopolipidosis III alpha/beta	Lyososomal storage disorder
GNPTG	84572	0090581	Q9UJJ9	275	Mucopolipidosis III gamma	Lyososomal storage disorder
GNS	2799	0135677	A0A024RBC5, P15586, Q7Z3X3	276	Mucopolysaccharidosis type IIID	Lyososomal storage disorder
GRN	2896	0030582	P28799	277	Neuronal ceroid lipofuscinosis 11 (CLN11), frontotemporal dementia	Lyososomal storage disorder
GUSB	2990	0169919	P08236	278	Mucopolysaccharidosis type VII	Lyososomal storage disorder
HEXA	3073	0213614	A0A0S2Z3W3, P06865, B4DVA7, H3BP20	279	Tay-Sachs disease	Lyososomal storage disorder

HEXB	3074	0049860	A0A024RAJ6, P07686, Q5URX0	280	Sandhoff disease	Lyosomal storage disorder
HGSNAT	138050	0165102	Q68CP4, Q8IVU6	281	Mucopolysaccharidosis type IIIC	Lyosomal storage disorder
HYAL1	3373	0114378	A0A024R2X3, Q12794, B3KUI5, A0A0S2Z3Q0	282	Mucopolysaccharidosis type IX	Lyosomal storage disorder
IDS	3423	0010404	P22304, B4DGD7	283	Mucopolysaccharidosis type II	Lyosomal storage disorder
IDUA	3425	0127415	P35475	284	Mucopolysaccharidosis type I	Lyosomal storage disorder
KCTD7	154881	0243335	Q96MP8, A0A024RDN7	285	Neuronal ceroid lipofuscinosis 14 (CLN14)	Lyosomal storage disorder
LAMP2	3920	0005893	P13473	286	Danon disease	Lyosomal storage disorder
MAN2B1	4125	0104774	O00754, A8K6A7	287	alpha-mannosidosis	Lyosomal storage disorder
MANBA	4126	0109323	O00462	288	beta-mannosidosis	Lyosomal storage disorder
MCOLN1	57192	0090674	Q9GZU1	289	Mucopolipidosis type IV	Lyosomal storage disorder
MFSD8	256471	0164073	Q8NHS3	290	Neuronal ceroid lipofuscinosis 7 (CLN7)	Lyosomal storage disorder
NAGA	4668	0198951	A0A024R1Q5, P17050	291	Schindler disease	Lyosomal storage disorder
NAGLU	4669	0108784	A0A140VJE4, P54802	292	Mucopolysaccharidosis IIIB	Lyosomal storage disorder
NEU1	4758	0204386,0227315,0227129,0223957,0234846,0184494,0228691,0234343	Q5JQI0, Q99519	293	Mucopolipidosis type I, Sialidosis I	Lyosomal storage disorder
NPC1	4864	0141458	O15118	294	Niemann-Pick type C	Lyosomal storage disorder
NPC2	10577	0119655	A0A024R6C0, P61916, G3V3E8	295	Niemann-Pick type C	Lyosomal storage disorder
SGSH	6448	0181523	P51688	296	Mucopolysaccharidosis IIIA	Lyosomal storage disorder
PPT1	5538	0131238	P50897	297	Neuronal ceroid lipofuscinosis 1 (CLN1)	Lyosomal storage disorder
PSAP	5660	0197746	P07602, A0A024QZQ2	298	Prosaposin deficiency, SapA deficiency (Krabbe variant), SapB deficiency (MLD variant), SapC deficiency (Gaucher variant)	Lyosomal storage disorder

SLC17A5	26503	0119899	Q9NRA2	299	Infantile sialic acid storage disease, Salla disease	Lyosomal storage disorder
SMPD1	6609	0166311	P17405, Q59EN6, E9LUE8, Q8IUN0, E9LUE9	300	Niemann Pick types A and B	Lyosomal storage disorder
SUMF1	285362	0144455	Q8NBK3	301	Multiple sulfatase deficiency	Lyosomal storage disorder
TPP1	1200	0166340	O14773	302	Neuronal ceroid lipofuscinosis 2 (CLN2)	Lyosomal storage disorder
AHCY	191	0101444	P23526, Q1RMG2	303	Hypermethionine mia	Aminoacidopathy
GNMT	27232	0124713	A0A0S2Z5F2, Q14749, V9HW60	304	Hypermethionine mia	Aminoacidopathy
MAT1A	4143	0151224	Q00266	305	Hypermethionine mia	Aminoacidopathy
GCH1	2643	0131979	A0A024R642, P30793, Q8IZH9	306	BH4 cofactor deficiency	Aminoacidopathy
PCBD1	5092	0166228	P61457	307	BH4 cofactor deficiency	Aminoacidopathy
PTS	5805	0150787	Q03393	308	BH4 cofactor deficiency	Aminoacidopathy
QDPR	5860	0151552	A0A140VKA9, P09417	309	BH4 cofactor deficiency	Aminoacidopathy
SPR	6697	0116096	P35270	310	BH4 cofactor deficiency	Aminoacidopathy
DNAJC12	56521	0108176	Q6IAH1, Q9UKB3	311	Phenylalanine, tyrosine, and tryptophan hydroxylases heat shock co-chaperone deficiency	Aminoacidopathy
ALDH4A1	8659	0159423	P30038, A0A024RAD8	312	Hyperprolinemia	Aminoacidopathy
PRODH	5625	0100033	O43272	313	Hyperprolinemia	Aminoacidopathy
HPD	3242	0158104	P32754	314	Tyrosinemia type II	Aminoacidopathy
GBA	2629	0177628,0262446	A0A068F658, P04062, B7Z6S9	315	Gaucher disease	
HGD	3081	0113924	Q93099, B3KW64	316	Alkaptonuria	
AMN	81693	0166126	Q9BXJ7, B3KP64	317	Combined Methylmalonic Acidemia and Homocystinuria	Organic acidemia

CD320	51293	0167775	Q9NPF0	318	Combined Methylmalonic Acidemia and Homocystinuria	Organic acidemia
CUBN	8029	0107611	O60494	319	Combined Methylmalonic Acidemia and Homocystinuria	Organic acidemia
GIF	2694	0134812	P27352	320	Combined Methylmalonic Acidemia and Homocystinuria	Organic acidemia
TCN1	6947	0134827	P20061	321	Combined Methylmalonic Acidemia and Homocystinuria	Organic acidemia
TCN2	6948	0185339	P20062	322	Combined Methylmalonic Acidemia and Homocystinuria	Organic acidemia
PREPL	9581	0138078	Q4J6C6	323	Cystinuria	Aminoacidopathy
PHGDH	26227	0092621	O43175	324	Disorders of Serine Biosynthesis	Aminoacidopathy
PSAT1	29968	0135069	A0A024R280, Q9Y617, A0A024R222	325	Disorders of Serine Biosynthesis	Aminoacidopathy
PSPH	5723	0146733	A0A024RDL3, P78330	326	Disorders of Serine Biosynthesis	Aminoacidopathy
AMT	275	0145020	A0A024R2U7, P48728	327	Glycine Encephalopathy	Aminoacidopathy
GCSH	2653	0140905	P23434	328	Glycine Encephalopathy	Aminoacidopathy
GLDC	2731	0178445	P23378	329	Glycine Encephalopathy	Aminoacidopathy
LIAS	11019	0121897	O43766, Q6P5Q6, B4E0L7, A0A024R9W0, A0A1W2PQE9, A0A1X7SBR7	330	Glycine Encephalopathy	Aminoacidopathy
NFU1	27247	0169599	Q9UMS0	331	Glycine Encephalopathy	Aminoacidopathy
SLC6A9	6536	0196517	P48067, B7Z3W8, B7Z589	332	Glycine Encephalopathy	Aminoacidopathy
SLC2A1	6513	0117394	P11166, Q59GX2	333	Glucose Transporter Type 1 Deficiency	Carbohydrate disorder
ATP7A	538	0165240	B4DRW0, Q04656, Q762B6	334	ATP7A-Related Disorders	Metal transport disorder

					Copper Metabolism Disorder	
AP1S1	1174	0106367	A0A024QYT6, P61966	335	Copper Metabolism Disorder	Metal transport disorder
CP	1356	0047457	A5PL27, P00450	336	Copper Metabolism Disorder	Metal transport disorder
SLC33A1	9197	0169359	O00400	337	Copper Metabolism Disorder	Metal transport disorder
PEX7	5191	0112357	O00628, Q6FGN1	338	Adult Refsum Disease Rhizomelic Chondrodysplasia Punctata Spectrum	Peroxisomal disorders
PHYH	5264	0107537	O14832	339	Adult Refsum Disease	Peroxisomal disorders
AGPS	8540	0018510	O00116, B7Z3Q4	340	Rhizomelic Chondrodysplasia Punctata Spectrum	Peroxisomal disorders
GNPAT	8443	0116906	O15228	341	Rhizomelic Chondrodysplasia Punctata Spectrum	Peroxisomal disorders
ABCD1	215	0101986	P33897	342	X-linked Adrenoleukodystrophy	Peroxisomal disorders
ACOX1	51	0161533	Q15067	343	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX1	5189	0127980	O43933, A0A0C4DG33, B4DER6	344	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX2	5828	0164751	P28328	345	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX3	8504	0034693	P56589	346	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX5	5830	0139197	A0A0S2Z480, P50542, B4DR50, A0A0S2Z4F3, A0A0S2Z4H1, B4E0T2	347	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX6	5190	0124587	A0A024RD09, Q13608	348	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX10	5192	0157911	A0A024R068, O60683, A0A024R0A4	349	X-linked Adrenoleukodystrophy	Peroxisomal disorders

PEX12	5193	0108733	O00623	350	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX13	5194	0162928	Q92968	351	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX14	5195	0142655	O75381	352	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX16	9409	0121680	Q9Y5Y5	353	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX19	5824	0162735	P40855, A0A0S2Z497	354	X-linked Adrenoleukodystrophy	Peroxisomal disorders
PEX26	55670	0215193	A0A024R100, Q7Z412, A0A0S2Z5M7, Q7Z2D7	355	X-linked Adrenoleukodystrophy	Peroxisomal disorders
AMACR	23600	0242110	Q9UHK6	356	Zellweger Spectrum Disorder	Peroxisomal disorders
ADA	100	0196839	A0A0S2Z381, P00813, F5GWI4	357	Purine Metabolism Disorder	Purine Metabolism Disorder
ADSL	158	0239900	P30566, X5D8S6, X5D7W4, A0A1B0GWJ0	358	Purine Metabolism Disorder	Purine Metabolism Disorder
AMPD1	270	0116748	P23109	359	Purine Metabolism Disorder	Purine Metabolism Disorder
GPHN	10243	0171723	Q9NQX3	360	Purine Metabolism Disorder	Purine Metabolism Disorder
MOCOS	55034	0075643	Q96EN8	361	Purine Metabolism Disorder	Purine Metabolism Disorder
MOCS1	4337	0124615	A0A024RD17, Q9NZB8	362	Purine Metabolism Disorder	Purine Metabolism Disorder
PNP	4860	0198805	P00491, V9HWH6	363	Purine Metabolism Disorder	Purine Metabolism Disorder
XDH	7498	0158125	P47989	364	Purine Metabolism Disorder	Purine Metabolism Disorder
SUOX	6821	0139531	A0A024RB79, P51687	365	Purine Metabolism Disorder	Purine Metabolism Disorder
OGDH	4967	0105953	A0A140VJQ5, Q02218, B4E3E9, E9PCR7, E9PDF2	366	2-Ketoglutarate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
SLC25A19	60386	0125454	Q5JPC1, Q9HC21	367	2-Ketoglutarate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
DHTKD1	55526	0181192	Q96HY7	368	2-Ketoglutarate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT

SLC13A5	284111	0141485	Q68D44, Q86YT5	369	Citrate Transporter Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
FH	2271	0091483	A0A0S2Z4C3, P07954	370	Fumarase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
DLAT	1737	0150768	P10515, Q86YI5	371	Pyruvate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
MPC1	51660	0060762	Q5TI65, Q9Y5U8	372	Pyruvate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
PDHA1	5160	0131828	A0A024RBX9, P08559	373	Pyruvate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
PDHB	5162	0168291	P11177	374	Pyruvate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
PDHX	8050	0110435	O00330	375	Pyruvate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
PDP1	54704	0164951	Q9P0J1, Q6P1N1, A0A024R9C0	376	Pyruvate Dehydrogenase Deficiency	PYRUVATE METABOLISM AND TRICARBOXYLIC ACID CYCLE DEFECT
ABCC2	1244	0023839	Q92887	377	Dubin-Johnson syndrome	
SLCO1B1	10599	0134538	A0A024RAU7, Q05CV5, Q9Y6L6	378	Rotor Syndrome	
SLCO1B3	28234	0111700	B3KP78, Q9NPD5	379	Rotor Syndrome	
HFE2	148738	0168509	Q6ZVN8, A8K466, A0A024R4F5	380	Hemochromatosis, type 2A	
ADAMTS13	11093	0160323,0281244	Q76LX8	381	Congenital thrombotic thrombocytopenic purpura due to ADAMTS-13 deficiency	
PYGM	5837	0068976	P11217	382	McArdle's Disease	
COL1A2	1278	0164692	A0A0S2Z3H5, P08123	383	Ehlers-Danlos syndrome, cardiac valvular type	
TNFRSF11B	4982	0164761	O00300	384	Juvenile Paget's disease	
TSC1	7248	0165699	Q86WV8, Q92574,	385	Tuberous sclerosis	

			X5D9D2, Q32NF0			
TSC2	7249	0103197	P49815, X5D7Q2, B3KWH7, Q5HYF7, H3BMQ0, X5D2U8	386	Tuberous sclerosis	
DHCR 7	1717	0172893	A0A024R5F7 , Q9UBM7	387	Smith-Lemli-Opitz Syndrome	
PGK1	5230	0102144	P00558, V9HWF4	388	D- glycericacidemia	
VLDL R	7436	0147852	P98155, Q5VVF5	389	Dysequilibrium syndrome	
KYNU	8942	0115919	Q16719	390	Encephalopathy due to hydroxykynurenin uria	
F5	2153	0198734	P12259	391	Factor V deficiency	
C3	718	0125730	B4DR57, P01024, V9HWA9	392	Atypical hemolytic uremic syndrome with C3 anomaly	
COL4 A1	1282	0187498	A5PKV2, F5H5K0, P02462	393	Autosomal dominant familial hematuria - retinal arteriolar tortuosity - contractures	
CFH	3075	0000971	A0A024R962 , P08603, A0A0D9SG8 8	394	Atypical hemolytic uremic syndrome	
SLC12 A2	6558	0064651	P55011, Q53ZR1, B7ZM24	395	Bartter syndrome type I (neonatal)	
GK	2710	0198814	B4DH54, P32189	396	Glycerol kinase deficiency	
SFTPC	6440	0168484	A0A0A0MT C9, P11686, A0A0S2Z4Q 0, E5RI64	397	Chronic respiratory distress with surfactant metabolism deficiency	
CRTA P	10491	0170275	O75718	398	Osteogenesis Imperfecta VII	
P3H1	64175	0117385	Q32P28	399	Osteogenesis Imperfecta VIII	
COL7 A1	1294	0114270	Q02388, Q59F16	400	Autosomal recessive dystrophic epidermolysis bullosa	
PKLR	5313	0143627	P30613	401	Pyruvate Kinase deficiency	
TALD O1	6888	0177156	A0A140VK5 6, P37837	402	Transaldolase deficiency	

TF	7018	0091513	A0PJA6, P02787, Q06AH7	403	Atransferrinemia (familial hypotransferrinemia)	
EPCAM	4072	0119888	P16422	404	Intestinal epithelial dysplasia	
VHL	7428	0134086	A0A024R2F2, P40337, A0A0S2Z4K1	405	Familial erythrocytosis type 2; von Hippel Lindau disease	
GC	2638	0145321	P02774	406	Vitamin D deficiency	
SERPIN A1	5265	0197249,0277 377	E9KL23, P01009	407	Alpha-1 antitrypsin deficiency	
ABCC6	368	0091262,0275 331	O95255	408	Pseudoxanthoma elasticum	
F8	2157	0185010	P00451	409	Hemophilia A	
F9	2158	0101981	P00740	410	Hemophilia B	
ApoB	338	0084674	P04114	411	Familial hypercholesterolemia	
PCSK9	255738	0169174	Q8NBP7	412	Familial hypercholesterolemia	
LDLRAP1	26119	0157978	B3KR97, Q5SW96	413	Familial hypercholesterolemia	
ABCG5	64240	0138075	Q9H222	414	Sitosterolemia	
ABCG8	64241	0143921	Q9H221	415	Sitosterolemia	
LCAT	3931	0213398	A0A140VK24, P04180	416	Lecithin cholesterol acyltransferase deficiency	
SPINK5	11005	0133710	Q9NQ38	417	Netherton syndrome	
GNE	10020	0159921	Q9Y223	418	Inclusion body myopathy 2	

[0314] In some embodiments, the targeted lipid particle or lentiviral vector contains an exogenous agent that is capable of targeting a T cell. In some embodiments, the exogenous agent capable of targeting a T cell is a chimeric antigen receptor (CAR), a T cell receptor, an integrin, an ion channel, a pore forming protein, a Toll-Like Receptor, an interleukin receptor, a cell adhesion protein, or a transport protein.

[0315] In some embodiments, the CAR is or comprises a first generation CAR comprising an antigen binding domain, a transmembrane domain, and signaling domain (e.g., one, two or three signaling domains). In some embodiments, the CAR comprises a third generation CAR comprising an antigen binding domain, a transmembrane domain, and at least three signaling domains. In some embodiments, a fourth generation CAR comprising an antigen binding domain, a transmembrane domain, three or four signaling domains, and a domain which upon successful signaling of the CAR induces expression of a cytokine gene. In some embodiments, the antigen binding domain is or comprises an scFv or Fab.

[0316] In some embodiments, a CAR antigen binding domain is or comprises an antibody or antigen-binding portion thereof. In some embodiments, a CAR antigen binding domain is or comprises an scFv or Fab. In some embodiments a CAR antigen binding domain comprises an scFv or Fab fragment of a T-cell α chain antibody; T-cell β chain antibody; T-cell γ chain antibody; T-cell δ chain antibody; CCR7 antibody; CD3 antibody; CD4 antibody; CD5 antibody; CD7 antibody; CD8 antibody; CD11b antibody; CD11c antibody; CD16 antibody; CD19 antibody; CD20 antibody; CD21 antibody; CD22 antibody; CD25 antibody; CD28 antibody; CD34 antibody; CD35 antibody; CD40 antibody; CD45RA antibody; CD45RO antibody; CD52 antibody; CD56 antibody; CD62L antibody; CD68 antibody; CD80 antibody; CD95 antibody; CD117 antibody; CD127 antibody; CD133 antibody; CD137 (4-1 BB) antibody; CD163 antibody; F4/80 antibody; IL-4Ra antibody; Sca-1 antibody; CTLA-4 antibody; GITR antibody GARP antibody; LAP antibody; granzyme B antibody; LFA-1 antibody; MR1 antibody; uPAR antibody; or transferrin receptor antibody.

[0317] In some embodiments, a CAR binding domain binds to a cell surface antigen of a cell. In some embodiments, a cell surface antigen is characteristic of one type of cell. In some embodiments, a cell surface antigen is characteristic of more than one type of cell.

[0318] In some embodiments, the antigen binding domain of the CAR targets an antigen characteristic of a T cell. In some embodiments, the antigen characteristic of a T cell is selected from a cell surface receptor, a membrane transport protein (e.g., an active or passive transport protein such as, for example, an ion channel protein, a pore-forming protein, etc.), a transmembrane receptor, a membrane enzyme, and/or a cell adhesion protein characteristic of a T cell. In some embodiments, an antigen characteristic of a T cell may be a G protein-coupled receptor, receptor tyrosine kinase, tyrosine kinase associated receptor, receptor-like tyrosine phosphatase, receptor serine/ threonine kinase, receptor guanylyl cyclase, histidine kinase

associated receptor, AKT1; AKT2; AKT3; ATF2; BCL10; CALM1; CD3D (CD3 δ); CD3E (CD3 ϵ); CD3G (CD3 γ); CD4; CD8; CD28; CD45; CD80 (B7-1); CD86 (B7-2); CD247 (CD3 ζ); CTLA4 (CD152); ELK1; ERK1 (MAPK3); ERK2; FOS; FYN; GRAP2 (GADS); GRB2; HLA-DRA; HLA-DRB1; HLA-DRB3; HLA-DRB4; HLA-DRB5; HRAS; IKBKA (CHUK); IKBKB; IKBKE; IKBKG (NEMO); IL2; ITPR1; ITK; JUN; KRAS2; LAT; LCK; MAP2K1 (MEK1); MAP2K2 (MEK2); MAP2K3 (MKK3); MAP2K4 (MKK4); MAP2K6 (MKK6); MAP2K7 (MKK7); MAP3K1 (MEKK1); MAP3K3; MAP3K4; MAP3K5; MAP3K8; MAP3K14 (NIK); MAPK8 (JNK1); MAPK9 (JNK2); MAPK10 (JNK3); MAPK11 (p38 β); MAPK12 (p38 γ); MAPK13 (p38 δ); MAPK14 (p38 α); NCK; NFAT1; NFAT2; NFKB1; NFKB2; NFKBIA; NRAS; PAK1; PAK2; PAK3; PAK4; PIK3C2B; PIK3C3 (VPS34); PIK3CA; PIK3CB; PIK3CD; PIK3R1; PKCA; PKCB; PKCM; PKCQ; PLCY1; PRF1 (Perforin); PTEN; RAC1; RAF1; RELA; SDF1; SHP2; SLP76; SOS; SRC; TBK1; TCRA; TEC; TRAF6; VAV1; VAV2; or ZAP70.

[0319] In some embodiments, the antigen binding domain of the CAR targets an antigen characteristic of a disorder. In some embodiments, the disease or disorder is associated with CD4⁺ T cells. In some embodiments, the disease or disorder is associated with CD8⁺ T cells.

[0320] In some embodiments, the CAR transmembrane domain comprises at least a transmembrane region of the alpha, beta or zeta chain of a T cell receptor, CD28, CD3 epsilon, CD45, CD4, CD5, CD8, CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD134, CD137, CD154, or functional variant thereof. In some embodiments, the transmembrane domain comprises at least a transmembrane region(s) of CD8 α , CD8 β , 4-1BB/CD137, CD28, CD34, CD4, Fc ϵ RI γ , CD16, OX40/CD134, CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , TCR α , TCR β , TCR ζ , CD32, CD64, CD64, CD45, CD5, CD9, CD22, CD37, CD80, CD86, CD40, CD40L/CD154, VEGFR2, FAS, and FGFR2B, or functional variant thereof.

[0321] In some embodiments, the CAR comprises at least one signaling domain selected from one or more of B7-1/CD80; B7-2/CD86; B7-H1/PD-L1; B7-H2; B7-H3; B7-H4; B7-H6; B7-H7; BTLA/CD272; CD28; CTLA-4; Gi24/VISTA/B7-H5; ICOS/CD278; PD-1; PD-L2/B7-DC; PDCD6; 4-1BB/TNFSF9/CD137; 4-1BB Ligand/TNFSF9; BAFF/BLyS/TNFSF13B; BAFF R/TNFRSF13C; CD27/TNFRSF7; CD27 Ligand/TNFSF7; CD30/TNFRSF8; CD30 Ligand/TNFSF8; CD40/TNFRSF5; CD40/TNFSF5; CD40 Ligand/TNFSF5; DR3/TNFRSF25; GITR/TNFRSF18; GITR Ligand/TNFSF18; HVEM/TNFRSF14; LIGHT/TNFSF14; Lymphotoxin-alpha/TNF-beta; OX40/TNFRSF4; OX40 Ligand/TNFSF4; RELT/TNFRSF19L;

TACI/TNFRSF13B; TL1A/TNFSF15; TNF-alpha; TNF RII/TNFRSF1B);
 2B4/CD244/SLAMF4; BLAME/SLAMF8; CD2; CD2F-10/SLAMF9; CD48/SLAMF2;
 CD58/LFA-3; CD84/SLAMF5; CD229/SLAMF3; CRACC/SLAMF7; NTB-A/SLAMF6;
 SLAM/CD150); CD2; CD7; CD53; CD82/Kai-1; CD90/Thy1; CD96; CD160; CD200;
 CD300a/LMIR1; HLA Class I; HLA-DR; Ikaros; Integrin alpha 4/CD49d; Integrin alpha 4 beta
 1; Integrin alpha 4 beta 7/LPAM-1; LAG-3; TCL1A; TCL1B; CRTAM; DAP12; Dectin-
 1/CLEC7A; DPPIV/CD26; EphB6; TIM-1/KIM-1/HAVCR; TIM-4; TSLP; TSLP R;
 lymphocyte function associated antigen-1 (LFA-1); NKG2C, a CD3 zeta domain, an
 immunoreceptor tyrosine-based activation motif (ITAM), CD27, CD28, 4-1BB, CD134/OX40,
 CD30, CD40, PD-1, ICOS, lymphocyte function-associated antigen-1 (LFA-1), CD2, CD7,
 LIGHT, NKG2C, B7-H3, a ligand that specifically binds with CD83, or functional fragment
 thereof.

[0322] In some embodiments, the CAR comprises a CD3 zeta domain or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof. In some embodiments, the CAR comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; and (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof. In some embodiments, the CAR comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; and (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof. In some embodiments, the CAR comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain, or a 4-1BB domain, or functional variant thereof, and/or (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof. In some embodiments, the CAR comprises (i) a CD3 zeta domain, or an immunoreceptor tyrosine-based activation motif (ITAM), or functional variant thereof; (ii) a CD28 domain or functional variant thereof; (iii) a 4-1BB domain, or a CD134 domain, or functional variant thereof; and (iv) a cytokine or costimulatory ligand transgene.

[0323] In certain embodiments, the intracellular signaling domain comprises a CD28 transmembrane and signaling domain linked to a CD3 (e.g., CD3-zeta) intracellular domain. In some embodiments, the intracellular signaling domain comprises a chimeric CD28 and CD137 (4-1BB, TNFRSF9) co-stimulatory domains, linked to a CD3 zeta intracellular domain.

[0324] In some embodiments, the CAR encompasses one or more, e.g., two or more, costimulatory domains and an activation domain, e.g., primary activation domain, in the cytoplasmic portion. Exemplary CARs include intracellular components of CD3-zeta, CD28, and 4-1BB.

[0325] In some embodiments the intracellular signaling domain includes intracellular components of a 4-1BB signaling domain and a CD3-zeta signaling domain. In some embodiments, the intracellular signaling domain includes intracellular components of a CD28 signaling domain and a CD3zeta signaling domain.

[0326] In some embodiments, the CAR comprises an extracellular antigen binding domain (e.g., antibody or antibody fragment, such as an scFv) that binds to an antigen (e.g. tumor antigen), a spacer (e.g. containing a hinge domain, such as any as described herein), a transmembrane domain (e.g. any as described herein), and an intracellular signaling domain (e.g. any intracellular signaling domain, such as a primary signaling domain or costimulatory signaling domain as described herein). In some embodiments, the intracellular signaling domain is or includes a primary cytoplasmic signaling domain. In some embodiments, the intracellular signaling domain additionally includes an intracellular signaling domain of a costimulatory molecule (e.g., a costimulatory domain). Examples of exemplary components of a CAR are described in Table 6. In provided aspects, the sequences of each component in a CAR can include any combination listed in Table 6.

Table 6: CAR components and Exemplary Sequences		
Component	Sequence	SEQ ID NO
<i>Extracellular binding domain</i>		
Anti-CD19 scFv (FMC63)	DIQMTQTTSSLSASLGDRVTISCRASQDISKY LNWYQQKPDGTVKLLIYHTSRLHSGVPSRFS GSGSGTDYSLTISNLEQEDIATYFCQQGNTLP YTFGGGTKEITGSTSGSGKPGSGEGSTKGE VKLQESGPGLVAPSQSLSVTCTVSGVSLPDY GVSWIRQPPRKGLEWLGVIWGSETTYNSA LKSRLTIKDNSKSQVFLKMNSLQTDDTAIYY CAKHYYYGGSYAMDYWGQGTSTVTVSS	419
Anti-CD19 scFv (FMC63)	DIQMTQTTSSLSASLGDRVTISCRASQDISKY LNWYQQKPDGTVKLLIYHTSRLHSGVPSRFS GSGSGTDYSLTISNLEQEDIATYFCQQGNTLP YTFGGGTKEITGGGGSGGGGSGGGGSEVK	420

	LQESGPGLVAPSQSLSVTCTVSGVSLPDYGV SWIRQPPRKGLEWLGVIWGSETTYNSALKS RLTIKDNSKSQVFLKMNSLQTDDTAIYYCA KHYYYGGSYAMDYWGQGTSTVTVSS	
<i>Spacer (e.g. hinge)</i>		
IgG4 Hinge	ESKYGPPCPPCP	421
CD8 Hinge	TTTPAPRPPTPAPTIASQPLSLRPE	422
CD28	IEVMYPPPPYLDNEKSNGTIIHVKGKHLCPSP FPGPSKP	423
<i>Transmembrane</i>		
CD8	ACRPAAGGAVHTRGLDFACDIYWAPLAGT CGVLLLSLVITLYC	424
CD28	FWVLVVVGVLACYSLLVTVAFIIFWV	425
CD28	FWVLVVVGVLACYSLLVTVAFIIFWV	426
<i>Costimulatory domain</i>		
CD28	RSKRSRLHSDYMNMTPRRPGPTRKHYQPY APPRDFAAYRS	427
4-1BB	KRGRKLLYIFKQPFMRPVQTTQEEDGCSCR FPEEEEGGCEL	428
<i>Primary Signaling Domain</i>		
CD3zeta	RVKFSRSADAPAYQQGQNQLYNELNLGRRE EYDVLDKRRGRDPEMGGKPRRKNPQEGLY NELQKDKMAEAYSEIGMKGERRRGKGHDG LYQGLSTATKDTYDALHMQALPPR	429
CD3zeta	RVKFSRSADAPAYKQGQNQLYNELNLGRRE EYDVLDKRRGRDPEMGGKPRRKNPQEGLY NELQKDKMAEAYSEIGMKGERRRGKGHDG LYQGLSTATKDTYDALHMQALPPR	430

[0327] In some embodiments, the CAR further comprises one or more spacers, e.g., wherein the spacer is a first spacer between the antigen binding domain and the transmembrane domain. In some embodiments, the first spacer includes at least a portion of an immunoglobulin constant region or variant or modified version thereof. In some embodiments, the spacer is a second spacer between the transmembrane domain and a signaling domain. In some embodiments, the second spacer is an oligopeptide, e.g., wherein the oligopeptide comprises glycine-serine doublets.

[0328] In addition to the CARs described herein, various chimeric antigen receptors and nucleotide sequences encoding the same are known and would be suitable for fusosomal delivery and reprogramming of target cells in vivo and in vitro as described herein. See, e.g., WO2013040557; WO2012079000; WO2016030414; Smith T, et al., Nature Nanotechnology. 2017.(DOI: 10.1038/NNANO.2017.57), the disclosures of which are herein incorporated by reference in their entirety.

[0329] In some embodiments a targeted lipid particle comprising a CAR or a nucleic acid encoding a CAR (e.g., a DNA, a gDNA, a cDNA, an RNA, a pre-mRNA, an mRNA, an miRNA, an siRNA, etc.) is delivered to a target cell. In some embodiments the target cell is an effector cell, e.g., a cell of the immune system that expresses one or more Fc receptors and mediates one or more effector functions. In some embodiments, a target cell may include, but may not be limited to, one or more of a monocyte, macrophage, neutrophil, dendritic cell, eosinophil, mast cell, platelet, large granular lymphocyte, Langerhans' cell, natural killer (NK) cell, T lymphocyte (e.g., T cell), a Gamma delta T cell, B lymphocyte (e.g., B cell) and may be from any organism including but not limited to humans, mice, rats, rabbits, and monkeys.

E. Methods of Generating Targeted lipid particles

[0330] Provided herein is a targeted lipid particle comprising a lipid bilayer, a lumen surrounded by the lipid bilayer, a targeted envelope protein, and a fusogen, in which the targeted envelope protein and fusogen are embedded within the lipid bilayer. In some embodiments, the targeted lipid particle can be a viral particle, a virus-like particle, a nanoparticle, a vesicle, an exosome, a dendrimer, a lentivirus, a viral vector, an enucleated cell, a microvesicle, a membrane vesicle, an extracellular membrane vesicle, a plasma membrane vesicle, a giant plasma membrane vesicle, an apoptotic body, a mitoparticle, a pyrenocyte, a lysosome, another membrane enclosed vesicle, or a lentiviral vector, a viral based particle, a virus like particle (VLP) or a cell derived particle.

1. Virus-like Particles

[0331] Provided herein are targeted lipid particles that are derived from virus, such as viral particles or virus-like particles, including those derived from retroviruses or lentiviruses. In some embodiments, the targeted lipid particle's bilayer of amphipathic lipids is or comprises the viral envelope. In some embodiments, the targeted lipid particle's bilayer of amphipathic lipids

is or comprises lipids derived from a producer cell. In some embodiments, 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 targeted lipid particle's lumen or cavity comprises a viral nucleic acid, e.g., a retroviral nucleic acid, e.g., a lentiviral nucleic acid. In some embodiments, the viral nucleic acid may be a viral genome. In some embodiments, the targeted lipid particle further comprises one or more viral non-structural proteins, e.g., in its cavity or lumen. In some embodiments, the targeted lipid particles is or comprises a virus-like particle (VLP). In some embodiments, the VLP does not comprise an envelope. In some embodiments, the VLP comprises an envelope.

[0332] In some embodiments, the viral particle or virus-like particle, such as retrovirus or retrovirus-like particle, comprises one or more of gag polyprotein, polymerase (e.g., pol), integrase (e.g., a functional or non-functional variant), protease, and a fusogen. In some embodiments, the targeted lipid particle 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 targeted lipid particle nucleic acid (e.g., retroviral nucleic acid) comprises one or more 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 targeted lipid particle nucleic acid further comprises one or more insulator element. In some embodiments, the recognition sites are situated between the poly A tail sequence and the WPRE.

[0333] In some embodiments, the targeted lipid particle comprises supramolecular complexes formed by viral proteins that self-assemble into capsids. In some embodiments, the targeted lipid particle is a viral particle or virus-like particle derived from viral capsids. In some embodiments, the targeted lipid particle is a viral particle or virus-like particle derived from viral nucleocapsids. In some embodiments, the targeted lipid particle comprises nucleocapsid-derived that retain the property of packaging nucleic acids. In some embodiments, the viral particles or virus-like particles comprises only viral structural glycoproteins. In some embodiments, the targeted lipid particle does not contain a viral genome.

[0334] In some embodiments, the targeted lipid particle packages nucleic acids from host cells during the expression process. In some embodiments, the nucleic acids do not encode any genes involved in virus replication. In particular embodiments, the targeted lipid particle is a virus-like particle, e.g. retrovirus-like particle such as a lentivirus-like particle, that is replication defective.

[0335] In some cases, the targeted lipid particle is a viral particle that is morphologically indistinguishable from the wild type infectious virus. In some embodiments, the viral particle presents the entire viral proteome as an antigen. In some embodiments, the viral particle presents only a portion of the proteome as an antigen.

[0336] In some embodiments, the viral particle or virus-like particle is produced utilizing proteins (e.g., envelope proteins) from a virus within the Paramyxoviridae family. In some embodiments, the Paramyxoviridae family comprises members within the Henipavirus genus. In some embodiments, the Henipavirus is or comprises a Hendra (HeV) or a Nipah (NiV) virus. In particular embodiments, the viral particles or virus-like particles incorporate a targeted envelope protein and fusogen as described in Section I.A. and 1.B.

[0337] In some embodiments, viral particles or virus-like particles may be produced in multiple cell culture systems including bacteria, mammalian cell lines, insect cell lines, yeast and plant cells.

[0338] In some embodiments, the assembly of a viral particle or virus-like particle is initiated by binding of the core protein to a unique encapsidation sequence within the viral genome (e.g. UTR with stem-loop structure). In some embodiments, the interaction of the core with the encapsidation sequence facilitates oligomerization.

[0339] In some embodiments, the targeted lipid particle is a virus-like 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 some embodiments, this may be achieved by using an endogenous packaging signal binding site on gag. In some embodiments, the endogenous packaging signal binding site is on pol. In some embodiments, the RNA which is to be delivered will contain a cognate packaging signal. In some embodiments, 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. In some embodiments, the heterologous sequence could be non-viral or it could be viral, in which case it may be derived from a different virus. In some embodiments, the vector

particles could be used to deliver therapeutic RNA, in which case functional integrase and/or reverse transcriptase is not required. In some embodiments, the vector particles could also be used to deliver a therapeutic gene of interest, in which case pol is typically included.

a. Transfer vectors

[0340] 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.

[0341] 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.

[0342] In some embodiments the retrovirus is a Gammaretrovirus. 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.

[0343] 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.

[0344] 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.

[0345] In some embodiments, a viral vector comprises 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). In some embodiments, a viral vector comprises 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). In some embodiments, a viral vectors and transfer plasmids 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.

[0346] 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.

[0347] In some embodiments, in the 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. In some embodiments, 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.

[0348] In some embodiments, 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). In some embodiments, the LTRs are involved in proviral integration and transcription. In some embodiments, LTRs serve as enhancer-promoter sequences and can control the expression of the viral genes. In some embodiments, encapsidation of the retroviral RNAs occurs by virtue of a psi sequence located at the 5' end of the viral genome.

[0349] In some embodiments, LTRs are similar 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.

[0350] In some embodiments, 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. In some embodiments, retroviruses comprise any one or more of the following genes that code for proteins that are involved in the regulation of gene expression: tat, rev, tax and rex.

[0351] In some embodiments, the structural genes gag, pol and env, gag encodes the internal structural protein of the virus. In some embodiments, Gag protein is proteolytically processed into the mature proteins MA (matrix), CA (capsid) and NC (nucleocapsid). In some embodiments, the pol gene encodes the reverse transcriptase (RT), which contains DNA polymerase, associated RNase H and integrase (IN), which mediate replication of the genome. In some embodiments, 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. In some embodiments, the interaction promotes infection by fusion of the viral membrane with the cell membrane.

[0352] In some embodiments, a replication-defective retroviral vector genome gag, pol and env may be absent or not functional. In some embodiments, the R regions at both ends of the RNA are typically repeated sequences. In some embodiments, U5 and U3 represent unique sequences at the 5' and 3' ends of the RNA genome respectively.

[0353] In some embodiments, 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. In some embodiments, 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).

[0354] In some embodiments, in addition to protease, reverse transcriptase and integrase, non-primate lentiviruses contain a fourth pol gene product which codes for a dUTPase. In some embodiments, this a role in the ability of these lentiviruses to infect certain non-dividing or slowly dividing cell types.

[0355] 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. In some embodiments, infection of the target cell can comprise reverse transcription and integration into the target cell genome. In some embodiments, the RLV typically carries non-viral coding sequences which are to be delivered by the vector to the target cell. In some embodiments, an RLV is incapable of independent replication to produce infectious retroviral particles within the target cell. In some embodiments, the RLV lacks a functional gag-pol and/or env gene and/or other genes involved in replication. In some embodiments, 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.

[0356] 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.

[0357] In some embodiments, a minimal lentiviral genome may comprise, e.g., (5')R-U5-one or more first nucleotide sequences-U3-R(3'). In some embodiments, 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. In some embodiments, the 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. In some embodiments, lentiviral genomes comprise additional sequences to promote efficient virus production. In some embodiments, in the case of HIV, rev and RRE sequences may be included. In some embodiments, 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. In some embodiments, alternative sequences which perform a similar or the same function as the rev/RRE system may also be used. In some embodiments, a functional analogue of the rev/RRE system is found in the Mason Pfizer monkey virus. In some embodiments, 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. In some embodiments, CTE may be used as an alternative to the rev/RRE system. In some embodiments, the Rex protein of HTLV-I can functionally replace the Rev protein of HIV-I. Rev and Rex have similar effects to IRE-BP.

[0358] 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.

[0359] 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.

[0360] In some embodiments, the deletion of additional genes may permit vectors to be produced without the genes associated with disease in lentiviral (e.g. HIV) infections. In some embodiments, tat is associated with disease. In some embodiments, the deletion of additional genes permits the vector to package more heterologous DNA. In some embodiments, 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.

[0361] 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.

[0362] 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. In some embodiments, 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.

[0363] In some embodiments, different cells differ in their usage of particular codons. In some embodiments, this codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. In some embodiments, 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. In some embodiments, 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. In some embodiments, 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.

[0364] In some embodiments 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.

[0365] In some embodiments, codon optimization has a number of other advantages. In some embodiments, 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.

[0366] 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.

[0367] 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.

[0368] In some embodiments, 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.

[0369] 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.

[0370] In some embodiments, 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.

[0371] In some embodiments, 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.

[0372] 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.

[0373] In some embodiments, 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.

[0374] In some embodiments, 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 protease (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.

[0375] In some embodiments, the lentiviral vector is integration-deficient. In some embodiments, the pol is integrase deficient, such as by encoding due to mutations in the integrase gene. For example, the pol coding sequence can contain an inactivating mutation in the integrase, such as by mutation of one or more of amino acids involved in catalytic activity, i.e. mutation of one or more of aspartic 64, aspartic acid 116 and/or glutamic acid 152. In some embodiments, the integrase mutation is a D64V mutation. In some embodiments, the mutation in the integrase allows for packaging of viral RNA into a lentivirus. In some embodiments, the mutation in the integrase allows for packaging of viral proteins into a lentivirus. In some embodiments, the mutation in the integrase reduces the possibility of insertional mutagenesis. In some embodiments, the mutation in the integrase decreases the possibility of generating replication-competent recombinants (RCRs) (Wanisch et al. 2009. *Mol Ther.* 1798):1316-1332). In some embodiments, 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.

[0376] In some embodiments, 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.

[0377] 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.

[0378] In some 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.

[0379] In some embodiments, 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).

[0380] In some embodiments, 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.

[0381] 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).

[0382] 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.

[0383] 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.

[0384] 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.

[0385] In some embodiments, 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.

[0386] In some embodiments, 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.

[0387] 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.

[0388] 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.

[0389] In some embodiments, a retroviral nucleic acid described herein lacks or does not comprise a posttranscriptional regulatory element such as a WPRE or HPRE.

[0390] In some embodiments, elements directing the termination and polyadenylation of the heterologous nucleic acid transcripts may be included, e.g., to increase 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 (r β gpA), or another suitable heterologous or endogenous polyA sequence.

[0391] In some embodiments, a retroviral or lentiviral vector further comprises one or more insulator elements, e.g., an insulator element described herein.

[0392] 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 or 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.

[0393] 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).

b. Packaging vectors and producer cells

[0394] 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.

[0395] In some 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 producer cell, and are introduced into the cell via transfection, transduction or infection. A

retroviral, e.g., lentiviral, transfer vector can be introduced into a producer 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.

[0396] In some embodiments, producer 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.

[0397] In some embodiments, a source cell line includes a cell line which is capable of producing recombinant retroviral particles, comprising a producer 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 producer 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.

[0398] 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.

[0399] 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.

[0400] 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.

[0401] 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.

[0402] 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.

[0403] 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.

[0404] 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.

2. Cell-derived particles

[0405] Provided herein are targeted lipid particles that comprise a naturally derived membrane. In some embodiments, the naturally derived membrane comprises membrane vesicles prepared from cells or tissues. In some embodiments, the targeted lipid particle comprises a vesicle that is obtainable from a cell. In some embodiments, the targeted lipid particle comprises a microvesicle, an exosome, a membrane enclosed body, an apoptotic body (from apoptotic cells), a particle (which may be derived from e.g. platelets), an ectosome (derivable from, e.g., neutrophils and monocytes in serum), a prostasome (obtainable from prostate cancer cells), or a cardiosome (derivable from cardiac cells).

[0406] 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., HeEa, 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.

[0407] In some embodiments, the targeted lipid particle 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. In some embodiments, the targeted

lipid particle 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 having a functional nucleus.

[0408] In embodiments, the targeted lipid particle has a size, or the population of targeted lipid particles have 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.

[0409] In some embodiments the targeted lipid particle comprises an extracellular vesicle, e.g., a cell-derived vesicle comprising a membrane that encloses an internal space and has a smaller diameter than the cell from which it is derived. In embodiments the extracellular vesicle has a diameter from 20 nm to 1000 nm. In embodiments the targeted lipid particle comprises an apoptotic body, a fragment of a cell, a vesicle derived from a cell by direct or indirect manipulation, a vesiculated organelle, and a vesicle produced by a living cell (e.g., by direct plasma membrane budding or fusion of the late endosome with the plasma membrane). In embodiments the extracellular vesicle is derived from a living or dead organism, explanted tissues or organs, or cultured cells.

[0410] In embodiments, the targeted lipid particle comprises a nanovesicle, e.g., a cell-derived small (e.g., between 20-250 nm in diameter, or 30-150 nm in diameter) vesicle comprising a membrane that encloses an internal space, and which is generated from said cell by direct or indirect manipulation. The production of nanovesicles can, in some instances, result in the destruction of the source cell. The nanovesicle may comprise a lipid or fatty acid and polypeptide.

[0411] In embodiments, the targeted lipid particle comprises an exosome. In embodiments, the exosome is a cell-derived small (e.g., between 20-300 nm in diameter, or 40-200 nm in diameter) vesicle comprising a membrane that encloses an internal space, and which is generated from said cell by direct plasma membrane budding or by fusion of the late endosome with the plasma membrane. In embodiments, production of exosomes does not result in the destruction of the source cell. In embodiments, the exosome comprises lipid or fatty acid and polypeptide. Exemplary exosomes and other membrane-enclosed bodies are also described in WO/2017/161010, WO/2016/077639, US20160168572, US20150290343, and US20070298118, each of which is incorporated by reference herein in its entirety.

[0412] In some embodiments, the targeted lipid particle is derived from a source cell with a genetic modification which results in increased expression of an immunomodulatory agent. In

some embodiments, the immunosuppressive agent is on an exterior surface of the cell. In some embodiments, the immunosuppressive agent is incorporated into the exterior surface of the targeted lipid particle. In some embodiments, the targeted lipid particle comprises an immunomodulatory agent attached to the surface of the solid particle by a covalent or non-covalent bond.

c. a. Generation of cell-derived particles

[0413] In some embodiments, targeted lipid particles 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.

[0414] In some embodiments, targeted lipid particles are generated by inducing cell enucleation. Enucleation may be performed using assays such as genetic, chemical (e.g., using Actinomycin D, see Bayona-Bafaluy et al., “A chemical enucleation method for the transfer of mitochondrial DNA to ρ^0 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.

[0415] In some embodiments, the targeted lipid particles 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.

[0416] In some embodiments, the targeted lipid particle is a microvesicle. In some embodiments the microvesicle has a diameter of about 100 nm to about 2000 nm. In some embodiments, a targeted lipid particle comprises a cell ghost. In some embodiments, a vesicle is a plasma membrane vesicle, e.g. a giant plasma membrane vesicle.

[0417] In some embodiments, the source cell used to make the targeted lipid particle will not be available for testing after the targeted lipid particle is made.

[0418] In some embodiments, a characteristic of a targeted lipid particle 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 targeted

lipid particle 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.

III. PHARMACEUTICAL COMPOSITIONS

[0419] The present disclosure also provides, in some aspects, a pharmaceutical composition comprising the targeted lipid particle composition described herein and pharmaceutically acceptable carrier. The pharmaceutical compositions can include any of the described targeted lipid particles.

[0420] In some embodiments, the targeted lipid particle meets a pharmaceutical or good manufacturing practices (GMP) standard. In some embodiments, the targeted lipid particle was made according to good manufacturing practices (GMP). In some embodiments, the targeted lipid particle has a pathogen level below a predetermined reference value, e.g., is substantially free of pathogens. In some embodiments, the targeted lipid particle has a contaminant level below a predetermined reference value, e.g., is substantially free of contaminants. In some embodiments, the targeted lipid particle has low immunogenicity.

[0421] In some embodiments, provided herein are the use of pharmaceutical compositions of the invention or salts thereof to practice the methods of the invention. Such a pharmaceutical composition may consist of at least one compound or conjugate of the invention or a salt thereof in a form suitable for administration to a subject, or the pharmaceutical composition may comprise at least one compound or conjugate of the invention or a salt thereof, and one or more pharmaceutically acceptable carriers, one or more additional ingredients, or some combination of these. In some embodiments, the compound or conjugate of the invention may be present in the pharmaceutical composition in the form of a physiologically acceptable salt, such as in combination with a physiologically acceptable cation or anion, as is well known in the art.

[0422] In some embodiments, the pharmaceutical compositions useful for practicing the methods of the invention may be administered to deliver a dose of between 1 ng/kg/day and 100 mg/kg/day. In another embodiment, the pharmaceutical compositions useful for practicing the invention may be administered to deliver a dose of between 1 ng/kg/day and 500 mg/kg/day.

[0423] In some embodiments, the relative amounts of the active ingredient, the pharmaceutically acceptable carrier, and any additional ingredients in a pharmaceutical composition of the invention will vary, depending upon the identity, size, and condition of the subject treated and further depending upon the route by which the composition is to be

administered. In some embodiments, the composition may comprise between 0.1% and 100% (w/w) active ingredient.

[0424] In some embodiments, pharmaceutical compositions that are useful in the methods of the invention may be suitably developed for oral, rectal, vaginal, parenteral, topical, pulmonary, intranasal, buccal, ophthalmic, or another route of administration. In some embodiments, a composition useful within the methods of the invention may be directly administered to the skin, vagina or any other tissue of a mammal. In some embodiments, formulations include liposomal preparations, resealed erythrocytes containing the active ingredient, and immunologically based formulations. In some embodiments, the route(s) of administration will be readily apparent to the skilled artisan and will depend upon any number of factors including the type and severity of the disease being treated, the type and age of the veterinary or human subject being treated, and the like.

[0425] In some embodiments, formulations of the pharmaceutical compositions described herein may be prepared by any method known or hereafter developed in the art of pharmacology. In some embodiments, preparatory methods include the step of bringing the active ingredient into association with a carrier or one or more other accessory ingredients, and then, if necessary or desirable, shaping or packaging the product into a desired single- or multi-dose unit.

[0426] In some embodiments, a “unit dose” is a discrete amount of the pharmaceutical composition comprising a predetermined amount of the active ingredient. In some embodiments, the amount of the active ingredient is generally equal to the dosage of the active ingredient that would be administered to a subject or a convenient fraction of such a dosage such as, for example, one-half or one-third of such a dosage. In some embodiments, the unit dosage form may be for a single daily dose or one of multiple daily doses (e.g., about 1 to 4 or more times per day). In some embodiments, when multiple daily doses are used, the unit dosage form may be the same or different for each dose.

[0427] In some embodiments, although the descriptions of pharmaceutical compositions provided herein are principally directed to pharmaceutical compositions that are suitable for ethical administration to humans, it will be understood by the skilled artisan that such compositions are generally suitable for administration to animals of all sorts. In some embodiments, modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various animals is well

understood, and the ordinarily skilled veterinary pharmacologist may design and perform such modification with merely ordinary, if any, experimentation. In some embodiments, subjects to which administration of the pharmaceutical compositions of the invention is contemplated include humans and other primates, mammals including commercially relevant mammals such as cattle, pigs, horses, sheep, cats, and dogs.

[0428] In some of any embodiments, the compositions of the invention are formulated using one or more pharmaceutically acceptable excipients or carriers. In one embodiment, the pharmaceutical compositions of the invention comprise a therapeutically effective amount of a compound or conjugate of the invention and a pharmaceutically acceptable carrier. In some embodiments, pharmaceutically acceptable carriers that are useful, include, but are not limited to, glycerol, water, saline, ethanol and other pharmaceutically acceptable salt solutions such as phosphates and salts of organic acids. Examples of these and other pharmaceutically acceptable carriers are described in Remington's Pharmaceutical Sciences (1991, Mack Publication Co., New Jersey).

[0429] In some embodiments, the carrier may be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. In some embodiments, the proper fluidity may be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. In some embodiments, prevention of the action of microorganisms may be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In some embodiments, it is preferable to include isotonic agents, for example, sugars, sodium chloride, or polyalcohols such as mannitol and sorbitol, in the composition. In some embodiments, prolonged absorption of the injectable compositions may be brought about by including in the composition an agent that delays absorption, for example, aluminum monostearate or gelatin. In one embodiment, the pharmaceutically acceptable carrier is not DMSO alone.

[0430] In some embodiments, formulations may be employed in admixtures with conventional excipients, i.e., pharmaceutically acceptable organic or inorganic carrier substances suitable for oral, vaginal, parenteral, nasal, intravenous, subcutaneous, enteral, or any other suitable mode of administration, known to the art. In some embodiments, the pharmaceutical preparations may be sterilized and if desired mixed with auxiliary agents, e.g., lubricants,

preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure buffers, coloring, flavoring and/or aromatic substances and the like. In some embodiments, pharmaceutical preparations may also be combined where desired with other active agents, e.g., other analgesic agents.

[0431] In some embodiments, “additional ingredients” include, but are not limited to, one or more of the following: excipients; surface active agents; dispersing agents; inert diluents; granulating and disintegrating agents; binding agents; lubricating agents; sweetening agents; flavoring agents; coloring agents; preservatives; physiologically degradable compositions such as gelatin; aqueous vehicles and solvents; oily vehicles and solvents; suspending agents; dispersing or wetting agents; emulsifying agents, demulcents; buffers; salts; thickening agents; fillers; emulsifying agents; antioxidants; antibiotics; antifungal agents; stabilizing agents; and pharmaceutically acceptable polymeric or hydrophobic materials. In some embodiments, “additional ingredients” that may be included in the pharmaceutical compositions of the invention are known in the art and described, for example in Genaro, ed. (1985, Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa.), which is incorporated herein by reference.

[0432] In some embodiments, the composition of the invention may comprise a preservative from about 0.005% to 2.0% by total weight of the composition. In some embodiments, the preservative is used to prevent spoilage in the case of exposure to contaminants in the environment. In some embodiments, examples of preservatives useful in accordance with the invention included but are not limited to those selected from the group consisting of benzyl alcohol, sorbic acid, parabens, imidurea and combinations thereof. In some embodiments, a particularly preferred preservative is a combination of about 0.5% to 2.0% benzyl alcohol and 0.05% to 0.5% sorbic acid.

[0433] In some embodiments, the composition preferably includes an anti-oxidant and a chelating agent that inhibits the degradation of the compound. In some embodiments, antioxidants for some compounds are BHT, BHA, alpha-tocopherol and ascorbic acid in the preferred range of about 0.01% to 0.3% and more preferably BHT in the range of 0.03% to 0.1% by weight by total weight of the composition. In some embodiments, the chelating agent is present in an amount of from 0.01% to 0.5% by weight by total weight of the composition. Particularly preferred chelating agents include edetate salts (e.g. disodium edetate) and citric acid in the weight range of about 0.01% to 0.20% and more preferably in the range of 0.02% to

0.10% by weight by total weight of the composition. In some embodiments, the chelating agent is useful for chelating metal ions in the composition that may be detrimental to the shelf life of the formulation. In some embodiments, other suitable and equivalent antioxidants and chelating agents may be substituted therefore as would be known to those skilled in the art.

[0434] In some embodiments, liquid suspensions may be prepared using conventional methods to achieve suspension of the active ingredient in an aqueous or oily vehicle. In some embodiments, aqueous vehicles include, for example, water, and isotonic saline. In some embodiments, oily vehicles include, for example, almond oil, oily esters, ethyl alcohol, vegetable oils such as arachis, olive, sesame, or coconut oil, fractionated vegetable oils, and mineral oils such as liquid paraffin. In some embodiments, liquid suspensions may further comprise one or more additional ingredients including, but not limited to, suspending agents, dispersing or wetting agents, emulsifying agents, demulcents, preservatives, buffers, salts, flavorings, coloring agents, and sweetening agents. In some embodiments, oily suspensions may further comprise a thickening agent. In some embodiments, suspending agents include, but are not limited to, sorbitol syrup, hydrogenated edible fats, sodium alginate, polyvinylpyrrolidone, gum tragacanth, gum acacia, and cellulose derivatives such as sodium carboxymethylcellulose, methylcellulose, hydroxypropylmethylcellulose. In some embodiments, dispersing or wetting agents include, but are not limited to, naturally-occurring phosphatides such as lecithin, condensation products of an alkylene oxide with a fatty acid, with a long chain aliphatic alcohol, with a partial ester derived from a fatty acid and a hexitol, or with a partial ester derived from a fatty acid and a hexitol anhydride (e.g., polyoxyethylene stearate, heptadecaethyleneoxycetanol, polyoxyethylene sorbitol monooleate, and polyoxyethylene sorbitan monooleate, respectively). Known emulsifying agents include, but are not limited to, lecithin, and acacia. Known preservatives include, but are not limited to, methyl, ethyl, or n-propyl-para-hydroxybenzoates, ascorbic acid, and sorbic acid. Known sweetening agents include, for example, glycerol, propylene glycol, sorbitol, sucrose, and saccharin. Known thickening agents for oily suspensions include, for example, beeswax, hard paraffin, and cetyl alcohol.

[0435] In some embodiments, liquid solutions of the active ingredient in aqueous or oily solvents may be prepared in substantially the same manner as liquid suspensions, the primary difference being that the active ingredient is dissolved, rather than suspended in the solvent. As used herein, an “oily” liquid is one which comprises a carbon-containing liquid molecule and which exhibits a less polar character than water. In some embodiments, liquid solutions of the

pharmaceutical composition of the invention may comprise each of the components described with regard to liquid suspensions, it being understood that suspending agents will not necessarily aid dissolution of the active ingredient in the solvent. In some embodiments, aqueous solvents include, for example, water, and isotonic saline. In some embodiments, oily solvents include, for example, almond oil, oily esters, ethyl alcohol, vegetable oils such as arachis, olive, sesame, or coconut oil, fractionated vegetable oils, and mineral oils such as liquid paraffin.

[0436] In some embodiments, powdered and granular formulations of a pharmaceutical preparation of the invention may be prepared using known methods. In some embodiments, formulations may be administered directly to a subject, used, for example, to form tablets, to fill capsules, or to prepare an aqueous or oily suspension or solution by addition of an aqueous or oily vehicle thereto. In some of any embodiments, formulations may further comprise one or more of dispersing or wetting agent, a suspending agent, and a preservative. Additional excipients, such as fillers and sweetening, flavoring, or coloring agents, may also be included in these formulations.

[0437] In some embodiments, a pharmaceutical composition of the invention may also be prepared, packaged, or sold in the form of oil-in-water emulsion or a water-in-oil emulsion. In some embodiments, the oily phase may be a vegetable oil such as olive or arachis oil, a mineral oil such as liquid paraffin, or a combination of these. In some embodiments, compositions further comprise one or more emulsifying agents such as naturally occurring gums such as gum acacia or gum tragacanth, naturally-occurring phosphatides such as soybean or lecithin phosphatide, esters or partial esters derived from combinations of fatty acids and hexitol anhydrides such as sorbitan monooleate, and condensation products of such partial esters with ethylene oxide such as polyoxyethylene sorbitan monooleate. In some embodiments, emulsions may also contain additional ingredients including, for example, sweetening or flavoring agents.

IV. METHODS OF TREATMENT

[0438] In some embodiments, the targeted lipid particles provided herein, or pharmaceutical compositions thereof as 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. In one embodiment, the subject has cancer. In one embodiment, the subject has an infectious disease. In some embodiments, the targeted lipid particle contains nucleic acid sequences encoding an exogenous

agent for treating the disease or condition in the subject. For example, the exogenous agent is one that targets or is specific for a protein of a neoplastic cells and the targeted lipid particle is administered to a subject for treating a tumor or cancer in the subject. In another example, the exogenous agent is an inflammatory mediator or immune molecule, such as a cytokine, and targeted lipid particle is administered to a subject for treating any condition in which it is desired to modulate (e.g. increase) the immune response, such as a cancer or infectious disease. In some embodiments, the targeted lipid particle is administered in an effective amount or dose to effect treatment of the disease, condition or disorder. Provided herein are uses of any of the provided targeted lipid particles in such methods and treatments, and in the preparation of a medicament in order to carry out such therapeutic methods. In some embodiments, the methods are carried out by administering the targeted lipid particle or compositions comprising the same, to the subject having, having had, or suspected of having the disease or condition or disorder. In some embodiments, the methods thereby treat the disease or condition or disorder in the subject. Also provided herein are uses of any of the compositions, such as pharmaceutical compositions provided herein, for the treatment of a disease, condition or disorder associated with a particular gene or protein targeted by or provided by the exogenous agent.

[0439] In some embodiments, the provided methods or uses involve administration of a pharmaceutical composition comprising oral, inhaled, transdermal or parenteral (including intravenous, intratumoral, intraperitoneal, intramuscular, intracavity, and subcutaneous) administration. In some embodiments, the targeted lipid particle may be administered alone or formulated as a pharmaceutical composition. In some embodiments, the targeted lipid particle or compositions described herein can be administered to a subject, e.g., a mammal, e.g., a human. In some of any 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 disease or disorder.

[0440] In some embodiments, the targeted lipid particles may be administered in the form of a unit-dose composition, such as a unit dose oral, parenteral, transdermal or inhaled composition. In some embodiments, the compositions are prepared by admixture and are adapted for oral, inhaled, transdermal or parenteral administration, and as such may be in the form of tablets, capsules, oral liquid preparations, powders, granules, lozenges, reconstitutable powders, injectable and infusable solutions or suspensions or suppositories or aerosols.

[0441] In some embodiments, the regimen of administration may affect what constitutes an effective amount. In some embodiments, the therapeutic formulations may be administered to the subject either prior to or after a diagnosis of disease. In some embodiments, several divided dosages, as well as staggered dosages may be administered daily or sequentially, or the dose may be continuously infused, or may be a bolus injection. In some embodiments, the dosages of the therapeutic formulations may be proportionally increased or decreased as indicated by the exigencies of the therapeutic or prophylactic situation.

[0442] In some embodiments, the administration of the compositions of the present invention to a subject, preferably a mammal, more preferably a human, may be carried out using known procedures, at dosages and for periods of time effective to prevent or treat disease. In some embodiments, an effective amount of the therapeutic compound necessary to achieve a therapeutic effect may vary according to factors such as the activity of the particular compound employed; the time of administration; the rate of excretion of the compound; the duration of the treatment; other drugs, compounds or materials used in combination with the compound; the state of the disease or disorder, age, sex, weight, condition, general health and prior medical history of the subject being treated, and like factors well-known in the medical arts. In some embodiments, the dosage regimens may be adjusted to provide the optimum therapeutic response. In some embodiments, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation. In some embodiments, the effective dose range for a therapeutic compound of the invention is from about 1 and 5,000 mg/kg of body weight/per day. One of ordinary skill in the art would be able to study the relevant factors and make the determination regarding the effective amount of the therapeutic compound without undue experimentation.

[0443] In some embodiments, the compound may be administered to a subject as frequently as several times daily, or it may be administered less frequently, such as once a day, once a week, once every two weeks, once a month, or even less frequently, such as once every several months or even once a year or less. In some embodiments, the amount of compound dosed per day may be administered, in non-limiting examples, every day, every other day, every 2 days, every 3 days, every 4 days, or every 5 days. In some embodiments, with every other day administration, a 5 mg per day dose may be initiated on Monday with a first subsequent 5 mg per day dose administered on Wednesday, a second subsequent 5 mg per day dose administered on Friday, and so on. The frequency of the dose will be readily apparent to the skilled artisan

and will depend upon any number of factors, such as, but not limited to, the type and severity of the disease being treated, the type and age of the animal, etc.

[0444] In some embodiments, dosage levels of the active ingredients in the pharmaceutical compositions of this invention may be varied so as to obtain an amount of the active ingredient that is effective to achieve the desired therapeutic response for a particular subject, composition, and mode of administration, without being toxic to the subject.

[0445] A medical doctor, e.g., physician or veterinarian, having ordinary skill in the art may readily determine and prescribe the effective amount of the pharmaceutical composition required. In some embodiments, the physician or veterinarian could start doses of the compounds of the invention employed in the pharmaceutical composition at levels lower than that required in order to achieve the desired therapeutic effect and gradually increase the dosage until the desired effect is achieved.

[0446] In some embodiments, it is especially advantageous to formulate the compound in dosage unit form for ease of administration and uniformity of dosage. In some embodiments, dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subjects to be treated; each unit containing a predetermined quantity of therapeutic compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical vehicle. In some embodiments, the dosage unit forms of the invention are dictated by and directly dependent on (a) the unique characteristics of the therapeutic compound and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding/formulating such a therapeutic compound for the treatment of a disease in a subject.

[0447] In some embodiments, the term “container” includes any receptacle for holding the pharmaceutical composition. In some embodiments, the container is the packaging that contains the pharmaceutical composition. In other embodiments, the container is not the packaging that contains the pharmaceutical composition, i.e., the container is a receptacle, such as a box or vial that contains the packaged pharmaceutical composition or unpackaged pharmaceutical composition and the instructions for use of the pharmaceutical composition. It should be understood that the instructions for use of the pharmaceutical composition may be contained on the packaging containing the pharmaceutical composition, and as such the instructions form an increased functional relationship to the packaged product. In some embodiments, instructions may contain information pertaining to the compound's ability to perform its intended function,

e.g., treating or preventing a disease in a subject, or delivering an imaging or diagnostic agent to a subject.

[0448] In some embodiments, routes of administration of any of the compositions disclosed herein include oral, nasal, rectal, parenteral, sublingual, transdermal, transmucosal (e.g., sublingual, lingual, (trans)buccal, (trans)urethral, vaginal (e.g., trans- and perivaginally), (intra)nasal, and (trans)rectal), intravesical, intrapulmonary, intraduodenal, intragastrical, intrathecal, subcutaneous, intramuscular, intradermal, intra-arterial, intravenous, intrabronchial, inhalation, and topical administration.

[0449] In some of any embodiments, suitable compositions and dosage forms include, for example, tablets, capsules, caplets, pills, gel caps, troches, dispersions, suspensions, solutions, syrups, granules, beads, transdermal patches, gels, powders, pellets, magmas, lozenges, creams, pastes, plasters, lotions, discs, suppositories, liquid sprays for nasal or oral administration, dry powder or aerosolized formulations for inhalation, compositions and formulations for intravesical administration and the like.

[0450] In some embodiments, the targeted lipid particle composition comprising an exogenous agent or cargo, may be used to deliver such exogenous agent or cargo to a cell tissue or subject. In some embodiments, delivery of a cargo by administration of a targeted lipid particle composition described herein may modify cellular protein expression levels. In certain embodiments, the administered composition directs upregulation of (via expression in the cell, delivery in the cell, or induction within the cell) of one or more cargo (e.g., a polypeptide or mRNA) that provide a functional activity which is substantially absent or reduced in the cell in which the polypeptide is delivered. In some embodiments, the missing functional activity may be enzymatic, structural, or regulatory in nature. In some embodiments, the administered composition directs up-regulation of one or more polypeptides that increases (e.g., synergistically) a functional activity which is present but substantially deficient in the cell in which the polypeptide is upregulated. In some of any embodiments, the administered composition directs downregulation of (via expression in the cell, delivery in the cell, or induction within the cell) of one or more cargo (e.g., a polypeptide, siRNA, or miRNA) that repress a functional activity which is present or upregulated in the cell in which the polypeptide, siRNA, or miRNA is delivered. In some of any embodiments, the upregulated functional activity may be enzymatic, structural, or regulatory in nature. In some embodiments, the administered composition directs down-regulation of one or more polypeptides that decreases (e.g.,

synergistically) a functional activity which is present or upregulated in the cell in which the polypeptide is downregulated. In some embodiments, the administered composition directs upregulation of certain functional activities and downregulation of other functional activities.

[0451] In some of any embodiments, the targeted lipid particle composition (e.g., one comprising mitochondria or DNA) 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 targeted lipid particle 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.

[0452] In some of any embodiments, the targeted lipid particle composition described herein is delivered ex-vivo to a cell or tissue, e.g., a human cell or tissue. In embodiments, the composition improves function of a cell or tissue ex-vivo, e.g., improves cell viability, respiration, or other function (e.g., another function described herein).

[0453] In some embodiments, the composition is delivered to an ex vivo tissue that is in an injured state (e.g., from trauma, disease, hypoxia, ischemia or other damage).

[0454] In some embodiments, the composition is delivered to an ex-vivo transplant (e.g., a tissue explant or tissue for transplantation, e.g., a human vein, a musculoskeletal graft such as bone or tendon, cornea, skin, heart valves, nerves; or an isolated or cultured organ, e.g., an organ to be transplanted into a human, e.g., a human heart, liver, lung, kidney, pancreas, intestine, thymus, eye). In some embodiments, the composition is delivered to the tissue or organ before, during and/or after transplantation.

[0455] In some embodiments, the composition is delivered, administered or contacted with a cell, e.g., a cell preparation. In some embodiments, the cell preparation may be a cell therapy preparation (a cell preparation intended for administration to a human subject). In embodiments, the cell preparation comprises cells expressing a chimeric antigen receptor (CAR), e.g., expressing a recombinant CAR. The cells expressing the CAR may be, e.g., T cells, Natural Killer (NK) cells, cytotoxic T lymphocytes (CTL), regulatory T cells. In embodiments, the cell preparation is a neural stem cell preparation. In embodiments, the cell preparation is a mesenchymal stem cell (MSC) preparation. In embodiments, the cell preparation is a hematopoietic stem cell (HSC) preparation. In embodiments, the cell preparation is an islet cell preparation.

[0456] In some embodiments, the targeted lipid particle 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).

[0457] In some embodiments, the source of targeted lipid particles are from the same subject that is administered a targeted lipid particle composition. In other embodiments, they are different. In some embodiments, the source of targeted lipid particles and recipient tissue may be autologous (from the same subject) or heterologous (from different subjects). In some embodiments, the donor tissue for targeted lipid particle compositions described herein may be a different tissue type than the recipient tissue. In some embodiments, 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.

[0458] In some embodiments, the targeted lipid particle composition described herein may be administered to a subject having a cancer, an autoimmune disease, an infectious disease, a metabolic disease, a neurodegenerative disease, or a genetic disease (e.g., enzyme deficiency). In some embodiments, the subject is in need of regeneration.

[0459] In some embodiments, the targeted lipid particle 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)). In some embodiments, the targeted lipid particle particles is co-administered with an inhibitor of syppressyn, e.g., a siRNA or inhibitory antibody.

V. EXEMPLARY EMBODIMENTS

Among the provided embodiments are:

1. A targeted lipid particle, comprising:
 - (a) a lipid bilayer enclosing a lumen,
 - (b) a henipavirus F protein molecule or biologically active portion thereof; and
 - (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) single domain antibody (sdAb) variable domain, wherein the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof,

wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.

2. The targeted lipid particle of embodiment 1, wherein the single domain antibody is attached to the G protein via a linker.
3. The targeted lipid particle of embodiment 2, wherein the linker is a peptide linker.
4. A targeted lipid particle, comprising:
 - (a) a lipid bilayer enclosing a lumen,
 - (b) a henipavirus F protein molecule or biologically active portion thereof; and
 - (c) a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or biologically active portion thereof attached to a single domain antibody (sdAb) variable domain via a peptide linker, wherein the single domain antibody binds to a cell surface molecule of a target cell,

wherein the F protein molecule or biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.

5. The targeted lipid particle of any of embodiments 1-4, wherein N-terminus of the F protein molecule or biologically active portion thereof is exposed on the outside of lipid bilayer.
6. The targeted lipid particle of any of embodiments 1-5, wherein the C-terminus of the G protein is exposed on the outside of the lipid bilayer.
7. The targeted lipid particle of any of embodiments 1-6, wherein the single domain antibody binds a cell surface molecule present on a target cell.
8. The targeted lipid particle of embodiment 7, wherein the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule.
9. The targeted lipid particle of embodiment 7, wherein the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells.
10. The targeted lipid particle of embodiment 9, wherein the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a

CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

11. The targeted lipid particle of any of the preceding embodiments, wherein the single domain antibody binds an antigen or portion thereof present on a target cell.

12. The targeted lipid particle of any of embodiments 3-11, wherein the peptide linker comprises up to 65 amino acids in length.

13. The targeted lipid particle of any of embodiments 3-11, wherein the peptide linker comprises from or from about 2 to 65 amino acids, 2 to 60 amino acids, 2 to 56 amino acids, 2 to 52 amino acids, 2 to 48 amino acids, 2 to 44 amino acids, 2 to 40 amino acids, 2 to 36 amino acids, 2 to 32 amino acids, 2 to 28 amino acids, 2 to 24 amino acids, 2 to 20 amino acids, 2 to 18 amino acids, 2 to 14 amino acids, 2 to 12 amino acids, 2 to 10 amino acids, 2 to 8 amino acids, 2 to 6 amino acids, 6 to 65 amino acids, 6 to 60 amino acids, 6 to 56 amino acids, 6 to 52 amino acids, 6 to 48 amino acids, 6 to 44 amino acids, 6 to 40 amino acids, 6 to 36 amino acids, 6 to 32 amino acids, 6 to 28 amino acids, 6 to 24 amino acids, 6 to 20 amino acids, 6 to 18 amino acids, 6 to 14 amino acids, 6 to 12 amino acids, 6 to 10 amino acids, 6 to 8 amino acids, 8 to 65 amino acids, 8 to 60 amino acids, 8 to 56 amino acids, 8 to 52 amino acids, 8 to 48 amino acids, 8 to 44 amino acids, 8 to 40 amino acids, 8 to 36 amino acids, 8 to 32 amino acids, 8 to 28 amino acids, 8 to 24 amino acids, 8 to 20 amino acids, 8 to 18 amino acids, 8 to 14 amino acids, 8 to 12 amino acids, 8 to 10 amino acids, 10 to 65 amino acids, 10 to 60 amino acids, 10 to 56 amino acids, 10 to 52 amino acids, 10 to 48 amino acids, 10 to 44 amino acids, 10 to 40 amino acids, 10 to 36 amino acids, 10 to 32 amino acids, 10 to 28 amino acids, 10 to 24 amino acids, 10 to 20 amino acids, 10 to 18 amino acids, 10 to 14 amino acids, 10 to 12 amino acids, 12 to 65 amino acids, 12 to 60 amino acids, 12 to 56 amino acids, 12 to 52 amino acids, 12 to 48 amino acids, 12 to 44 amino acids, 12 to 40 amino acids, 12 to 36 amino acids, 12 to 32 amino acids, 12 to 28 amino acids, 12 to 24 amino acids, 12 to 20 amino acids, 12 to 18 amino acids, 12 to 14 amino acids, 14 to 65 amino acids, 14 to 60 amino acids, 14 to 56 amino acids, 14 to 52 amino acids, 14 to 48 amino acids, 14 to 44 amino acids, 14 to 40 amino acids, 14 to 36 amino acids, 14 to 32 amino acids, 14 to 28 amino acids, 14 to 24 amino acids, 14 to 20 amino acids, 14 to 18 amino acids, 18 to 65 amino acids, 18 to 60 amino acids, 18 to 56 amino acids, 18 to 52 amino acids, 18 to 48 amino acids, 18 to 44 amino acids, 18 to 40 amino acids, 18 to 36 amino acids,

18 to 32 amino acids, 18 to 28 amino acids, 18 to 24 amino acids, 18 to 20 amino acids, 20 to 65 amino acids, 20 to 60 amino acids, 20 to 56 amino acids, 20 to 52 amino acids, 20 to 48 amino acids, 20 to 44 amino acids, 20 to 40 amino acids, 20 to 36 amino acids, 20 to 32 amino acids, 20 to 28 amino acids, 20 to 26 amino acids, 20 to 24 amino acids, 24 to 65 amino acids, 24 to 60 amino acids, 24 to 56 amino acids, 24 to 52 amino acids, 24 to 48 amino acids, 24 to 44 amino acids, 24 to 40 amino acids, 24 to 36 amino acids, 24 to 32 amino acids, 24 to 30 amino acids, 24 to 28 amino acids, 28 to 65 amino acids, 28 to 60 amino acids, 28 to 56 amino acids, 28 to 52 amino acids, 28 to 48 amino acids, 28 to 44 amino acids, 28 to 40 amino acids, 28 to 36 amino acids, 28 to 34 amino acids, 28 to 32 amino acids, 32 to 65 amino acids, 32 to 60 amino acids, 32 to 56 amino acids, 32 to 52 amino acids, 32 to 48 amino acids, 32 to 44 amino acids, 32 to 40 amino acids, 32 to 38 amino acids, 32 to 36 amino acids, 36 to 65 amino acids, 36 to 60 amino acids, 36 to 56 amino acids, 36 to 52 amino acids, 36 to 48 amino acids, 36 to 44 amino acids, 36 to 40 amino acids, 40 to 65 amino acids, 40 to 60 amino acids, 40 to 56 amino acids, 40 to 52 amino acids, 40 to 48 amino acids, 40 to 44 amino acids, 44 to 65 amino acids, 44 to 60 amino acids, 44 to 56 amino acids, 44 to 52 amino acids, 44 to 48 amino acids, 48 to 65 amino acids, 48 to 60 amino acids, 48 to 56 amino acids, 48 to 52 amino acids, 50 to 65 amino acids, 50 to 60 amino acids, 50 to 56 amino acids, 50 to 52 amino acids, 54 to 65 amino acids, 54 to 60 amino acids, 54 to 56 amino acids, 58 to 65 amino acids, 58 to 60 amino acids, or 60 to 65 amino acids.

14. The targeted lipid particle of any of embodiments 3-11, wherein peptide linker comprises a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64 or 65 amino acids in length.

15. The targeted lipid particle of any of embodiments 3-14, wherein the peptide linker is a flexible linker that comprises GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) or combinations thereof.

16. The targeted lipid particle of any of embodiments 3-15, wherein the peptide linker comprises (GGS)_n, wherein n is 1 to 10.

17. The targeted lipid particle of any of embodiments 3-15, wherein the peptide linker comprises (GGGGS)_n (SEQ ID NO:42), wherein n is 1 to 10.

18. The targeted lipid particle of any of embodiments 3-15, wherein the peptide linker comprises (GGGGGS)_n (SEQ ID NO:27), wherein n is 1 to 6.

19. The targeted lipid particle of any of embodiments 1-18, wherein the G protein or the biologically active portion thereof is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein.

20. The targeted lipid particle of any of embodiments 1-19, wherein the G protein or the biologically active portion thereof is a wild-type NiV-G protein or a functionally active variant or biologically active portion thereof.

21. The targeted lipid particle of embodiment 20, wherein the mutant NiV-G protein or functionally active variant or biologically active portion thereof comprises an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44.

22. The targeted lipid particle of embodiment 21, wherein the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

23. The targeted lipid particle of any of embodiments 1-18, wherein the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and has the sequence set forth in any of SEQ ID NOS: 10-15, 35-40 or 45-50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40 or 45-50.

24. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

25. The targeted lipid particle of embodiment 24, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 10 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10.

26. The targeted lipid particle of embodiment 24, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 35 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35.

27. The targeted lipid particle of embodiment 24, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 45 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45.

28. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9,SEQ ID NO:28 or SEQ ID NO:44).

29. The targeted lipid particle of embodiment 28, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 11 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about

91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11.

30. The targeted lipid particle of embodiment 28, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 36 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36.

31. The targeted lipid particle of embodiment 28, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 46 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

32. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9,SEQ ID NO:28 or SEQ ID NO:44).

33. The targeted lipid particle of embodiment 32, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 12 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12.

34. The targeted lipid particle of embodiment 32, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 37 or an amino acid sequence having at

least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37.

35. The targeted lipid particle of embodiment 32, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 47 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47.

36. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

37. The targeted lipid particle of embodiment 36, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 13 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13.

38. The targeted lipid particle of embodiment 36, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 38 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or

about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38.

39. The targeted lipid particle of embodiment 36, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 48 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48.

40. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

41. The targeted lipid particle of embodiment 40, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 14 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14.

42. The targeted lipid particle of embodiment 40, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 39 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39.

43. The targeted lipid particle of embodiment 40, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 49 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%,

at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49.

44. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

45. The targeted lipid particle of embodiment 44, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 15 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15.

46. The targeted lipid particle of embodiment 44, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 40 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40.

47. The targeted lipid particle of embodiment 44, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:50.

48. The targeted lipid particle of any of embodiments 21-23, wherein the NiV-G protein has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

49. The targeted lipid particle of embodiment 48, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 22 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:22.

50. The targeted lipid particle of embodiment 48, wherein the NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 53 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:53.

51. The targeted lipid particle any of embodiments 1-48, wherein the G-protein or the biologically active portion thereof is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3.

52. The targeted lipid particle of embodiment 51, wherein the mutant NiV-G protein comprises:

one or more amino acid substitutions corresponding to amino acid substitutions selected from the group consisting of E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

53. The targeted lipid particle of embodiment 51 or embodiment 52, wherein the mutant NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 16 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at

least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16.

54. The targeted lipid particle of embodiment 51 or embodiment 52, wherein the mutant NiV-G protein has the amino acid sequence set forth in SEQ ID NO: 51 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:51.

55. The targeted lipid particle of any of embodiments 1-54, wherein the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a Hendra virus F protein or is a functionally active variant or biologically active portion thereof.

56. The targeted lipid particle of any of embodiments 1-55, wherein the F protein or the biologically active portion thereof is a wild-type NiV-F protein or a functionally active variant or a biologically active portion thereof.

57. The targeted lipid particle of any of embodiments 1-56, wherein the NiV-F-protein or the functionally active variant or biologically active portion thereof comprises the amino acid sequenceset forth in SEQ ID NO: 2, or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 2.

58. The targeted lipid particle of any of embodiments 1-57, wherein the NiV-F protein is a a biologically active portion thereof that has a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2).

59. The targeted lipid particle of embodiment 58, wherein the NiV-F protein has an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about

82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 5.

60. The targeted lipid particle of any of embodiments 1-57, wherein the NiV-F protein is a biologically active portion thereof that comprises:

- i) a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2); and
- ii) a point mutation on an N-linked glycosylation site.

61. The targeted lipid particle of embodiment 60, wherein the NiV-F protein has an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 7.

62. The targeted lipid particle of any of embodiments 1-57, wherein the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2).

63. The targeted lipid particle of embodiment 62, wherein the NiV-F protein has an amino acid sequence that is encoded by a sequence of nucleotides encoding a sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 8.

64. The targeted lipid particle of embodiment 63, wherein the NiV-F protein has an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or

about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 23.

65. The targeted lipid particle of any of embodiments 1-57, wherein the F-protein or the biologically active portion thereof comprises an F1 subunit or a fusogenic portion thereof.

66. The targeted lipid particle of embodiment 65, wherein the F1 subunit is a proteolytically cleaved portion of the F0 precursor.

67. The targeted lipid particle of embodiment 66, wherein the F1 subunit comprises the sequence set forth in SEQ ID NO: 4, or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 4.

68. The targeted lipid particle of any of embodiments 1-67, wherein the lipid bilayer is derived from a membrane of a host cell used for producing a retrovirus or retrovirus-like particle.

69. The targeted lipid particle of any of embodiments 1-60, wherein the lipid bilayer is or comprises a viral envelope.

70. The targeted lipid particle of embodiment 68, wherein the retrovirus-like particle is replication defective.

71. The targeted lipid particle of any of embodiments 1-70, wherein the targeted lipid particle comprises one or more viral components other than the F protein molecule and the G protein.

72. The targeted lipid particle of embodiment 71, wherein the one or more viral components are from a retrovirus.

73. The targeted lipid particle of embodiment 72, wherein the retrovirus is a lentivirus.

74. The targeted lipid particle of any of embodiments 71-73, wherein the one or more viral components comprise a viral packaging protein selected from one or more of Gag, Pol, Rev and Tat.

75. The targeted lipid particle of any of embodiments 71-74, wherein the one or more viral components 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)/central termination sequence (CTS) (e.g. DNA flap), Poly A tail sequence, a posttranscriptional regulatory element (e.g. WPRE), a Rev response element (RRE), and 3' LTR (e.g., comprising U5 and lacking a functional U3).

76. The targeted lipid particle of any of embodiments 1-75, wherein the lipid particle further comprises an exogenous agent.

77. The targeted lipid particle of embodiment 76, wherein the exogenous agent is present in the lumen.

78. The targeted lipid particle of embodiment 77, wherein the exogenous agent is a protein or a nucleic acid, optionally wherein the nucleic acid is a DNA or RNA.

79. The targeted lipid particle of any of embodiments 76-78, wherein the exogenous agent encodes a therapeutic agent or a diagnostic agent.

80. The targeted lipid particle of any of embodiments 68-79, wherein the host cell is selected from the group consisting of 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.

81. The targeted lipid particle of any of embodiments 68-80, wherein the host cell comprises 293T cells.

82. A polynucleotide comprising a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a single domain antibody (sdAb) variable domain, wherein the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof.

83. The polynucleotide of embodiment 82, further comprising (iii) a nucleic acid sequence encoding a henipavirus F protein molecule or a biologically active portion thereof.

84. The polynucleotide of embodiment 82 or embodiment 83, further comprising at least one promoter that is operatively linked to control expression of the nucleic acid.

85. The polynucleotide of any of embodiments 83-84, wherein the promoter is a constitutive promoter.

86. The polynucleotide of any of embodiments 83-85, wherein the promoter is an inducible promoter.

87. The polynucleotide of any of embodiments 82-86, wherein the sdAb variable domain is attached to the G protein via an encoded peptide linker.

88. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises up to 65 amino acids in length.

89. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises from or from about 2 to 65 amino acids, 2 to 60 amino acids, 2 to 56 amino acids, 2 to 52 amino acids, 2 to 48 amino acids, 2 to 44 amino acids, 2 to 40 amino acids, 2 to 36 amino acids, 2 to 32 amino acids, 2 to 28 amino acids, 2 to 24 amino acids, 2 to 20 amino acids, 2 to 18 amino acids, 2 to 14 amino acids, 2 to 12 amino acids, 2 to 10 amino acids, 2 to 8 amino acids, 2 to 6 amino acids, 6 to 65 amino acids, 6 to 60 amino acids, 6 to 56 amino acids, 6 to 52 amino acids, 6 to 48 amino acids, 6 to 44 amino acids, 6 to 40 amino acids, 6 to 36 amino acids, 6 to 32 amino acids, 6 to 28 amino acids, 6 to 24 amino acids, 6 to 20 amino acids, 6 to 18 amino acids, 6 to 14 amino acids, 6 to 12 amino acids, 6 to 10 amino acids, 6 to 8 amino acids, 8 to 65 amino acids, 8 to 60 amino acids, 8 to 56 amino acids, 8 to 52 amino acids, 8 to 48 amino acids, 8 to 44 amino acids, 8 to 40 amino acids, 8 to 36 amino acids, 8 to 32 amino acids, 8 to 28 amino acids, 8 to 24 amino acids, 8 to 20 amino acids, 8 to 18 amino acids, 8 to 14 amino acids, 8 to 12 amino acids, 8 to 10 amino acids, 10 to 65 amino acids, 10 to 60 amino acids, 10 to 56 amino acids, 10 to 52 amino acids, 10 to 48 amino acids, 10 to 44 amino acids, 10 to 40 amino acids, 10 to 36 amino acids, 10 to 32 amino acids, 10 to 28 amino acids, 10 to 24 amino acids, 10 to 20 amino acids, 10 to 18 amino acids, 10 to 14 amino acids, 10 to 12 amino acids, 12 to 65 amino acids, 12 to 60 amino acids, 12 to 56 amino acids, 12 to 52 amino acids, 12 to 48 amino acids, 12 to 44 amino acids, 12 to 40 amino acids, 12 to 36 amino acids, 12 to 32 amino acids, 12 to 28 amino acids, 12 to 24 amino acids, 12 to 20 amino acids, 12 to 18 amino acids, 12 to 14 amino acids, 14 to 65 amino acids, 14 to 60 amino acids, 14 to 56 amino acids, 14 to 52 amino acids, 14 to 48 amino acids, 14 to 44 amino acids, 14 to 40 amino acids, 14 to 36 amino acids, 14 to 32 amino acids, 14 to 28 amino acids, 14 to 24 amino acids, 14 to 20 amino acids, 14 to 18 amino acids, 18 to 65 amino acids, 18 to 60 amino acids, 18 to 56 amino acids, 18 to 52 amino acids, 18 to 48 amino acids, 18 to 44 amino acids, 18 to 40 amino acids, 18 to 36

amino acids, 18 to 32 amino acids, 18 to 28 amino acids, 18 to 24 amino acids, 18 to 20 amino acids, 20 to 65 amino acids, 20 to 60 amino acids, 20 to 56 amino acids, 20 to 52 amino acids, 20 to 48 amino acids, 20 to 44 amino acids, 20 to 40 amino acids, 20 to 36 amino acids, 20 to 32 amino acids, 20 to 28 amino acids, 20 to 26 amino acids, 20 to 24 amino acids, 24 to 65 amino acids, 24 to 60 amino acids, 24 to 56 amino acids, 24 to 52 amino acids, 24 to 48 amino acids, 24 to 44 amino acids, 24 to 40 amino acids, 24 to 36 amino acids, 24 to 32 amino acids, 24 to 30 amino acids, 24 to 28 amino acids, 28 to 65 amino acids, 28 to 60 amino acids, 28 to 56 amino acids, 28 to 52 amino acids, 28 to 48 amino acids, 28 to 44 amino acids, 28 to 40 amino acids, 28 to 36 amino acids, 28 to 34 amino acids, 28 to 32 amino acids, 32 to 65 amino acids, 32 to 60 amino acids, 32 to 56 amino acids, 32 to 52 amino acids, 32 to 48 amino acids, 32 to 44 amino acids, 32 to 40 amino acids, 32 to 38 amino acids, 32 to 36 amino acids, 36 to 65 amino acids, 36 to 60 amino acids, 36 to 56 amino acids, 36 to 52 amino acids, 36 to 48 amino acids, 36 to 44 amino acids, 36 to 40 amino acids, 40 to 65 amino acids, 40 to 60 amino acids, 40 to 56 amino acids, 40 to 52 amino acids, 40 to 48 amino acids, 40 to 44 amino acids, 44 to 65 amino acids, 44 to 60 amino acids, 44 to 56 amino acids, 44 to 52 amino acids, 44 to 48 amino acids, 48 to 65 amino acids, 48 to 60 amino acids, 48 to 56 amino acids, 48 to 52 amino acids, 50 to 65 amino acids, 50 to 60 amino acids, 50 to 56 amino acids, 50 to 52 amino acids, 54 to 65 amino acids, 54 to 60 amino acids, 54 to 56 amino acids, 58 to 65 amino acids, 58 to 60 amino acids, or 60 to 65 amino acids.

90. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises a polypeptide that is 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64 or 65 amino acids in length.

91. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) and combinations thereof.

92. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises (GGS) n , wherein n is 1 to 10.

93. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises (GGGGS) n (SEQ ID NO:42), wherein n is 1 to 10.

94. The polynucleotide of any of embodiments 86-87, wherein the encoded peptide linker comprises (GGGGGS)_n (SEQ ID NO:27), wherein n is 1 to 4.

95. The polynucleotide of any of embodiments 86-87, wherein the nucleic acid sequence encoding the G protein is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein or is a variant thereof that exhibits reduced binding for the native binding partner.

96. The polynucleotide of any of embodiments 82-95, wherein the nucleic acid sequence encoding the G protein is a wild-type NiV-G protein.

97. The polynucleotide of any of embodiments 82-95, wherein the nucleic acid sequence encoding the G-protein is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3.

98. The polynucleotide of embodiment 97, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44.

99. The polynucleotide of any of embodiments 82-95 and 97, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the sequence set forth in any of SEQ ID NOS: 10-15, 35-40 or 45-50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40 or 45-50.

100. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44).

101. The polynucleotide of embodiment 100, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 10 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10.

102. The polynucleotide of embodiment 100, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 35 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35.

103. The polynucleotide of embodiment 100, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 45 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45.

104. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44).

105. The polynucleotide of embodiment 104, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 11 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least

at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11.

106. The polynucleotide of embodiment 104, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 36 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36.

107. The polynucleotide of embodiment 104, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 46 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

108. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises a 15 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44).

109. The polynucleotide of embodiment 108, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 12 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12.

110. The polynucleotide of embodiment 108, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 37 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37.

111. The polynucleotide of embodiment 108, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 47 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47.

112. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44).

113. The polynucleotide of embodiment 112, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 13 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13.

114. The polynucleotide of embodiment 112, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 38 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least

at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38.

115. The polynucleotide of embodiment 112, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 48 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48.

116. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44).

117. The polynucleotide of embodiment 116, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 14 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14.

118. The polynucleotide of embodiment 116, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 39 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39.

119. The polynucleotide of embodiment 116, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 49 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49.

120. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44).

121. The polynucleotide of embodiment 120, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 15 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15.

122. The polynucleotide of embodiment 120, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 40 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40.

123. The polynucleotide of embodiment 120, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least

at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 50.

124. The polynucleotide of any of embodiments 97-99, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises:

- i) a truncation at or near the N-terminus; and
- ii) point mutations selected from the group consisting of E501A, W504A, Q530A and E533A.

125. The polynucleotide of embodiment 124, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 16 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16.

126. The polynucleotide of embodiment 124, wherein the nucleic acid sequence encoding the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 51 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:51.

127. A vector, comprising the polynucleotide of any of embodiments 82-126.

128. The vector of embodiment 127, wherein the vector is a mammalian vector, viral vector or artificial chromosome, optionally wherein the artificial chromosome is a bacterial artificial chromosome (BAC).

129. A cell comprising the polynucleotide of any of embodiments 82-126 or the vector of embodiment 127 or embodiment 128.

130. A method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain comprising:

- a) providing a cell that comprises a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and a nucleic acid encoding a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain;
- b) culturing the cell under conditions that allow for production of a targeted lipid particle, and
- c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

131. A method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain, comprising:

- a) providing a cell that comprises the polynucleotide of any of embodiments 82-126 or the vector of embodiment 127 or embodiment 128;
- b) providing the cell a polynucleotide encoding a henipavirus F protein molecule or biologically active portion thereof;
- c) culturing the cell under conditions that allow for production of a targeted lipid particle, and
- d) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

132. The method of embodiment 130 or embodiment 131, wherein the cell is a mammalian cell.

133. The method of any of embodiments 130-131, wherein the cell is a producer cell and the targeted lipid particle is a viral particle or a viral-like particle, optionally a retroviral particle or a retroviral-like particle, optionally a lentiviral particle or lentiviral-like particle.

134. A producer cell comprising (i) a viral nucleic acid(s) and (ii) nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and (iii) a nucleic acid encoding a targeted envelope protein comprising a henipavirus envelope attachment

glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain, optionally wherein the viral nucleic acid(s) are lentiviral nucleic acids.

135. The producer cell of embodiment 134, wherein the viral nucleic acid(s) lacks one or more genes involved in viral replication.

136. The producer cell of embodiment 134 or embodiment 135, wherein the viral nucleic acid comprises a nucleic acid encoding a viral packaging protein selected from one or more of Gag, Pol, Rev and Tat.

137. The producer cell of any of embodiments 134-136, wherein the viral 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)/central termination sequence (CTS) (e.g. DNA flap), Poly A tail sequence, a posttranscriptional regulatory element (e.g. WPRE), a Rev response element (RRE), and 3' LTR (e.g., comprising U5 and lacking a functional U3);

138. The producer cell of any of embodiments 134-137, wherein the henipavirus F protein molecule or biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 2;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:2.

139. The producer cell of any of embodiments 134-137, wherein the henipavirus F protein molecule or biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 5;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:5.

140. The producer cell of any of embodiments 134-137, wherein the henipavirus F protein molecule or biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 7;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:7.

141. The producer cell of any of embodiments 134-137, wherein the henipavirus F protein molecule or biologically active portion thereof comprises:

(i) a sequence encoding by a nucleotide sequence encoding the sequence set forth in SEQ ID NO: 8;

(ii) a amino acid sequence encoded by a nucleotide sequence encoding a sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:8.

142. The producer cell of any of embodiments 134-137, wherein the henipavirus F protein molecule or biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 23;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:23.

143. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO:44;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO:44.

144. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 10;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:10.

145. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 35;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:35.

146. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 45;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:45.

147. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 11;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:11.

148. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 36;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:36.

149. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 46;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or

about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:46.

150. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 12;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:12.

151. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 37;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:37.

152. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 47;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%,

at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:47.

153. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 13;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:13.

154. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 38;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:38.

155. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 48;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:48.

156. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 14;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:14.

157. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 39;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:39.

158. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 49;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or

about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:49.

159. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 15;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:15.

160. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 40;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:40.

161. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 50;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about

93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:50.

162. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 16;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16.

163. The producer cell of any of embodiments 134-142, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 51;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:51.

164. A viral vector particle or viral-like particle produced from the producer cell of any of embodiments 134-163.

165. A composition comprising a plurality of targeted lipid particles of any of embodiments 1-81 and 173-176.

166. The composition of embodiment 165 further comprising a pharmaceutically acceptable carrier.

167. The pharmaceutical composition of embodiment 165 or embodiment 166, wherein the targeted lipid particles comprise an average diameter of less than 1 μm .

168. A method of delivering an exogenous agent to a subject (e.g., a human subject), the method comprising administering to the subject the targeted lipid particle of any of embodiments 1-81 and 173-176 or the composition of any of embodiments 165-167 and 177.

169. A method of treating a disease or disorder in a subject (e.g., a human subject), the method comprising administering to the subject a targeted lipid particle of any of embodiments 1-81 and 173-176 or the composition of any of embodiments 165-167 and 177.

170. A method of fusing a mammalian cell to a targeted lipid particle, the method comprising administering to the subject a targeted lipid particle of any of embodiments 1-81 and 173-176 or the composition of any of embodiments 165-167 and 177.

171. The method of embodiment 170, wherein the fusing of the mammalian cell to the targeted lipid particle delivers an exogenous agent to a subject (e.g., a human subject).

172. The method of embodiment 170 or embodiment 171, wherein the fusing of the mammalian cell to the targeted lipid particle treats a disease or disorder in a subject (e.g., a human subject).

173. The targeted lipid particle of any of embodiments 1-81, wherein the targeted lipid particle has greater expression of the targeted envelope protein compared to a reference lipid particle that has incorporated into a similar lipid bilayer the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting moiety is a single chain variable fragment (scFv).

174. The targeted lipid particle of embodiment 173, wherein the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more.

175. The targeted lipid particle of embodiment 173, wherein the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, preferably at or about or greater than 10-fold or more.

176. The targeted lipid particle of any of embodiments 1-81 and 173-175 or the viral vector particle or viral-like particle of embodiment 164, wherein the titer in target cells following transduction is at or greater than 1×10^6 transduction units (TU)/mL, at or greater than 2×10^6 TU/mL, at or greater than 3×10^6 TU/mL, at or greater than 4×10^6 TU/mL, at or greater than 5×10^6 TU/mL, at or greater than 6×10^6 TU/mL, at or greater than 7×10^6 TU/mL, at or

greater than 8×10^6 TU/mL, at or greater than 9×10^6 TU/mL, or at or greater than 1×10^7 TU/mL.

177. The composition of any of embodiments 165-167, wherein among the population of lipid particles in the composition, greater than at or about 50%, greater than at or about 55%, greater than at or about 60%, greater than at or about 65%, greater than at or about 70%, or greater than at or about 75% are surface positive for the targeted envelope protein.

178. The targeted lipid particle of any of embodiments 1-81 and 173-176, wherein the targeted envelope protein is present on the surface of the targeted lipid particle at a density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

179. A composition comprising a plurality of the targeted lipid particles of any of embodiments 1-81, 173-176 and 178, wherein the targeted envelope protein is present on the surface of the targeted lipid particles at an average density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

180. The producer cell of any one of embodiments 134-163, wherein the producer cell has greater membrane (e.g., plasma membrane) expression of the targeted envelope protein compared to a reference producer cell that has incorporated into its membrane (e.g. plasma membrane) the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting moiety is a single chain variable fragment (scFv).

181. The producer cell of embodiment 180, wherein the expression is increased by at or greater than 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 125%, 150%, 200%, 300%, 400%, 500% or more.

182. The producer cell of embodiment 180, wherein the expression is increased by at or greater than 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 15-fold, 20-fold, 30-fold or more, preferably at or about or greater than 10-fold or more.

183. The producer cell of any one of embodiments 134-163 and 180-182, wherein the producer cell has the expression of the targeted envelope protein on a membrane (e.g., plasma membrane) of the producer cell is at least 20 proteins (e.g., at least 50, 100, 200, 500, 1000, 2000, 5000, or 10,000 proteins) per square micron.

184. The producer cell of any one of embodiments 134-163 and 180-183, wherein the targeted envelope protein comprises at least 0.1% (e.g., at least 0.2%, 0.5%, 1%,

2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10%) of the total membrane (e.g., plasma membrane) proteins of the producer cell (e.g., by total protein weight).

EXAMPLES

[0460] The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1: GENERATION AND CHARACTERIZATION OF PRODUCER CELLS CONTAINING TARGETED BINDERS.

[0461] This Example describes generation and assessment of NiVG targeted binding sequences in which NiVG was linked to scFv or VHH binding modalities.

A. Binding Modalities Directed to CD4.

[0462] Exemplary retargeted NivG fusogen constructs were generated containing an scFv or VHH binding modality against human cellular receptor CD4. For each binding modality, four different sequences that contained a unique CDR3 were assessed. Each exemplary binder sequence was codon optimized and cloned into an expression vector as a fusion with a sequence encoding NiVG (GcΔ34; Bender et al. 2016 PLoS Pathol 12(6):e1005641). The resulting vectors encoded a NivG targeting domain containing NiVG (SEQ ID NO:16) a flexible linker and the binding domain, followed by a 6xHis-tag for detection (NivG-linker-scFv-6xHis).

[0463] After subcloning, 5μg of each exemplary construct was transfected into HEK 293 cells using a transfection reagent. A pcDNA3.1 plasmid (empty vector) and the expression vector without the binder domain (NiVG-linker-NoBinder) were used as negative controls.

[0464] At 48 hours post-transfection, cells were harvested and 100,000 cells were incubated for 1 hour at 4°C with either 50 nM or 300 nM of soluble human CD4 protein with a human Fc tag (hCD4-Fc). After incubation, cells were washed and co-stained with an anti-His antibody conjugated to Alexa-647 to detect surface expression of NivG-binders and an anti-human Fc antibody conjugated to Alexa-488 to detect binding to soluble hCD4-Fc protein.

[0465] Cells were analyzed by flow cytometry, and gates for His (surface expression) and Fc (CD4-protein binding) were set based on the negative control empty vector (pcDNA3.1). Evaluation of median fluorescence intensity (MFI) of cells transfected with constructs containing VHH binding modalities demonstrated higher surface expression as quantified by % of His+ cells (**FIG. 1A**) and higher binding to soluble hCD4-Fc protein as quantified by %Fc+ cell (**FIG. 1B**), than cells transfected with constructs containing scFv binding modalities.

B. Binding Modalities Directed to Multiple Cellular Receptors

[0466] Exemplary constructs were generated containing scFv and VHH binding modalities generally as described above, but containing unique sequences directed against other cellular receptors hCD8, CD4, ASGR2, TM4SF5, LDLR or ASGR1. Multiple sequences, each containing a unique CDR3, were assessed for each binding modality containing distinct cellular receptors. After subcloning into the NivG-linker- -6xHis expression vector as described above, 5µg of each exemplary construct was transfected into about HEK 293 cells. The pcDNA3.1 plasmid (empty vector) and the expression vector without the binding domain (NivG-linker-NoBinder) were used as negative controls.

[0467] At 48 hours post-transfection, cells were harvested and 100,000 cells were washed and stained with an anti-His antibody conjugated to Alexa-647 to detect surface expression of NivG-binders. Cells were analyzed by flow cytometry, and gates for His (surface expression) were set based on the negative control empty vector (pcDNA3.1). Median fluorescence intensity (MFI) was normalized to that of the NivG-NoBinder control set to 100. Cells transfected with constructs containing VHH binding modalities, compared to the scFv binding modalities, demonstrated higher surface expression of targeted binding sequences on 293 cells as quantified by % of His+ cells (**FIG. 1C**).

EXAMPLE 2: GENERATION AND CHARACTERIZATION OF LENTIVIRUSES PSEUDOTYPED WITH TARGETED BINDERS.

[0468] This Example describes generation of lentiviruses pseudotyped with NivG retargeted fusogens and assessment of transduction of primary human T cells.

A. Generation of NivG Pseudotyped Lentiviruses.

[0469] 293 cells were plated at 5.4×10^6 into 10 cm dishes and allowed to rest for 24 hours. At 24 hours after plating, cells were transfected using polyethylenimine (PEI) with the following plasmids: NivG pseudotyped vector containing hCD4 targeted binding sequences linked to scFv or VHH binding modalities (NivG-linker-hCD4-binding modality), vector containing a nucleotide sequence encoding the NivF sequence NivFdel22 (SEQ ID NO:8; or SEQ ID NO:23 without a signal sequence; Bender et al. 2016 PLoS), a packaging plasmid containing an empty backbone, an HIV-1 pol, HIV-1 gag, HIV-1 Rev, HIV-1 Tat, an AmpR promoter and an SV40 promoter and a lentiviral reporter plasmid encoding an enhanced green fluorescent protein (eGFP) under the control of a SFFV promoter pLenti-SFFV-eGFP. Positive control cells were generated using the plasmids described above along with 4 μ g of VSV-G.

B. NivG Pseudotyped Lentiviral Transduction Efficiency of Primary Human T Cells.

[0470] PanT cells from peripheral blood (StemCellTech, Vancouver, Canada) that were negatively selected to enrich for T cells were thawed and activated with anti CD3/ anti-CD28 for 2 days. Concentrated lentiviruses generated generally as described above were serially diluted 6-fold starting at 0.05 dilution with a total of 4 points in the dilution series. Lentiviruses were added to 100,000 PanT cells and transduced by spinfection for 90 minutes at 1000g at 25C. Transduced PanT cells were split on days 2 and 5 post-transduction, and on day 7 post-transduction, cells were harvested and stained with an Alexa-647 conjugated anti-human CD4 antibody. Cells were analyzed by flow cytometry, and titer was determined by % of CD4-positive cells that were GFP+. Cells transfected with constructs containing VHH binding modalities demonstrated a 10-fold increased titer over constructs containing scFv binding modalities on primary human T cells (**FIG. 2**).

EXAMPLE 3. IN VIVO DELIVERY OF LENTIVIRUSES PSEUDOTYPED WITH CD8 TARGETED BINDERS.

[0471] This Example describes generation of lentiviruses pseudotyped with a CD8 NivG retargeted fusogen and in vivo assessment of transduction of primary human T cells.

[0472] CD8 retargeted NivG fusogens were generated essentially as described in Example 2. The retargeted NivG pseudotyped fusogen contained a NivG targeting domain containing NivG

(SEQ ID NO:16) a flexible linker and an exemplary CD8 binding domain, either a VHH or scFv binding modality.

[0473] T cells from human peripheral blood mononuclear cells (PBMCs) were activated with anti CD3/ anti-CD28 for 3 days. After 3 days of incubation, 1×10^7 cells were injected intraperitoneally into NOD-scid-IL2 γ^{null} mice. One day post-injection, mice received 1×10^7 transducing units (TU) of CD8 NivG pseudotyped lentiviruses generated as described above, or no lenti-viral vector (LVV) control, through intraperitoneal injection. On day 7 post-CD8 NivG pseudotyped lentivirus injection, peritoneal cells were harvested and analyzed by flow cytometry, and titer was determined by % of CD8 positive or negative cells that were GFP+. The CD8 retargeted pseudotyped lentiviruses demonstrated significant in vivo transduction of CD8+ T cells (**FIG. 3A**) and minimal transduction of CD8- T cells (**FIG. 3B**). These results indicate that CD8 targeted pseudotyped lentiviral-mediated delivery permits specific delivery of a transgene to the intended cell type (e.g. CD8+ T cells).

EXAMPLE 4. IN VITRO ASSESSMENT OF CHIMERIC ANTIGEN RECEPTOR (CAR) CONTAINING PSEUDOTYPED LENTIVIRUSES WITH CD8 TARGETED BINDERS.

[0474] This Example describes the in vitro tumor killing activity of lentivirus pseudotyped with a CD8 retargeted fusogen and expressing a CD19-directed chimeric antigen receptor (CD19CAR). The lentiviruses were generated substantially as described in Example 3, except that a plasmid encoding either the eGFP or the CD19CAR were transfected into the 293 producer cells. The CD19CAR contained an anti-scFv directed against CD19 and an intracellular signaling domain containing intracellular components of 4-1BB and CD3-zeta.

[0475] Human peripheral blood mononuclear cells (PBMCs) were activated with anti CD3/ anti-CD28 reagent and were transduced with CD8 retargeted NivG lentiviruses expressing CD19+CAR or GFP at various concentration ranges (10-10,000 transducing units/well). RFP+ Nalm6 leukemia cells were added to cultures on day 3, and elimination of Nalm6 cells was evaluated at 18 hours by flow cytometry.

[0476] As shown in **FIG. 4A**, CD19+ CAR expression was detected specifically in CD8+ cells with both CD8 retargeted fusogens at 4 days after transduction. Transduced CD8+ T cells expressing the CD19CAR also mediated a potent and lentivirus dose-dependent increase in

killing of CD19+ Nalm6 leukemia cells, while in contrast, cells transduced to express GFP did not exhibit target cell killing (**FIG. 4B**).

[0477] These results demonstrate that CD8-retargeted pseudotyped lentiviruses with a transgene encoding a CD19CAR deliver CD19CAR to human CD8+ T cells to mediate a specific transduction of CD8+ T cells in a complex mixture of PBMCs and showed a dose-dependent anti-tumor response by killing of leukemic cells in vitro.

[0478] The present invention is not intended to be limited in scope to the particular disclosed embodiments, which are provided, for example, to illustrate various aspects of the invention. Various modifications to the compositions and methods described will become apparent from the description and teachings herein. Such variations may be practiced without departing from the true scope and spirit of the disclosure and are intended to fall within the scope of the present disclosure.

SEQUENCES

#	SEQUENCE	ANNOTATION
1	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIIVRV YFPILTEIQQ AYIQELLPV FNNNDSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLGNVIIISLG KYLGSVNYS EGIAIGPPVF TDKVDISSQI SSMNQLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTYSRLED RRVPTSSGD LYYIGT	Nipah virus NiV-F with signal sequence (aa 1-546) Uniprot Q9IH63
2	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIIVRV YFPILTEIQQ AYIQELLPV FNNNDSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLGNVIIISLG KYLGSVNYS EGAIGPPVF TDKVDISSQI SSMNQLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTYSRLED RRVPTSSGD LYYIGT	Nipah virus NiV-F F0 (aa 27- 546)
3	ILHYEKLSKIGLVKGVTRKYKIKSNPLTKDIVIKMIPNVSNMSQCTGSVME NYKTRLNGILTPIKGALEIYKNNTHDLVGDVR	Nipah virus NiV-F F2 (aa 27- 109)
4	LAGVIMAGVAIGIATAAQITAGVALYEAMKNADNINKLKSSIESTNEAVVK LQETAECTVYVLTALQDYINTNLVPTIDKISCKQTELSLDLALSKYLSL FVFGPNLQDPVSNSMTIQAISQAFGGNYETLLRTLGYATEDFDDLLESDSI TGQIIYVDLSSYYIIIVRVYFPILTEIQQAYIQELLPVFNNNDSEWISIVP NFILVRNTLISNIEIGFCLITKRSVICNQDYATPMTNNMRECLTGSTTEKCP RELVS SHVPRFALSNGVLFANCISVTCQCQTGRAISQSGETLLMIDNT TCPTAVLGNVIIISLGKYLGSVNYSNSEGIAIGPPVFDTDKVDISSQISSMNQS LQSKDYIKEAQLLDVTNPSLISMLSMIILYVL SIASLCIGLITFISFII VEKKRNTYSRLED RRVPTSSGD LYYIGT	Nipah virus NiV F F1 (aa 110- 546)
5	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIIVRV YFPILTEIQQ AYIQELLPV FNNNDSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLGNVIIISLG KYLGSVNYS EGAIGPPVF TDKVDISSQI SSMNQLQQS	Nipah virus NiV-F F0 T234 truncation (aa 525-544)

	KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK KRNTGT	
6	LAGVIMAGVAIGIATAAQITAGVALYEAMKNADNINKLKSSIESTNEAVVK LQETAECTVYVLTALQDYINTNLVPTIDKISCKQTELSLDLALSXYLSDLL FVFGPNLQDPVSNMTIQAISQAFGGNYETLLRTLGYATEDFDDLLESDSI TGQIIYVDLSSYYIIIVRVYFPILTEIQQAYIQELLPVSNNDNSEWISIVP NFILVRNTLISNIEIGFCLITKRSVICNQDYATPMTNNMRECLTGSTEKCP RELVSSSHVPFALSNGVLFANCISVTCQCQTTGRAISQSGETLLMIDNT TCPTAVLGNVVISLGYLGSVNYNSEGIAIGPPVFTDKVDISSQISSMNQS LQQSKDYIKEAQRLDVTNPSLISMLSMIILYVLSIASLCIGLITFISFII VEKKRNTGT	Nipah virus NiV F F1 (aa 110- 546) truncation (aa 525-544)
7	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNQ THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSXYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLES SITGQIIYVD LSSYYIIIVRV YFPILTEIQQ AYIQELLPV SNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTTGRA ISQSGETLL MIDNTTCPTA VLGNVVISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLLQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK KRNTGT	Nipah virus NiV-F F0 T234 truncation (aa 525-544) AND mutation on N- linked glycosylation site
8	MVVIDLKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSXYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLES SITGQIIYVD LSSYYIIIVRV YFPILTEIQQ AYIQELLPV SNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTTGRA ISQSGETLL MIDNTTCPTA VLGNVVISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLLQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK KRNT	Truncated NiV fusion glycoprotein (FcDelta22) at cytoplasmic tail (with signal sequence)
9	MGPAENKKVR FENTTSKGGK IPSKVIKSY GTMDIKKINE GLLDKILSA FNTVIALLS IVIIVNMIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNECKFT LPPLKIHEN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVG F LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein (602 aa)

10	MGKVR FENTTSKDGK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALLGs IVIIVNMIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNshyIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLg QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ5
11	MGNTTSKDGK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALLGs IVIIVNMIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNshyIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLg QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ10
12	MGK GK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALLGs IVIIVNMIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNshyIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLg QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ15
13	MGSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALLGs IVIIVNMIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS	NiVG protein attachment glycoprotein Truncated Δ20

	IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
14	MGSYY GTMDIKKINE GLLDSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ25
15	MGTMDIKKINE GLLDSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ30
16	MKKINEGLLSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PAICAEVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated and mutated (E501 A, W504A, Q530A, E533A) NiV G protein (Gc Δ 34)
17	MATQEVRLKC LLCGIIVLVL SLEGLGILHY EKLSKIGLVK GITRKYKIKS	Hendra virus F protein

	NPLTKDIVIK MIPNVSNSVK CTGTVMENYK SRLTGILSPI KGAIELYNNN THDLVGDVKL AGVVMAGIAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDQI SCKQTEALD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SIAGQIVYVD LSSYYIIVRV YFPILTEIQQ AYVQELLPVS FNNDNSEWIS IVPNFVLIRN TLISNIEVKY CLITKKSVC NQDYATPMTA SVRECLTGST DKCPRELVS SHVPRFALSG GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCTTV VLGNIISLG KYLGSINYS ESIAVGPPVY TDKVDISSQI SSMNQSLQQS KDYIKEAQKI LDTVNPISLIS MLSMIILYVL SIAALCIGLI TFISFVIVEK KRGYSRLDD RQVRPVSNGD LYYIGT	Uniprot O89342 (with signal sequence)
18	MMADSKLVSL NNNLSGKIKD QGKVIKYYG TMDIKKINDG LLDSKILGAF NTVIALLGSI IIIVMNIMII QNYTRTTDNQ ALIKESLQSV QQQIKALTDK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL PPLKIHENI SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP RLISYTLFIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII GVGEVLDGRD KVPSPFMTNV WTPPNPSTIH HCSSTYHEDF YYTLCAVSHV GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP YGPSGIKQGD TLYFPAVGFL PRTEFQYND S NCPHCKYS KAENCRLSMG VNSKSHYILR SGLLKYNLSL GGDIIILQFIE IADNRLTIGS PSKIYNLSGQ PVFYQASYSW DTMIKLGVDV TVDPLRVQWR NNSVISRPGQ SQCPRFNVCP EVCWEGTYND AFLIDRLNWV SAGVYLSNQ TAENPVFAVF KDNEILYQVP LAEDDTNAQK TITDCFLLEN VIWCISLVEI YDTGDSVIRP KLFVAKIPAQ CSES	Hendra virus G protein Uniprot O89343
19	MNVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLGNIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQL LDTVNPISLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTGT	Nipah virus NiV-F F0 T234 truncation (aa 525-544) (with signal sequence)

20	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNQ THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNTGT	Nipah virus NiV-F F0 T234 truncation (aa 525-544) AND mutation on N- linked glycosylation site (with signal sequence)
21	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	Truncated NiV fusion glycoprotein (FcDelta22) at cytoplasmic tail (with signal sequence)
22	MKKINEGLLDKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALRS IEKGRYDKVM PYGPSGIKQG DTLYFFAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated (Gc Δ 34)
23	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRNLGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVS SHVPRFALS GVLFANCISV TCQCQTGRA ISQSGETLL MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK KRNT	Truncated mature NiV fusion glycoprotein (FcDelta22) at cytoplasmic tail

24	MSNKRRTTVLIIISYTLFYLNNAATVGFDFDKLNKIGVVQGRVLNYKIKGDP MTKDLVLKFIPNIVNITECVREPLSRYNETVRRLLLPIHNMLGLYLNNTNA KMTGLMIAGVIMGGIAIGIATAAQITAGFALYEAKKNTENIQKLTDSIMKT QDSIDKLTDSVGTSLILNLKLTQYINNQLVPLELLSCRQNKIEFDLMLTK YLVDLMTVIGPNINNPVNKDMTIQSLSLFLDGNNDIMMSELGYTPQDFLDL IESKSITGQIIYVDMENLYVVIRTYLPTLIEVPDAQIYEFNKITMSSNGGE YLSTIPNFILIRGNYSNIDVATCYMTKASVICNQDYSLPMSQNLRSYQG ETEYCPVEAVIASHSPRFALTNGVIFANCINTICRCQDNGKTITQINQFV SMIDNSTCNDVMVDKFTIKVGKYMGRKDINNINIQIGPQIIIDKVDLSNEI NKMNQSLKDSIFYLREAKRILDSVNISLISPSVQLFLIIISVLSFIILLII IVYLYCKSKHSYKYNKFIDDPDYNDYKRERINGKASKSNNIYYVGD	gb:JQ001776:61 29- 8166 Organism: Cedar virus Strain Name:CG1a Prot ein Name:fusion glycoprotein Gen e Symbol:F (with signal sequence)
25	MALNKNMFSSFLGYLLVYATTVQSSIHYSLSKVGVIKGLTYNYKIKGSP STKLMVVKLIPNIDSVKNCTQKQYDEYKNLVRKALEPVKMAIDTMLNNVKS GNNKYRFAGAIMAGVALGVATAATVTAGIALHRSNENAAIANMKSATQNT NEAVKQLQLANKQTLAVIDTIRGEINNNIIPVINQLSCDTIGLSVGIRLTQ YYSEIITAFGPALQNPVNTRITIQAISSVFNGFDELLKIMGYTSGDLYEI LHSELIRGNIIDVDVDAGYIALEIEFPNLTLVPNAVQELMPISYNIDGDE WVTLVPRFVLTRTLLSNIDTSRCTITDSSVICDNDYALPMSHELIGCLQG DTSKAREKVVSSYVPKFALSDGLVYANCLNTICRCMDTDTPIQSGLGATV SLLDNKRCSVYQVGDVLISVGSYLGDEYNADNVELGPPIVIDKIDIGNQL AGINQTLQEAEDYIEKSEEFKGVNPSIITLGSMMVLYIFMILIAIVSVIA LVLSIKLTVKGNVVRQQFTYQHVPSMENINYVSH	gb:NC_025352:5 950- 8712 Organism: Mojiang virus Strain Name:Tongguan 1 Protein Name:fusion protein Gene Symbol:F (with signal sequence)
26	MKKKTDNPTISKRGHNHSRGIKSRALLRETDNYSNGLIVENLVRNCHHPSK NNLNYTKTQKRDSTIPYRVEERKGHPKIKHLIDKSYKHIKRGKRRNGHNG NIITIIILLILILKTMSEGAHYETLSKIGLIKGITREYKVKGTPSSKDI VIKLIPNVTGLNKCTNISMENYEQLDKILIPINNIIELYANSTKSAPGNA RFAGVIIAGVALGVAAAAQITAGIALHEARQNAERINLLKDSISATNNAVA ELQEATGGIVNVITGMQDYINTNLVPQIDKLQCSQIKTALDISLSQYYSEI LTVFGPNLQNPVTTSMISQAISSQSGGNIDLLNLLGYTANDLLDLESKS ITGQITYINLEHYFMVIRVYYPIMTTISNAYVQELIKISFNVDGSEWVSLV PSYILIRNSYLSNIDISECLITKNSVICRHFAMPMSYTLKECLTGDEK PREAVVTSYVPRFAISGGVIYANCLSTTCQCYQTGKVIAQDGSQTLMMIDN QTCSIVRIEELISTGKYLGSQFYNTMHVSVGNPVFTDKLDITSQISNINQ SIEQSKFYLDKSKAILDKINLNIGSVPIISILFIIAILSLILSIITFVIM IIVRRYNYKTYPLINSDPSSRRSTIQDVYIIPNPGEHSSIRSAARSIDRDRD	gb:NC_025256:6 865- 8853 Organism: Bat Paramyxovirus Eid_hel/GH- M74a/GHA/200 9 Strain Name:BatPV/Ei d_hel/GH- M74a/GHA/200 9 Protein Name:fusion protein Gene Symbol:F (with signal sequence)
27	(GGGGGS)n wherein n is 1 to 6	Peptide Linker
28	MPAENKKVRFENTTSKDGKIPSKVIKSYSGTMDIKKINEGLLDSKILSAFN TVIALLGSIIVMNIMIIONYTRSTDNQAVIDKDALQGIQQQIKGLADKIG TEIGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTLPPL KIHCECNISCPNPLPFREYRPQTEGVSNLVGLPNNICLQKTSNQILKPKLIS YTLFPVVGSGTCTIDPLLAMDEGYFAYSHLERIGSCSRGVSKQRIIGVGEV LDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTVGDPILN STYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMPYGPSPGIK QGDITLYFPVAVGLVRTEFKYNDSNCPITKQYSKPENCRLSMGIRPNSHYI LRSGLLKYNLSDGENPKVVFIEISDQRLSIGSPSKIYDSLQGPVIFYQASFS WDTMIKFGDVLTVNPLVVNWRNNTVISRPGSQCPRFNTCPEICWEGVYND AFLIDRINWISAGVFLDSNQTAENPVFTVFKDNEILYRAQLASEDTNAQKT ITNCFLLKNKIWCISLVEIYDTGDNVIRPKLFAVKIPEQCT	gb:AF212302 Or ganism:Nipah virus Strain Name:UNKNO WN- AF212302 Protei n Name:attachmen t glycoprotein Gen e Symbol:G (Uniprot Q9IH62)

29	<p> MLSQLQKNYLDNSNQGDKMNNPDKKLSVNFNPLELDKGQKDLNKSYYVKN KNYNVSNLLNESLHDIKFCIYCIFSLIIITIIITISIVITRLKVHEEN NGMESPNLQSIQDSLSSLTNMINTETPRIGILVTATSVTLSSSINYVGTK TNQLVNELKDYITKSCGFKVPELKLHECNISCADPKISKSAMYSTNAYAE AGPPKIFCKSVSKDPDFRLKQIDYVIPVQQDRSICMNNPLLDISDGFFTYI HYEGINSCKKSDSFKVLLSHGEIVDRGDYRPSLYLLSSHYHPYSMQVINC PVTCTNQSSVFCHISNNTKTLDNSDYSSDEYYITYFNGIDRPKTKKIPIN MTADNRYIHFTFSGGGVCLGEEFIIPVTTVINTDVFTHDYCESFNCSVQT GKSLKEICSESLRSPTNSSRYNLNGIMII SQNNMTDFKIQLNGITYNKLSF GSPGRLSKTLGQVLYYQSSMSWDTYLKAGFVEKWKPFTPNWMNNTVISRP QGNCPRYHKCPEICYGGTYNDIAPLDLGKDMYVSVILDSDQLAENPEITVF NSTTILYKERVSKDELNTRSTTTSCFLFLDEPWCISVLETNRFNGKSIRPE IYSYKIPKYC </p>	<p> gb:JQ001776:81 70- 10275 Organism: Cedar virus Strain Name:CG1a Prot ein Name:attachmen t glycoprotein Gen e Symbol:G </p>
30	<p> MPQKTVEFIMNNSPLERGVSTLSDKKTLNQSKITKQGYFGLGSHSERNWKK QKNQNDHYMTVSTMILEILVVLGIMFNLIIVLTVMVYQNDNINQMAELTSN ITVLNLNLNQLTNKIQREIIPRITLIDTATTITIPSAITYILATLTTRISE LLPSINQKCEFKTPTLVLNDCRINCTPPLNPSDGVKMSSLATNLVAHGSP CRNFSVPTIYYRIPGLYNRTALDERCILNPRLTISSTKFAYVHSEYDKN CTRGFKYYELMTFGEILEGPEKEPRMFSRSFYSPNAVNYHSCPTIVTVNE GYFLCLECTSSDPLYKANLSNSTFHLVILRHNDKDEKIVSMPSFNLSTDQ EYVQIIPAEGGGTAESGNLYFPCIGRLLHKRVTHPLCKKSNCSRTDDE SCLKSYYNQGGSPQHVVNCLIRIRNAQRDNPTWDVITVDLTNTYPGSR SRIFGSFSKPMLYQSSVSWHTLLQVAEITDLDKYQLDWLDTPIYSRPGG SEC PF GNYCPTVCWEGTYNDVYSLTPNNDLFVTVYLKSEQVAENPYFA IFSRDQILKEFPLDAWISSARTTTISCFMFNNEIWCIAALEITRLNDDI IRPIYYSFWLPTDCRTPYPHTGKMTRVPLRSTYNY </p>	<p> gb:NC_025256:9 117- 11015 Organism: Bat Paramyxovirus Eid_hel/GH- M74a/GHA/200 9 Strain Name:BatPV/Ei d_hel/GH- M74a/GHA/200 9 Protein Name:glycoprote in Gene Symbol:G </p>
31	<p> MATNRDNTITSAEVSQEDKVKKYYGVETAEKVADSISGNKVFILMNTLLIL TGAIITITLNLTAAKSQQNMLKI IQDDVNAKLEMVNLQDLVKGEIKP KVSLINTAVSVSIPGQISNLQTKFLQKYVYLEESITKQCTCNPLSGIFPTS GPTYPTDKPDDDTTDDDKVDTTIKPIEYKPDGCRNTGDHFTMEPGANFY TVPNLGPASSNSDECYTNPSFSIGSSIYMF SQEIRKTDCTAGEILSIQIVL GRIVDKGQQGPQASPLLVWAVPNPKI INSCAVAAGDEMGWVLC SVTLTAAS GEPIPHMFDGFWLYKLEPDTEVVSRYITGYAYLLDKQYDSVFIGKGGGIQK GNDLYFQMYGLSRNRQSFKALCEHGSCLGTGGGGYQVLC DRAVMSFGSEES LITNAYLKVNDLASGKPVIIIGQTFPPSDSYKGSNGRMYTIGDKYGLYLAPS SWNRYLRFGITPDISVRSTTWLKSQDPIMKILSTCTNTDRDMCPEICNTRG YQDIFPLSEDSEYYTYIGITPNNGGTKNFVAVRSDGHIASIDILQNYYSI TSATISCFMYKDEIWCIAITEGKKQKDNQPQRIYAHSYKIRQMCYNMKSATV TVGNAKNITIRRY </p>	<p> gb:NC_025352:8 716- 11257 Organism: Mojiang virus Strain Name:Tongguan 1 Protein Name:attachmen t glycoprotein Gen e Symbol:G </p>
32	<p> FNTVIALLGSIIVIMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALARS IEKGRYDKVM PYGPSGIKQG DTLYFFAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPN SHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL G QPVFYQASF WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN </p>	<p> NivG protein attachment glycoprotein Without cytoplasmic tail Uniprot Q9IH62 </p>

	QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
33	FNTVIALLSGI IIIVMNIMI QNYTRTTDNQ ALIKESLQSV QQQIKALTDK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL PPLKIHCECN SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP RLISYTLPIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII GVGEVLDRGD KVPSMFMTNV WTPPNPSTIH HCSSTYHEDF YYTLCVSHV GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERKGYDKVMP YGPSGIKQGD TLYFPAVGFL PRTEFQYND NCPPIHCKYS KAENCRLSMG VNSKSHYILR SGLLKYNLSL GGDIILOFIE IADNRLTIGS PSKIYNSLGQ PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRFNVCP EVCWEGTYND AFLIDRLNWV SAGVYLSNQ TAENPVFAVF KDNEILYQVP LAEDDTNAQK TITDCFLLEN VIWCISLVEI YDTGDSVIRP KLFPAVKIPAQ CSES	Hendra virus G protein Uniprot O89343 Without cytoplasmic tail
34	MVVILDKRCY CNLLILILMI SECSVG	signal sequence
35	MKVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLDLSKILSA FNTVIALLSGI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHCECN ISCPNPLPFR EYRPQTEGVN NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVG LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated Δ5
36	MNTTSDKDGK IPSKVIKSY GTMDIKKINE GLDLSKILSA FNTVIALLSGI IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHCECN ISCPNPLPFR EYRPQTEGVN NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVG LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated Δ10

37	MKGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQOIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated Δ15
38	MSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQOIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated Δ20
39	MSYY GTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQOIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPN SHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated Δ25
40	MTMDIKKINE GLLDSKILSA FNTVIALLG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQOIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM	NiVG protein attachment glycoprotein Truncated Δ30

	PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	
41	GGGGGS	Peptide linker
42	(GGGGGS) _n wherein n is 1 to 10	Peptide linker
43	GGGGGS	Peptide linker
44	PAENKKVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPIINSTY WSGSLMMTRL AVKPKSNGGG YNQHQALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein (602 aa) Without N- terminal methionine
45	KVR FENTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPIINSTY WSGSLMMTRL AVKPKSNGGG YNQHQALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ5 Without N- terminal methionine
46	NTTSKDGK IPSKVIKSY GTMDIKKINE GLLDSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPIINSTY WSGSLMMTRL AVKPKSNGGG YNQHQALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN	NiVG protein attachment glycoprotein Truncated Δ10 Without N- terminal methionine

	QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	
47	KGK IPSKVIKSY GTMDIKKINE GLDLSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ15 Without N- terminal methionine
48	SKVIKSY GTMDIKKINE GLDLSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ20 Without N- terminal methionine
49	SYT GTMDIKKINE GLDLSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG protein attachment glycoprotein Truncated Δ25 Without N- terminal methionine
50	TMDIKKINE GLDLSKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN	NiVG protein attachment glycoprotein Truncated Δ30 Without N-

	VWTPPNPNTV YHCSAVYNNE FYYVLCVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	terminal methionine
51	KKINEGLLDKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW RNNTVISRPG QSQCPRFNTC PAICAEGVYN DAFLIDRINW ISAGVFLDSN ATAANPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG protein attachment glycoprotein Truncated and mutated (E501 A, W504A, Q530A, E533A) NiV G protein (Gc Δ 34) Without N- terminal methionine
52	MADSKLVSL NNNLSGKIKD QGKVIKNYYG TMDIKKINDG LLDSKILGAF NTVIALLSGI IIIIVMNIMII QNYTRTTDNQ ALIKESLQSV QQIKALTDK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL PPLKHECN SCNPPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP RLISYTLPIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII GVGEVLD RG KVP SMFMTNV WTPPNPSTIH HCSSTYHEDF YYTLCAVSHV GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP YGPSGIKQGD TLYFPAVGFL PRTEFQYND NCP IIHCKYS KAENCRLSMG VNSKSHYILR SGLLKYNLSL GGDIIILQFIE IADNRLTIGS PSKIYNLSGQ PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRFNVCP EVCWEGTYND AFLIDRLNWV SAGVYLSNQ TAENPVFAVF KDNEILYQVP LAEDDTNAQK TITDCFLEN VIWCISLVEI YDTGDSVIRP KLF AVKIP AQ CSES	Hendra virus G protein Uniprot O89343 Without N-terminal methionine
53	KKINEGLLDKILSA FNTVIALLS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWTPPNPNTV YHCSAVYNNE FYYVLCVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVF I EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW	NiVG protein attachment glycoprotein Truncated (Gc Δ 34) Without N- terminal methionine

	RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	
54	LSQLQKNYLDNSNQGDKNMNPDKKLSVNFNPLELDKGQKDLNKSYYVKNK NYNVSNLLNESLHDIKFCIYCIFSLLIIITIIITISIVITRLKVHEENN GMESPNLQSIQDSLSSLTNMINTEITPRIGILVTATSVTLSSSINYVGTKT NQLVNELKDYITKSCGFKVPELKLHECNISCADPKISKSAMYSTNAYELA GPPKIFCKSVSKDPDFRLKQIDYVIPVQQDRSICMNNPLLDISDGFFTYIH YEGINSCKKSDSFKVLLSHGEIVDRGDYRPSLYLLSSHYHPYSMQVINCVP VTCNQSSSFVCHISNNTKTLDNSDYSSDEYYITYFNGIDRPKTKKIPINNM TADNRYIHFTFSGGGGVCLGEEFIIPVTTVINTDVFTHDYCESFNCSVQTG KSLKEICSESLRSPNTSSRYNLNGIMIISQNNMTDFKIQLNGITYNKLSTFG SPGRLSKTLGQVLYYQSSMSWDTYLKAGFVEKWKPFPTNWMNNTVISRPNQ GNCPRYHKCEICYGGTYNDIAPLDLGKDMYVSVILDSQLAENPEITVFN STTILYKERVSKDELNTRSTTTSCFLFLDEPWCISVLETNRNFNGKSIRPEI YSYKIPKYC	gb:JQ001776:81 70- 10275 Organism: Cedar virus Strain Name:CG1a Prot ein Name:attachmen t glycoprotein Gen e Symbol:G Without N- terminal methionine
55	PQKTVEFINMNSPLERGVSTLSDKKTLNQSKITKQGYFGLGSHSERNWKKQ KNQNDHYMTVSTMLEILVVLGIMFNLIIVLTMVYYQNDNINQMAELTSNI TVLNLNLNLQTLNKIQREIIPRITLIDTATTITIPSAITYILATLTTRISEL LPSINQKCEFKTPTLVNDCRINCTPPLNPSDGVKMSSLATNLVAHGPPSPC RNFSSVPTIYYRIPGLYNRTALDERCILNPRLTISSTKFAYVHSEYDKNC TRGFKYELMTFGEILEGPEKEPRMFSRSFYSPNAVNYHSCTPIVTVNEG YFLCLECTSSDPLYKANLSNSTFHLVILRHKNDEKIVSMPSFNLSTDQEYV QIIPAEGGGTAESGNLYFPCIGRLLHKRVTHPLCKKSNCSTDDDESCLSY YNQGSPPQHVVNCLIRIRNAQRDNPTWDVITVDLTNTYPGSRSRIFGSFSK PMLYQSSVSWHTLLQVAEITDLDKYQLDWDTPYISRPGGSECPFGNYCPT VCWEGTYNDVYSLTPNNDLFVTYVLKSEQVAENPYFAIFSRDQILKEFPLD AWISSARTTTISCFMFNNEIWCIAALEITRLNDDIIRPIYYSFWLPTDCRT PYPHTGKMTRVPLRSTYNY	gb:NC_025256:9 117- 11015 Organism: Bat Paramyxovirus Eid_hel/GH- M74a/GHA/200 9 Strain Name:BatPV/Ei d_hel/GH- M74a/GHA/200 9 Protein Name:glycoprote in Gene Symbol:G Without N- terminal methionine
56	ATNRDNTITSAEVSQEDKVKKYYGVETAEKVADSISGNKVFI LMNTLLILT GAIITITLNLITNLTAAKSQQNMLKIIQDDVNAKLEMFVNLDQLVKGEIKPK VSLINTAVSVSIPGQISNLQTKFLQKYVYLEESITKQCTCNPLSGIFPTSG PTYPTDPKDDDDTTDDDKVDTTIKPIEYKPKDGCNRTGDHFTMEPGANFYT VPNLGPASSNSDECYTNPSFSIGSSIMFYSQEI RKTDC TAGEILSIQIVLG RIVDKGQQGPQASPLLVWAVPNPKIINSCAVAAGDEMGWVLC SVTLTAASG EPIPHMFDGFWLYKLEPDTEVVSRYITGYAYLLDKQYDSVFIGGGGIQKG NDLYFQMYGLSRNRQSFKALCEHGSLGTGGGGYQVLC DRAVMSFGSEESL ITNAYLKVNDLASGKPVIIIGQTFPPSDSYKGSNGRMYTIGDKYGLYLAPSS WNRYLRFGITPDISVRSTTWLKSQDPIMKILSTCTNTDRDMCPEICNTRGY QDIFPLSEDEYYTYIGITPNNGGTKNFVAVRDSGDGHASIDILQNYYSIT SATISCFMYKDEIWCIAITEGKKQKDNPQRIYAHSYKIRQMCYNMKSATVT VGNAKNITIRRY	gb:NC_025352:8 716- 11257 Organism: Mojiang virus Strain Name:Tongguan 1 Protein Name:attachmen t glycoprotein Gen e Symbol:G Without N- terminal methionine
57	DFDKLNKIGVVQGRVLNYKIKGDPMTKDLVLKFIPNIVNITECVREPLSRY NETVRRLLLP IHNMLGLYLNNNTAKMTGLMIAGVIMGGIAIGIATAAQITA GFALYEAKKNTENIQKLTD SIMKTQDSIDKLTD SVGTSILILNKLQTYINN QLVPLELLSCRQNKIEFDLMLTKYLVLDLMTVIGPNINNPVNKDMTIQSLS	gb:JQ001776:61 29- 8166 Organism: Cedar

	LLFDGNYDIMMSELGYTPQDFLDLIESKSITGQIIYVDMENLYVVIRTYLP TLIEVPDAQIYEFNKITMSSNGGEYLSTIPNFILIRGNYSNIDVATCYMT KASVICNQDYSPLMSQNLRSYQGETEYCPVEAVIASHSPRFALTNGVIFA NCINTICRCQDNGKTTITQINQFVSMIDNSTCNDVMVDKFTIKVGKYMGRK DINNINIQIGPQIIIDKVDLSNEINKMNQSLKDSIFYLREAKRILDSVNIS LISPSVQLFLIIISVLSFIILLIIIVYLYCKSKHSYKYNKFIDDPDYNDY KRERINGKASKSNNIYYVGD	virus Strain Name:CG1a Prot ein Name:fusion glycoprotein Gen e Symbol:F (without signal sequence)
58	SRALLRETDNYSNGLIVENLVRNCHHPSKNNLNYTKTQKRDSTIPYRVEER KGHPKIKHLIDKSYKHIKRGKRRNGHNGNIITIIILLILILKTQMSEGAI HYETLSKIGLIKITREYKVGKTPSSKDVIKLI PNVTGLNKCTNISMENY KEQLDKILIPINNI IELYNSTKSAPGNARFAGVIIAGVALGVAAAQITA GIALHEARQNAERINLLKDSISATNNAVAELQEATGGIVNVITGMQDYINT NLVPQIDKLQCSQIKTALDISLSQYYSEILTVFGPNLQNPVTTSMISQAI QSFGGNIDLLLNLGYTANDLLDLESKSITGQITYINLEHYFMVIRVYYP IMTTISNAYVQELIKISFNVDGSEWVSLVPSYILIRNSYLSNIDISECLIT KNSVICRHDFAFMPMSYTLKECLTGDTEKCPREAVVTSYVPRFAISGGVIYA NCLSTTCQCYQTGKVIAQDGSQTLMMIDNQTCISIVRIEELISTGKYLGSQ EYNTMHVSVGNPVTDFDKLITSQISNINQSIEQSKFYLDKSKAILDKINLN LIGSVPI SILFIIAILSLILSIITFVIVMIIVRRYNYKTYPLINSDPSSRRS TIQDVYIIPNPGEHSIRSAARSIDRDRD	gb:NC_025256:6 865- 8853 Organism: Bat Paramyxovirus Eid_hel/GH- M74a/GHA/200 9 Strain Name:BatPV/Ei d_hel/GH- M74a/GHA/200 9 Protein Name:fusion protein Gene Symbol:F (without signal sequence)
59	ILHY EKLSKIGLVK GITRKYKIKS NPLTKDIVIK MIPNVSNVSK CTGTVMENYK SRLTGILSPI KGAIELYNIN THDLVGDKL AGVVMAGIAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDQI SCKQTEALD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SIAGQIVYVD LSSYYIIIVRV YFPILTEIQQ AYVQELLPVS FNNDNSEWIS IVPNFVLIRN TLISNIEVKY CLITKKSVC NQDYATPMTA SVRECLTGST DKCPRELVS SHVPRFALSG GVLFANCISV TCQCQTGRA ISQSGEQTL MIDNTTCTTV VLGNIISLG KYLGSINYN ESIAVGPPVY TDKVDISSQI SSMNQSLOQS KDYIKEAQKI LDTVNP SLIS MLSMIILYVL SIAALCIGLI TFISFVIVEK KRGNYSRLDD RQVRPVSNGD LYYIGT	Hendra virus F protein Uniprot O89342 (without signal sequence)
60	IHYDSLKVGVIKGLTYNYKIKGSPSTKLMVVKLIPNIDSVKNCTQKQYDE YKNLVRKALEPVKMAIDTMLNNVKSNGNRYRFAAIMAGVALGVATAATVT AGIALHRSNENAQAIANMKSATQNTNEAVKQLQLANKQTLAVIDTIRGEIN NNIIPVINQLSCDTIGLSVGIRLTQYYSEIITAFGPALQNPVNTRITIQAI SSVFNGNFDELLKIMGYTSGDLYEILHSELIRGNIIDVDVDAGYIALEIEF PNLTLVPNAVQELMPI SYNIDGDEWVTLVPRFVLTRTTLLSNIDTSRCTI TDSSVICDNDYALPMSHELIGCLQGDTSKCAREKVVS SYVPKFALSDGLVY ANCLNTICRCMDTDPISQSLGATVSLLDNKRCSVYQVGDLISVGSYLG GEYNADNVELGPPIVIDKIDIGNQLAGINQTLQEAEDYIEKSEEFKGVNP SIITLGS MVVLYIFMILIAIVSVIALVLSIKLTVKGNVVRQQTYYTQHVP MENINYVSH	gb:NC_025352:5 950- 8712 Organism: Mojiang virus Strain Name:Tongguan 1 Protein Name:fusion protein Gene Symbol:F

		(without signal sequence)
61	MLFNLRLILLNNAAFRNGHNFVRNFRCGQPLQNKVQLKGRDLLTL KNFTGEEIKYMLWLSADLKFRKQKGEYLP LLQGKSLGMIFEKRSTR TRLSTETGFALLGGHPCFLT TQDIHLGVNESLTD TARVLSSMADAVL ARVYKQSDLDTLAKEASIPINGLSDLYHPIQILADYLT LQEHYSSLK GLT LSWIGDGNNILHSIMMSAAKFGMHLQAATPKGYEPDASVTKL AEQYAKENGTKLLLTNDPLEAAHGGNVLTDTWISMGQEEEEKKR LQAFQGYQVTMKTAKVAASDWTFLHCLPRKP EVDDEVFYSRSL VFPEAENRKWTIMAVMVSLLDYSPQLQKPKF	OTC
62	MTRILTA FKVVRTLKTGF GFTNVTAHQKWKFSRPGIRLLSVKAQTA HIVLEDGTKMKGYSFGHPSSVAGEVVFNTGLGGYPEAITDPAYKGQ ILTMANPIIGNGGAPDTTALDELGLSKYLESNGIKVSGLLVLDYSKD YNHWLATKSLGQWLQEEKVPAIYGVDTRMLTKIIRDKGTMLGKIEF EGQPVDFVDPNKQNLIAEVSTKDVKVYGGKGNPTKVVAVDCGIKNN VIRLLVKRGAEVHLVPWNHDFTKMEYDGILIAAGPGNPALAEPLIQ NVRKILESDRKEPLFGISTGNLITGLAAGAKTYKMSMANRGQNQPV LNITNKQAFITAQNHGYALDNTLPAGWKPLFVNVNDQTNNEGIMHES KPFFAVQFHPEVTPGPIDTEYLFDSFFSLIKKGKATTITSVLPKPALVA SRVEVSKVLILGSGLSIGQAGEFDYSGSQAVKAMKEENVKTVLMN PNIASVQTNEVGLKQADTVYFLPITPQFVTEVIKAEQPDGLILGMGG QTALNCGVELFKRGVLKEYGVKVLGTSVESIMATEDRQLFSDKLE INEKIAPSF AVESIEDALKAADTIGYPVMIRSAYALGGLSGICPNRE TLMDLSTKAFAMTNQILVEKSVTGWKEIEYEVVRDADDNCVTVCN MENVDAMGVHTGDSVVVAPAQTL SNAEFQMLRRTSINVVRHLGIV GECNIQFALHPTSMEYCIIEVNARLSRSSALASKATGYPLAFIAAKIA LGIPLPEIKNVVSGKTSACFEP SLDYMVTKIPRWDLDRFHGTSSRIGS SMKSVGEVMAIGRTFEESFQKALRMCHPSIEGFTPR LPMNKEWPSN LDLRKELSEPSSTRIYAIKAIDDNMSLDEIEKLT YIDKWFLYKMRDI LNMEKTLKGLNSESMT EETLKRAKEIGFSDKQISKCLGLTEAQTR RLKKNIHVPVKQIDTLAAEYPSVTNYLYVTYNGQEHDVNFDDHGM MVLGCGPYHIGSSVEFDWCAVSSIRTLRQLGKKT VVVCNPNPETVST DFDECDKLYFEELS LERILDIYHQEACGGCIISVGGQIPNNLAVPLYK NGVKIMGTSPLQIDRAEDRSIFSAVLDELKVAQAPWKA VNTLNEAL EFAKSVDYPCLLRPSYVLSGSAMNVVFSEDEMKKFLEEATRVSQEH PVVLTKFVEGAREVEMDAVGKDGRVISHAISEHVEDAGVHSGDAT LMLPTQTISQGAIEKVKDATR KIAKAF AISGPFNVQFLVKGNDVLVI ECNLRASRSFPFVSKTLGVDFIDVATKVMIGENVDEKHLPTLDHPIIP ADYVAIKAPMF SWPRLRDADPILRCEMASTGEVACFGEGIH TAFLK AMLSTGFKIPQKGILIGIQQSFRPRFLGVAEQLHNEGFKLFATEATSD WLNANNVPATPVAWPSQEGQNPSLSSIRKLIRDG SIDLVINLPNNNT KFVHDNYVIRRTAVDSGIPLL TNFQVTKLFAEAVQKSRKVDSKSLF HYRQYSAGKAA	CPS1
63	MATALMAVVLRAAAVAPRLRGRGGTGGARRLSCGARRRAARGTS PGRRLSTAWSQPQPPPEEYAGADDVSQSPVAE EPSWVPSRPPVPHE SPEPPSGRSLVQRDIQAFLNQC GASPGEARHWLTQFQTCHHSADKPF AVIEVDEEVLKCQQGVSSLAFLALFLQRM DMKPLVVLGLPAPTAPS GCLSFWEAKAQLAKSCKVLVDALRHNA AAAVPFFGGGSVLRAAEP APHASYGGIVSVETDLLQWCLESGSIPILCPIGETAARRSVLLDSLEV TASLAKALRPTKII FLNNTGGLRDSSHKVLSNVNLPADLDLVCNAE WVSTKERQQMRLIVDVL SRLPHHSSAVITAASTLL TELFSNKGSGTL FKNAERMLRV RSLDKLDQGR LVDLVNASFGKKLRDDYLASLRPRL HSIYVSEGYNAAILTMEPVLG GTPYLDKFVVS SRQGQSGQMLW	NAGS

	ECLRRDLQTLFWRSRVTPINPWYFKHSDGSFSNKQWIFFWFGLAD IRDSYELVNHAAGLPDSFHKPASDPGS	
64	MAVAIAAARVWRLNRGLSQAALLLLRQPGARGLARSHPPRQQQF SSLDDKPQFPGASAEFIDKLEFIQPNVISGIPIYRVMDRQGGIINPSEDP HLPKEKVLKLYKSMTLLNTMDRILYESQRQGRISFYMTNYGEEGTH VGSAAALDNTDLVFGQYREAGVLMYRDYPLELFMAQCYGNISDLG KGRQMPVHYGCKERHFVTISSPLATQIPQAVGAAYAANKRANANRV VICYFGEAASEGDAHAGFNFAATLECPHFFCRNNGYAISTPTSEQY RGDGIAARGPGYGIMSIRVDGNDVFAVYNATKEARRRAVAENQPF LIEAMTYRIGHHSTSDDSSAYRSVDEVNYWDKQDHPISRLRHILLS QGWWDEEQEKAWRKQSRKVMFAFEQAERKPKPNPNNLLFSDVYQ EMPAQLRKQQESLARHLQTYGEHYPLDHFDDK	BCKDHA
65	MAVVAAAAGWLLRLRAAGAEGHWRLPGAGLARGFLHPAATVE DAAQRRQVAHFTFQPDPEPREYQGTQKMNLFQSVTSALDNSLAKD PTAVIFGEDVAFGGVFRCTVGLRDKYGKDRVFNTPLCEQGIVGFGIG IAVTGATAIAEIQFADYIFPAFDQIVNEAAKYRYSGLDFNCGSLTIR SPWGCYVGHGALYHSQSPEAFFAHCPGKIVVIPRSPFQAKGLLLSCIE DKNPCIFFEPKILYRAAAEEVPIEPYNIPLSQAQEVIEGSDVTLVAVG TQVHVIREVASMAKEKLGVSCEVIDLRTIIPWDVDTICKSVIKTGRL ISHEAPLTGGFASEISSTVQEECFNLNLEAPISRVCGYDTPFPHIFEPFYI PDKWKCYDALRKMINY	BCKDHB
66	MAAVRMLRTWSRNAGKLCVRYFQTCGNVHVLKPNYVCFGGYPSF KYSHPHHFLKTTAALRGQVQVQKLSDIGEGIREVTVKEWYVKEGDT VSQFDSICEVQSDKASVTITSRYDGVKIKLYNLDDIAYVGKPLVDI ETEALKDSEEDVVETPAVSHDEHTEHKEIKGRKTLATPAVRRLAMEN NIKLEVVGSKGDKGRILKEDILNYLEKQTGAILPPSPKVEIMPPPPKP KDMTVPILVSKPPVFTGKDKTEPIKGFQKAMVKTMSAALKIPHFGY CDEIDLTELVLKREELKPIAFARGIKLSFMPFFLKAASLGLLQFPILNA SVDENCQNITYKASHNIGIAMDTEQGLIVPNVKNVQICSIFDIATELN RLQKLGSGVQLSTTDLTGGTFTLSNIGSIGGTFAKPVIMPPEVAIGAL GSIKAIPRFNQKGEVYKAQIMNVSWADHRVIDGATMSRFSNLWKS YLENPAFMLLDLK	DBT
67	MQSWSRVYCSLAKRGHFNRIHGLQGLSAVPLRTYADQPIDADVT IGSGPGGYVAAIKAAQLGFKTVCKNETLGGTCLNVGCIPSKALLN NSHYHMAHGKDFASRGIESEVRLNLDKMMQKSTAVKALTGGI AHLFKQNKVVHVNGYGKITGKNQVTATKADGGTQVIDTKNLIATG SEVTPFPGITIDEDTIVSSTGALSLLKVPKEMVIGAGVIGVELGSVW QRLGADVTAVEFLGHVGGVGIDMEISKNFQRIKQKQGFKFLNTKV TGATKKSDBGKIDVSIEAASGGKAEVITCDVLLVCIGRRPFTKNLGL ELGIELDPRGRIPVNTFRQTKIPNIYAIGDVVAGPMLAHKAEDGEIIC VEGMAGGAVHIDYNCVPSVIYTHPEVAWVGKSEEQLKEEGIEYKV GKFPFAANSRAKTNADTDGMVKILGQKSTDRVLGAHILGPGAGEM VNEAALALEYGASCEDIARVCHAHPTLSEAFREANLAASFSGKSINF	DLD
68	MLRAKNQLFLLSPHYLRQVKESGSRLLIQRLHQQQLHPPEWAAL AKKQLKGKNPEDLIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTR GPYPTMYTFRPWTIRQYAGFSTVEESNKFYKDNIAKQQGLSVAFD LATHRGYDSNPRVRGDVGMAGVAIDTVEDTKILFDGIPLEKMSVS MTMNGAVIPVLANFIVTGEEQGVPEKLTGTIQNDILKEFMVRNTYI FPPEPSMKIADIFEYTAHMPKFNSISISGYHMQEAGADAILELAYT LADGLEYSRTGLQAGLTIDEFAPRLSFFWGIGMNFYMEIAKMRAGR RLWAHLIEKMFQPKNSKSLLLRAHCQTSWLSLTEQDPYNNIVRTAI EAMAAVFGGTQSLHTNSFDEALGLPTVKSARIARNTQIIIQEEGIPK	MUT

	VADPWGGSYMMECTNDVYDAALKLINEIEMGGMAKAVAEGIP KLRIECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVLAINTSVR NRQIEKLKKIKSSRDQALAECLAALTECAASGDGNILALAVDASR ARCTVGEITDALKKVFGEHKANDRMVSGAYRQEFGESKEITSAIKR VHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGFADLGFVDIG PLFQTPREVAQQAVDADVHAVGISTLAAGHKTLVPELIKELNSLGRP DILVMCGGVIPPQDYEFLEFVGVSNNVFGPGTRIPKAAVQVLDDIEKC LEKKQQSV	
69	MPMLLPHPHQHFLKGLLRAPFRCYHFIFHSSTHLGSGIPCAQPFNSL GLHCTKWMLLSDDLKRLCVQTTLKDHTGLSDKEQRFVDKLYTG LIQGGQACLAEAITLVESTHSRKKELAQVLLQKVLLYHREQECSNK GKPLAFRVGLSGPPGAGKSTFIEYFGKMLTERGHKLSVLAVDPSSCT SGGSLLGDKTRMTELSRDMNAYIRPSPTRGTGGVTRTTNEAILLCE GAGYDIILIEVGVGQSEFAVADMVDMFVLLPPAGGDELQGIKRG IEMADLVAVTKSDGDLIVPARRIQAEYVSALKLLRKRQVWPKVI RISARSGEGISEMWDKMKDFQDLMLASGELTAKRRKQKQVWMWN LIQSVLEHFRTHPTVREQIPLEQKVLIGALSPGLAADFLLKAFKSR D	MMAA
70	MAVCGLSRRLGLSRLGLRGCFGAARLLYPRFQSRGPQGVEDGDR PQPSSKTPRIPIYTKTGDKGFSSTFTGERPKDDQVFEAVGTTDELS SAIGFALELVTEKGHTFAEELQKIQCTLQDVGSALATPCSSAREHL KYTTFKAGPILELEQWIDKYTSQPLPLTAFILPSGGKISSALHFCRAV CRRRAERRVPLVQMGETDANVAKFLNRLSDYLFLLARYAAMKEG NQEKIYMKNDPSAESEGL	MMAB
71	MFDRALKPFLQSCHLRMLTDPVDQCVAYHLGRVRESLPELQIEIAD YEVHPNRRPKILAQTAHVAGAAAYYQQRQDVEADPWGNQRISGVC IHPRFGGWFAIRGVVLLPGIEVPDLPPRKPHDCVPTRADRIALLEGFN FHWRDWTYRDAVTPQERYSEEQKAYFSTPPAQRALLGLAQPSKPK SSPSPDLPTTPAPKPGNPSRARSWLSPRVSPASPSP	MMACHC
72	MANVLCNRRARLVSYLPGFCSLVKRVVNPKAFTAGSSGSDESHVA AAPPDICSRVWPDETMGPFPGPDQRFQLPGNIGFDCHLNGTASQK KSLVHKTLPDVLAEPLESSERHEFVMAQYVNEFQGNDAPEVEQINSA ETYFESARVECAIQTCPELLRKDFESLFPEVANGKLMILTVTQKTKN DMTVWSEEVEIEREVLLKFFINGAKEICYALRAEGYWADFIDPSSGL AFFGPYTNNTLFETDERYRHLGFSVDDLGCCKVIRHSLWGTHVVVG SIFTNATPDSSHIMKKLSGN	MMADHC
73	MARVLKAAAANAVGLFSRLQAPIPTVRASSTSQPLDQVTGSVWNL GRLNHVAIAVPDLEKAAAFYKNILGAQVSEAVPLPEHGVSVVFNVL GNTKMELLHPLGRDSPIAGFLQKNKAGGMHHICIEVDNINAAMVMDL KKKKIRSLSEEVKIGAHGKPVIFLHPKDCGGVLVELEQA	MCEE
74	MAGFWVGTAPLVAAAGRRGRWPPQQMLLSAALRTLKHVLYYSRQC LMVSRNLGSGVYDPNEKTFDKILVANRGEIACRVIRTCKKMGIKTV AIHSDVDASSVHVKMADEAVCVGPAPTSKSYLNMDAIMEAIKKTR AQAVHPGYGFLSENKEFARCLAAEDVVFIGPDTHAIQAMGDKIESK LLAKKAEVNTIPGFDGVVKDAEEAVRIAREIGYPVMIKASAGGGGK GMRIAWDDEETRDGFRLSSQEAASSFGDDRLLIEKFIDNPRHIEIQVL GDKHGNAWLNERECIQRRNQKVVEEAPSIFLDAETRRAMGEQA VALARAVKYSSAGTVEFLVDSKKNFYFLEMNTRLQVEHPVTECITG LDLVQEMIRVAKGYPLRHKQADIRINGWAVECRVYAEOPYKSFGPL SIGRLSQYQEPLHPLPGVRVDSGIQPGSDISIYDPMISKLITYGSDRTE ALKRMADALDNYVIRGVTHNIALREVIINSRFBKGDISTKFLSDVY PDGFKGHMLTKSEKNQLLAASSLFVAFQLRAQHFQENSRRMPVIKP	PCCA

	DIANWELSVKLHDKVHTVVASNNGSVFSVEVDGSKLNVSTWNLA SPLLSVSVDDGTQRTVQCLSREAGGNMSIQFLGTVYKVNILTRLAAEL NKFMLEKVTEDTSSVLRSPMPGVVAVSVKPGDAVAEGQEICVIEA MKMQNSMTAGKTGTVKSVCQAGDTVGEGLLVELE	
75	MAAALRVAAVGARLSVLASGLRAAVRSLCSQATSVNERIENKRRT ALLGGGQRRIDAQHKRGKLTARERISLLDPSGFSVESDMFVEHRCA DFGMAADKNKFPGDSVVTGRGRINGRLVYVFSQDFTVFGGSLSGA HAQKICKIMDQAITVGAPVIGLNDSSGARIQEGVESLAGYADIFLRN VTASGVIPQISLIMGPCAGGAVYSPALTDFTFMVKDTSYLFITGPDV VKSVTNEDVTQEELGGAKTHTTMSGVAHRAFENDVDALCNLRDFF NYLPLSSQDPAPVRECHDPSDRLVPELDTIVPLESTKAYNMVDIIHSV VDEREFFEIMPNYAKNIIVGFARMNGRTVGIVGNQPKVASGCLDINS SVKGARFVRFCDAFNIPLITFVDVPGFLPGTAQEYGGIIRHGAKLLY AFAEATVPKVTVITRKAYGGAYDVMSSKHLCDTNYAWPTAEIAV MGAKGAVEIIFKGHENVEAAQAEYIEKFANFPFAAVRGFVDDIIQPS STRARICCDLDVLASKKVQRPWRKHANIPL	PCCB
76	MAVESQGGRLVGLLLCVLGPVVSAGKILLIPVDGSHWLSMLGA IQQLQQRGHVIVLAPDASLYIRDGAFTLKTYPVPFQREDVKESFV SLGHNVFENDSFLQRVIKTYKKIKKDSAMLLSGCSHLLHNKELMAS LAESSFDVMLTDPFLPCSPIVAQYLSLPTVFFLHALPCSLEFEATQCP NPFSYVPRPLSSHSDHMTFLQRVKNMLIAFSQNFLCDVVYSPYATL ASEFLQREVTVDLLSSASVWLFSDVFKDYPRPIMPNMVFVGGIN CLHQNPQSFEFEAYINASGEHGIVVFSLGSMVSEIPEKKAMAIADAL GKIPQTVLWRYTGTRPSNLANTILVKWLPQNDLLGHPMTRAFITH AGSHGVYESICNGVPMVMPLFGDQMDNAKRMETKGAGVTLNVL EMTSEDLENALKAVINDKSYKENIMRLSSLHKDRPVEPLDLAVFWV EFVMRHKGAPHLRPAADHLDWYQYHSLDVIGFLLAVVLTVAFITFK CCAYGYRKCLGKKGRVKKAHKSKTH	UGT1A1
77	MSSKGSVVLAYSGLDTSCLVWLKEQGYDVIAYLANIGQKEDFEE ARKKALKLGAKKVFIEDVSREFVEFIWPAIQSSALYEDRYLLGTSL ARPCIARKQVEIAQREGAKYVSHGATGKGNDQVRFELSCYSLAPQI KVIAPWRMPEFYNRFKGRNDLMEYAKQHGIPIPVTPKNPWSMDEN LMHISYEAGILENPKNQAPPGLYTKTQDPAKAPNTPDILEIEFKKGVP VKVTNVKDGTTHTQTSLELFMYLNEVAGKHGVGRIDIVENRFIGMKS RGIYETPAGTILYHAHLDIEAFTMDREVRKIKQGLGLKFAELVYTGf WHSPECEFVRHCAIKSQERVEGKVQSVLKGQVYILGRESPLSLYN EELVSMNVQGDYEPTDATGFININSLRLKEYHRLQSKVTAK	ASS1
78	MSTAVLENPGLGRKLSDFGQETSYIEDNCNQNGAISLIFSLKEEVGA LAKVLRLEFENDVNLTHIESRPSRLKKDEYEFFTHLDRSLPALTNII KILRHDIGATVHELSDRDKKDTVPWFPRTIQELDRFANQILSYGAEL DADHPGFKDPVYRARRKQFADIAYNRYRHGQPIPRVEYMEEEEKKTW GTVFKTLKSLYKTHACYEYNHIFPILLEKYCGFHEDNIPQLEDVVSQFL QTCTGFRLRPVAGLLSSRDFLGGLAFRVFHTCTQYIRHGSKPMYTPEP DICHELLGHVPLFSDRSFAQFSQEIGLASLGAPDEYIEKLATYWFYTV EFGLCQGDSEIKAYGAGLLSSFGELQYCLSEKPKLLPLELEKTAIQN YTVTEFQPLYYYVAESFNDAKEKVRNFAATIPRPFVRYDPYTQRIEV LDNTQQLKILADSINSEIGILCSALQKIK	PAH
79	MAKTLQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGT LVSLTNNTDILQGIQASCDYINNAVESGEPIYGVTSFGGGMANVAIS REQASELQTNLVWFLKTGAGNKLPLADVRAAMLLRANSHMRGAS GIRLELIKRMIFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDP SFKVDFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGI	PAL

	AANCVYDTQILTAIAMGVHALDIQALNGTNQSFHPFIHNSKPHPGQL WAADQMISLLANS QLVRDELDDGKHLDYRDHELIQDRYSLRCLPQYLGPIVDGISQIAKQIEI EINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDHLRYYIGLLAK HLDVQIALLASPEFSNGLPPSLLGNRERKVNMGKGLQICGNSIMPL LTFYGNISADRFPTHAEQFNQNSQGYTSATLARRSVDFQNYVAI ALMFGVQAVDLRTYKKTGHYDARASLSPATERLYSAVRHVVGQKP TSDRPYIWNDEQGLDEHIARISADIAAGGVIVQAVQDILPSLH	
80	MSTERDSETTFDEDSQPNDEVVPYSDDTEDELDQGS AVEPEQNR VNREAEENREPFRKECTWQVKANDRKYHEQPHFMNTKFLCIKESK YANNAIKTYKYNAFTFIPMNLFEQFKRAANLYFLALLILQAVPQIST LAWYTTLVPLLVLVGVTAIKDLVDDVARHKMDKEINNRTCEVIKD GRFKVAKWKEIQVGDVIRLKKNDVFPADILLSSSEPNSLCYVETAE LDGETNLKFKMSLEITDQYLQREDTLATFDGFIECEEPNNRLDKFTG TLFWRNTSFPLDADKILLRGCVRINTDFCHGLVIFAGADTKIMKNSG KTRFKRTKIDYLMNYMVYITFVVLILLSAGLAIGHAYWEAQVGNSS WYLYDGEDDTPSYRGFLIFWGYIIVLNTMVPISLYVSVEVIRLGQSH FINWDLQMYAEDKTPAKARTTTLNEQLGQIHIFSDKTGTLTQNI MTFFKCCINGQIYGDHRDASQHNHNKIEQVDFSWNTYADGKLAFY DHYLIEQIQSGKEPEVRQFFFLAVCHTVMVDRTDGQLNYQAASPD EGALVNAARNFGFAFLARTQNTITISELGTERTYNVLAIDLDFNSDRK RMSIIVRTPEGNIKLCKGADTVIYERLHRMNPTKQETQDALDIFAN ETLRTLCLCYKEIEEKEFTWNNKKFMAASVASTNRDEALDKVYEEI EKDLILLGATAIEDKLQDGVPETISKLAKADIKIWVLTGDKKETAENI GFACELLTEDTTICYGEDINSLLHARMENQRNRGGVYAKFAPPVQE SFFPPGGNRALITGSWLNEILLEKKTNRNKLKLFPRTEEEERRMRT QSKRRLEAKKEQRQKNFVDLACECSAVICCRVTPKQKAMVVDLVK RYKKAITLAIGDGANDVNMIKTAHIGVVISGQEGMQAVMSSDYSFA QFRYLQRLLLVHGRWSYIRMCKFLRYFFYKNFAFTLVHFWYSFFNG YSAQTA YEDWFITLVNLYTSLPVLLMGLLDQDVSDKLSLRFPGLY IVGQAYDLFNKYRFFVSLHGVLTSMILFFIPLGAYLQTVGQDGEAP SDYQSFAVTIASALVITVNFQIGLDTSYWTFVNAFSIFGSIALYFGIMF DFHSAGIHVLFPSAFQFTGTASNALRQPYIWLTIILAVAVCLLPVVAI RFLSMTIWPSESDKIQKHRKRLKAEQWQRRQQVFRRGVSTRRSAY AFSHQRGYADLISSGRSIRKKRSPDLAIVADGTAEYRRTGDS	ATP8B1
81	MSDSVILRSIKKFGGEENDGFESDKSYNNDKKSRLQDEKKGDGVRVG FFQLFRFSSSTDIWLMFVGSLSLCAFLHGIAQPGVLLIFGTMTDVFIDYD VELQELQIPGKACVNNTIVWTNSSLNQNMTNGTRCGLLNIESEMIKF ASYYAGIAVAVLITGYIQICFWVIAAARQIQKMRKFYFRRIMRMEIG WFDCNSVGELNTRFSDDINKINDAIADQMALFIQRMSTICGFLGFG FRGWKLTLVHSVSPLIGIGAATIGLSVSKFTDYELKAYAKAGVVAD EVISSMRTVAAFGGEKREVERYEKNLVFAQRWGIKRGIVMGFFTGF VWCLIFLCYALAFWYGSTLVLDEGEYTPGTLVQIFLSVIVGALNLGN ASPCLEAFATGRAAATSIFETIDRKPIIDCMSEDGYKLDRIKGEIEFHN VTFHYPSPREVKILNDLNMVIKPGEMTALVGPSGAGKSTALQLIQR YDPCCEGMVTVDGHDIRSLNIQWLRDQIGIVEQEPVLFSTTIAENIRYG REDATMEDIVQAAKEANAYNFIMDLPQQFDTLVGEGGGQMSGGQ KQRVAIARALIRNPKILLDMATSALDNESEAMVQEVLSKIQHGHTII SVAHRLSTVRAADTIIGFEHGTAVERGTHEELLERKGVYFTLVTLQS QGNQALNEEDIKDATEDDMLARTFSRGSYQDSLRSIRQRSKSQLS YLVHEPPLAVVDHKSTYEEDRKDKDIPVQEEVEPAPVRRILKFSAP WPYMLVGSVGA AVNGTVTPLYAFLFSQILGTF SIPDKEEQRSQINGV CLLFVAMGCVSLFTQFLQGYAFAKSGELLTKRLRKFGFRAMLGQDI	ABCB11

	<p>AWFDDLRNSPGALTTRLATDASQVQGAAGSQIGMIVNSFTNVTVA MIIAFSFSWKLSLVILCFFPFLALSGATQTRMLTGFASRDKQALEMV GQITNEALSNIRTVAGIGKERRFIEALETELEKPFKTAIQKANIYGFCF AFAQCIMFIANSASYRYGGYLISNEGLHFSYVFRVISAVVLSATALG RAFSYTPSYAKAKISAARFFQLLDRQPPISVYNTAGEKWDNFQGKID FVDCKFTYPSRPDSQVLNGLSVSISPGQTLAFVGSSGCGKSTSIIQLE RFYDPDQGKVMIDGHDSKKVNVQFLRSNIGIVSQEPVLFACSIMDNI KYGDNTKEIPMERVIAAAKQAQLHDFVMSLPEKYETNVGSQGSQLS RGEKQRIARIAIVRDPKILLDEATSALDTESEKTVQVALDKAREG RTCIVIAHRLSTIQNADIHAVMAQGVVIEKGTHEELMAQKGAYYKLV TTGSPIS</p>	
82	<p>MDLEAAKNGTAWRPTSAEGDFELGISSKQKRKKTKTVKMIGVLTFL RYSWDQDKLFMSLGTIMAIHAGSGLPLMMIVFGEMTDKFVDTAGN FSFPVNFSLSLNPGKILEEEMTRYAYYYSGLGAGVLVAAYIQVSFW TLAAGRQIRKIRQKFFHAILRQEIGWFDINDTTELNTRLTDDISKISEG IGDKVGMFFQAVATFFAGFIVGFIRGWKLTLMVMAISPILGLSAAVW AKILSAFSDKELAAAYAKAGAVAEALGAIRTVIAFGGQNKELERYQ KHLENAKEIGIKKAISANISMGIALLIYASYALAFWYGSTLVISKEY TIGNAMTVFFSILIGAFSVGQAAPCIDAFANARGAAYVIFDIIDNNPKI DSFSERGHKPDSEIKGNLEFNDVHFSYPSRANVKILKGLNLKVQSGQT VALVGSSGCGKSTTVQLIQLYDPDEGTINIDGQDIRNFNVNLYREII GVVSQEPVLFSTTIAENICYGRGNVTMDEIKKAVKEANAYEFIMKLP QKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLDEATSALDTE SEAEVQAALDKAREGRTTIVIAHRLSTVRNADVIAGFEDGVIVEQGS HSELMKKEGVYFKLVNMQTSQSIIQSEEFELNDEKAATRMAPNGW KSRLFRHSTQKNLKNQMCQKSLDVETDGLEANVPPVSFLKVLKLN KTEWPYFVVGTVCAIANGGLQPAFSVIFSEIIAIFGPGDDAVKQKQC NIFSLIFLFLGIISFFTFLLQGFTFGKAGEILTRRLRSMAFKAMLRQDM SWFDDHKNSTGALSTRLATDAAQVQGATGTRLALIAQNANLGTGII ISFIYQWQLTLLLLAVVPIHAVSGIVEMKLLAGNAKRDKKELEAAGK IATEAIENIRTVVSLTQERKFESMYVEKLYGPYRNSVQKAHIYGITFS ISQAFMYFSYAGCFRFGAYLIVNGHMRFRDVILVFSIVFGAVALGH ASSFAPDYAKAKLSAAHLFMLFERQPLIDSYSEEGLKPKDFEGNITF NEVVFNYPTRANVPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQL LERFYDPLAGTVFVDFGFQLLDGQEAKKLVQWLRAQLGIVSQEPI LFDCSIAENIAYGDNRSRVVSQDEIVSAAKAANIHPFIETLPHKYETRV GDKGTQLSGGQKQRIAIARALIRQPQILLDEATSALDTESEKVVQE ALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQLLA QKGIYFSMVSQAGTQNL</p>	ABCB4
83	<p>MPVRGDRGFPPRRELSGWLAPGMEELIWEQYTVTLQKDSKRGFGI AVSGGRDNPHFENGETSIVISDVLPGGPADGLLQENDRVVMVNGTP MEDVLHSFAVQQLRKSGKVAIVVKRPRKVQVAALQASPPLDQDD RAFEVMDEFDGRSFRSGYSERSRLNSHGGRSRWSWEDSPERGRPHR ARSERDLSRDRSRGRSLERGLDQDHARTDRSRGRSLERGLDHDF GPSRDRDRDRSRGRSIDQDYERAYHRAYPDYERAYSPEYRRGAR HDARSRGPRSRSRREHPSRSPSPPEPRGRPGPIGVLLMKSRANEYGL RLGSQIFVKEMTRTGLATKDGNLHEGDIIKINGTVTENMSLTDARK LIEKSRGKLQLVVLKDSQQTILINPSLNDSDSEIEDISEIESNRSFSPEE RRHQYSYDYHSSSEKLKERPSSREDTPSRLSRMGATPTPFKSTGDI AGTVVPETNKEPRYQEDPPAPQPKAAPRTFLRPSPEDEAIYGPNTKM VRFKKGDSVGLRLAGGNDVGIFVAGIQEGTSAEQEGLQEGDQILKV NTQDFRGLVREDAVLILLEIPKGEMVTILAQSRADVYRDILACGRG DSFFIRSHFECEKETPQSLAFTGRGEVFRVVDTLVDGKLGWNLAVRIG</p>	TJP2

	NELEKGLIPNKSRAEQMASVQNAQRDNAGDRADFWRMRGQRSGV KKNLRKSREDLTAVVSVSTKFPAYERVLLREAGFKRPVVLFGPIADI AMEKLANELPDWFQTAKEPKDAGSEKSTGVVRLNTRVQIIEQDKH ALLDVTQKAVDLLNYTQWFPVIVFFNPDSRQGVKTMQRQLNPTS SSRKLFDQANKLKKTKCAHLFTATINLNSANDSWFGSLKDTIQHQGG EAVVWVSEGMMEGMDDDPEDRMSYLTAMGADYLSCDRLISDFEDT DGEAGAYTDNELDEPAEEPLVSSITRSSEPVQHEESIRKPSPEPRAQM RRAASSDQLRDNSPPPAFKPEPPKAKTQNKESYDFSYSYKSNPS AVAGNETPGASTKGYPPPVAAPKPTFGRSILKPSTPIPPQEGEEVGESS EEQDNAPKSVLGKVKIFEKMDHKARLQRMQELQEAQNARIEIAQK HPDIYAVPIKTHKPDPTQHTSSRPPEPQKAPSRPYQDTRGSYGSD AEEEEYRQQLSEHSKRGYYGQSARYRDTL	
84	MATATRLGWRVASWRLRPPLAGFVSQRAHSLLPVDDAINGLSEE QRQLRQTMAKFLQEHLAPKAQEIDRSNEFKNLREFWKQLGNLGV GITAPVQYGGSGLYLEHVLVMEISRASGAVGLSYGAHNSLNCINQ LVRNGNEAQKEKYLKLSGEYIGALAMSEPNAGSDVVSMLKAE KKGNYHILNGNKFWITNGPDADVLIVYAKTDLAAPASRGITAFIVE KGMGPFSTSKKLDKLGMRGNTCELIFEDCKIPAANILGHENKGVY VLMSGLDLERLVLGGPLGLMQAVLDHTIPYLHVREAFGQKIGHFQ LMQGMADMYTRLMACRQYVYNVAKACDEGHCTAKDCAGVILY SAECATQVALDGIQCFGGNGYINDFPMGRFLRDAKLYEIGAGTSEV RRLVIGRAFNAFHF	IVD
85	MALRGVSVRLLSRGPGLHVLRTWVSSAAQTEKGGRTQSQLAKSSR PEFDWQDPLVLEEQLTTDEILIRDTFRITYCQERLMPRILLANRNEVF HREIISEMGELGVLGPTIKGYGCAGVSSVAYGLLARELERVDSGYRS AMSVQSSLVMHPIYAYGSEEQRQKYLPLAKGELLGCFGLTEPNSG SDPSSMETRAHYNSSNKSITLNGTKTWITNSPMADLFVWWARCED GCIRGFLEKGMRLSAPRIQKGFSLRASATGMIIMDGEVPEENVL PGASSLGGPFGLNARYGIAWGVLGASEFCLHTARQYALDRMQF GVPLARNQLIQKKLADMLTEITLGLHACLQLGRLKDQDKAAPMV SLLKRNNCGKALDIARQARDMLGGNGISDEYHVRHAMNLEAVNT YEGTHDIHALILGRAITGIQAFASK	GCDH
86	MFRAAAPGQLRRAASLLRFQSTLVIAEHANDSLAPITLNTITAATRL GGEVSVCLVAGTKCDKVAQDLCKVAGIAKVLVAQHDVYKGLLPEEL TPLILATQKQFNTHICAGASAFGKNLLPRVAAKLEVAPISDIIAISP DTFVRTIYAGNALCTVKCEKVKVFSVRGTSFDAAATSGGSASSEK ASSTSPVEISEWLDQKLTKSDRPELTGAKVVVSGGRGLKSGENFKLL YDLADQLHAAVGASRAAVDAGFVPNDMQVGQTGKIVAPELYIAV GISGAIQHLAAGMKDSKTIVAINKDPEAPIFQVADYGIVADLFKVVPE MTEILKKK	ETFA
87	MAELRVLVAVKRVIDYAVKIRVKPDRTGVVTDGVKHSMPNPFCEIA VEEAVRLKEKKLVKEVIAVSCGPAQCQETIRALAMGADRGHVEV PPAEERLGLPLQVARVLAKLAEKEKVDLVLLGKQAIDDDCNQTGQ MTAGFLDWPGQTFASQVTLEGDKLKVREIDGGLETTLRLKLPVVT ADLRLNEPRYATLPNIMKAKKKKIEVIKPGDLGVDLTSKLSVISVED PPORTAGVKVETTEDLVAKLKEIGRI	ETFB
88	MLVPLAKLSCLAYQCFHALKIKKNYLPLCATRWSSTSTVPRITHTYT IYPRDKDKRWEGVNMERFAEEADVIVGAGPAGLSAAVRLKQLAV AHEKDIRVCLVEKAAQIGAHTLSGACLDPGAFKELFPDWKEKGAPL NTPVTEDRFGLTEKYRIPVPILPGLPMNNHGNYIVRLGHLVSWMGE QAEALGVEVYPGYAAAEVLFHDDGSVKGIAATNDVGIQKDGAPKAT FERGLELHAKVTIFAEGCHGLAKQLYKKFDLRANCEPQTYGIGLK	ETFDH

	ELWVIDEKNWKPRVDHTVGVPLDRHTYGGSFYHLNEGEPLVAL GLVVGLDYQNPYLSPFREFQRWKHHPSIRPTLEGGKRIAYGARALN EGGFQSIPKLTFFPGLLIGCSPGFMNVPIKGTHTAMKSGILAAESIF NQLTSENQSKTIGLHVTEYEDNLKNSWVWKELYSVRNIRPSCHGV LGVEYGGMIYTGIFYWILRGMEPWTLKHKGSDFERLKPDKDCTPIEY PKPDGQISFDLLSSVALSGTNHEHDQPAHLTLRDDSIPVNRNLSIYDG PEQRFCPAGVYEFVPEQGDGFRLQINAQNCVHCKTCDIKDPSQNIN WVVEGGGGPAYNGM	
89	MASESGKLWGGRFVGAVDPIMEKFNASIAFYDRHLWEVDVQGSKA YSRGLEKAGLLTKAEMDQILHGLDKVAEEWAQGTFLNSNDEDIH TANERRLKELIGATAGKLHTGRSRNDQVVTDLRLWMRQTCSTLSG LLWELIRTMVDRAEAERDVLFPGYTHLQRAQPIRWSHWILSHAV TRDSERLLEVRKRINVLPLGSGAIAGNPLGVDRELLRAELNFGAITL NSMDATSERDFVAEFLFWASLCMTHLSRMAEDLILYCTKEFSFVQL SDAYSTGSSLMPPQKKNPDSLELIRSKAGRFGRCAGLLMTLKGPS TYNKDLQEDKEAVFEVSDTMSAVLQVATGVISTLQIHQENMGQAL SPDMLATDLAYYLVKGMPPFRQAHEASGKAVFMAETKGVALNQL SLQELQTISPLFSGDVICVWDYGHVSVEQYGALGGGTARSSVDWQIRQ VRALLQAQQA	ASL
90	MVGGSVPVFDEILSTARMNRVLSFHSVSGILVCQAGCVLEELSRYV EERDFIMPLDLGAKGSCHIGGNVATNAGGLRFLRYGSLHGTVLGLE VVLADGTVLDCLTSLRKDNTGYDLKQLFIGSEGTGHIITVSILCPPK PRAVNVAFGLGCPGFAEVLQTFSTCKGMLGEILSAFEMDAVCMQLV GRHLHLASPVQESPFYVLIETSGSNAGHDAEKLGHFLEHALGSGLV DGTMATDQRKVKMLWALRERITEALS RDGYVYKYDLSLPVERLYD IVTDLRARLGPAAKHVVGYGHLGDGNLHLNVTAEAFSPSLLAALEP HVYEWTAGQQGSVSAEHGVGFRKRDLVGYSKPPGALQLMQQLKA LLDPKGILNPYKTLPSQA	D2HGDH
91	MAAMRKALPRRLVGLASLRAVSTSSMGTLPKRVKIVEVGPRDGLQ NEKNIVSTPVKIKLIDMLSEAGLSVIETTSFVSPKWVPQMGDHTVL KGIQKFPGINYPVLTPLNKGFEAAVAAGAKEVVIFGAASELFTKKNI NCSIEESFQRFDAILKAAQSANISVRGYVSCALGCPYEGKISPAKVAE VTKKFYSMGCYEISLGDITGVGTPGIMKDMLSAVMQEVPLAALAV HCHDITYGQALANTLMALQMGSVVDSSVAGLGCPYAQQASGNL ATEDLVYMLEGLGIHTGVNLQKLLEAGNFICQALNRKTSSKVAQAT CKL	HMGCL
92	MAAASAVSVLLVAAERNRWHRLPSLLLPPRTWVWRQRTMKYTTA TGRNITKVLIANRGEIACRVMRTAKKLGVTAVYSEADRNSMHV DMADEAYSIGPAPSQQSYLSMEKIIQVAKTSAAQAIHPGCGFLSEN EFAELCKQEGHIFIGPPPSAIRDMGIKSTSKSIMAAAGVPVVEGYHGE DQSDQCLKEHARRIGYPVMIKAVRGGGGKGMRIVRSEQEFQEQLS ARREAKKSFNDAMLIEKFVDTPRHVEVQVFGDHHGNAVYLFERD CSVQRRHQKIIIEAPAPGIKSEVRKKLGEAAVRAAKAVNYVGAGTV EFIMDSKHNFCFMMENTRLQVEHPVTEMITGTDLVEWQLRIAAGEK IPLSQEEITLQGHAFEARIYAEDPSNNFMPVAGPLVHLSTPRADPSTR IETGVRQGDVSVHYDPMIAKLVVWAADRQAALTKLRYSLRQYNI VGLHTNIDFLLNLSGHPEFEAGNVHTDFIPQHHKQLLSRKAAAKES LCQAALGLILKEKAMTDTFTLQAHDQFSPFSSSSGRRLNISYTRNMT LKDGKNNVAIAVTYNHDGSYSMQIEDKTFQVLGNLYSEGDCYTLK CSVNGVASKAKLIILENTIYLFSEKGSIEIDIPVPKYLSSVSSQETQGG PLAPMTGTIEKVFKAGDKVKAGDSLMMVMIAMKMEHTIKSPKDGT VKKVFYREGAQANRHTPLVEFEESDRESE	MCCC1

93	MWAVLRLALRPCARASPAGPRAYHGDSVASLGTQPDLGSALEYQEN YKQMKALVNQLHERVEHIKLGGEKARALHISRGKLLPRERIDNLI DPGSPFLELSQFAGYQLYDNEEVPGGGIITGIGRVSGVECMIIANDAT VKGGAYYPVTVKKQLRAQEIAMQNRLPCIYLVDSGGAYLPRQADV FPDRDHFGRTFYNQAIMSSKNIAQIAVVMGSCATAGGAYVPAMADE NIIVRKQGTIFLAGPPLVKAATGEEVSAEDLGGADLHCRKSGVSDH WALDDHHALHLTRKVVRNLNYQKKLDVTIEPSEEPLFPADELYGIV GANLKRSDVDREVIARIVDGSRFTEFKAFYGDTLVTGFARIFGYPVGI VGNNGVLFSesakKGTHFVQLCCQRNIPLLFQNTITGFMVGREYEA EGIAKDGAKMVA AVACAQVPKITLIIGGSYGAGNYGMCGRAYSPR FLYIWPNARISVMGGEQAANVLATITKDQRAREGKQFSSADEAALK EPIIKKFEEEGNPYYSSARVWDDGIIDPADTRLVLGLSFSAAALNAPIE KTDFGIFRM	MCCC2
94	MAVAGPAPGAGARPRDLQFLQRFQILKVLFPWSWSSQNALMFLTL LCLTLLEQFVIYQVGLIPSQYYGVLGNKDLEGFKTLTFLAVMLIVLN STLKSFDQFTCNLLYVSWRKDLTEHLHRLYFRGRAYYTNLVLRDDI DNPQDRISQDVERFCRQLSSMASKLIISPTLVYYTYQCFQSTGWL PVSIFGYFILGTVVNKTLMGPVIMKLVHQEKLEGDFRFXHMQIRVN AEPAAFYRAGHVEHMRDRLQRLQTQRELMSKELWLYIGINTFD YLGSIYSVVIAIPIFSGVYGDLSPAELSTLVSKNAFVCIYLISCFTQLI DLSTTLSDVAGYTHRIGQLRETLLDMSLKSQDCEILGESEWGLDTP GWPAAPADTAFLLERVSISAPSSDKPLIKDLSLKISEGQSLITGNTG TGKTSLLRVLGGLWTSTRGSVQMLTDFGPHGVFLFPQKPFFTDGT REQVIYPLKEYVPDSSGSADDERILRFLELAGLSNLVARTEGLDQQVD WNWYDVLSPGEMQRLSFARLFYLPKYAVLDEATSALTEEVESEL YRIGQQLGMTFISVGHRSLEKFHSLVLKLCGGGRWELMRIKVE	ABCD4
95	MASAVSPANLPAVLLQPRWKRVVWGSVPVPRPRHGHRAVAIKELI VVFSGGNEGIVDELHVYNTATNQWFIPAVRGDIPPGCAAYGFVCDG TRLLVFGGMVEYGYKYSNDLYELQASRWKRLKAKTPKNGPPPCP RLGHSFSLVGNKCYLFGGLANDSEDPKNNIPRYLNDLYILELRPGSG VVAWDIPITYGVLPPPRESHTAVVYTEKDNKSKLVYGGMSGCRL GDLWTLIDITLTWNKPSLSGVAPLPRSLHSATTIGNKMYVFGGWVP LVMDDVKVATHEKEWKCTNTLAACNLDTMAWETILMDTLEDNIPR ARAGHCAVAINTRLYIWSGRDGYRKAWNNQVCCDLWYLETEKP PPPARVQLVRANTNSLEVSWGAVATADSLLQLQKYDIPATAATAT SPTPNPVSPVANPPKSPAPAAAAPAVQPLTQVGITLLPQAAPAPPTT TTIQVLPVPGSSISVPTAARTQGVPAVLKVTGPQATTGTPLVTMRP ASQAGKAPVTVTSLPAGVRMVVPTQSAQGTIVGSSPQMSGMAALA AAAAATQKIPSSAPTIVLSVPAGTTIVKTMAVTPGTTTTLPATVKVAS SPVMVSNPATRMLKTAQAQVGTSSATNTSTRPIITVHKSGTVTV AQQAQVVTTVVGGVTKTITLVKSPISVPGGSALISNLGKVM SVVQT KPVQTS AVTGQASTGPVTQIIQTKGPLAGTILKLVT SADGKPTTIIT TQASGAGTKPTILGISSVSPSTTKPGTTTIKTIPMSAITQAGATGVT SPGIKSPITITTKVMTSGTGAPAKIITAVPKIATGHGQQGVTVVVLK GAPGQPGTILRTVPMGGVRLVTPVTVSAVKPAVTTLVVKGTTGVT LGTVTGTVSTSLAGAGGHSTASLATPITTLGTIATLSSQVINPTAITV SAAQTTLTAAGGLTPTITMQPVSQPTQVTLITAPSGVEAQPVHDL VSILASPTTEQPTATVTIADSGQGDVQPGTVTLVCSNPPCETHETGTT NTATTTVVANLGGHPQPTQVQFVCDRQEAAASLVSTVGGQNGSV VRVCSNPPCETHETGTTNTATTATSNMAGQHGCNPPCETHETGTT NTATTAMSSVGANHQRDARRACAAGTPAVIRISVATGALEAAQGS KSQCQTRQTSATSTMTVMATGAPCSAGPLLGPMSAREPGGRSPA	HCFC1

	VQLAPLSSKVRLSSPSIKDLPAGRHSHAVSTAAMTRSSVGAGEPRM APVCESLQGGSPSTTVTVTALEALLCPSATVTQVCSNPPCETHETGT TNTATTSNAGSAQRVCSNPPCETHETGTTHTATTATSNNGGTGQPEG GQQPPAGRPCETHQTTSTGTTMSVSVGALLPDATSSHRTVESGLEV AAAPSVTPQAGTALLAPFPTQRVCSNPPCETHETGTTHTATTVTSN MSSNQDPPPAASDQGEVESTQGDSVNITSSSAITTTVSSTLTRA VTTV TQSTPVPGPSVPPPEELQVSPGPRQQLPPRQLQSASTALMGESAEV LSASQTPELPAAVDLSTGEPSSGQESAGSAVVATVVVQPPPPTQSE VDQLSLPQELMAEAQAGTTTLMVTGLTPEELAVTAAAEAAAQAAA TEEAQALAIQAVLQAAQAVMGTGEPMDTSEAAATVTQAE LGHLS AEGQEQGQATTIPIVLTQQELAALVQQQQLQEAQAQQHHHLPTAL APADSLNDPAIESNCLNELAGTVPSTVALLPSTATESLAPSNTFVAPQ PVVVASPAKLQAAATL TEVANGIESLGVKPDLPPPPSKAPMKKENQ WFDVGVIGKTNVMVTHYFLPPDDAVPSDDDLGTVPDYNQLKKQEL QPGTAYKFRVAGINACGRGPFSEISAFKTC L PGFPGAPCAIKISKSPD GAHLTWEPPSVTSKGII EYSVYLAIQSSQAGGELKSSTPAQLAFMRV YCGPSPSCLVQSSSLNAHIDYTKPAIIFRIARNEKGYGPATQVRW LQETSKDSSGTKPANKRPMSSPEMKSAPKSKADGQ	
96	MATSGAASAELVIGWCIFGLLLLAILAFCWIYVRKYQSRRESEVVST ITAIFSLAIALITSALLPVDIFLVSYMKNQNGTFKDWANANVSRQIED TVLYGYTLYSVILFCVFFWIPFVYFYEEKDDDDTSKCTQIKTALK YTLGFVVICALLLL VGAFVPLNVPNNKNSTEWKVKSLFEELGSSH GLAALSFSISSLTIGMLAAITYTAYGMSALPLNLIKGTRSAAYERLE NTEDIEEVEQHIQTIKSKSKDGRPLPARDKRALKQFEERLRTLKKRE RHLEFIENSWWTKFCGALRPLKIVWGIFFILVALLFVISLFLSNLDKA LHSAGIDSGFIIFGANLSNPLNMLLPLLQTVFPLDYILITHIMYFIFTSM AGIRNIGIWWFFWIRLYKIRRGTRTRPQALLFLCMILLIVLHTSYMIYSL APQYVMYGSQNYLIETNITSDNHKGNSTLSVPKRCADADAPEDQCTV TRTYLFLHKFWFFSAAYYFGNWAFLGVFLIGLIVSCCKGKKSVIEGV DESDISDDEPSVYSA	LMBRD1
97	MSAKSRTIGHGAPFSKGQPRGGVEEGPTVLRKAGLLEKLKEQECDV KDYGDLPFADIPNDSPFQIVKNPRSVGKASEQLAGKVAEVKKNGRIS LVLGGDHSLAIGSISGHARVHPDLGVIWVDAHTDINTPLTTTSGNLH GQPVSFLLKELKGKIPDVPGFSWVTPCISAKDIVYIGLRDVPGEHYI LKTGLIKYFSMTEVDRLGIGKVMETLSYLLGRKKRPIHLSFDVDGL DPSFTPATGTPVVGGLTYREGLYITEEIYKTGLLSGLDIMEVNP SLGK TPEEVTRTVNTAVAITLACFLAREGNHKPIDYLNPPK	ARG1
98	MKSNPAIQAAIDLTAGAAGGTACVLTGQPFDTMKVKMQTFPDLYR GLTDCCLKTYSQVGFGRFYKGTSPALIANIAENSVLFMCYGCQQV VRKVAGLDKQAKLSDLQNAAGSFASAFALVLCPTELVKCRLQT MYEMETSGKIAKSQNTVWSVIKSILRKDGPLGFYHGLSSTLLREVPG YFFFFGGYELSRSFASGRSKDELGPVPLMLSGGVGGICLWLAVYPV DCIKSRIQVLSMSGKQAGFIRTFINVVKNEGITALYSGLKPTMIRAFP ANGALFLAYEYSRKLMMNQLEAY	SLC25A15
99	MAAAKVALTKRADPAELRTIFLKYASIEKNGEFFMSPNDFVTRYLNI FGESQPNPKTVELLSGVVDQTKDGLISFQEFVAFESVLCAPDALFMV AFQLFDKAGKGEVTFEDVKQVFGQTTIHQHIFPNWDSEFVQLHFGK ERKRHLTYAEFTQFLLEIQLEHAKQAFVQRDNARTGRVTAIDFRDI MVTIRPHVLT P FVEECLVAAAGGTTSHQVSFSYFNGFN SLLNNMELI RKIYSTLAGTRKDVEVTKEEFVLAAQKFGQVTPMEVDILFQLADLY EPRGRMTLADIERIAPLEEGTL PFNLAE AQ RQKASGDSARPVLLQVA ESAYRFGLGSVAGAVGATAVYPIDLVKTRMQNQ RSTG S FVGELMY	SLC25A13

	KNSFDCFKKVLRYEGFFGLYRGLLPQLLGVAPEKAIKLTVNDFVRD KFMHKDGSVPLAAEILAGGCAGGSQVIFTNPLEIVKIRLQVAGEITT GPRVSALSVVRDLGFFGIYKGAKACFLRDIPFSAIYFPCYAHVKASF ANEDGQVSPGSLLLAGAIAGMPAASLVTPADVIKTRLQVAARAGQT TYSGVIDCFRKILREEGPKALWKGAGARVFRSSPQFGVTLLTYELLQ RWFYIDFGGVKPMGSEPVPKSRINLPAPNPDHVGGYKLAVATFAGI ENKFGLYLPLFKPSVSTSKAIGGGP	
100	MQPQSVLHSGYFHPLLRAWQTATTTLNASNLIYPIFVTDPDDIQPIT SLPGVARYGVKRLLEMLRPLVEEGLRCVLIFGVPSRVPKDERGSAA DSEESPAIEAIHLLRKTFPNLLVACDVCLCPYTSHGHCGLLSENGAF RAEESRQRLAEVALAYAKAGCQVVAPSDMMDGRVEAIKEALMAH GLGNRVSVMSYSAKFASCFYGPFRDAAKSSPAFGDRRCYQLPPGAR GLALRAVDRDVREGADMLMVKPGMPYLDIVREVKDKHPDLPLAV YHVSGEFAMLWHGAQAGAFDLKAAVLEAMTAFRRAGADIITYYT PQLLQWLKEE	ALAD
101	MALQLGRLSSGPCWLVARGGCGGPRAWSCQCGGGGLRAWQSRSAA GRVCRPPGAGTEQSRGLGHGSTRGGPWVGTGLAAALAGLVGLA TAAFGHVQRAEMLPKTSGTRATSLGRPEEEDELAHRCSSFMAPPV TDLGELRRRPGDMKTKMELLILETQAQVCQALAQVDGGANFSVDR WERKEGGGGISCVLQDGCVFEEKAGVSISVVHGNLSEEAQKMRSR GKVLKTKDGKLPFCAMGVSSVIHPKNPHAPTIHFNYRYFEVEEADG NKQWWFGGGCDLTPTYLNQEDAVHFHRTLKEACDQHGPDLYPKF KKWCDDYFFIAHRGERRGIGGIFDDLDSPSKEEVFRFVQSCARAVV PSYIPLVKKHCDDSFPTQEKLWQQLRRGRYVEFNLLYDRGTFGLF TPGSRIESILMSLPLTARWEYMHSPSENSKEAEILEVLRHPRDWVR	CPOX
102	MSGNGNAAATAEENSPKMRVIRVGTRKSQLARIQTDSVVATLKAS YPGLQFEIHAMSTTGDKILDTALSKIGEKSFLTKELEHALEKNEVDLV VHSLKDLPTVLPFGFTIGAICKRENPHDAVVFPKFKVGTLETLPK SVVGTSSLRRAAQLQRKFPHEFRSIRGNLNTRLRLKLDEQQEFSAIL ATAGLQRMGWHRNVGQILHPEECMYAVGQGALGVEVRAKDQDIL DLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVAVHTAMKDGQLY LTGGVWSLDGSDSIQETMQATIHVPAQHEDGPEDDPQLVGITARNIP RGPQLAAQNGLISLANLLLSKGAKNILDVARQLNDAH	HMBS
103	MGRTVVVLGGGISGLAASYHLSRAPCPPKVVLVESSERLGGWIRSV RGPNGAIFELGPRGIRPAGALGARTLLLSELGLDSEVLVVRGDHPA AQNRFLYVGGALHALPTGLRGLLRSPPPFSKPLFWAGLRELTKPRG KEPDET VHSFAQRRLGPEVASLAMDSLGRGVFAGNSRELSIRSCFPS LFQAEQTHRSILLGLLLGAGRTPQPDSALIRQALAEERWSQWSLRGG LEMLPQAETHLTSRGVSVLRGQPVCGLSLQAEGRWKVSLRDSLE ADHVISAI PASVLSSELLPAEAAPLARALSAITAVSVAVVNLQYQGAH LPVQGFGLHVPSSDPGVLGIVYDSVAFPEQDGSPPGLRVTVM LGG SWLQTLEASGCVLSQELFQQRQAQEAATQLGLKEMPSHCLVHLHK NCIPQYTLGHWQKLESARQFLTAHRLPLTLGASYESGVAVNDICIES GRQAAVSVLGTEPNS	PPOX
104	MAHAHIQGGRRAKSRFVVCIMSGARSKLALFLCGCYVVALGAHTG EESVADHHEAEYYVAAYEHPSILSLNPLALISRQEALMNMNQNDI YEQQVMTAAQKDVQIIVFPEDGIHGFNFTRTSIYPFLDFMPSPQVVR WNPCLEPHRFNDTEVLQRLSCMAIRGDMFLVANLGTKEPCHSSDPR CPKDGRYQFNTNVVFSNNGTLVDYRKHNLVFEAAFDVPLKVDLIT FDTPFAGRFGIFTCFDILFFDPAIRVLRDYKVKHVYPTAWMNQLPL LAAIEIQKAFVAFGINVLAANVHHPVLGMTGSGIHTPLESFYHHD MENPKSHLIIAQVAKNPVGLIGAENATGETDPSHSHKFLKILSGDPYC	BTD

	EKDAQEVHCDEATKWNVNAPPTFHSEMMYDNFTLVPVWGKEGYL HVC SNGLCCYLLYERPTLSKELYALGVFDGLHTVHGTYIIQV CALV RCGGLGFDTCGQEITEATGIFEHLWGNFSTSYIFPLFLTSGMTLEVP DQLGWENDHYFLRKSRLLSSGLVTAALYGRLYERD	
105	MEDRLHMDNGLVPQKIVSVHLQDSTLKEVKDQVSNKQAQILEPKP EPSLEIKPEQDGMHEHVRDDPKALGEEPQRRGSASGSEPAGDS DR GGGPVEHYHLHLSSCHECLELENSTIESVKFASAENIPDLPYDYSSSL ESVADETSPEREGRRVNLTGKAPNILLYVGSDSQEALGRFHEVRSVL ADCVDIDSYILYHLLSDALRDPWTDNCLLLVIATRESIPEDLYQKF MAYLSQGGKVLGLSSSFTFGGFQVTSKGALHKTQNLVFSKADQSE VKLSVLSSGCRYQEGPVRLSPGRLQGHLENEDKDRMIVHVPFGTRG GEAVLCQVHLELPSSNIVQTPEDFNLLKSSNFRRYEVLREILTTLGL SCDMKQVPALTPLYLLSAAEEIRDPLMQWLGHVDSGEIKSGQLS LRFVSSYVSEVEITPSCIPVVTNMEAFSSEHFNLEIYRQNLQTKQLGK VILFAEVTPTTMRLLDGLMFQTPQEMGLIVIAARQTEGKGRGGNVW LSPVGCALSTLLISIPLSQLGQRIPFVQHLMSVAVVEAVRSIPEYQDI NLRVKWPNDIYYSDLMKIGGVLVNSTLMGETFYILIGCGFNVTNSN PTICINDLITEYNKQHKAEKPLRADYLIARVVTVLEKLIKEFQDKGP NSVLPLYRYWVHSGQQVHLGSAEGPKVSIVGLDDSGFLQVHQEG GEVTVVHPDGN SFDMLRNILPKRR	HLCS
106	MLKFRTVHGGRLRLGIRRTSTAPAA SPNVRRLEYKPIKKVMVANRG EIAIRVFRACTELGIRTVAIYSEQDTGQMHRQKADEAYLIGRGLAPV QAYLHIPDIKVAKENNVDAVHPGYGFLSERADFAQACQDAGVRFI GPSPEVVRKMGDKVEARAIAIAAGVPVVP GTDAPITSLHEAHEFSNT YGFPIIFKAA YGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGA LFVEKFIEKPRHIEVQILGDQYGNILHLYERDCSIQRRHQKVVEIAPA AHLDPQLRTRLTSDSVKLAKQVGYENAGTVEFLVDRHGKHYFIEV NSRLQVEHTVTEETD VDLVHAQIHVAEGRSLPDLGLRQENIRINGC AIQCRVTTEDPARSFQPD TGRIEVRFSGEGMGIRLDNASAFQGA VISP HYDSLLVKVIAHGKDHTAATKMSRALAEFRVRGVKTNIAFLQNV LNNQFLAGTVD TQFIDENPELFQLRPAQNRAQKLLHYLGHV MN GPTTPIPVKASPSPTDPVPAVPIGPPPAGFRDILLREGPEGFARAVRN HPGLLLMDTTFRDAHQSLLATRV RTHDLKKIAPYVAHNFSKLFSME NWGGATFDVAMRFLYEC PWRLQELRELIPNIPFQMLLRGANAVG YTNPDPNVVFKFCEVAKENGMDVFRVFD SLNYLPNMLLGMEAAG SAGGVVEAAISYTG DVADPSRTKYSLQYYMGLAEELVRAGTHILCI KDMAGLLKPTACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLA CAQAGADVVDVAADSMGMSGTSQPSMGALVACTRGTPLDTEVPME RVFDYSEYWE GARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNL HFQAHSMLGSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGDLAQ FMVQNGLSRAEAEAEAEELSFPRSVVEFLQGYIGVPHGGFPEPFRSK VLKDLPRVEGRPGASLPDLQALEKELVDRHGEEVTPEDVLSAAM YPDVFAHFKDFTATFGPLDSLNTLRLFLQGPKIAEEFEVELERGKTLHI KALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAMKEMHFHPKAL KDVKGQIGAPMPGKVIDIKVVAGAKVAKGQPLCVLSAMKMETVVT SPMEGTVRKVHVTKDMTLEGD DLILEIE	PC
107	MVDSTEYEVASQPEVETSPLGDGASPGPEQVKLKEISLLNGVCLIV GNMIGSGIFVSPKGVLIIYSASFGLSLVIWAVGGLFSVFGALCYAELG TTIKSGSASYAYILEAFGGFLAFIRLWTSLLIIEPTSQAIHITFANYMV QPLFPSCFAPYAASRLAAACICLLTFINCA YVKWGTLVQDIFTYAK VLALIAVIVAGIVRLGQGASTHFENSFEGSSFAVG DIALALYSALFSY SGWDTLNYVTEEIKNPERNLPLSIGISMPIVTIYILT NVAYYTVLDM RDILASDAVAVTFADQIFGIFNWIPLSV ALSCFGGLNASIVAASRLFF	SLC7A7

	VGSREGHLPDAICMIHVERFTPVPSSLFNGIMALIYLCVEDIFQLINY YSFSYWWFFVGLSIVGQLYLRWKEPDRPRPLKLSVFFPIVFCLCTIFLV AVPLYSDTINSLIGIAIALSGLPFYFLIIRVPEHKRPLYLRRIVGSATRY LQVLCMSVAAEMDLEDGGEMPKQRDPKSN	
108	MVPRLLLRWPRGPAVGPAGPSRPLSAGSGPGQYLQRSIVPTMHYQ DSLPRLLPIPKLEDITIRRYLSAQKPLLNDGQFRKTEQFCKSFENGIGKE LHEQLVALDKQNKHTSYISGPWFDMYLSARDSVVLNFPNPFMAFNP DPKSEYNDQLTRATNMTVSAIRFLKTLRAGLLEPEVFHLNPAKSDTI TFKRLIRFVPSLSWYGAYLVNAYPLDMSQYFRLFNSTRLPKPSRDE LFTDDKARHLLVLRKGNFYIFDVLDDQDGNIVSPSEIQAHLKYILSDSS PAPEFPLAYLTSENNDIWAELRQKLMSSGNEESLRKVDSAVFCLCLD DFPIKDLVHLSHNMHLHGDGTNRWFDKSFNLIADKDGSTAVHFEHSW GDGVAVLRFNEVFKDSTQTPAVTPQSQPATTDSTVTVQKLNFEFT DALKTGITAAKEKFDATMKTLTIDCVQFQRRGGKEFLKKQKLSFDAV AQLAFQMAFLRQYGGTVATYESCSTAFAKHGRTETIRPASVYTKRC SEAFVREPSRHSAGELQMMVECSKYHGQLTKEAAMGQGFDRLHF ALRHAAAKGIILPELYLDPAYGQINHNVLSTSTLSSPAVNLLGGFAP VVSDGFGVGYAVHDNWIGCNVSSYPGRNAREFLQCVEKALEDMMFD ALEGKSIKS	CPT2
109	MAAGFGRCCRVLRISRFHWRSQHTKANRQREPGLGFSFEFTEQQK EFQATARKFAREEIIPVAAEYDKTGEYPVPLIRRAWELGLMNTHIPE NCGGLGLGTFDAKLISEELAYGCTGVQTAIEGNSLGQMPIIIAGNDQ QKKKYLGRMTEEPLMCAYCVTEPGAGSDVAGIKTKAEKKGDEYII NGQKMWITNGGKANWYFLLARSDPDPKAPANKAFTGFIVEADTPG IQIGRKELNMGQRCSDTRGIVFEDVKVPKENVLIGDGAGFKVAMGA FDKTRPVVAAGAVGLAQRALDEATKYALERKTFGKLLVEHQAI SF MLAEMAMKVELARMSYQRAAWEVDSGRRNTYYASIAKAFAGDIA NQLATDAVQILGGNGFNTEYPVEKLMRDAKIYQIYEGTSQIQLIVA REHIDKYKN	ACADM
110	MAAALLARASGPARRALCPRAWRLHTIYQSVELPETHQMLLQTC RDFAEKELFPAAQVDKEHLFPAAQVKKMGGLGLLAMDVPEELGG AGLDYLAIAAMEEISRGCASTGVIMSVNNSLYLGPILKFGSKEQKQ AWVTPFTSGDKIGCFALSEPGNGSDAGAASTTARAEGDSWVLNGT KAWITNAWEASAAVVFASDTRALQNKGISAFVPMPTPGLTLGKKE DKLGIRGSSTANLIFEDCRIPKDSILGEPGMGFKIAMQTLDMGRIGIA SQALGIAQTALDCAVNYAENRMAFGAPLTKLQVIQFKLADMALAL ESARLLTWRAAMLKDNKKPFIKEAAMAKLAASEAATAISHQAIQIL GGMGYVTEMPAERHYRDARITEIYEGTSEIQRLLVIAGHLLRSYRS	ACADS
111	MQAARMAASLGRQLRLGGGSSRLTALLGQPRPGPARRPYAGGAA QLALDKSDSHPSDALTRKKPAKESKSAVGMFKGQLTTDQVFPYP SVLNEEQTQFLKELVEPVSRFFEEVNDPAKNDALMVEETTWQGLK ELGAFGLQVPSELGGVGLCNTQYARLVEIVGMHDLGVGITLGAHQ IGFKGILLFGTKAQKEKYLPLASGETVAAFCLTEPSSGSDAASIRTS AVPSPCGKYITLNGSKLWISNGGLADIFTVFAKTPVTDPATGAVKE KITAFVVERGFGGITHGPPEKKMGIKASNTAEVFFDGVVRVPSENVLG EVGSGFKVAMHILNNGRFGMAAALAGTMRGIIAKAVDHATNRTQF GEKIHNFLIQEKLARMVMLQYVTESMAYMV SANMDQGATDFQIE AAISKIFGSEAAWKVTDECIQIMGGMGMFMKEPGVERVLRDLRIFRIF EGTNDILRLFVALQGCMDKGKELSGLSALKNPFGNAGLLLGAG KQLRRRAGLGSGLSLSGLVHPELSRSGELAVRALEQFATVVEAKLIK HKKGIVNEQFLQLRLADGAIDLYAMVVVLSRASRSLSEGHPTAQHE	ACADVL

	KMLCDTWCI EAAARIREGMAALQSDPWQEL YRNFKSISKALVER GGVVT SNPLGF	
112	MGHSKQIRILLNEMEKLEKTLFRLEQGYELQFRLGPTLQGKAVTV YTNYPFPGETFNREKFRSLDWENPTEREDDSKYCKLNLQQSGSFQ YYFLQGNESGGGYIVVDPILRVGADNHVLPDLCVTLQTFLAKCLG PFDEWESRLRVAKESGYNMIHFTPLQTLGLSRSCYSLANQLELNP SRPNRKYTWNDVGQLVEKLKKEWNVICITDVVYNHTAANSKWIQE HPECAYNLVNSPHLKPAWVLDRALWRFSCDVAEGKYKEKGIPALIE NDHHMNSIRKIIWEDIFPKLKLWEFFQVDVNKA VEQFRRLLTQENR RVTKSDPNQHLTIIQDPEYRRFGCTVDMNIALTTFIPHDKGPA AIEEC CNWFHKKRMEELNSEKHRLINYHQEQAVNCLLGNV FYERLAGHGPK LGPVTRKHPLVTRYFTFPFEEIDFSMEESMIHL PNKACFLMAHNGW VMGDDPLRNFAEPGSEVYLRRELICWGDSVKLRYGNKPEDCPYLW AHMKKYTEITATYFQGVRLDNCHSTPLHVAEYMLDAARNLQPNLY VVAELFTGSEDLDNV FVTRLGISSLIREAMSAYNSHEEGRLVYRYG GEPVGSFVQPCLRPLMPAIAHALFMDITHDNECPIVHRSAYDALPST TIVSMACCSGSTRGYDELVPHQISVVSEERFYTKWNPEALPSNTGE VNFQSGIIARCAISKLHQELGAKGFIQVYVDQVDEDIVAVTRHSPSI HQSVVAVSRTAFRNPKTSFY SKEVPQMCIPGKIEEVVLEARTIERNT KPYRKDENSINGTPDITVEIREHIQLNESKIVKQAGVATKGPNEYIQEI EFENLSPGSVIIFRVSLDPAHQVAVGILRNHLTQFSPHFKSGSLAVDN ADPILKIPFASLASRLTLAELNQILYRCESEEKEDGGGCYDIPNWSAL KYAGLQGLMSVLA EIRPKNDLGHPFCNNLRSGDWMIDYVSNRLISR SGTIAEVGKWLQAMFFYLKQIPRYLIPC YFDAILIGAYTTLLDTAWK QMSSFVQNGSTFVKHLSLGSVQLCGVGKFPSPILSPALMDVPYRLN EITKEKEQCCVSLAAGLPHFSSGIFRCWGRDTFIALRGILLITGRYVE ARNIILAFAGTLRHGLIPNLLGEGIYARYNCRDAVWWWLQCIQDYC KMVPNGLDILKCPVSRMYPTDDSA PLPAGTLDQPLFEVIQEQAMQKH MQGIQFRERNAGPQIDRNMKDEGFNITAGVDEETGFVYGGNRNFC GTWMDKMGESDRARNRGIPATPRDGS AVEIVGLSKSAVRWLELS KKNIFPYHEVTVKRHGKAIVSYDEWNRKIQDNFEKLFHVSEDPSD LNEKHPNLVHKRGIYKDSYGASSPWCDYQLRPNFTIAMVVAPELFT TEKAWKALEIAEKKLLGPLGMKTLDPDDMVYCGIYDNALDNDNY NLAKGFNYHQGPEWLWPIGYFLRAKLYFSRLMGPETTAKTIVLVKN VLSRHYVHLERSPWKGLPELTNENAQYCPFS CETQAWSIATILETLY DL	AGL
113	MEEGMNVLHDFGIQSTHYLQVNYQDSQDWFILVSVIADLRNAYFV LFPIWFHLQEAVGIKLLWVAVIGDWLNLVFKWILFGQRPYWWVLD TDYYSNTSVPLIKQFPVTCETGPGSPSGHAMGTAGVYYVMVTSTLSI FQGKIKPTYRFRCLNVILWLGFWAVQLNVCLSRIYLA AHFPHQVVA GVLSGIAVAETFSHIHSIYNASLKKYFLITFFLFSFAIGFYLLLKGLGV DLLWTLEKAQRWCEQPEWVHIDTTPFASLLKNLGTFLGLALNSS MYRESCKGKLSKWL PFRLLSSIVASLVLLHVFD SLKPPSQVELVFYVL SFCKSAVVPLASVSVIPYCLAQVLGQPHKKS L	G6PC
114	MAAPMTPAARPEDYEAALNAALADVPELARLLEIDPYLKPYAVDF QRRYKQFSQILKNIGENEGGIDKFSRGYESFGVHRCADGGLYCKEW APGAEGVFLTGDFNGWNPFSYPYKKLDY GKWELYIPPKQNKSVLV PHGSKLKVVITSKSGEILYRISPAKYV VREGDNVNYDWIHWDP SYEFKHSRPPKPRSLRIYESHVGISSHEGKVASYKHFTCNVLPRIKGL GYNCIQLMAIMEHAYYASFGYQITSFFAASSRYGTPEELQELVDTAH SMGIIVLLDVVHSHASKNSADGLNMF DGTDCYFHS GPRGTHDLW DSRLFAYSSWEILRFLLSNIRWWLEEYRFDGFRFDGVTSMLYHHHG VGQGFSGDYSEYFGLQVDEDALTYLMLANHLVHTLCPDSITIAEDV	GBE1

	SGMPALCSPISQGGGGFDYRLAMAIPDKWIQLLKEFKDEDWNMGDI VYTLTNRRYLEKCIAYAESHDAQALVGDKSLAFWLMDAEMYTNMS VLTPFTPVIDRGIQLHKMIRLITHGLGGEGYLNFMGNEFGHPEWLD PRKGNNESYHYARRQFHLTDDDLLRYKFLNNFDRDMNRLEERYG WLAAPQAYVSEKHEGNKIIAFERAGLLFIFNFHPSKSYTDYRVGTAL PGKFKIVLDSDAAEYGGHQRLDHSTDFSEAFEHNGRPYSLLVYIPS RVALILQNVDLPN	
115	MRSRSNSGVRLDGYARLVQQTILCHQNPVTGLLPASYDQKDAWVR DNVYSILAVWGLGLAYRKNADRDEDKAKAYELEQSVVKLMRGLL HCMIRQVDKVESFKYSQSTKDSLHAKYNTKTCATVVGDDQWGH QLDATSVYLLFLAQM TASGLHIIHSLDEVNFIQNLVFIYAAYKTAD FGIWERGDKTNQGISELNASSVGMAKAALEALDELDFGVKGGPQS VIHVLADEVQHCQSILNSLLPRASTSKEVDASLLSVVSFPFAVEDS QLVELTKQEITKLQGRYGCCRFLRDGYKTPKEDPNRLYYEPAELKL FENIECEWPLFWTYFILDGVFSGNAEQVQEYKEALEAVLIKGNV PLLPELYSVPPDRVDEEYQNPHTVDRVPMGKLPBMWGGQSLYILGSL MAEGLAPGEIDPLNRRFSTVPKPDVVVQVSILAETEEIKTILKDKGI YVETIAEVYPIRVQPARILSHIYSSLGCNNRMKLSGRPYRHMGVLT SKLYDIRKTIFTFTPQFIDQQQFYALDKNMIVEMLRDLSYLCSRW RMTGQPTITFPISHSMLDEDGTSLSNLSILAALRKMQDGYFGGARVQT GKLSEFLTTSCTHLSFMDPGPEGKLYSEYDDNYDYLESNWMN DYDSTSHARCGDEVARYLDHLLAHTAPHPKLAPTSQKGGDRFQA AVQTTCDLMSLVTKAKELHVQNVHMYLPTKLFQASRPSFNLLDSP HPRQENQVPSVRVEIHLPRDQSGEVDFKALVLQLKETSSLQEADIL YMLYTMKGPDWNTELYNERSATVRELLTELYGKVGEIRHWGLIRYI SGILRKKVEALDEACTDLLSHQKHLTVGLPPEPREKTISAPLPYEALT QLIDEASEGDMISISILTQEIMVYLAMYMRTQPGLFAEMFRLRIGLIQ VMATELAHSLRCSAEATEGLMNLSPSAMKNLLHHILSGKEFGVER SVRPTDSNVSPAISIHEIGAVGATKTERTGIMQLKSEIKQVEFRRLSIS AESQSPGTSMTSPSSGSFPSAYDQSSKDSRQGWQRRRRRLDGA NRPVPGFYQVWKVLQKCHGLSVEGFVLPSTTREMTPEIKFSVHVE SVLNRVPQPEYRQLLVEAILVLTMLADIEIHSIGSIIAVEKIVHIANDL FLQEQLTLGADDTMLAKDPASGICTLLYDSAPSGRFGTMTYLSKAA ATYVQEFLPHSICAMQ	PHKA1
116	MRSRSNSGVRLDGYARLVQQTILCYQNPVTGLLSASHEQKDAWVR DNIYSILAVWGLGMAYRKNADRDEDKAKAYELEQNVVKLMRGLL QCMMRQVAKVEKFKHTQSTKDSLHAKYNTATCGTVVGDDQWGH LQVDATSLFLLFLAQM TASGLRIIFTLDEVAFIQNLVFIYAAYKVA DYGMWERGDKTNQGIPELNASSVGMAKAALEAIDELDFGAHGG KSVIHVLPDEVEHCQSILFSMLPRASTSKEIDAGLLSIISFPFAVEDV NLVNVTKNEIISKLQGRYGCCRFLRDGYKTPREDPNRLHYDPAELK LFENIECEWPFVFTYFIIDGVFSGDAVQVQEYREALEGILIRGKNIR LVPELYAVPPNKVDEEYKNPHTVDRVPMGKVPHLWGQSLYILSSLL AEGFLAAGEIDPLNRRFSTSVKPDVVVQVTVLAENNHKIDLLRKHG VNVQSIADHPIQVQGRILSHIYAKLGRNKNMNLGRPYRHIGVLG TSKLYVIRNQIFTFTPQFTDQHFFYLALDNEMIVEMLRIELAYLCTC WRMTGRPTLTFPISRTMLTNDGSDIHS AVLSTIRKLEDGYFGGARVK LGNLSEFLTTSFYTYLTFLDPDCDEKLFDNASEGTFSPDSDSL VG LEDTCNQESQDELHYINHLLQSTSLRSYLPPLCKNTEDRHVFS AIHSTRDILSVMAKAKGLEVPFVPMTLPTKVLSAHRKSLNLVDSPQPLLE KVPESDFQWPRDDHGDVDCLEKLVQLKDCSNLQDQADILYILYVIK GPSWDTNLSGQHGVTVQNLLGELYGKAGLNQEWGLIRYISGLLRK KVEVLAEACTDLLSHQKQLTVGLPPEPREKIISAPLPPEELTKLIYEA	PHKA2

	SGQDISIAVL TQEIVVYLAMYVRAQPSLFVEMLRRLRIGLIQVMATEL ARSLNCSGEEASESLMNLSPFDMKNLLHHILSGKEFGVERSVRPIHS STSSPTISIEVGHGTGVTKTERSGINRLRSEM KQMTRRFSAD EQFFSV GQAASSAHSSKSARSSTPSSPTGTSSSDSGHHIGWGERQQQWLRR RRLDGAINRVPVGFYQVRWKILQKCHGLSIDGYVLPSTTREMTPH EIKFAVHVESVLNRVPQPEYRQLLVEAIMVLTLLSDTEMTSIGGIIHV DQIVQMASQLFLQDQVSIGAMDTLEKDQATGICHFFYDSAPSGAYG TMTYLTRAVASYLQELLPNNSGCQMQ	
117	MAGAAGLTAEVSWKVLERRARTKRSGSVYEPLKSINLPRPDNETL WDKLDHYYRIVKSTLLLYQSPPTGLFPTKTCGGDQKAKIQDSL YCA AGAWALALAYRRIDDDKGRTHELEHSAIKCMRGILYCYMRQADKV QQFKQDPRPTTCLHSVFNVHTGDELLSYEEYGH LQINAVSLYLLYL VEMISSGLQIIYNTDEV SFIQNLVFCVERVYRVPDFGVWERGSKYNN GSTELHSSSVGLAKAALEAINGFNLFGNQGCSWSVIFVDLDAHNRN RQTLCSLLPRESRSHNTDAALLPCISYPAFALDDEV LFSQTLDKVVR KLKGKYGFKRFLRDGYRTSLEDPNRCYYKPAEIKLFDGIECEFPFFL YMMIDGVFRGNPKQVQEYQDLLTPVLHHTTEGYPPVVPKYYYVPAD FVEYEKNPNPGSQKRFP SNCGRDGKLF LWGQALYIIAKLLADELISPK DIDPVQRYVPLKDQRNVSMRFSNQGLENDLVVHV ALIAESQRLQV FLNTYGIQTQTPQQVEPIQIWPQQELVKAYLQLGINEKLGLSGRPDR PIGCLGTSKIYRILGKTVVCYPIIFDLSDFYMSQDVFLIDDIKNALQF IKQYWKMHGRPLFLVLIREDNIRGSRFNPILDMLAALKKG IIGGVKV HVDRLQTLISGAVVEQLDFLRISDTEELPEFKSFEELEPPKHSKVQRQ SSTPSAPELGQQPDVNISEWKDKPTHEILQKLND CCLASQAILLGIL LKREGPNFITKEGTVSDHIERVYRRAGSQKLWLAVRYGAAFTQKFS SSIAPHITTFLVHGKQVTLGAFGHEEEVISNPLSPRVIQNIYYKCNTH DEREAVIQQELVIHIGWIISSNNPELFSGMLKIRIGWIIHAMEYELQIRG GDKPALDLYQLSPSEVKQLLLDILQPQQNGRCWLNRRQIDGSLNRT PTGFYDRVWQILERTPNGIIVAGKHLPPQPTLSDMTMYEMNFSLLV EDTLGNIDQPQYRQIVVELLMVVSIVLERNPELEFQDKVDLDRLVKE AFNEFQKQDQSRLEIEKQDDMTSFYNTPLGKRGTC SYLTKAVMNL LLEGEVKPNNDPCLIS	PHKB
118	MTLDVGPEDLPDWA AAKEFYQKYDPKDVIGRGVSSVVRRCVHRA TGHEFAVKIMEVTAERLSPEQLEEVREATRRETHILRQVAGHPHIITL IDSYESSFMFLVFDLMRKGELFDYLTEKVALSEKETRSIMRSLLEA VSFLHANNIVHRDLKPENILLDDNMQIRLSDFGFSCHLEPGEKLREL CGTPGYLAPEILKCSMDETHPGYGKEVDLWACGVILFTLLAGSPPF WHRRQILMLRMIMEGQYQFSSPEWDDRSSTVKDLISRLQLVDPEAR LTAEQALQHPFFERCEGSQPWNLTQRFRVAVWTVLAAGRVALS THRVRPLTKNALLRDPYALRSVRHLIDNCAFRLYGHWWKKGEQQN RAALFQHRPPGPFPIMGPEEEGDSAAITEDEAVLVLG	PHKG2
119	MAAQGYGYRTVIFSAMFGGYSLYYFNKRTFSFVMPSLVEEIPLDK DDLGFITSSQSAAYAISKFVSGVLSQMSARWLFSSGLLLVGLVNIF FAWSSTVPVFAALWFLNGLAQGLGWPPCGKVLKWFEP SQFGTW WAILSTSMNLAGGLGPILATILAQSYSWRSTLALS GALCVVVSFLCL LLIHNEPADVGLRNLDPMPSSEGKKGSLKEESTLQELLLSPYLWVLST GYLVVFGVKTCCTDWGQFFLIQEKQSALVGSSYMSALEVGGLVG SIAAGYLSDRAMAKAGLSNYGNPRHGLLLFMMAGMTVSMYLFVRV TVTSDSPKLWILVLGAVFGFSSYGPIALFGVIANESAPPNLCGTSHAI VGLMANVGGFLAGLPFSTIAKHYSWSTAFWVAEVICA ASTAAFFLL RNIRTKMGRVSKKAE	SLC37A4

120	MAAPGPALCLFDVDGTLTAPRQKITKEMDDFLQKLRQKIKIGVVGG SDFEKVQEQLGNDVVEKYDYVFPENGLVAYKDGKLLCRQNIQSHL GEALIQDLINYCLSYIAKIKLPKKGRTFIEFRNGMLNVSPIGRSCSQEE RIEFYELDKKENIRQKFVADLRKEFAGKGLTFSIGGQISFDVFPDQW DKRYCLRHVENDGYKTIYFFGDKTMPGGNDHEIFTDPRMTMGYSVT APEDTRRICELLS	PMM2
121	MPSETPQAEVGPTGCPHRSGPHSAKGSLEKSPEDKEAKEPLWIRPD APSRCTWQLGRPASESPHHHTAPAKSPKILPDILKKIGDTPMVRINKI GKKFGLKCELLAKCEFFNAGGSVKDRISLRMIEDAERDGTLPKPGDTI IEPTSGNTGIGLALAAVVRGYRCIIVMPEKMSSEKVDVLRALGAEIV RTPTNARFDSPEHVGVAVRLKNEIPNSHILDQYRNASNPLAHYDT TADEILQQCDGKLDMLVASVGTGGTITGIARKLKEKCPGCRIGVDP EGSILAEPEELNQTEQTTYEVEGIGYDFIPTVLDRTVVVDKWFKSND EAFTFARMLIAQEGLLCGGSAGSTVAVAVKAAQELQEGQRCVVILP DSVRNYMTKFLSDRWMLQKGFLKEEDLTEKKPWWHHLRVQELGL SAPLTVLPTITCGHTIEILREKGFQAPVVDEAGVILGMVTLGNMLS SLLAGVQPSDQVGVKVIYKQFKQIRLTDTLGRLSHILEMDHFALVV HEQIQYHSTGKSSQRQMVFGVVT AIDLLNFVAAQERDQK	CBS
122	MSFIPVAEDSDFPPIHNLPGYVFSTRGDPRPRIGVAIGDQILDLSIIKHLF TGPVLSKHQDVFNQPTLNSFMGLGQAAWKEARVFLQNLLSVSQAR LRDDTELKRCAFISQASATMHLPATIGDYTDYSSRQHATNVGIMFR DKENALMPNWLHLPVGYHGRASSVVSGTPIRRPMGQMKPDDSKP PVY GACKLLDMELEMAFFVGPNGRLGEPIPIKAHEHIFGMVLMND WSARDIQWEYVPLGPFLGKSGFTTVSPWVPM DALMPFAVPNPK QDPRPLPYLCHDEPYTFDINLSVNLKGEGMSQAATICKSNFKYMYW TMLQQLTTHSVNGCNLRPGDLLASGTISGPEPENFGSMLELSWKGT KPIDLGNGQTRKFLLDGDEVIITGYCQGDGYRIGFGQCAGKVLPA PS	FAH
123	MDPYMIQMSSKGNLPSILDVHVNVGGRSSVPGKMKGRKARWSVRP SDMAKKTFFNPIRAIVDNMVKPNPNKTMISLSIGDPTVFGNLPTDPE VTQAMKDALDSGKYNGYAPSIGFLSSREEIASYYHCPEAPLEAKDVI LTSGCSQAIDLCCLAVLANPGQNILVPRPGFSLYKTLAESMGIEVKLY NLLPEKSWEIDLKQLEYLIDEKTA CLIVNNPSNPCGSVFSKRHLQKIL AVAARQCVPILADEIYGDMVFSCKYEPLATLSTDVPILSCGGLAKR WLVPGWRLGWILHRRDIFGNEIRDGLVKLSQRILGPCTIVQGALK SILCRTPGEFYHNTLSFLKSNADLCYGALAAIPGLRPVRPSGAMYLM VGIEMEHPFEFENDVEFTERLVAEQSVHCLPATCFEYPNFIRVVITVP EVMML EACSRIQEFCEQHYHCAEGSQEECDK	TAT
124	MSRSGTDPQQRQQASEADAAAATFRANDHQHIRYNPLQDEWVLVS AHRMKRPWQGQVEPQLLKTVP RHDPLNPLCPGAIRANGEVNPQYD STFLFDNDFPALQPDAPSPGSDHPLFQA KSARGVCKVMCFHPWSD VTLPLMSVPEIRAVVDAWASVTEELGAQYPWVQIFENKGAMMGCS NPHPHCQVWASSFLPDIAQREERSQQA YKSQHGEPLLMEYSRQELL RKERLVL TSEHWLVLPFWATWPYQTLLLP RRHVRRLPELTPAERD DLASIMKKLLTKYDNLFETSFYPSMGWHGAPTGEAGANWNHWQ LHAHYYPPLRSATVRKFMVGYEMLAQAQRDLTPEQAAERLRALP EVHYHLGQKDRETATIA	GALT
125	MAALRQPQVAELLAEARRAFREFGAEP ELAVSAPGRVNLIGEHTD YNQGLVLPMALELMTVLVGSPRKDGLVSLTTSEGADEPQRLQFPL PTAQRSLEPGTTPRWANYVKGVIIQYYPAAPLP GFSAVVVSSVPLGGG LSSASLEVATYTF LQQCPDSGTIAARAQVCQQA EHSFAGMPCGI MDQFISLMGQKGHALLIDCRSLETSLVPLSDPKLAVLITNSNVRHSL	GALK1

	ASSEYPVRRRQCEEVARALGKESLREVQLEEEAARDLVSKEGFRR ARHVVGIEIRRTAQAAAAALRRGDYRAFGRMLMVESHRSRLRDDYEVSC PELDQLVEAALAVPGVYGSRTMTGGGFGGCTVTLLLEASAAPHAMRH IQEHYGGTATFYLSQAADGAKVLCL	
126	MAEKVLVTGGAGYIGSHTVLELLEAGYLPVVIDNFHNAFRGGGSLP ESLRRVQELTGRSVEFEEMDILDQGALQRLFKKYSFMAVIHFAGLK AVGESVQKPLDYRVNLTGTIQLLEIMKAHGVKNLVFSSSATVYGN PQYLPLDEAHPTGGCTNPYGKSKFFIEEMIRDLCQADKTWNAVLLR YFNPTGAHASGCIGEDPQGIPNNLMPYVSQVAIGRREALNVFGNDY DTEDGTGVRDYIHVVDLAKGHIAALRKLKEQCGCRIYNLGTGTGYS VLQMVQAMEKASGKKIPYKVVARREGDVAACYANPSLAQEELGW TAALGLDRMCEDLWRWQKQNPSTGFGTQA	GALE
127	MAEQVALSRTQVCGILREELFQGDFAHQSDTHIFIIMGASGDLAKKK IYPTIWWLFRDGLLPENTFIVGYARSRLTVADIRKQSEPFKATPEEK LKLEDDFFARNSYVAGQYDDAASYQRLNSHMNALHLGSQANRLFYL ALPPTVYEAATKNIHESCMSQIGWNRIIVEKPFGRDLQSSDRLSNHIS SLFREDQIYRIDHYLGKEMVQNLMLVLRFANRIFGIWNRDNIACVIL TFKEPFQTEGRGGYFDEFGIIRDVMQNHLLQMLCLVAMEKPASTNS DDVRDEKVKVLKCISEVQANNVVLGQYVGNPDGEGEATKGYLDD PTVPRGSTTATFAAVVLYVENERWDGVPFILRCGKALNERKAEVRL QFHDVAGDIFHQCKRNELVIRVQPNEAVYTKMMTKKPGMFFNPE ESELDTYGNRYKNVKLPDAYERLILDVFCGSQMHFVRSDELREA WRIFTPLLHQIELEKPKPIPIYIGSRGPTEADELMKRVGFQYEGTYK WVNPBKL	G6PD
128	MAEDKSKRDSIEMSMKGCQTNNGFVHNEDILEQTPDPSSTDNLKH STRGILGSQEPDFKGVQPYAGMPKEVLFQFSGQARYRIPREILFWLT VASVLVLIAATIAIIALSPKCLDWWQEGPMYQIYPRSFKDSNKGNG DLKGIQDKLDYITALNIKTWITSFYKSSLKDFRYGVEDFREVDPIFG TMEDFENLVAIHDKGLKLIIDFIPNHTSDKHIWFQLSRTRTGKYTD YYIWHDCETHENGKTIPPNNWLSVYGNSSWHFDEVNRQCYPHFQFMK EQPDLNFRNPDVQEEIKEILRFWLTKGVDGFSLDVAVKFLLEAKHLR DEIQVNKTQIPDVTQYSELYHDFTTTQVGMHDIRSFRQTMQYS TEPGRYRFMGTEAYAESIDRTVMYYGLPFIQEADFPFNLYLSMLDT VSGNSVYEVITSWMENMPEGKWPWNWMIGGPDSSRLTSRLGNQYVN VMNMLLFTLPGTPITYYGEEIGMGNIVAAANLNEYSINTLRKSPMQ WDNSSNAGFSEASNTWLPTNSDYHTVNVDVQKTQPRSAKLKYQDL SLLHANELLNLRGWFCFLRNDSHYVVYTRELDGIDRIFIVVLNFGES TLLNLHNMISGLPAKMIRLSTNSADKGSKVDTSIGIFLDKGEGLIFE HNTKNLLHRQTAFRDRCFVSNRACYSSVLNLYTSC	SLC3A1
129	MGDTGLRKRREDEKSIQSQEPKTTSLQKELGLISGISIIVGTIIGSGIFV SPKSVLSNTEAVGPCLIWAACGVLATLGALCFELGTMITKSGGEY PYLMEAYGPIPAYLFSWASLIVIKPTSFAIICLSFSEYVCAPFYVGCKP PQIVVKCLAAAILFISTVNSLSVRLGSYVQNIFTAAKLIVAIHISGL VLLAQGNTKNFDFNSFEGAQLSVGAISLAFYNGLWAYDGNQNLNYI TEELRNYPYRNPLAIIIGIPLVTACYILMNVSYFTVMTATELLQSQAV AVTFGDRVLYPASWIVPLFVAFSTIGAANGTCFTAGRLIYVAGREGH MLKVLSYISVRRLTPAPAIIFYGIIATIIYIIPGDINSLVNYFSFAAWLFY GLTILGLIVMRFRKERLERPIKVPVVIPVLMTLISVFLVLAPIISKPTW EYLYCVLFILSGLLFYFLFVHYKFGWAQKISKIPITMHLQMLMEVVP EEDPE	SLC7A9
130	MVNEARGNSSLNPCLEGSASSGSESSKDSSRCSTPGLDPERHERLRE KMRRRLESGDKWFSLEFFPPRTAEGAVNLISRFDMAAGGPLYIDV	MTHFR

	<p>TWHPAGDPGSDKETSSMMIASTAVNYCGLETILHMTCCRQRLEEIT GHLHKAKQLGLKNIMALRGDPIGDQWEEEEGGFNAYVDLVKHRS EFGDYFDICVAGYPKGHPEAGSFEADLKHLKEKVSAGADFIITQLFF EADTFFRFVKACTDMGITCPIVPGIFPIQGYHSLRQLVKLSKLEVPQE IKDVIEPIKDNDAAIRNYGIELAVSLCQELLASGLVPGLHFYTLNREM ATTEVLKRLGMWTEPRRPLPWALSAHPKRREEDVRPIFWASRPKS YIYRTQEWDEFNGRWGNSSPAFGEKDYLYFYLSKSPKEELLK MWGEELTSEESVFEVFLYLSGEPNRNGHKVTCLPWNDPLAAETS LLKEELLRVNRQGILTINSQPNINGKPSSDPVGVGPGSGGYVFQKAY LEFFTSRETAEALLQVLKKYELRVNYHLVNVKGENITNAPELQPN VTWGIIFGREIIQPTVVDPVSMFWKDEAFALWIERWGKLYEESPS RTIIQYIHDNYFLVNLVDNDFPLDNCLWQVVEDTLELLNRPTQNAR ETEAP</p>	
131	<p>MSPALQDLSQPEGLKKTTLRDEINAILQKRIMVLDGGMGTMIQREKL NEEHFRGQEFKD HARPLKGNNDILSITQPDVIYQIHKEYLLAGADII TNTFSSTSIAQADYGLEHLAYRMNMCSAGVARKAAEEVTLQTGIKR FVAGALGPTNKTLSVSPSVERPDYRNITFDELVEAYQEQA KGLLDG GVDILLIETIFDTANAKAALFALQNLFEKYAPRPFIISGTIVDKSGRT LSGQTGEGFVISVSHGEPLCIGLNCALGAAEMRPFIEIIGKCTTAYVL CYPNAGLPNTFGDYDETPSMMAKHLKDFAMDGLVNIVGGCCGSTP DHIREIAEAVKNCKPRVPPATAFEGHMLLSGLEPFRIGPYTNFVNIGE RCNVAGSRKFAKLIMAGNYEEALCVAKVQVEMGAQVLDVNMD GMLDGPSAMTRFCNLIASEPDIKVPCLCIDSSNFAVIEAGLKCCQK CIVNSISLKEGEDDFLEKARKIKKYGAAMVVMFAFDEEGQATETDTK IRVCTRAYHLLVKKLGFNPNDIIFDPNLTIGTGMEEHNL YAINFIHAT KVIKETLPGARISGGLSNLSFSFRGMEAMHGVFLYHAIKSGMD MGIVNAGNLPVYDDIHKELLQLCEDLIWNKDPEATEKLLRYAQ TQG TGGKKVIQTDEWRNGPVEERLEYALVKGIEKHIEDTEEARLNQKK YPRPLNIIIEGLPMNGMKIVGDLFGAGKMFLPQVIKSARVMKKA VGH LIPFMEKEREETRVLNGTVEEEDPYQGTIVLATVKGDVHDIGKNIVG VVLGCNNFRVIDLGVMTPCDKILKAALDHKADIIGLSGLITPSLDEMI FVAKEMERLAIRIPLLIGGATTSTHTAVKIAPRYSAPVIHVLDASKS VVVCSQLLDENLKDEYFEEIMEEYEDIRQDHYESLKERRYLP LSQAR KSGFQMDWLSEHPVKPTFIGTQVFEDYDLQKLVDYIDWKPFDFV WQLRGKYPNRGFPKIFNDKTVGGEARKVYDDAHNMLNTLISQKKL RARGVVGFWPAQSIQDDIHL YAEAAVPQAAEPIATFYGLRQQA EKD SASTEPYYCLSDFIAPLHSGIRDYLG LFAVACFGVEELSKAYEDDGD DYSSIMVKALGDRLAEAFAEELHERVRRELWAYCGSEQLDVADLR RLRYKGIRPAPGYPSQPDHTEKLT MWRLADIEQSTGIRL TESLAMAP ASAVSGLYFSNLKSKYFAVGKISKDQVEDYALRKNISVAEVEKWL G PILGYDTD</p>	MTR
132	<p>MGAASVRAGARLVEVALCSFTVTCLEVMRRFLLLYATQQGQAKAI AEEICEQAVVHGFSADLHCISESDKYDLKTETAPLVVVVSTTGTGDP PDTARKFVKEIQNTLPVDFFAHLRYGLLGLGDSEYTYFCNGGKIID KRLQELGARHFYDTGHADDCVGLLEVVEPWIAGLWPALRKHFSS RGQEEISGALPVASPASSRTDLVKSELLHIESQVELLRFDSSGRK DSE VLKQNAVNSNQSNVIEDFESSLTRSVPLSQASLNIPGLPPEYLQVH LQESLGQEEESQVS VTSADPVFQVPISKAVQLTTNDAIKTTLLVELDIS NTDFSYPQGD AFSVICPNSDSEVQSLLQRLQLEDKREHCVLLKIKAD TKKKGATLPQHIPAGCSLQFIFTWCLEIRAIPKKAFLRALVDYTSDSA EKRRQLQELCSKQGAADYSRFVRDACACLLDLLAFPSCQPPLSLLLE HLPKLQPRPYSCASSSLFHPGKLHFVFNIVEFLSTATTEVLRKGVCTG WLALLVASVLQPNIHASHEDSGKALAPKISISPRTTNSFHL PDDPSIPI</p>	MTRR

	IMVGP GTGIAPFIGFLQHREKLQEHPDGNFGAMWLF FGC R HKDRD YLF RKELRHFLKHGILTHLKV SFSRDAPVGE E EAPAKYVQDN IQLH GQQVARILLQENGHIYVCGDAKNMAKDVH DALVQIISKEVGVEKL EAMKTLATLKEEKRYLQDIWS	
133	MPEQERQITAREGASRKILSKLSLPTRA WEPAMKKSF AFDNVGYEG GLDGLGPSSQVATSTVRILGMT CQSCVKSIEDRISNLKG IISMKVSL E QGSATVKYVPSV VCLQQVCHQIGDMGF EASIAEGKAASWPSRSLPA QEAVVKLRVEGMT CQSCVSSIEGKVRKLQGVVRVKVSLSNQEAVIT YQPYLIQPEDLRDHVNDMGFEAAIKSKVAPLSLGPIDIERLQSTNPK RPLSSANQNFNNSETLGHQGS HVVTLQLRIDGMHCKSCV LNIEENIG QLLGVQSIQVSLENKTAQVKYDP SCTSPVALQRAIEALPPGNFKVSL PDGAEGSGTDHRSSSSHSPGSPPRNQVQGT CSTTLIAIAGMTCASCV HSIEGMISQLEGVQQISVSLAEGTATVLYNPSVISPEELRAAIEDMGF EASV VSESCSTNPLGNHSAGNSMVQTTDGTPTSVQEVAPHTGR LPA NHAPDILAKSPQSTRAVAPQKCF LQIKGMT CASCVSNIERNLQKEAG VLSVLVALMAGKAEIKYDPEVIQPLEIAQFIQDLGF EAAVMEDYAG SDGNIELTITGMT CASCVHNIESKLTRTNGITYASVALATSKALVKF DPEIIGPRDIIKIIIEIGFHASLAQRNPNAHHL DHKMEIKQWKKSFLCS LVFGIPVMALMIYMLIPSNEPHQSMVLDHNIIPGLSILNLIFFILCTFV QLLGGWYFYVQAYKSLRHR SANMDVLIVLATSIA YVYSLVILVVA VAEKAERSPVTF FDTPPMLFVFIALGRWLEHLAKSKTSEALAKLMS LQATEATVVT LGEDNLIREEQVPMELVQRGDIVKVVP GGKFPVDG KVLEGNTMADESLITGEAMPVTKKPGSTVIAGSINAHG SVLIKATHV GNDTTLAQIVKLVEEAQMSKAPIQQLADRFSGYFVPFIIMSTLTLVV WIVIGFIDFGVVQRYFPNPNKHISQTEVIRFAFQTSITVLCIACPCSLG LATPTAVMVG TGVAQAQNGILIKGGKPLEMAHKIKTVMF DKTGTITH GVPRVMRVL LLDG VATLPLRKVLAVVGTAEASSEHPLGVAVTKYC KEELGTETLGYCTDFQAVPGCGIGCKVSNVEGILAHSERPLSAPASH LNEAGSLPAEKDAVPQTFSVLIGNREWLRRNGLTISSDVSDAMTDH EMKGQTALVAIDGVLCGMIAIADAVKQEAALAVHTLQSMGVDDV LITGIDNRKTARAIATQVGINKVFAEVLPSHKVAKVQELQNKGGKVA MVG DGVNDSPALAQADMGVAIGTGT DVAIEAADVVLIRNDLLDVV ASIHLSKRTVRRIRINLV LALIYNLVGIPIAAGVFMPIGIVLQPWMGS AAMAASSVS VVSSLQLKCYKKPDLERYEAQAHGHMKPLTASQVS VHIGMDDRWRDSPRATPWDQVSYSVQVSLSSLTSDKPSRHSAAAD DDGDKWSLLLNGRDEEQYI	ATP7B
134	MATRSPGVVISDDEPGYDLDFCIPNHYAEDLERVFIPHGLIMDRTE RLARDVMKEMGGHHIVALCVLKG GYKFFADLLDYIKALNRNSDRS IPMTVD FIRLKS YCNDQSTGDIKVIGGDDLSTLTGKNVLIVEDIIDTG KTMQTLLSLVRQYNPKMVKVASLLVKRTPRSVGYKPDFVGF EIPDK FVVGYALDYNEYFRDLNHVCV ISETGKAKYKA	HPRT1
135	MGE PGQSPSPRSSHGSPPTLSTL TLLLLCGHAHSQCKILRCNAEYVS STLSLRGGGSSGALRGGGGGGRGGGVSGGLCRALRSYALCTRRT ARTCRGDLAFHSAVHGIEDLMIQHNC SRQGPTAPPPRGPALPGAGS GLPAPDPCDYEGRFSRLHGRPPGFLHCASF GDPHVRSFHHHFHTCR VQGA WPLLDNDFLFVQATSSPMALGANATATRKLTIIFKNMQECID QKVYQAEVDNLPVAFEDGSINGGDRPGGSSLSIQTANPGNHVEIQA AYIGTTIIIRQTAGQLSFSIKVAEDVAMAFSAEQDLQLCVGGCPPSQR LSRSERNRRGAITIDTARRLCKEGLPVEDAYFHSCVFDVLISGDPNFT VAAQALEDARAFLPDLEKLHLF PSDAGVPLSSATLLAPLLSGLFVL WLCIQ	HJV

136	MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGAR ASWMPMFQRRRRRDTHFPICIFCCGCCCHRSKCGMCCKT	HAMP
137	MRSPTTRGRSGRPLSLLLALLCALRAKVCASGQFELEILSMQNVN GELQNGNCCGGARNPGDRKCTRDECPTYFKVCLKEYQSRVTAGGP CSFGSGSTPVIGGNTFNLKASRGNDRNRIVLPFSFAWPRSYTLLVEA WDSSNDTVQPDSIIEKASHSGMINPSRQWQTLKQNTGV AHFEYQIR VTCDDYYYGFGCNKFCRPRDDFFGHYACDQNGNKTCEGWMGPE CNRAICRQGCSPKHGSKLPDGCRCQYGWQGLYCDKCIPHPGCVH GICNEPWQCLCETNWGGQLCDKDLNYCGTHQPCLNGGTCSNTGPD KYQCSCPEGYS GPNCEIAEHACLSDPCHNRGSKETS LGFECECSPG WTGPTCSTNIDDCSPNNCSHG GTCQDLVNGFKVCPPQWTGKTCQ LDANECEAKPCVNAKSCKNLIASYYCDCLPGWMGQNC DININDCL GQCQNDASCRDLVNGYRCICPPGYAGDHCERDIDECASNPLNGG HCQNEINRFQCLCPTGFSGNLCQLDIDYCEPNPCQNGAQCYNRASD YFCKCPEDYEGKNCSHLKDHCRTPCEVIDSCTVAMASNDTPEGVR YISSNVC GPHGKCKSQSGGKFTCDCNKGFTGT YCHENINDCESNPC RNGGT CIDGVNSYKCICSDGWEGAYCETNINDCSQNPCHNGGT CRD LVNDFYCDCKNGWKGTCHSRDSQCDEATC NNGGTCYDEGD AFK CMCPGGWEGTTCNIARNSSCLPNPCHNGGTCVVNGESFTCVCKEG WEGPICAQNTNDCSPHPCYNSGTCVDGDNWYRCECAPGFAGPDCR ININECQSSPAFGATCVDEINGYRCVCPPGHSGAKCQEVSGRPCIT MGSVIPDGAKWDDDCNTCQCLNGRIACSKVWCGPRPCLLHKGHSE CPSGQSCIPILDDQCFVHPCTGVGECRSSSLQPVKTKCTSDSY YQDN CANITFTFNKEMMSPGLTTEHICSELRLN LNILKNVSAEYSIYIACEPSP SANNEIHVAISAEDIRDDGNPIKEITDKIIDLVSKRDGNSSLIAAVA EV RVQRRPLKNRTDFLVPLLSSVLTVAWICCLVTA FYWCLRRRKPGS HTHSASEDNTTNVREQLNQIKNPIEKHGANTVPIKDYENKNSKMS KIRTHNSEVEEDMDKHQQKARFAKQPAYTLVDREEKPPNGTPTK HPNWTNKQDNRDLESAQSLNRMEYIV	JAG1
138	MASHRLLLLCLAGLVFVSEAGPTGTGESKCPMLVKVLDAVRGSPAI NVAVHVFRKAADDTWEPFASGKTSSESGELHGLTTEEEFVEGIYKVE IDTKSYWKALGISPFHEHA EVVFTANDSGPRRYTIAALLSPYSYSTT AVVTNPKE	TTR
139	MASHKLLVTPPKALLKPLSIPNQLLLGPGPSNLPPRIMAAGGLQMIG SMSKDMYQIMDEIKEGIQYVFQTRNPLTLVISGSGHCALEAALVNV LEPGDSFLVGANGIWGQRAVDIGERIGARVHPMTKDPGGHYTLQEV EEGLAQHKPVLLFLTHGESSTGVLQPLDGFGE LCHRYKCLLLVDSV ASLGGTPLYMDRQGIDILYSGSQKALNAPPGTSLISFSDKAKKKMYS RKTKPFSFYLDIKWLANFWGCDDQPRMYHHTIPVISLYSLRESLALI AEQGLENSWRQHREAAAYLHGRLQALGLQLFVKDPALRLPTVTTV AVPAGYDWRDIVSYVIDHFDIEIMGGLGPSTGKVLRI GLLGCNATRE NVDRVTEALRAALQHCPKKKL	AGXT
140	MKMRFLGLVVCLVLWTLHSESGGKLTAVDPETNMNVSEIISYWG FPSEEYL VETEDGYILCLNRIPHGRKNHSDKGPKPVVFLQHGLLADS SNWVTNLANSSLGFIADAGFDVWMGNSRGNTWSRKHKTL SVSQD EFWAFSYDEMAKYDLPASINFILNKTGQE QVYVYVGHSGQTTIGFIAF SQIPELAKRIKMFFALGPVASVAFCTSPMAKLGRLPDHLIKDLFGDK EFLPQSAFLKWLGHVCTHVILKELCGNLCFLLCGFNERNLNMSRV DVYTTHSPAGTSVQNMLHWSQAVKFQKQAFDWGSSAKNYFHYN QSYPTTYNVKDMLVPTAVWSGGHDWLADVYDVNILLTQITNLVFH ESIPEWEHLDFIWGLDAPWRLYNKIINLMRKYQ	LIPA

141	MASRLTLLTLLLLLAGDRASSNP NATSSSSQDPESLQDRGEGKVAT TVISKMLFVEPILEVSSLPTNSTTNSATKITANTTDEPTTQPTTEPTT QPTIQPTQPTTQLPTDSPTQPTTGSFCGPGVTLCSDLSESHSTEAVLGD ALVDFSLKLYHAFSAMKKVETNMAFSPFSIASLLTQVLLGAGENTK TNLESILSYPKDFTCVHQALKGFTTKGVTSVSQIFHSPDLAIRDFTVN ASRTLYSSSPRVLSNNSDANLELINTWVAKNTNNKISRLLDSLPSDT RLVLLNAIYLSAKWKTTFDPKKTRMEPFHFKNSVIKVPMMNSKKYP VAHFIDQTLKAKVGQLQLSHNLSLVILVPQNLKHRLEDMEQALSPS VFKAIMEKLEMSKFQPTLLTLPRIKVTTSQDMLSIMEKLEFFDFS YD LNLCLGTEDPDLQVSAMQHQT VLELTETGVEAAAASAVISARTLLV FEVQQPFLFVLWDQQHKFPVFMGRVYDPRA	SERPING1
142	MGSPLRFDGRVVLVTGAGAGLGRAYALAFERGA LVVVNDLGGD FKGVGKGLAADKVVEIRRRGGKAVANYDSVEEGEKVVKTALDA FGRIDVVVNAGILRDRSFARISDEDWDIIHRVHLRGSFQVTRAWE HMKKQKYGRIIMTSSASGIYGNFGQANYSAAKLGLLGLANSLAIEG RKSNIHCNTIAPNAGSRMTQTVMPEDLVEALKPEYVAPLVWLCH E SCEENGLFEVGAGWIGKLRWERTLGAIVRQKNHPMTPEAVKANW KKICDFENASKPQSIQUESTGSIIEVLSKIDSEGGVSANHTSRATSTATS GFAGAIGQKLPPFSYAYTELEAIMYALGVGASIKDPKDLKFIYEGSS DFSCLPFTGVIIGQKSMGGGLAEIPGLSINFVKVLHGEQYLELYKP LPRAGKLKCEAVVADVLDKSGGVVIMDVYSYSEKELICHNQFSLF LVGSGGFGGKRTSDKVKVAVAI PNRPDAVLTDTTSLNQAALYRLS GDWNPLHIDPNFASLAGFDKPIHLGLCTFGFSARRVLQQFADNDVS RFKAIAKARFAKPVYPGQTLQTEMWKEGNRIHFQTKVQETGDIVISN AYVDLAPTSGETSAKTPSEGGKLQSTFVFEEIGRRLKDIGPEVVKKVN AVFEWHITKGGNIGAKWTIDLKSGSGKVYQGPAGKAADTTIILSDE DFMEVVLGKLDPOKAFPSGRLKARGNIMLSQKLQMLKDYAKL	HSD17B4
143	MEANGLGPQGFPELKNDTFLRAAWGEETDYTPVWCMRQAGRYLP EFRETAAQDFSTCRSPEACCELTLQPLRRFPLDAAIIFSDILVVPQA LGMEVTMVPGKGPSFPEPLREEQDLERLRDPEVVASELGYVFQAITL TRQRLAGRVPLIGFAGAPWTLMTYMEVGGGSSTMAQAKRWLYQR PQASHQLLRILTALVPYLVGVVAGAQAQLQLFESHAGHLGPQLFN KFALPYIRDVAKQVKARLREAGLAPVPMIIFAKDGHFALEELAQAG YEVVGLDWTVA PKKARECVGKTVTLQGNLDPICALYASEEEIGQLV KQMLDDFGPHRYIANLGHGLYPDMDPEHVGA FVDVAVHKHSRLLR QN	UROD
144	MGPRARPALLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGLSL FEALGYVDDQLFVFDHESRRVEPRTPWVSSRISSQMWLQLSLSLK GWDHMFVTDFWTIMENHNHSHKESHTLQVILGCEMQEDNSTEGYW KYGYDGDHLEFCPDTLDWRAAEPRAPWPTKLEWERHKIRARQNR AYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTTSSVTTLRCR ALNYYPNITMKWLKDKQPMDAKEFEKPDVLPNGDGTYYQGWITL AVPPGEEQRYTCQVEHPGLDQPLIVIEWPSPSGTLVIGVISGIAVFVVI LFIGILFIILRKRQGSRGAMGHYVLAERE	HFE
145	MESKALLVLTAVWLQSLTASRGGVAAADQRRDFIDIESKFALRTP EDTAEDTCHLIPGVAESVATCHFNHSSKTFMVIHGWTVTGMYESW VPKLVAALYKREPDSNVIVVDWLSRAQEHYPVSAGYTKLVGQDVA RFINWMEEEFNYPLDNVHLLGYSLGAHAAGIAGSLTNKKVNRTGL DPAGPNFEYAEAPSRLSPDDADFDVLTFTTRGSPGRSIGIQKPVGH VDIYPNGGTGTFQPGCNIGEAIRVIAERGLGDVDQLVKCSHERSIHLFID SLLNEENPSKAYRCSSKEAFEKGLCLSCRKNRCNNLGYEINKVRAK RSSKMYLKTRSOMPYKVFHYQVKIHFSGTESETHNQAFEISLYGT	LPL

	VAESENIPFTLPEVSTNKTYSFLLIYTEVDIGELMLKLKWKSDSYFS WSDWWSSPGFAIQKIRVKAGETQKKVIFCSREKVSHLQKGKAPAVF VKCHDKSLNKKSG	
146	MRPVRLMKVFVTRRIPAEGRV ALARAADCEVEQWDSDEPIPAKELE RGVAGAHGLCLLSDHVDKRILDAAGANLKVISTMSVGIDHLALDE IKKRGRIRVGYTPDVLDTTAEALVSLLLTTCRRLPEAIEEVKNGGWT SWKPLWLCGYGLTQSTVGIIGLGRIGQAIARRLKPFQVQRFLYTGRQ PRPEEAAEFQAEFVSTPELAAQSDFIVVACSLTPATEGLCNKDFQK MKETA VFINISRGDVVNQDDLYQALASGKIAAAGLDVTSPEPLPTN HPLLTLKNCVILPHIGSATHRTRNTMSLLAANNLLAGLRGEPMPSEL KL	GRHPR
147	MLGPQVWSSVRQGLSRSLSRNVGVWASGEGKKVDIAGIYPPVTTPF TATAEVDYDKLEENLHKLGTFFRFGFVVQGSNGEFPFLTSSERLEV SRVRQAMPKNRLLLAGSGCESTQATVEMTVSMAQVGADAAMVVT PCYYRGRMSSAALIHHTYTKVADLSPIPVVLYSVPANTGLDLPVDAV VTLSQHPNIVGMKDSGGDVTRIGLIVHKTRKQDFQVLAGSAGFLMA SYALGAVGGVCALANVLGAQVCQLERLCCTGQWEDAQKLQHRLE PNAAVTRRFIPGLKKIMDWFGYYGGPCRAPLQELSPAEEEEALRMD FTSNGWL	HOGA1
148	MGPWGWKLRTVALLLAAAGTAVGDRCERNEFQCQDGKCSISK WVCDGSAECQDGSDESQETCLSVTCKSGDFSCGGRVNRCPQFWRC DGQVDCDNGSDEQGCPPKTCQDEFCHDGKCSIRQFVCDSDRDCL DGSDEASCPVLTGCPASFQCNSSTCIPQLWACDNDPDCEDGSDEWP QRCRGLYVFQGDSSPCSAFEFHCLSGECIHSSWRCDGGPDCKDKSD EENCAVATCRPDEFQCSGNCIHGSRQCDREYDCKDMSDEVGCVN VTLCGPKNFKCHSGECITLDKVCNMARDCRDWSDEPIKECGTNEC LDNNGGCSHVCNDLKIGYECLCPDGFQLVAQRRCEDIDECQDPDTC SQLCVNLEGGYKQCCEEGFQLDPHTKACKAVGSIAYLFFTNRHEVR KMTLDRSEYTSIPNLRNVVALDTEVASNRIYWSDLQRMICSTQLD RAHGVSSTYDVISRDIQAPDGLAVDWIHSNIYWTDSVLGTVSVADT KGVKRKTLFRENGSKPRAIVVDPVHGFMWYTDWGTAPAKIKKGLN GVDIYSLVTENIQWPNGITLDDLGRLYWVDSKLHSSIDVNGNR KTILEDKRLAHPFSLAVFEDKVFWDIINEAIFSANRLTGSDVNLLA ENLLSPEDMVLFNLTQPRGVNWCERTTSLNNGGCQYLCLPAPQINP HSPKFTCACPDGMILLARDMRSCLTEAEAAVATQETSTVRLKVSSTA VRTQHTTTTRPVPDTSRLPGATPGLTTVEIVTMSHQALGDVAGRGNE KKPSSVRALSIVLPIVLLVFLCLGVFLLWKNWRLKNINSINFDPVY QKTTEDEVHICHNQDGYSYPSRQMVSLEDDVA	LDLR
149	MLWSGCRRFGARLGCLPGGLRVLVQTGHR LTSCIDPSMGLNEEQKEFQKVAFDFAAREM APNMAEWDQKELFPVDVMRKAALQGFGGVY IQTVDVGGSGLSRLDTSVIFEALATGCTSTT AYISIHNMCAWMIDSFSGNEE QRHKFCPLCTMEKFASYCLTEPGSGSDAA SLLTSAKKQGDHYILNGSKAFISGAGESDI YVVMCRTGGPGPKGISCIVVEKGTPLSFG KKEKKVGWNSQPTRAVIFEDCAVPVANRIG SEGQGLIAVRGLNGGRINIASCSLGAHA SVILTRDHLNVRKQFGEPLASNQYLQFTLADMATRLVAARLMVRN AAVALQEERKDAVALCSMAKLFATDECFAICNQUALQMHGGYGYL KDYAVQQYVRDSRVHQILEGSNEVMRILISRSLLQE	ACAD8

150	MEGLAVRLLRGSRLRRNFLTCLSSWKIPPHVSKSSQSEALLNITNN GIHFAPLQTFDEEMMIKSSVKKFAQEIQIAPLVSTMDENSKMEKSVI QGLFQQGLMGIEVDPEYGGTGASFLSTVLVIEELAKVDASVAVFCEI QNTLINTLIRKHGTEEQKATYLPQLTTEKVGFSFCLSEAGAGSDSFAL KTRADKEGDYYVLNGSKMWISSAEHAGLFLVMANVDPTIGYKGIT SFLVDRDTPGLHIGKPENKLGLRASSTCPLTFENVKVPEANILGQIGH GYKYAIGSLNEGRIGIAAQMGLAQGCDFYTIPIYKERIQFGKRLDFD QGLQHQAHVATQLEAARLLTYNAARLLEAGKPFKEASMAKYA SEIAGQTTSKCIEWMGGVGYTKDYPVEKYFRDAKIGTIYEGASNIQL NTIAKHIDAAY	ACADSB
151	MAVLAALLRSGARSRSPLLRRLVQEIRYVERSYVSKPTLKEVVIVSA TRTPIGSFLGSLSPATKLGSAIQGAIEKAGIPKEEVKEAYMGNVL QGGEGQAPTRQAVLGAGLPSTPCTTINKVCASGMKAIMMASQSLM CGHQDVMVAGGMESMSNVPMNRGSTPYGGVKLEDLIVKDGLT DVYNKIHMGSCAENTAKKLNIARNEQDAYAINSYTRSKAAWEAGK FGNEVIPVTVTVKGQPDVVVKEDEEYKRVDFSKVPKLKTVFQKEN GTVTAANASTLNDGAAALVMTADAARKRLNVTPLARIVAFADA EPIDFPIAPVYAASMVLDKDVGLKKEDIAMWEVNEAFSLVVLANIKM LEIDPQKVNINGGAVSLGHPIGMSGARIVGHLTHALKQGEYGLASIC NGGGGASAMLIQKL	ACAT1
152	MLPHVVLTFRRLGALASCRLAPARHRGSGLLHTAPVARSDRSAPV FTRALAFGDRIALDQHGRHTYRELYSRSLRSLQEICRLCGCVGGDLR EERVVSFLCANDASYVVAQWASWMSGGVAVPLYRKHPAAQLEYVI CDSQSSVVLASQEYLELLSPVVRKLGVPPLPLTPAIYTGAVEEPAEV PVPEQGWRNKGAMIIYTS GTTGRPKGVLS THQNIRAVVTGLVHKW AWTKDDVILHVLPLHHVHGVVNALLCPLWVGATCVMMPFESPQQ VWEKFLSSETPRINVFMAVPTIYTKLMEYYDRHFTQPHAQDFLRV CEEKIRLMVSGSAAALPLVLEKWKNTGHTLLERYGMTEIGMALSG PLTTAVRLPGSVGTPLPGVQVRIVSENQREACSYTHAEGDERGTK VTPGFEEKEGELLVRGPSVFREYWNKPEETKSAFTLDGWFKTGDTV VFKDGQYWIRGRTSVDIKTGGYKVSALVEVHLLAHPSITDVAVIG VPDMTWGQRVTA VVTLREGHSLSHRELKEWARNVLAPYAVPSELV LVEEIPRNQMGKIDKKALIRHFHPS	ACSF3
153	MTSCHIAEEHIQKVAIFGGTHGNELTG VFLVKHWLENGAEIQR TGLE VKPFITNPRAVKKCTRYIDCDLNRIFDLENLGKKMSEDLPYEVRR AQ EINHFLFGPKDSEDSYDIIFDLHNTTSNMGCTLILED SRNNFLIQMFHYI KTS LAPLPCYVYLIEHPSLKYATTRSIAKYPVGIEVGPQPQGVLRADI LDQMRKMIKHALDFIHHFNEGKEFP PCAIEVYKII EKVDYPRDENG E IA AIIHPNLQDQDWKPLHPGDP MFLTLDGKTIPLGGDCTVYPV FVNEA AYYEKKEAFAKTTKLT LNAKSIRCC LH	ASPA
154	MAAAVAAAPGALGSLHAGGARLVAACSAWLCPGLRLPGSLAGRR AGPAIWAQGWVPAAGGPAPKRGYSSEMKTEDLRVRHLEENRGI VVLGINRAYGKNSLSKNLIKMLSKAVDALKSDKKVRTIIRSEVPGIF CAGADLKERAKMSSEVGPFVSKIRAVINDIANLPVPTIAAIDGLAL GGGLELALACDIRVAASSAKMGLVETKLAII PGGGGTQRLPRAIGMS LAKELIFSARVLDGKEAKAVGLISHVLEQNQEGDAAYRKALDLARE FLPQGPVAMRVAKLAINQGMEVDLVTGLAIEEACYAQTIPTKDRLE GLLAFKEKRPPRYKGE	AUH

155	MASTVVAVGLTIAAAGFAGRYVLQAMKHMPEQVKQVFQSLPKSAF SGGYRGGFEPKMTKREAALILGVSP TANKGKIRDAHRRIMLLNHP DKGGSPYIAAKINEAKDLLEGQAKK	DNAJC19
156	MAEAVLRVARRQLSQRGGSGAPILLRQMFEPVSCTFTYLLGDRESR EAVLIDPVLETAPRDAQLIKELGLRLLYAVNTHCHADHITGSGLLRS LLPGCQSVISRLSGAQADLHIEDGDSIRFGRFALETRASPGHTPGCVT FVLNDHSMMAFTGDALLIRGCGRTDFQQGCAKTLYHSVHEKIFTLP DCLIYPAHDYHGFTVSTVEEERTLNPRLTLSCEEVVKIMGNLNLKP QQIDFAVPANMRCGVQTPTA	ETHE1
157	MADQAPFDTDVNTLTRFVMEEGRKARGTGELTQLLNSLCTAVKAIS SAVRKAGIAHLYGIAGSTNVTGDQVKKLDVLSNDLVMNMLKSSFA TCVLVSEEDKHAHIVEPEKRGKYVVCDFPLDGSSNIDCLVSVGTIFGI YRKKSTDEPSEKDALQPGRNVAAGYALYGSATMLVLAMDCGVN CFMLDPAIGEFILVDKDVKIKKKGKIYSLNEGYARDFDPAVTEYIQR KKFPPDNSAPYGARYVGSMVADVHRTL VYGGIFLYPANKKSPNGK LRLLYECNPMAYVMEKAGGMATTGKEAVLDVIPTDIHQRAPVILGS PDDVLEFLKVYEKHSQAQ	FBP1
158	MSQLEVCVPNFSEGNQEVIDAISGAITQTPGCVLLDVDAGPSTNRT VYTFVGPPECVVEGALNAARVASRLIDMSRHQGEHPRMGALDVC FIPVRGVSVDECVLCAQAFGQRLAEELDVPVYLYGEAARMDSRRTL PAIRAGEYEALPKKLQQADWAPDFGPSSFVPSWGATATGARKFLIA FNINLLGTKEQAHRIALNLREQGRGKDQPGRLKKVQGIGWYLDEKN LAQVSTNLLDFEVTALHTVYEETCREAQELSLPVVGSQVLVGLVPLK ALLDAAAFYCEKENLFILEEEQRI RLVVSRLGLDSLCPFSPKERIIEYL VPERGPERGLGSKSLRAFVGEVG ARSAAPGGGSVAAAAAAMGAALGSMVGLMTYGRRQFQSLDTTMR RLIPPFREASAKLTTLVDADAEFTAYLEAMRLPKNTPEEKDRRTA ALQEGLRRAVSVPPLTAETVASLWPALQELARCGNLACRSDLQVA AKALEMGVFGAYFNVLINLRDITDEAFKDQIHHRVSSLLQEAKTQA ALVLDCCLETRQE	FTCD
159	MATNWGSLLQDKQQLLEELARQAVDRALAEGLVLLRTSQEPTSSEVV SYAPFTLFPVSLVPSALLEQAYAVQMDFNLLVDAVSQNAAFLEQTL STIKQDDFTARLFDIHKQVLKEGIAQTVFLGLNRSYMFQRSADGSP ALKQIEINTISASFGGLASRTPAVHRHVLVLSKTKEAGKILSNNPSK GLALGIAKAWELYGSPNALVLLIAQEKERNIFDQRAIENELLARNIH VIRRTFEDISEKGSLDQDRRLFVDGQEIADVYFRDGYMPPRQYSLQN WEARLLLSHAACPDIAQLAGTKKVQQELSRPGMLEMLLPQG PEAVARLRATFAGLYSLDVGEEDQAIAEALAAPSRFVLKPQREGG GNNLYGEEMVQALKQLKDSEERASYILMEKIEPEPFENCLLRPGSPA RVVQCISELGIFGVYVRQEKTLMNKHVGHLLRTKAIEHADGGVA AGVAVLDNPYPV	GSS
160	MGQREMWRLMSRFNAFKRTNTILHHLRMSKHTDAAEEVLLEKKG CTGVITLNRPKFLNALTLMIRQIYPQLKKWEQDPETFLIIKGAGGK AFCAGGDIRVISEAEKAKQKIAPVFFREEYMLNNAVGSQKPYVALI HGITMGGGVGLSVHGQFRVATEKCLFAMPETAIGLFPDVGGGYFLP RLQGLGYFLALTGFRLKGRDVYRAGIATHFVDSEKLAMLEEDLLA LKSPSKENIASVLENYHTESKIDRDKSFILEEHMDKINSCFSANTVEEI	HIBCH

	IENLQQDGSSFALEQLKVINKMSPTSLKITLRQLMEGSSKTLQEVLT MEYRLSQACMRGHDHFEHGVRAVLIDKDQSPKWKPADLKEVTEEDL NNHFKSLGSSDLKF	
161	MAGYLRVVRSRCRASGRPAWAPAALTAPTSQEQRPRHYADKRIK VAKPVVEMDGDDEMTRIOWQFIKEKLILPHVDIQLKYFDLGLPNRDQT DDQVTIDSALATQKYSVAVKCATITPDEARVEEFKLKKMWKSPNG TIRNILGGTVFREPIICKNIPRLVPGWTKPITIGRHAHGDQYKATDFV ADRAGTFKMVFTPDKGSGVKEWEVYNFPAGGVGMGMYNNTDESIS GFAHSCFYQAIQKKWPLYMSTKNTILKAYDGRFKDIFQEIFDKHYK TDFDKNKIWEHRLIDDMVAQVLKSSGGFVWACKNYDGDVQSDIL AQGFGSLGLMTSVLVCPDGKTI EAEEAAHGTVTRHYREHQKGRPTST NPIASIFAWTRGLEHRGKLDGNQDLIRFAQMLEKVCVETVESGAMT KDLAGCIHGLSNVKNLNEHFLNTTDFLDTIKSNLDRALGRQ	IDH2
162	MVPALRYLVGACGRARGLFAGGSPGACGFASGRPRPLCGGSRASST SSFDIVIVGGGIVGLASARALILRHPSLSIGVLEKEKDLAVHQTGHNS GVIHSGIYYKPESLKAALCVQGAALLYEYCCQKGISYKQCCKLIVA VEQEEIPRLQALYEKGLQNGVPLRLIQQEDIKKKEPYCRGLMAIDC PHTGIVDYRQVALSFAQDFQEAGGSVLTNFEVKGIEMAKESPSRSID GMQYPIVINKTKGEEIRCQYVVTCAGLYSDRIELSGCTPDPRIVPFR GDYLLLKPEKCYLVKGNIPVPDSRFPFLGVHFTPRMDGSIWLGPN AVLAFKREGYRPFDFSATDVMIIINSGLIKLASQNFSYGVTEMYKA CFLGATVKYLQKFIPEITISDILRGPAGVRAQALDRDGNLVEDFVFD AGVGDIGNRILHVRNAPSPAATSSIAISGMIADDEVQQRFE	L2HGDH
163	MRGFGPGLTARRLLPLRLPPRPPGPRLASGQAAGALERAMDELLRR AVPPTPAYELREKTPAPAEGQCADFVSFYGGLAETAQRAELLGRLA RGFGVDHGGVVAEQSAGVLHLRQQQREAAVLLQAEDRLRYALVPR YRGLFHHISKLDGGVRFLVQLRADLLEAQALKLVEGPDVREMNGV LKGMLSEWFSSGFLNLERVTWHSPCEVLQKISEAEAVHPVKNWMD MKRRVGPYRRCYFFSHCSTPGEPLVVHLVALTGDISSNIQAIVKEHP PSETEENKITAIFYSISLTQQGLQG VELGTFLIKRVVKELQREFPHLGVFSSLSPIPGFTKWLLGLLNSQTKE HGRNELFTDSECKEISEITGGPINETLKLSSSEWVQSEKLVRALQT PLMRLCAWYLYGEKHRGYALNPVANFHLQNGAVLWRINWMADV SLRGITGSCGLMANYRYFLEETGPNSTSYLGSKIHKASEQVLSLVAQF QKNSKL	MLYCD
164	MVVGAFPMAKLLYLGIQVSKPLANRIKEAARRSEFFKTYICLPPAQ LYHWVEMRTKMRIMGFRGTVIKPLNEEAAAELGAELLGEATIFIVG GGCLVLEYWRHQAQQRHKEEQRAAWNALRDEVGHLALALEALQ AQVQAAPPQGALEELRTELQEVRAQLCNPGRSASHAVPASKK	OPA3
165	MGSPEGRFHF AIDRGGTFTDVFAQCPGGHVRVLKLLSEDPANYADA PTEGIRRIEQUEAGMLLPRDQPLDSSHASIRMGTTVATNALLERKGE RVALLVTRGFRDLLHIGTQARGDLFDLAVPMPEVLYEEVLEVDERV VLHRGEAGTGTPVKGRGTGDLLEVQQPVDLGALRGKLEGLLSRGIRS LAVVLMHSYTWAEHQVGVLAELGFTHVSLSSEAMPVVRIVPR GHTACADAYLTPAIQRYVQGFCRGFQQLKDVQVLFMRSDGGLAP MDTFSGSSAVLSGPAGGVVGYSATTYQEGGQPVIGFDMGGTSTD VSRYAGEFEHVFEASTAGVTLQAPQLDINTVAAGGGSRLFFRSGLF VVGPEASAGAHGPACYRKGGPVTVTDANLVLGRLLPASFPICFGPG	OPLAH

	<p>ENQPLSPEASRKALEAVATEVNSFLTNGPCPASPLSLEEVAMGFVRV ANEAMCRPIRALTQARGHDPSAHVLACFGGAGGQHACAIARALGM DTVHIHRHSGLLSALGLALADVVEAHEAQEPSCSLLYAPETFVQLDQRL SRLEEQCVDALQAQGFPRSQISTESFLHLRYQGTDCALMVSAHQHP ATA RSPRAGDFGAAFVERYMREFGFVIPERPVVDDVRVRGTGRSGLRL EDAPKAQTGPPRVDKMTQCYFEGGYQETPVYLLAELGYGHKLHGP CLIIDNSTILVEPGCQAEVTKTGDISVGAEPGTVPQDPIQLSIF SHRFMSIAEQMGRILQRTAISTNIKERLDFSCALFGPDGGLVSNAPHI PVHLGAMQETVQFQIQHLGADLHPGDVLLSNHPSAGGSHLPDLTVI TPVFWPGQTRPVFYVASRGHHADIGGITPGSMPPHSTMLQQEGAVF LSFKLVQGGVFQEEAVTEALRAPGKVPNCSGTRNLHDNLSDLRAQ VAANQKGIQLVGELIGQYGLDVVQAYMGHIQANAELAVRDMRLRAF GTSRQARGLPLEVSEDMDDGSPIRLRVQISLSQGSVDFDSGTGP EVFGNLNAPRAVTLASALIYCLRCLVGRDIPLNQGCLAPVRVVIPRGS LDPSPEAAVVGGNVLTSSQRVVDVILGAFGACAASQGCMMNNVTLGN AHMGYYETVAGGAGAGPSWHGRSGVHSHMTNTRITDPEILESRYP VILRRFELRRSGGRGRFRGGDGVTRRELLFREEALLSVLTERRAFRP YGLHGGEPGARGNLIRKNGRTVNLGGKTSVTVPYPGDVFLHTPG GGGYGDPEDPAPPPGSPQALAFPEHGSVYVEYRRAQEA</p>	
166	<p>MAALKLLSSGLRLCASARGSGATWYKGCVCFSSTSAHRHTKFYTD PVEAVKDIPDGATVLVGGFGLCGIPENLIDALLKTGVKGLTAVSNN AGVDNFGLLGLLSKQIKRMVSSYVGENAEFERQYLSGELEVELTP QGTLAERIRAGGAGVPAFYTPTGYGTLVQEGGSPIKYNKDGSVAIA SKPREVREFNGQHFILEEAITGDFALVKAWKADRAGNVIFRKSARN FNLPMCKAAETTVVEVEEIVDIGAFAPEDIHIPQIYVHRLIKGEKEYE RIERLSIRKEGDGEAKSAKPGDDVRERIIRKRAALEFEDGMYANLIGI PLLASNFI PNITVHLQSENGVLGLGPYPRQHEADADLINAGKETVTI LPGASFFSSDES FAMIRGGHVDLTMLGAMQVSKYGD LANWMIPGK MVKGMGGAMD LVSSAKTKVVVTMEHSAKGNNAHKIMEKCTPLTG KQCVNRIITEKAVFDVDKKGGLTIELWEGLTVDDVQKSTGCDFAV SPKLMMPMQIAN</p>	OXCT1
167	<p>MSRLLWRKVAGATVGPVPAPGRWVSSSVPASDPSDGQRRRQQQ QQQQQQQQQQPQQPQVLSSEGGQLRHNPLDIQMLSRGLHEQIFGQG GEMPGEAAVRRSVEHLQKHGLWGQPAVPLPDVELRLPPLYGDNLD QHFRLLAQKQSLPYLEAANLLLQAQLPPKPPAWAWAEGWTRYGPE GEAVPVAIPEERALVFDVEVCLAEGTCPTLAVAISSPAWYSWCSQR LVEERYSWTSQLSPADLIPLEVPTGASSPTQRDWQEQLVVGHNVSF DRAHIREQYLIQGSRMFLDTMSMHMAISGLSSFQRLWIAAKQKG HKVQPPTKQGGKSQRKARRGPAISSWDWLDISSVNSLAEVHRLYV GGPPLEKEPRELFVKGTMKDIRENFQDLMQYCAQDVWATHEVFQQ QLPLFLERC PHPVTLAGMLEMGVSYLPVNQNWERYLAEAQGTYYE LQREMKKSLMDLANDACQLLSGERYKEDPWLWDLEWDLQEFKQK KAKKVKKEPATASKLPIEGAGAPGDPMDQEDLGPCSEEEEFQQDV MARACLQKLKGTTELLPKRPQHLPGHPGWYRKLCPRLDDPAWTPG PSLLSLQMRVTPKLMALTWDGFPLHYSERHWGYLVPGRRDNLAK LPTGTTLESAGVVC PYRAIESLYRKHCLEQKQQLMPQEAGLAEF LLTDNSAIWQTVEELDYLEVEAEAKMENLRAAVPGQPLALTARGG PKDTQPSYHHGNGPYNDVDIPGCWFFKLPHKDGNSCNVGSFPAKDF LPKMEDGTLQAGPGGASGPRALEINKMISFWRNAHKRISSQM VVW LPRSALPRAVIRHPDYDEEGLYGAILPQVV TAGTITRRAVEPTWLTA SNARPDRVGSELKAMVQAPPGYTLVGADVDSQELWIAAVLGDAHF</p>	POLG

	AGMHGCTAFGWMTLQGRKSRGTDLHSKTATTVGISREHAKIFNYG RIYGAGQPFARLLMQFNHRLTQQEAAEKAQQMYAATKGLRWYR LSDEGEWL VREL NLPVDRTEGGWISLQDLRKVQRETARKSQWKKW EVVAERAWKGGTESEMFNKLESIATSDIPRTPVLGCCISRALEPSAV QEEFMTSRVNWVVQSSAVDYHLHMLVAMKWLFEFEAIDGRFCISIH DEVRYLVREEDRYRAALALQITNLLTRCMFAYKLGLNDLPQSVAFF SAVDIDRCLRKEVTMDCKTPSNPTGMERRYGIPQGEALDIYQHIELT KGSLEKRSQPGP	
168	MSTAALITLVRSGGNQVRRRVLLSSRLLQDDRRVTPTCHSSTSEPRC SRFDPDGS SPATWDNFGIWDNRIDEPIILLPSIKY GKPIPKISLENVG CASQIGKRKENEDRFDFACLTDVLYFAVYDGHGGPAAADFCHTH MEK CIMDLLPKEKNLETLTLLAFLEIDKAFSSHARLSADATLLTSGT TATVALLRDGIELVVASVGDSRAILCRKGKPMKLTIDHTPERKDEKE RIKKCGGFVAWNSLGQPHVNGRLAMTRSIGDLDLKTSGVIAEPETK RIKLHHADDSFLVLTDDGINFMVNSQEICDFVNQCHDPNEAAHAVT EQAIQYGTEDNSTAVVVPFGAWGKYKNSEINFSSRSFASSGRWA	PPMIK
169	MSLAA YCVICCRIGTSTSPKSGTHWRDIRNIIKFTGSLILGGSFLT YEV LALKKAVTLDTQVVEREKMSYIYVHTVSLDKGENHGIWQA RKELHKAVRKVLATS AKILRNPFADPFSTVDIEDHECAVWLLRKS KSDDKTTRLEAVREMSETHHWHDYQYRIIAQACDPKTLIGLARSEE SDLRFFLLPPPLPSLKEDSSTEEELRQLLASLPQTELDECIQYFTSLAL SESSQ SLAAQKGGWLWCFGGNGLPYAESFGEVPSATVEMFCLEAIVKHSEIST HCDKIEANGGLQLLQRLYRLHKDCPKVQRNIMRVIGNMALNEHLH SSIVRSGWVSIMAEAMKSPHIMESSHAARILANLDRET VQEKYQDG VYVLHPQYRTSQPIKADVLFIHGLMGA AFKTRWQQDSEQAVIEKPM EDEDRYTTCWPKTWLAKDCPALRIISVEYDTSLSDWRARCPMERKS IAFRSNELLRLKRAAGVGDRPVVWISHSMGGLLVKKMLLEASTKPE MSTVINNTRGIIFYSVPHHGSRLAEYSVNIRYLLFPSLEVKELSKDSP ALKTLQDDFLEFAKDKNFQVLNFVETLPTYIGSMIKLHVVPVESADL GIGDLIPVDVNHLNICKPKKKDAFLYQRTLQFIREALAKDLEN	SERAC1
170	MPAPRAPRALAAAAPASGKAKLTHPGKAILAGGLAGGIEICITFPT YVK TQLQLDERSHPPRYRGIGDCVRQTVRSHGVLGLYRGLSSLLYG SIPKAAVRFGMFEFLSNHMRDAQGRLDSTRGLLCGLGAGVAEAVV VVC PMETIKVKFIHDQTSNP KYRGFFHGVREIVREQGLKGT YQGLT ATVLKQGSNQAIRFFVMTSLRNWYRGDNPKNPMNPLITGVFGAIA AASVFGNTPLDVIKTRMQGLEAHKYRNTWDCGLQILKKEGLKAFY KGTVPRLGRVCLDVAIVFVIYDEV VKLLNKVWKTD	SLC25A1
171	MAASMFYGRVLVAVATLRNHRPRTAQRAAAQVLGSSGLFNNHGLQ VQQQQQRNLSLHEYMSMELLQEAGVSVPKGYVAKSPDEAYAIKK LGSKDVVIKAQVLAGGRGKGT FESGLKGGVKIVFSPEEAKAVSSQM IGKKLFTKQTGEKGRICNQVLVCERKYPRREYYFAITMERSFQGPVL IGSSHGGVNIEDVAESPEAIIKEPIDIEEGIKKEQALQLAQKMGFPPN IVESAAENMVKLYSLFLKYDATMIEINPMVEDSDGAVLCMDAKINF DSNSAYRQKKIFDLQDWTQEDERDKDAAKANLNYIGLDGNIGCLV NGAGLAMATMDIIKLHGGTPANFLDVGGGATVHQVTEAFKLITSDK KVLAILVNIFGGIMRCDVIAQGIVMAVKDLEIKIPVVRLQGTRVDD AKALIADSGLKILACDDLDEAARMVVKLSEIVTLAKQAHVDVKFQL PI	SUCLA2
172	MTATLAAAADIATMVSGSSGLAAARLLSRSFLLPQNGIRHCSYTAS RQHLYVDKNTKIICQGFTGKQGT FHSQQALEYGT KLVG GTTPGKGG	SUCLG1

	QTHLGLPVFNTVKEAKEQTGATASVIYVPPPFAAAAINEAIEAEIPLV VCITEGIPQQDMVRVKHKLLRQEKTRLIGNPCPGVINPGECKIGIMP GHIHKKGRIGIVSRSGTLTYEAVHQTTQVGLGQSLCVGIGGDPFNGT DFIDCLEIFLNDSETEGIIIGGNAEENAAEFLKQHNSGPNKSPVV SFIAGLTAPPGRRMGHAGAIAGGKGGAKEKISALQSAGVVVSMSP AQLGTTIYKEFEKRKML	
173	MPLHVKWPPFPAVPLTWTLASSVVMGLVGTYSFVTKYMNHLTV HNREVL YELIEKRGPATPLITVSNHQSCMDPHLWGILKLRHIWNLK LMRWTPAAADICFTKELHSHFFSLGKCVPCRGAEFFQAENEGKGV LDTGRHMPGAGKRREKGDGVYQKGMDFILEKLNHGDWVHIFPEG KVNMSSEFLRFKWGIGRLIAECHLNPIILPLWHVGMNDVLPNSPPYF PRFGQKITVLIGKPFSAIPVLERLRAENKSAVEMRKALTDIFIQEEFQ HLKTQAEQLHNHLQPGR	TAZ
174	MTVFFKTLRNHWKKTATAGLCLLTWGGHWLYGKHCDNLLRRAACQ EAQVFGNQLIPPNAQVKKATVFLNPAACKGKARTLFEKNAAPILHL SGMDVTIVKTDYEGQAKKLELEMENTDVIIVAGGDGTLQEVVTGV LRRTDEATFSKIPIGFIPLGETSSLSHTLFAESGNKVQHITDATLAIV GETVPLDVLQIKGEKEQPVFAMTGLRWGSFRDAGVKVSKYWYLG LKIKAAHFFSTLKEWPQTHQASISYTGPTERPPNEPEETPVQRPSLYR RILRRLASYWAQPQDALSQEVSPEVWKDVQLSTIELSITTRNNQLDP TSKEDFLNICIEPDTISKGDFITIGSRKVRNPKLHVEGTECLQASQCTL LIEGAGGSFSIDSEEYEAMPVEVKLLPRKLQFFCDPRKREQMLTSPT Q	AGK
175	MLGSLVLRKALAPRLLLRLLRSPTLRGHGGASGRNVTTGSLGEPQ WLRVATGGRPGTSPALFSGRGAATGGRQGGRFDTKCLAAATWGR PGPEETLPQDQSWNGVPSRAGLGMALAAALVVHCYSKSPSNKDA ALLEAARANNMQEVSRLSEGADVNAKHRLGWTALMVAAINRNN SVVQVLLAAGADPNLGGDFSSVYKTAKEQGIHSLEDGGQDGASRHI TNQWTSALFRRWLGLPAGVLITREDDFNNRLNNRASFKGCTALH YAVLADDYRTVKELLDGGANPLQRNEMGHTPLDYAREGEVMKLL RTSEAKYQEKQRKREAEERRRFPLEQRLKEHIIIGQESAIATVGAA IRRKENGWYDEEHPLVFLFLGSSGIGKTELAKQTAKYMHKDAKKG FIRLDMSEFQERHEVAKFIGSPPGYVGHEEGGQLTKKLKQCPNAV LFDEVDKAHPDVLTIMLQLFDEGRITDGKGTIDCKDAIFIMTSNVA SDEIAQHALQLRQEALEMSRNRIENLGDVQISDKITISKNFKENVIR PILKAHFRRDEFLGRINEIVYFLPFCHSELIQLVNKELNFWAKRAKQR HNITLLWDREVADVLVDGYNVHYGARSIKHEVERRVNVQLAAAYE QDLLPGGCTLRITVEDSDKQLLKSPQLPSPQAEKRLPKLRLEHDKDS KTRRLDIRAPLHPEKVCNTI	CLPB
176	MLFLALGSPWAVELPLCGRRALCAAAALRGPRASVSRASSSGPS GPVAGWSTGPSGAARLLRRPGRAQIPVYWEGYVRFLNTPSKSEDG RLIYTGNMARAVFGVKCFSSYSTSLIGLTFLPYIFTQNNAISESVP LIQ IIFYGIMGSFTVITPVLLHFITKGYVIRLYHEATTDYKAITYNAMLA ETSTVFHQNDVKIPDAKHVFTTFYAKTKSLLVNPVLPNREDYIHL MYDKEEFILYMEETSEEKRRHKDDK	TMEM70
177	MLSQVYRCGFQPFNQHLLPWVKCTTVFRSHCIQPSVIRHVRWSNIP FITVPLSRTHGKSFAHRSELKHAKRIVVKLGSAVVTRGDECGLALGR LASIVEQVSVLQNGQREMMLVTSGAVAFGKQRLRHEILLSQSVRQA LHSGQNQLKEMAIPVLEARACAAAGQSGLMALYEAMFTQYSICAA QILVTNLDHFHDEQKRRNLNGTLHELLRMNIVPIVNTNDVPPAEP NSDLQGVNVISVDNDSLAARLAVEMKTDLLIVLSDVEGLFDSPPG SDDAKLIDIFYPGDQQSFTFGTKSRVGMGGMEAKVKAALWALQGG	ALDH18A1

	TSVVIANGTHPKVSGHVITDIVEGKKVGTFFSEVKPAGPTVEQQGE MARSGGRMLATLEPEQRAEIIHHLADLLTDQRDEILLANKKDLEEA EGRLAAPLLKRLSLSTSKLNSLAIGLRQIAASSQDSVGRVLRRTRIAK NLELEQVTVPIGVLLVIFESRPDCLPQVAALAIASGNGLLLKGGKEA AHSNRILHLLTQEALSIHGVKEAVQLVNTREEVEDLCRLDKMIDLIIP RGSSQLVRDIQKAAKGIPVMGHSEGICHMYVDSEASVDKVRTLVLD SKCEYPAACNALETLLIHRDLLRTPLFDQIIDMLRVEQVKIHAGPKF ASYLTFSPSEVKSLRTEYGDLELCIEVVDNVQDAIDHIHKYGSSTHD VIVTEDENTAEFFLQHVDSACVFWNASTRFSDGYRFGLGAEVGISTS RIHARGPVGLEGLLTTKWLLRGKDHVVSDFSEHGSLKYLHENLPIIP QRNTN	
178	MFSKLAHLQRFVLSRGVHSSVASATSVATKKTVQGPPTSDDIFERE YKYGAHNYHPLPVALERGKGIYLDVVEGRKYFDLSSYSANQGH CHPKIVNALKSQVDKLTLSRAFYNNVLGEYEEYITKLFNYHKVLP MNTGVEAGETACKLARKWGYTVKGIQKYKAKIVFAAGNFWGRTL SAISSSTDPTSVDGFGPFMPGFDIIPYNDLPALERALQDPNVA AFMVE PIQGEAGVVPDPGYLMGVRELCTRHQVLFIADEIQTGLARTGRWL AVDYENVRPDIVLLGKALSGGLYPVSAVLCDDDIMLTIKPGEHGST YGGNPLGCRVAIAALEVLEENLAENADKLGILRNELMKLPDVT AVRGKGLLNAIVIKETKDWDAWKVCLRLRDNGLLAKPTHGDIIIRFA PPLVIKEDELRESIEIINKTILSF	OAT
179	MLGRNTWKTSAFSFLVEQMWAPLWSRSMRPGRWCSQRSCAWQTS NNTLHPLWTVPVSVPGGTRQSPINIQWRDSVYDPQLKPLRVSYEAA SCLYIWNTGYLFQVEFDDATEASGISGGPLENHRYRLKQFHFHWGAV NEGGSEHTVDGHAYPAELHLVHWNSVKYQNYKEAVVGENGLAVI GVFLKLGAAHQTLQRLVDILPEIKHKDARAAMRPFDPSTLLPTCWD YWTYAGSLTTPPLTESVTWIIQKEPVEVAPSQLSAFRTLLFSALGEEE KMMVNYYRPLQPLMNRKVWASFQATNEGTRS	CA5A
180	MYRYLGEALLSRRAGPAALGSASADSAALLGWARQPAAAPQPGL ALARRHYSEAVADREDDPNFFKMVEGFFDRGASIVEDKLVEDLRT RESEEQKRNVRGILRIKPCNHVLSLSFPIRRDDGSWEVIEGYRAQH SQHRTCKGGIRYSTDVSVDEVKALASLMTYKCAVVDVPFGGAKA GVKINPKNYTDNELEKITRRFTMELAKKGFIGPIDVPAPDMSTGER EMSWIADTYASTIGHYDINAHACVTGKPISQGGIHRISATGRGVFH GIENFINEASYMSILGMTPGFG DKTFVQVQFGNVGLHSMRYLHRFGAKCIAVGESDGSIWNPDGIDPK ELEDKFLQHGSLGFPKAKPYEGSILEADCILIPAASEKQLTKSNAP RVKAKIIAEGANGPTTPEADKIFLERNIMVIPDLYLNAGGVTVSYFE WLKNLNHVSYGRLTFKYERDSNYHLLMSVQESLERKFGKHGGTIPI VPTAEFQDRISGASEKDIVHSGLAYTMERSARQIMRTAMKYNLGLD LRTAAAYVNAIEKVFKVYNEAGVTFT	GLUD1
181	MTTSASSHLNKGKQVYMSLPQGEKVQAMYIWDGTGEGLRCKTR TLDSEPKCVELPEWNFDGSSTLQSEGSNSDMYLVAAMFRDPFRK DPNKLVLCEVFKYNNRPAETNLRHTCKRIMDMVSNQHPWFGMEQE YTLMGTDGHPFGWPSNGFPGPGPYPCGVGADRAYGRDIVEAHYR ACLYAGVKIAGTNAEVMQAQWQFQIGPCEGISMGDHLWVARFILH RVCEDFGVIAFTDPKPIPGNWNAGCHTNFSTKAMREENGLKYIEE AIEKLSKRHYHIRAYDPKGGLDNARRLTGPHETSNINDFSAGVAN RSASIRIPRTVGQEKKGYPEDRRPSANCDPFSVTEALIRTCLLNETGD EPFQYKN	GLUL
182	MAVARAALGPLVTGLYDVQAFKFGDFVLKSGLSPIYIDLRGIVSRP RLLSQVADILFQTAQNAISFDTVCGVPYTALPLATVICSTNQIPMLI	UMPS

	RRKETKDYGTKRLVEGTINPGETCLIIEDVVTSGSSVLETVEVLQKE GLKVTDAIVLLDREQGGKDKLQAHGIRLHSVCTLSKMLEILEQQKK VDAETVGRVKRFIQENVFVAANHNGSPLSIKEAPKELSFGARAEPR IHPVA SKLLRLMQKKETNLCLSADVSLARELLQLADALGPSICMLKTHVDI LNDFTLDVMKELITLAKCHEFLIFEDRKFDIGNTVKKQYEGGIFKIA SWADLVNAHVPGSGVVKGLQEVGLPLHRGCLLIAEMSSTGSLAT GDYTRAAVRMAEEHSEFVVGFISGSRVSMKPEFLHLPVGVQLEAGG DNLGQQYNPQEVIGKRGSDIIVGRGIISAADRLEAAEMYRKAWE AYLSRLGV	
183	MRDYDEVTAFLGEWGPFQRLIFFLLSASIIPNGFTGLSSVFLIATPEHR CRVPDAANLSSAWRNHTVPLRLRDGREVPHSCRRYRLATIANFSAL GLEPGRDVDLGQLEQESCLDGWEFSQDVYLSSTIVTEWNLVCEDDW KAPLTISLFFVGVLLGSFISGQLSDRFGRKNVLFVTMGMQTGFSFLQI FSKNFEMFVVLVFLVGMGQISNYVAAFVLGTEILGKSVRIIFSTLGV CIFYAFGYMVLPLFAYFIRDWRMLLVALTMPGVLCVALWWFIPESP RWLISQGRFEEAEVIIRKAAKANGIVVPSTIFDPSELQDLSSKKQOSH NILDLLRTWNIRMVTIMSIMLWMTISVGYFGLSLDTPNLHGDIFVNC FLSAMVEVPAYVLAWLLLQYLPRRYSMATALFLGGSVLLFMQLVP PDLYLATVLMVGMKFGVTAAFSMVYVYTAELYPTVVRNMGVGV SSTASRLGSILSPYFVYLGA YDRFLPYILMGS LTILTAILTLFLPESFGT PLPDTIDQMLRVKGMKHKRKTPSHTR MLKDGQERPTILKSTAF	SLC22A5
184	MAEAHQAVAFQFTVTPDGIDLRLSHEALRQIYLSGLHSWKKKFIRF KNGIITGVYPASPSSWLIVVGVMTTMYAKIDPSLGIIAKINRTLETA NCMSSQTKNVVSGVLFGTGLWVALIVTMRYSLKVLLSYHGWMTFE HGKMSRATKIWMGMVKIFSGRKPMLYSFQTSPLRPLVPAVKDTVN RYLQSVRPLMKEEDFKRMTALAQDFAVGLGPRLQWYLLKLSWWA TNYVSDWWEYIYLRGRGPLMVNSNYAMDLLYILPHIQAARAG NAIHAILLYRRKLDREEIKPIRLLGSTIPLCSAQWERMFNTRSIPGEET DTIQHMRDSKHIVVYHRGRYFKVWLYHDGRLLKPREMEQQMQRIL DNTSEPQPGEARLAALTAGDRVPWARCRQAYFGRGKKNKQSLDAVE KAAFFVTLDETEEGYRSEDPTSMDSYAKSLLHGRCYDRWFDKSFT FVVFKNGKMGLNAEHSWADAPIVAHLWEYVMSIDSLQLGYAEDG HCKGDINPNIPYPTRLQWDIPGECQEVIELNTANLLANDVDFHSFP FVAFGKGIKKCRTSPDAFVQLALQLAHYKDMGKFCLTYEASMTL FREGRTETVRSCCTESCDVFRAMVDPAQTVEQRLKLFKLASEKHQH MYRLAMTGSIDRHLFCLYVVS KYLAVESPFLKEVLSEPWRLSTSQ TPQQQVELFDLENNPEYVSSGGGFGPVADDGYGVSYILVGENLINF HISSKFSCPETDSHRFGRHLKEAMTDIITLFLGLSSNSKK	CPT1A
185	MVACRAIGILSRFSAFRILRSRGYICRNFTGSSALLTRTHINYGVKGD VAVVRINSPNSKVNTLSKELHSEFSEVMNEIWASDQIRSAVLISKPG CFIAGADINMLAACKTLQEVTLQSQAQRIVEKLEKSTKPIVAAING SCLGGGLEVAISCQYRIATKDRKTVLGTPEVLLGALPGAGGTQRLP KMVGVPAAALDMMLTGRSIRADRAKKMGLVDQLVEPLGPGLKPPEE RTIEYLEEVAITFAKGLADKKISPKRDKGLVEKLTAYAMTIPFVRQQ VYKKVEEKVRKQTKGLYPAPLKIIDVVKTGIEQGS DAGYLCESQKF GELVMTKESKALMGLYHGQVLCKKNKFGAPQKDVKHLAILGAGL MGAGIAQVSVDKGLKTILKDATLTALDRGQQQVFKGLNDKVKKKA LTSFERDSIFSNTLGQLDYQGF EKADMVIEAVFEDLSLKHRLKEVE AVIPDHCIFASNTSALPISEIAAVSKRPEKVIGMHYFSPVDKMQLLEII TTEKTSKDTASAVAVGLKQGVIIIVK	HADHA

	DGPGFYTTTRCLAPMMSEVIRILQEGVDPKKLDSLTTSGFPVGAATL VDEVGVDAKHVAEDLGKVFGERFGGNGPELLTQMVSFGFLGRKS GKGFYIYQEGVKRKDLNSDMDSILASLKLPPKSEVSSDEDIQFRLVT RFVNEAVMCLQEGILATPAEGDIGAVFGLGFPPCLGGPFRFVDLYG AQKIVDRLKKYEAAYGKQFTPCQLLADHANSPNKKFYQ	
186	MAFVTRQFMRSVSSSTASASAKKIIVKHVTVIGGGMLMGAGIAQVA AATGHTVVLVDQTEDILAKSKKIEESLRKVAKKKFAENLKAGDEF VEKTLSTIATSTDAASVHSTDLVVEAIVENLKVKNELFKRLDKFAA EHTIFASNTSSLQITSIANATTRQDRFAGLHFFNPVPMKLVEVIKTP MTSQKTFESLVDFSKALGKHPVSKDTPGFIVNRLLVPYLMEAIRLY ERGDASKEDIDTAMKLGAGYPMGPFELLDYVGLDTTKFIVDGWHE MDAENPLHQSPSLNKLVAENKFGKKTGEGFYKYK	HADH
187	MAAPTLGRLVLTHLLVALFGMGSWAAVNGIWWELPVVVKDLPEG WSLPSYLSVVVALGNLGLLVTLWRQLAPGKGEQVPIQVVQVLSV VGTALLAPLWHHVAPVAGQLHSVAFLTLALVLAMACCTSNVTFLP FLSHLPPFLRSFFLGQGLSALLPCVLALVQGVGRLECAPPAPTNGTSG PPLDFPERFPASTFFWALTALLVTSAAAFRGLLLLLPSLPSVTTGGSG PELQLGSPGAEEEEKEEEEEALPLQEPPSQAAGTIPGPDPEAHQLFSAH GAFLGLMAFTSAVTNGVLPVSVQSFSCLPYGRLAYHLAVVLGSAAN PLACFLAMGVLCRSLAGLVGLSLLGMLFGAYLMALAILSPCPPLVG TTAGVVLVLSWVLCVFSYVKVAASSLLHGGGRPALLAAGVAI QVGSLLGAGAMFPPTSIYHVFQSRKDCVDPCGP	SLC52A1
188	MAAPTPARPVLTHLLVALFGMGSWAAVNGIWWELPVVVKELPEG WSLPSYVSVLVVALGNLGLLVTLWRRLAPGKDEQVPIRVVQVLGM VGTALLASLWHHVAPVAGQLHSVAFLALAFVLALACCASNVTFLP FLSHLPPFLRSFFLGQGLSALLPCVLALVQGVGRLECAPPAPINGTPG PPLDFLERFPASTFFWALTALLVASAAAFQGLLLLLPPPPSVPTGELG SGLQVGAPGAEEVEESSPLQEPPSQAAGTTPGPDPAKAYQLLSARSA CLLGLLAATNALTNGVLPVAVQSFSCLPYGRLAYHLAVVLGSAANPL ACFLAMGVLCRSLAGLGLSLLGVFCGGYLMALAVLSPCPPLVGT AGVVLVLSWVLCVFSYVKVAASSLLHGGGRPALLAAGVAIQV GSLLGAVAMFPPTSIYHVFHSRKDCADPCDS	SLC52A2
189	MAFLMHLLVCVFGMGSWVTINGLWVELPLLVMELPEGWYLPSTLT VVIQLANIGPLLVTLLHHFRPSCLSEVPIIFTLLGVGTVTCHFAFLWN MTSWVLDGHHIAFLVLTFFLALVDCTSSVTFLPFMSRLPTYLTTF FVGEGLSGLLPALVALAQSGSLTTCNVNTEISDSVPSPVPTRETDIAQ GVPRALVSALPGMEAPLSHLESRYLPAHFSPLVFFLLLSIMMACCLV AFFV LQRQPRCWEASVEDLLNDQVTLHSIRPREENDLGPAGTVDSSQGGQ YLEEKAAPCCPAHLAFIYTLVAFVNALTNGMLPSVQTYSCLSYGPV AYHLAATLSIVANPLASLVSMFLPNRSLLFLGVLSVLGTFCGGYNM AMAVMSPCPLLQGHWGGEVLIVASWVLFSGCLSYVKVMLGVVLR DLRSALLWCGAAVQLGSLLGALLMFPLVNVLRFLFSSADFCNLHCP A	SLC52A3
190	MTILTYPFKNLPTASKWALRFSIRPLSCSSQLRAAPAVQTKTKKTLA KPNIRNVVVVDGVRTPFLSGTSYKDLMPHDLARAALTGLLHRTSV PKEVVYDIIFGTVIQEVKTSNVAREAAAGAGFSDKTPAHTVTMACIS ANQAMTTGVGLIASGQCDVIVAGGVELMSDVPIRHSRKMRLMLD LNKAKSMGQRLSLISKFRFNFLAPELLPAVSEFSTSETMGHSADRLAA AFAVSRLEQDEYALRSHSLAKKAQDEGLLSDVVPFKVPGKDTVTK DNGIRPSSLEQMAKLKPAFIKPY	HADHB

	GTVTAANSSFLTDGASAMLIMAEKALAMGYKPKAYLRDFMYVSQ DPKDQLLLGPTYATPKVLEKAGLTMNDIDAFEFHEAFSGQILANFK AMDSDFWFAENYMGRKTKVGLPPELKFNNWGGSLSLGHPFGATGC RLVMAAANRLRKEGGQYGLVAACAAGGQGHAMIVEAYPK	
191	MLRGRSLSVTSLGGLPQWEVEELPVEELLLFEVAWEVTNKVGGIYT VIQTKAKTTADEWGENYFLIGPYFEHNMKTQVEQCEPVNDVRRRA VDAMNKHGCVHFGRWLIEGSPYVVLFDIGYSAWNLDRWKGDLW EACSVGIPYHDREANDMLIFGSLTAWFLKEVTDHADGKYVVAQFH EWQAGIGLILSRARKLPIATIFTTHATLLGRYLCAANIDFYNHLDKFN IDKEAGERQIYHRYCMERASVHCAHVFTTVSEITAIEAEHMLKRP DVVTPNGLNVKKFSVHEFQNLHAMYKARIQDFVRGHFYGHLDFFD LEKTLFLFIAGRYEFSNKGADIFLESLSRLNLLRMHKSDITVMVFFI MPAKTNNFNVTETLKGQAVRKQLWDVAHSVKEKFGKKLYDALLRG EIPDLNDILDRDDLTKMRAIFSTQRQSLPPVTTNMMIDDSTDPISTI RRIGLFNNRTDRVKVILHPEFLSSTSPLLPMDYEEFVRGCHLGVPFSY YEPWGYTPAECTVMGIPSVTTNLSGFGCFMQEHVADPTAYGIYIVD RRFRSPDDSCNQLTKFLYGFCKQSRQRRIQRNRTERLSDDLWDWRYL GRYYQHARHLTLRAFPDKFHVLTSPPTTEGFKYPRSSVPPSPSGS QASSPQSSDVEDEVEDERYDEEEEAERDRLNIKSPFSLSHVPHGKKK LHGEYKN	GYS2
192	MAKPLTDQEKRRQISIRGIVGVENV AELKKSFNRLHFTLVKDRNV ATTRDYFALAHTRVDRHLVGRWIRTQQHYDCKPKRVYYLSLEFY MGRTLQNTMINLGLQNACDEAIYQLGLDIEELEEIEEDAGLGNGGL GRLAACFLDSMATLGLAAYGYGIRYEGIFNQKIRDGWQVEEADD WLRYGNPWKESRPEFMLPVHFGYKVEHTNTGTKWIDTQVVLALPY DTPVPGYMNNTVNTMRLWSARAPNDFNLRDFNVGDYIQAVALDRN LAENISRVLYPNDFNFFEGKELRLKQYFVVAATLQDIIRRFKASKFG STRGAGTVFADFDPQVAIQLNDTHPALAIPELMRIFVDIEKL PWSKAWELTQKTFAYTNHTVLPEALERWPVDLVEKLLPRHLEIIEI NQKHLDRIVALFPKDVDRRLRMSLIEEESKRINMAHLICVGSNAV NGVAKIHSDIVKTKVFKDFSELEPKFKQNTNGITPRRWLLCNPG AELIAEKIGEDYVKDLSQLTKLHSFLGDDVFLRELAKVKQENKLFKS QFLETEYKVKINPSSMFDVQVKRIHEYKRQLLNCLHVITMYNRIKK DPKKLFVPRTVIIGGKAAPGYHMAKMIKLITSVADVNNNDPMVGS KLKVIFLENYRVSLAEKVIPATDLSEQISTAGTEASGTGNMKFMLNG ALTIGTMDGANVEMAEAGEENLFIFGMRIDDVAALDKKGYEAKE YYEALPELKLVIDQIDNGFFSPKQPDLFKDIIINMLFYHDFKVFADY EAYVKCQDKVSQLYMNPKAWNTMVLKNIAASGKFSSDRTIKEYAQ NIWNVEPSDLKISLSNESNKVNGN	PYGL
193	MTEDKVTGTLVFTVITAVLGSFQFGYDIGVINAPQQVIISHYRHVLG VPLDDRKAANNYVINSTDELPTISYSMNPKPTPWAEETVAAAQLIT MLWSLSVSSFAVGGMTASFFGGWLGDTLGRIKAMLVANILSLVGA LLMGFSKLGPSHILIIAGRSISGLYCGLISGLVPMYIGEIAPTALRGAL GTFHQLAIVTGILISQIIGLEFILGNYDLWHILLGLSGVRAILQSLLFF CPESPRYLYIKLDEEVKAKQSLKRLRGYDDVTKDINEMRKEREAS SEQKVSIIQLFTNSSYRQPILVALMLHVAQQFSGINGIFYYSTSIFQTA GISKPVYATIGVGAVNMVFTAVSVFLVEKAGRRLSLFLIGMSGMFVC AIFMSVGLVLLNKFSWMSYVSMIAIFLVSSFEIGPGIPWFMVAEFF SQGPRPAALAAAFSNWTCNFIVALCFQYIADFCGPYVFFLFAGVLL AFTLFTFFKVPETKGKSFEIEIAAEFQKKSGSAHRPKAAVEMKFLGAT ETV	SLC2A2

194	MAASCLVLLALCLLLPLLLGGWKRWRRGRAARHVAVVLGDVG RSPRMQYHALSLAMHGFSVTLLGFCNSKPHDELLQNNRIQIVGLTE LQSLAVGPRVFQYGVKVVQAMYLLWKLWREPGAYIFLQNPFG LPSIAVCWFVVGCLCGSKLVIDWHNYGYSIMGLVHGPNHPLVLLAK WYEKFFGRLSHLNLCTNAMREDLADNWHIRAVTVYDKPASFFKE TPLDLQHRLFMKLGSMHSPFRARSEPDPVTERSAFTERDAGSGLVT RLRERPALLVSSTSWTEDEDFSILLAALEKFEQLTLDGHNLPVLVCVI TGKGPLREYYSRILHQHFIHQVCTPWLEAEDYPLLLGSADLGV LHTSSSGLDLPKVVDMFGCCLPVCANFKCLHELKHEENGLVF EDSEELAAQLQMLFSNFPDPAGKLNQFRKNLRESQQLRWDES WVQTVLPLVMDT	ALG1
195	MAEEQGRERDSVPKPSVLFLHPDLGVGGAERLVLDAALALQARGC SVKIWTAHYDPGHCAESRELPRCAGDWLPRGLGWGGRGAAVC AYVRMVFLALYVFLADEEFDVVVCDQVSACIPVFRRLARRRKKILF YCHFPDLLLLTKRDSFLKRLYRAPIDWIEEYTTGMADCILVNSQFTAA VFKETFKSLSHIDPDVLYPSLNVTSFDSVVPKLDLVPKGGKFLLL SINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVL ENVEHYQELKKMVQQSDLGQYVTFRLSFSDKQKISLLHSCTCVLYT PSNEHFIVPLEAMYMQCPVIAVNSGGPLESIDHSVTGFLCEPDVH FSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAFTEQLYRYVTCLL V	ALG2
196	MAAGLRKRGRSGSAAQAEGLCKQWLQRAWQERRLLLREPRYTLL VAACLCCLAEVGITFWVIHRVAYTEIDWKAYMAEVEGVINGTYDYT QLQGDGTPLVYPAGFVYIFMGLYYATSRGTDIRMAQNIFAVLYLAT LLLVLFIYHQTCVPPFVFFMCCASYRVHSIFVLRLFNDPVA MVLLFLSINLLLAQRWGWGCCFFSLAVSVKMNVLFPAGLLFLLLTQFGF RGALPKLGICAGLQVVLGLPFLLENPSGYLSRSFDLGRQFLFHWTVN WRFLPEALFLHRAFLHALLTAHLTL LLLFALCRWHRTGESILLSLRDPSKRKVPPQPLTPNQIVSTLFTSNFIG ICFSRSLYQFYVWYFHTLPYLLWAMPARWLTHLLRLLVLGLIELS WNTYPTSCSSAALHICHA VILLQLWLGPQFPKSTQHSSKKAH	ALG3
197	MEKWYLMTVVVLIGLTVRWTVSLNSYSGAGKPPMFGDYEAQRHW QEITFNLPVKQWYFNSSDNNLQYWGLDYPLTAYHSLLCAYVAKFI NPDWIALHTSRGYESQAHKLFMRRTTVLIADLLIYIPAVVLYCCCLKE ISTKKKIANALCILLYPGLILIDYGHFQYNSVSLGFALWGVLGISCDC DLLGSLAFCLAINYKQMELYHALPFFCFLGKCFKKGLKKGKGFVLL VKLACIVVASFVLCWLPFFTEREQLQVLRRLFPVDRGLFEDKVANI WCSFNVFLKIKDILPRHIQLIMSFCSTFLSLLPACIKLILQPSSKGFKFT LVSCALSFFLFSFQVHEKSILLVSLPVCLVLSEIPFMSTWFLLVSTFSM LPLLLKDELLMPSVVTMAFFIACVTSFSIFEKTSEEELQLKSFSISVR KYLPCFTFLSRHIIQYLFLISVITMVLLTLMTVTLDPQKLPDLFSVLVC FVSCNLNLFLLVYFNIIIMWDSKSGRNQKKIS	ALG6
198	MAALTATGTGNWFSALALGVTLKCLLIPTYHSTDFEVHRNWLAI THSLPISQWYYEATSEWTLDPFFFAWFEYILSHVAKYFDQEMLN VHNLNYSSTRTLLFQRFVSIFMDVLFVYAVRECKCIDGKKVGKELTE KPKFILSVLLWNFGLLIVDHIHFQYNGFLFGLMLLSIARLFQKRHM EGAFLFAVLLHFKHIYLYVAPAYGVYLLRSYCFTANKPDGSIRWKS FSFVRVISLGLVFLVSALSGLPFLALNQLPQVFSRLFPFKRGLCHAY WAPNFWALYNALDKVLSVIGLKLKFLDPNNIPKASMTSGLVQQFQ HTVLPSTVPLATLICTLIAILPSIFCLWFKPQGPRGFLRCLTLCALSSF MFGWHVHEKAILLAILPMSLLSVGKAGDASIFLILTTTGHYSLFPLLF TAPELPIKILLMLLFTIYSSISLKTFRKEKPLFNWMETFYLLGLGPLE	ALG8

	VCCEVFVPFTSWKVKYPIPLLLTSVYCAVGITYAWFKLYVSVLIDS AIGKTKKQ	
199	MASRGARQLKKGSGASSGDTAPAADKLRELLGSREAGGAEHRTTEL SGNKAGQVWAPEGSTAFKCLLSARLCAALLSNISDCDETFNYWEPT HYLIYGEFGQTWEYSPAYAIRSYAYLLLHAWPAAFHARILQTNKILV FYFLRCLLAFVSCICELFYKAVCKKFGHLVSRMMLAFLVLSTGMF CSSSAFLPSSFCMYTTLIAMTGWYMDKTSIAVLGVAAGAILGWPFS AALGLPIAFDILLVMKHRWKSFFHWSLMALILFLVPVVVIDSYYYGK LVIAPLNIVLYNVFTPHGPDLYGT EPWYFYLINGFLNFNVAALALLVLPLTSLMEYLLQRFHVQNLGHP YWLTLAPMYIWFIFFIQPHKEERFLFPVYPLICLCGAVALSALQKCY HFVFQRYRLEHYTVTSNWLALGTVFLFGLLSFSRSVALFRGYHGPL DLYPEFYRIATDPTIHTVPEGRPVNVCVGKEWYRFPSSFLLPDNWQL QFIPSEFRGQLPKPFAEGPLATRIVPTDMNDQNLEEPSRYIDISKCHY LVDLDTMRETREPKEYSSNKEEWISLAYRPFLDASRSSKLLRAFYVP FLSDQYTVYVNYTILKPRKAKQIRKKSGG	ALG9
200	MAAGERSWCLCKLLRFFYSLFFPGLIVCGTLCVCLVIVLWGIRLLLQ RKKKL VSTSKNGKNQMVI AFFHPYCNAGGGGERVLWCALRALQK KYPEAVYVVYTG DVNVNGQQILEGAFRRFNIRLIHPVQFVFLRKRY LVEDSLYPHFTLLGQSLGSIFLGWEALMQCVPDVYIDSMGYAFTLPL FKYIGGCQVGSYVHYPTISTDMLS VVKNNQIGFNNAAFITRNPFLSK VKLIYYLFAFIYGLVGSCSDVVMVNSSWTLNHILSLWKVGNCNTNI VYPPCDVQTFDLIPLHEKKMTPGHLLVSVGQFRPEKNHPLQIRAFAK LLNKKMVESPPSLKVLIGGCRNKDDEL RVNQLRRLSED LGVQEYV EFKINIPFDELKNYLSEATIGLHTMWNEHFGIGVVECMAAGTILAH NSGGPKLDIVVPHEGDITGFLAESEEDY AETIAHILSMSAEKRLQIRK SARASVSRFSDQEFEVTFLLSSVEKLFK	ALG11
201	MAGKGSSGRPRLLGLLVAVATVHLVICPYTKVEESFNLQATHDLL YHWQDLEQYDHLEFPGVVPRTFLGPVVIAVFSSPAVYVLSLLEMSK FYSQLIVRGVLGLGVIFGLWTLQKEVRRHFGAMVATMFCWVTAM QFHLMFYCTRTPNVLALPVL LALAAWLRHEWARFIWLSAFAIIV FRVELCLFLGLLLLLALGNRKVS VVRALRHAVPAGILCLGLTVAVD SYFWRQLTWPEGKVLWYNTVLNKSSNWGTSPLLWYFYFYSALPRGL GCSLLFIPLGLVDRRTHAPT V LALGFMALYSLLPHKELRFIYAFPM NITAARGCSYLLNNYKKS WL YKAGSLLVIGHLVVNAAYSATALYV SHFNYPGGVAMQRLHQLVPPQTDVLLHIDVAAAQTGVSRLQVNS AWRYDKREDVQPGTGMLAYTHILMEAPGLLALYRDTHRVLASV VGTTGVSLNLTQLPPFNVHLQTKLVLLERLPRPS	ALG12
202	MKCVFVTVGTTSFDDLIACVSAPDSLQKIESLGYNRLILQIGRGT VV PEPFSTESFTLDVYRYKDSLKEDIQKADLVISHAGAGSCLETLEK GK PLVVVINEKLMNNHQLELAKQLHKEGHLFYCTCRVLTCPGQAKSIA SAPGKCQDSAALTSTAFSGLD FGLLSGYLHKQALVTATHPTCTLLFP SCHAFFPLPLTPTLYKMHKGWKNYCSQKSLNEASMDEYLGSLGLFR KLTA KDASCLFRAISEQLFCSQVHHLEIRKACVSYMRENQQT FESYV EGSF EKYLERLGD PKESAGQLEIRALS LIYNRDFILYRFP GK PPT YVT DNGYEDKILLCYSSSGHYDSVYSKQFQSSAAVCQAVLYEILYKDV VVDEEELKTAIKLFRSGSKKNRNNAVTGSEDAHTDYKSSNQNRME EWGACYN AENIPEGYNKGTEETKSPENPSKMPFPYKVLKALDPEIY RNVEFDVWLDSRKELQKSDYMEYAGRQYYLGDKCQVCLESEGRY YNAHIQEVGNENNSVTVFIEELAEKHVVPLANLKPVTQVMSVPAW NAMPSRKGRGYQKMPGGYVPEIVISEMDIKQKKMFKKIRGKEYV M	ALG13

	<p>TMAYGKGDP LLPRLQHSMHYGHDPPMHYSQTAGNVMSNEHFHP QHPSPRQGRGYGMPNRSSRFINRHNMPGPKVDFYPGPGKRCCQSYD NFSYRSRFRSHRQMSCVNKESQYGFPGNGQMPRGLEETITFYE VEEGDETA YPTLPNHGGPSTMVPATSGYCVGRRGHSSGKQTLNLEE GNGQSENGRYHEEYLYRAEPDYETSGVYSTTASTANLSLQDRKSCS MSPQDVTVTSYNYPQKMMGNIAAVAASCANNVPAPVLSNGAAANQ AISTTSVSSQNAIQPLFVSPPTHGRPVIASPSYPCHSAIPHAGASLPPPP PPPPPPPPPPPPPPPPPPPPALDVGETSNLQPPPPPLPPPPYSCDPSGS DLPQDTKVLQYYFNLGLQCYHSHYWSHSMVYVPQMQQQLHVENYP VYTEPPLVDQTVPCYSEVRREDGIQAEASANDTFPNADSSSVPHG AVYYPVMSDPYGPPLPGFDSCLPVVPDYSCVPPWHPVGTAYGGSS QIHGAINPGPIGCIAPSPASHYVPQGM</p>	
203	<p>MGSLFRSETMCLAQLFLQSGTAYECLSALGEKGLVQFRDLNQNVSS FQRKFVGEVKRCEELERILVYLVEINRADIPLPEGEASPPAPPLKQV LEMQEQLQKLEVELREVTKNKEKLRKNLLEIEYTHMLRVTKTFVK RNVEFEPTYEEFPSLESDSLDDYSCMQRLGAKLGFVSGLINQGVKA FEKMLWRVCKGYTIVSYAELDESLEDPETGEVIKWYVFLISFWGEQI GHKVKKICDCYHCHVYPYPNTAEERREIQEGLNTRIQDLYTVLHKT EDYLRQVLCKAAESVYSRVIVKMKKAIYHMLNMCSEFVTNKCLI AEVWCPEADLQDLRRALEEGSRESGATIPSMNIPTKETPPTRIRTN KFTEGFQNIVDAYGVGSYREVNPAFTIITFPFLFAVMFGDFGHGFV MFLFALLLVLENHPRLNQSQEIMRMFFNGRYILLMLGLFSVYTGLI YNDCFKSVNLFGSGWNVSAMYSSSHPPAEHKKMVLWNDSSVVRH NSILQLDPSIPGVFRGPYPLGIDPIWNLATNRLTFLNSFKMKMSVILGI IHMTFGVILGIFNHLHFRKKFNIYLVSIPELLFMLCIFGYLIFMIFYKW LVFSAETSRVAPSILIEFINMFLFPASKTSGLYTGQEYVQRVLLVTA LSVPVLFLGKPLFLLWLHNGRSCFGVNRSGYTLIRKDSEEEVSLGGS QDIEEGNHQVEDGCREMACEEFNFGELMTQVIHSIEYCLGCISNTA SYLRLWALS LAHAQLSDVLWAMLMRVGLRVDTTYGVLLLLPVIAL FAVLTIFILLIMEGLSAFLHAIRLHWVEFQNKFYVGAGTKFVPPF SFSLLSSKFNNDDSA</p>	ATP6V0A2
204	<p>MRPPACWWLLAPPALLALLTCSLAFGLASEDTKKEVKQSQDLEKS GISRKNIDIDLK GIVFVIQSQNSFHA KRAEQLKKSILKQAADLTQELP SVLLLHQLAKQEGAWTILPLLPHFSVTYSRNSSWIFFCEEETRIQIPK LLETLLRRYDPSKEWFLGKALHDEEATIIHHYAFSENPTVFKYPDFAA GWALSIPLVNKLTKRLKSESLKSDFTIDLKHEIALYIWDKGGGPPLTP VPEF CTNDVDFYCATTFSFLPLCRKPVKKKDIFVAVKTCKKFHGDRIPIV KQTWESQASLIEYYSDYTENSIPTVDLGPNTDRGHCGKTFAILERFL NRSQDKTAWLVIVDDDTLISISRLQHLLSCYDSGEPVFLGERYGYGL GTGGYSYITGGGGMVFSREAVRLLASKCRCYSNDAPDDMVLGMC FSGLGIPVTHSPLFHQARPVDYPKDYLSHQVPISFHKHWNIDPVKVY FTWLAPSEDEKARQETQKGFREEL</p>	B3GLCT
205	<p>MFPRPLTPLAAPNGAEPLGRALRRAPLGRARAGLGGPPLLLPSMLM FAVIVASSGLLLMIERGILAEMKPLPLHPPGREGTAWRGKAPKPGGL SLRAGDADLQVRQDVNRNRTLRAVCGQPGMPRDPWDLVPGQRRTL LRHILVSDRYRFLYCYVPKVACSNWKRVKVLAVLDSVDVRLK MDHRSDLVFLADLRPEEIRYRLQHYFKFLFVREPLERLLSAYRNKFG EIREYQORYGAEIVRRYRAGAGPSPAGDDVTFPEFLRYLVDEDPER MNEHWMPVYHLCQPCAVHYDFVGSYERLEADANQVLEWVRAPP</p>	CHST14

	VRFPARQAWYRPASPESLHYHLC SAPRALLQDVL PKYILDFSLFAYP LPNVTKEACQQ	
206	MATAATSPALKRLDLRDP AALFETHGAE EIRGLERQVRAEIEHKKE ELRQMVGERYRDLIEAADTIGQMRRCAVGLVDAVKATDQYCARLR QAGSAAPRPPRAQQPQQPSQEFYSMAAQIKLLEIPEKIWSSMEAS QCLHATQLYLLCCHLHSLQLDSSSSRYSPVLSRFPILIRQVAAASHF RSTILHESKMLLKCGVSDQAVAEALCSIMLLEESSPRQALTDFLLA RKATIQLLNQPHHGAGIKAQICSLVELLATTCLKAHALFYTLPEGL LPDPALPCGLLFSTLETITGQHPAGKGTGVLQEEMKLCSWFKHLPAS IVEFQPTLRTLAHPISQEYKDTLQKWIHMCNEDIKNGITNLLMYVK SMKGLAGIRDAMWELLTNESTNHSWDVLCRRILLEKPLLFWEDMM QQFLDRLQTLTKEGFDSISSSSKELLVSALQELESSTSNSPSNKHHIF EYNMSLFLWSESPNDLPSDAAWVSVANRGQFASSGLSMKAQAISPC VQNFCSALDSKLVKLDLLAYLPSSD SSLPKDVSPTQAKSSAFDRYADAGTVQEMLRTQSVACIKHIVDCIRA ELQSIEEGVQGGQDALNSAKLHSLVFMARLCQSLGELCPHLKQCIL GKSESSEKPAREFRALRKQGVKTQEIIPTQAKWQEVKEVLLQSSV MGYQVWSSAVVKVLIHGFTQSLLLDDAGSVLATATSWDELEIQEEA ESGSSVTSKIRLPAQPSWYVQSFLSLCQEINRVGGHALPKVTLQEM LKSCMVQVVAAYEKLSEEKQIKKEGAFPVTQNRALQLLYDLRYLNI VLTAKGDEVKSGRSKPDSRIEK VTDHLEALIDPFDLDVFTPHLNSNLHRLVQRTSVLFGVLTGTENQLA PRSSFTNSQEPHNILPLASSQIRFGLPLSMTSTRKAKSTRNIETKAQV VPPARSTAGDPTVPGSLFRQLVSEEDNTSAPSLFKLGWLSMTK	COG1
207	MEKSRMNLPKGPDTLCFDKDEFMKEDFDVDHFVSDCRKRVQLEEL RDDLELYYKLLKTAMVELINKDYADFNVLSTNLVGMDKALNQLSV PLGQLREEVLSLRSSVSEGIRAVDERMSKQEDIRKKKMCVLRILQVI RSVEKIEKILNSQSSKETSALASSPLLTGQILERIAFEFNQLQFHAVQ SKGMPLLDKVRPRIAGITAMLQQSLEGLLLEGLQTSDDVDIIRHCLRT YATIDKTRDAEALVGQVLVKPYIDEVIEQFVESHNPGLQVMYNKLL EFVPHHCRLLREVTTGGASSEKGNTPGYDFLVNSVWPQIVQGLEE KLPSLFNPGNPDAFHEKYTISMDFVRRLERQCGSQASVKRLRAHPA YHSFNKKWNLPVYFQIRFREIAGSLEAALTDVLEDAPAESPCLLAS HRTWSSLRRCWSDMFLPLLHRLWRLTLQILARYSVFVNLSLRPI SNESPKEIKKPLVTGSKEPSITQGNTEQSGSPSETKPVVSISRTQLV YVVADLDKLQEQLPELLEIIKPKLEMIGFKNFSSISAALEDSQSSFSA CVPSLSKKIQLSDSCFGFLKSALEVPRLYRRTNKEVPTTASSYVDS ALKPLFQLQSGHKDKLKQAIQQWLEGTLSESTHKYYETVSDVLNS VKKMEESLKRLLKQARKTTPANPVGPSGGMSDDDKIRLQLALDVEY LGEQIQKLGLQASDIKSFSALAEVAAAKDQATAEQP	COG2
208	MADLDSPPKLSGVQQPSEG VGGRCSEISAELIRSLTELQELEAVYE RLCGEEKVVERELDALLEQQNTIESKMVTLHRMGPNLQIEGDAKQ LAGMITFTCNLAENVSSKVRQLDLAKNRLYQAIQRADDILDKFCM DGVQTLRSEDYEQAAAHTRYLCLDKSVIELSRQGKEGSMIDANL KLLQEAQRLKAIVAEKFAIATKEGDLQPVERFFKIFPLGLHEEGLR KFSEYLCKQVASKAEENLLMVLGTDMSDRRAAVIFADTLTLLFEG ARIVETHQPIVETYYGPGRLYTLIKYLQVECDRQVEKVVDKFIKQRD YHQQFRHVQNNLMRNSTTEKIEPRELDPILTEVTLMNARSELYLRF KKRISSDFEVGDSMASEEVKQEHQCLDKLLNNCLLSCTMQELIGL YVTMEEYFMRETVNKAVALDITYEKGQLTSSMVDDVFYIVKKCIGR ALSSSIDCLCAMILNATTELESDFRDVLCNKLRMGFPATTFQDIQR GVTSVAVNIMHSSLQQGFDTKGIESTDEAKMSFLVTLNNVEVCSENI STLKKTTLES DCTKLF SQIGGEQAQAKFDSCLSDLA AVSNKFRDLLQ	COG4

	EGLTELNSTAIKPQVQPWINSFFSVSHNIEEEEFNDYEANDPWVQQFI LNLEQQMAEFKASLSPVIYDSLTLGMLTSLVAVELEKVVVKSTFNRL GGLQFDKELRSLIAYLTTVTTWTIRDKFARLSQMATILNLERVTEILD YWGPNSGPLTWRLTPAEVRQVLAALRIDFRSEDIKRLRL	
209	MGWVGGRRRDSASPPGRSRSAAADDINPAPANMEGGGGSVAVAGL GARGSGAAAATVRELLQDGCYSDFLNEDFDVKTYTSQSIHQAVIAE QLAKLAQGISQLDRELHLQVVARHEDLLAQATGIESLEGVLQMMQ TRIGALQGAVDRIKAKIVEPYNKIVARTACLARLQVACDLLRRIIRIL NLSKRLQGQLQGGREITKAAQSLNELDYLSQGIDLSGIEVIENDLLF IARARLEVENQAKRLLQGLETONPTQVGTALQVFYNLGTKDITITS VVDGYCATLEENINSALDIKVLTQPSQSAVRGGPGRSTMPPTGNTA ALRASFWTNMEKLMDDHIYAVCGQVQHLQKVLAKKRPVSHICFIE EIVKDGQPEIFYTFWNSVTQALSSQFHMATNSSMFLKQAFEGEYPK LLRLYNDLWKRLQQYSQHIQGNFNASGTTDLVVDLQHMEDDAQDI FIPKKPDYDPEKALKDSLQPYEAAVLSKSLSRFLDPINLVFPFGGRNP PSSDELGGIITIASLNVAAVDTNLTAVSKNVAKTIQLYSVKSEQL LSTQGDASQVIGPLTEGQRRNVAVVNSLYKLHQSVTKAIHALMENA VQPLLTSGDAIEAIIITMHQEDFSGSLSSSGKPDVPCSLYMKELQGF IARVMSDYFKHFECLDFVFDNTEAIAQRAVELFIRHASLIRPLGEGG KMRLAADFAQMELAVGPFCCRVSDDLKSYRMLRSFRPLLQASEH VASSPALGDVIPFSIIIQFLFTRAPAEKSPFQRAEWSHTRFSQWLDD HPSEKDRLLLIRGALEYVQSVRSREGKEFAPVYPIMVQLLQKAMS ALQ	COG5
210	MAEGSGEVVAVSATGAANGLNNGAGGTSATTNPLSRKLHKILET RLDNDKEMLEALKALSTFFVENSRLRTRRNLRGDIERKSLAINEEFVSI FKEVKEELESISEDVQAMSNCCQDMTSRLQAAKEQTQDLIVKTTKL QSESQKLEIRAQVADAFSLKFQLTSDMSLLRGTRREGPITEDFFKAL GRVKQIHNDVKVLLRTNQQTAGLEIMEQMALLQETAYERLYRWAQ SECRTLTQESCDVSPVLTQAMEALQDRPVLYKYTLDEFGTARRSTV VRGFIDALTRGGPGGTPRPIEMHSHDPLRYVGDMLAWLHQATASE KEHLEALLKHVTTQGVENIQEVVGHITEGVCRPLKVRIEQVIVAEP GAVLLYKISNLLKFYHHTISGIVGNSATALLTTIEEMHLLSKKIFFNS LSLHASKLMDKVELPPPDLGPSSALNQTMLMLLREVLASHDSSVPL DARQADVFQVLSCVLDPLLQMCTVSASNLTADMATFMVNSLYM MKTTLALFEFTDRRLEMLQFQIEAHLDTLINEQASYVLTRVGLSYIY NTVQQHKPEQGLANMPNLDVTLKAAMVQFDRYLSAPDNLLIPQ LNFLLSATVKEQIVKQSTELVCRAYGEVYAAMNPINEYKDPENIL HRSPQQVQTLLS	COG6
211	MDFSKFLADDFDVKEWINAAFRAGSKEAASGKADGHAATLVMKL QLFIQEVNHAVEETSHQALQNMPKVLVDVEALKQEASFLKEQMILV KEDIKKFEQDTSQSMQVLVEIDQVKSRMQLAAESLQEADKWLSTLSA DIEETFKTQDIAVISAKLTGMQNSLMMMLVDPDYSEKCVHLEALKN RLEALASPQIVAAFTSQAVDQSKVFVKVFTEIDRMPQLLAYYYKCH KVQLLAAWQELCQSDLSLDRQLTGLYDALLGAWHTQIQWATQVF QKPHEVVMVLLIQTLGALMPSLPSCLSNNGVERAGPEQELTRLLEFY DATAHFAKGLEMALLPHLHEHNLVKVTELVDAVYDPYKPYQLKY GDMEESNLLIQMSAVPLEHGEVIDCVQELSHSVNKLFLGLASAAVDR CVRFTNGLGTGCLLSALKSLFAKYVSDFTSTLQSIRKKCKLDHIPPNS LFQEDWTAQNSIRIIATCGELLRHCGDFEQQLANRILSTAGKYLSDS CSPRSLAGFQESILTDKKNKSAKNPWQEYNYLQKDNPAEYASLMEIL YTLKEKGSSNHNLLAAPRAALTRLNQQAHQLAFDSVFLRIKQQLLLI SKMDSWNTAGIGETLTDELPAFSLTPLEYISNIGQYIMSLPLNLEPFV TQEDSALELALHAGKLPFPPEQGDPELDNMADNWLGSARATM	COG7

	QTYCDAILQIPELSPHSAKQLATDIDYLINVMDALGLQPSRTLQHIVT LLKTRPEDYRQVSKGLPRRLATTVATMRSVNY	
212	MATAATIPSVATATAAALGEVEDEGLLASLFRDRFPEAQWRERPDV GRYLRELSGSGLERLRREPERLAEEAQLLQQTRDLAFANYKTFIRG AECTERIHRLFGDVEASLGRLLDRLPSFQQSCRNFVKEAEEISSNRR MNSLTNLNRHTEILEILEIPQLMDTCVRNSYEEALELAAYVRRLERK YSSIPVIQIVNEVRQSMQLMLSQLIQQLRTNIQLPACLRVIGYLRRM DVFTAEALRVKFLQARDAWLRSILTAPNDDPYFHITKTIEASRVHLF DIITQYRAIFSDEPLLPPAMGEHTVNESAIFHGWLQKVSQFLQVL ETDLYRGIGGHLDSLLGQCMYFGLSFSRVGADFRGQLAPVFQRVAI STFQKAIQETVEKFQEEMNSYMLISAPAILGTSNMPAAVPATQPGTL QPPMVLLDFPPLACFLNNILVAFNDLRLCCPVALAQDVTGALEDAL AKVTKIILAFHRAEEAAAFSSGEQELFVQFCTVFLEDLVPYLNRLQV LFPPAQIAQTLGIPPTQLSKYGNLGHVNIGAIQEPLAFILPKRETLFTL DDQALGPELTAPAPEPPAEPRLEPAGPACPEGRAETQAEPPSVGP	COG8
213	DRLLQQGSAVFQFRMSANSGLLPASVMVPLLGLVMKERCQTAGNP FFERFGIVVAATGMAVALFSSVLALGITRPVPTNTCVILGLAGGVIIY IMKHSLSVGEVIEVLEVLIFVYLNLMILLYLLPRCFTPGEALLVLGGI SFVLNQLIKRSLTLVESQGDVPDFFLLVVVVGVMVLMGIFSTLFFVM DSGTWASSIFFHLMTCVLSLGVVLPWLHRLIRRNPLLWLLQFLFQTD TRIYLLAYWSLLATLACLVVLYQNAKRSSSESCKKHQAPTARKYFH LIVVATYIPGIHFDRLPLYVAATVCLAVFIFLEYVRYFRIKPLGHTLRS FLSLFLDERDSGPLILTHIYLLLGMSLPIWLIPRPTQKGSLLGGARAL VPYAGVLAVGVGDTVASIFGSTMGEIRWPGTKKTFEGTMTSIFAQII SVALILIFDSGVDLNYSYAWILGSISTVSLLEAYTTQIDNLLLPLYLLI LLMA	DOLK
214	MSWIKEGELSLWERFCANIIKAGPMPKHIAFIMDGNRRYAKKCQVE RQEGHSQGFNKLAEATLRWCLNLGILEVTVYAFSIENFKRSKSEVDGL MDLARQKFSRLMEEKEKLQKHGVCIRVLGDLHLLPLDLQELIAQAV QATKNYNKCFNLVCFAYTSRHEISNAVREMAWGVEQGLDPSDISE SLLDKCLYTNRSPHPDILRTSGEVRLSDFLLWQTSLSCLVFQPVW PEYTFWNLFMAILQFQMNHSLVQKARDMYAEERKQQLERDQATV TEQLLREGLQASGDAQLRRLRLHKLARSREERVQGFLLQALELKRAD WLARLGTASA	DHDDS
215	MWAFSELPMPLINLIVSLLGFVATVTLIPAFRGHFIAARLCGQDLN KTSRQQIPESQGVISGAVFLILFCFIPFPFLNCFVKEQCKAFPHHEFV ALIGALLAICCMIFLGFADDVLNLRWRHKLKLLPTAASLPLLMVYFTN FGNTTIVVPKPRPILGLHLDLGLILYYVYMGLLAVFCTNAINILAGIN GLEAGQSLVISASIIVFNLVELEGDCRDDHVFSLYFMIPFFFTLGLL YHNWYPSRVFVGDTFCYFAGMTFAVVGILGHFSKTMILLFFMPQVF NFLYSLPQLLHIIPCPRHRIPLNIKTGKLEMSYSKFKTKSLSFLGTFIL KVAESLQLVTVHQSETEDGEFTECNMTLINLLLKVLGPIHERNLTL LLLLLQILGSAITFSIRYQLVRLFYDV	DPAGT1
216	MASLEVSRSRPRSRRLEVRSPRQNKYSVLLPTYNERENLPLIVWLL VKSFSSEGINYEIIIDDGSPDGTRDVAEQLEKIYGS DRILLRPREKKL GLGTAYIHGMKHATGNYIIIMDADLSHHPKFIPEFIRKQKEGNFDIVS GTRYKGNNGGVYGWDLKRKIISRGANFLTQILLRPGASDLTGSRFLY RKEVLEKLEKCVSKGYVFQMEMIVRARQLNYTIGEVPISFVDRVY GESK LGGNEIVSFLKGLLTLFATT	DPM1
217	MATGTDQVVGLGLVAVSLIFTYYTAWVILLPFIDSQHVHVKYFLPR AYAVAIPLAAGLLLLLVGLFISYVMLKTKRVTKKAQ	DPM2

218	MTKLAQWLWGLAILGSTWVALTTGALGLELPLSCQEVLPWPAYL LVSAGCYALGTVGYRVATFHDCEDAARELQSQIQEARADLARRGL RF	DPM3
219	MESTLGAGIVIAEALQNQLAWLENVWLWITFLGDPKILFLFYFPAAY YASRRVGIAVLWISLITEWLNLIKWFLLFGDRPFWVHESGYYSQA PAQVHQFPSSCETGPGSPSGHCMITGAALWPIMTALSSQVATRARSR WVRVMPSLAYCTFLLAVGLSRIFILAHFPHQVLAGLITGAVLGWLM TPRPVMERELSFYGLTALALMLGTSLIYWTFTLGLDLSWSISLAFK WCERPEWIHVDSPFASLSRDSGAALGLGIALHSPCYAQVRRALG NGQKIACLVLAMGLLGPLDWLGHPQISLFYIFNFKYTLWPCLVL ALVPWAVHMFSAQEAPPIHSS	G6PC3
220	MCGIFAYLNYHVPRTREILETLIKGLQRLEYRGYDSAGVGFDDGN DKDWEANACKIQLIKKKGKVKALDEEVHKQQDMDLDIEFDVHLGI AHTRWATHGEPSPVNSHPQRSKDNNEFIVIHNGIITNYKDLKKFLES KGYDFESETDTETIAKLKMYMDNRESQDTSFTTLVERVIQLEGAF ALVFKSVHFPQGAVGTRRGSPLLIGVRSEHKLSTDHIPILYRTARTQI GSKFTRWGSQGERGKDKKGCNSLRVDSTCLFPVEEKAVEYYFAS DASAVIEHTNRVIFLEDDDDVAADVDDGRLSIHRIKRTAGDHPGRAVQ TLQMELOQIMKGNFSSFMQKEIFEQPEVVNTMRGRVNFDDYTVNL GGLKDHIEIQRCCRLLIACGTSYHAGVATRQVLEELTELPVMVEL ASDFLDNRTPVFRDDVCFFLSQSGETADTLMGLRYCKERGALT VGI TNTVGSISRETDCGVHINAGPEIGVASTKAYTSQFVSLVMFALMM CDDRISMQERRKEIMLGLKRLPDLIKEVLSMDDEIQKLATELYHQKS VLIMGRGYHYATCLEGALKIKEITYMHSEGILAGELKHGPLALVDK LMPVIMIIMRDHTYAKCQNALQQVVARQGRPVICDKEDTETIKNT KRTIKVPHSVDCLOGILSVIPLQLLAFHLAVLRGYDVDFPRNLAKSV TVE	GFPT1
221	MLKAVILIGGPQKGTRFRPLSFEVPKPLFPVAGVPMIQHHIEACAQV PGMQEILLIGFYQPDEPLTQFLEAAQQEFNLPVRYLQEFAPLGTGGG LYHFRDQILAGSPEAFFVLNADVCSDFPLSAMLEAHRQRHPFLLLG TTANRTQSLNYGCIVENPQTHEVLHYVEKPSTFISDIHNCGIYLSPEA LKPLRDVQQRNQDQGQLEDSPGLWPGAGTIRLEQDVFSALAGQGQI YVHL TDGIWSQIKSAGSALYASRLYLSRYQDTHPERLAKHTPGGPWIRGN VYIHPTAKVAPSAVLGPNVSGKGVTVGEGVRLRESIVLHGATLQEH TCVLHSIVGWGSTVGRWARVEGTPSDPNPNDRARMDSESLFKDG KLLPAITILGCRVRIPAEVLILNSIVLPHKELSRSTNQIIL	GMPPA
222	MKALILVGGYGTRLRPLTLSTPKPLVDFCNKPILLHQVEALAAAGV DHVILAVSYMSQVLEKEMKAQEQLGIRISMSHEEEPLGTAGPLAL ARDLLSETADPFFVLNSDVICDFPFQAMVQFHRHHGQEGSILVTKE EPSKYGVVVCEADTGRIHRFVEKPQVFVSNKINAGMYILSPAVLQRI QLQPTSIEKEVFPIMAKEGQLYAMELQGFWM DIGQPKDFLTGMCLF LQSLRQKQPERLCSPGIVGNVLVDPSARIGQNC SIGPNVSLGPGVV VEDGVCIRRCTVLRDARIRSHSWLESCIVGWRCRVGQWVRMENV VLGEDVIVNDELYLNGASVLPKHSIGESVPEPRIIM	GMPPB
223	MAARWRFWCVSVTMVVALIVCDVPSASAQRKKEMVLSEKVSQ MEWTNKRVPVIRMGDKFRRLVKAPPRNYSVIVMFTALQLHRQCVV CKQADEEFQILANSWRYSSAFTNRIFAMVDFDEGSDVFQMLNMNS APTFINFPAGKPKRGDTYELQVRGFSAEQIARWIADRTD VNIRVIR PPNYAGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFALCFVL AMTSGQMWNHIRGPPYAHKNPHTGHVNYIHGSSQAQFVAETHIVL	MAGT1

	LFNGGVTLGMVLLCEAATSDMDIGKRKIMCVAGIGLVVLFSSWML SIFRSKYHGYPSFLMS	
224	MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPH RDFISVTLFGENYDNSKSWRRRSCWRKWKQLSRLQRNMILFLLAF LLFCGLLFYINLADHWKALAFRLEEEQKMRPEIAGLKPANPPVLPAP QKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEAT KRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPP LPPARTQGTPVHLNRYQKGVIDVFLHAWKGYRKFAWGHDELKPV RSFSEWFGGLTLIDALDTMWILGLRKEFEEARKWVSKKLHFEKDV DVNLFESTIRILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI PYSDVNIGTGVAHPPRWTSSTVAEVTISIQLFRELSTLTDGKKFQE AVEKVTQHIHGLSGKKDGLVPMFINTHSGLFTHLGVFTLGARADSY YEYLLKQWIQGGKQETQLLEDYVEAIEGVTRTHLLRHSEPSKLT ELAHGRFSAKMDHLVCFPLGTALGVYHGLPASHMELAQELMETC YQMNRMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLPETVE SLFYLYRVGTDRKYQDWGWEILQSFSRFRVPSGGYSSINNVQDPQ KPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLDAYVFNTAEHPLPI WTPA	MAN1B1
225	MRFRYKRVKVLITLVVAACGFVLWSSNGRQRKNEALAPLLDAEP ARGAGGRGGDHPSVAVGIRRVSNVSAASLVPAVPQPEADNLT RSLVYQLNFDQTLRNVDKAGTWAPRELVLVQVHNRPEYLRLLLD SLRKAQGDNLVIFSHDFWSTEINQLIAGVNFPCVLQVFFPSIQLY PNEFPDSDPRDCPRDLPKNAALKLGCINAEYPDSFGHYREAKFSQTK HHWWWKLHFVWERVKILRDYAGLILFLEEDHYLAPDFYHVFKKM WKLKQCECDVLSLGTYSASRSF YGMADKVDVKTWKSTEHNMGALTRNAYQKLIECTDTFCTYDDY NWDWTLQYLTVSCLPKFWKVLVPQIPRIFHAGDCGMHHKKT TQSAQIESLLNNNKQYMPETLTISEKFTVVAISPPRKNNGWGD IRDHELCKSYRRLQ	MGAT2
226	MARGERRRAVPAEGVRTAERAARGGPGRRDGRGGGPRSTAGGV ALAVVVLALGMSGRWVLAWYRARRAVTLHSAPPVLPADSSSPA VAPDLFWGTYRPHVYFGMKTRSPKPLLTGLMWAQQGTTPTPKLR HTCEQGDGVGPYGWEFHDGLSFGQRHIQDGAALRLTTEFVKRPGGQ HGGDWSWRVTVEPQDSGTSALPLVSLFFYVTDGKEVLLPEVGAK GQLKFISGHTSELGDFRFTLLPPTSPGDTAPKYGSYNVFWTSNPGLP LLTEMVKSRLNSWFQHRPPGAPPERYLGLPGSLKWEDRGPSGQGGQ GQFLIQVTLKIPISIEFVFESGSAQAGGNQALPRLAGSLLTQAL ESHAEGRFRERFEKTFQLKEKGLSSGEQVLGQAALSGLLGGIGYFYGQGL VLPDIGVEGSEQKVDPALFPPVPLFTAVPSRSFFPRGFLWDEGFHQL VVQRWDPSLTREALGHWLGLLNADGWIGREQILGDEARARVPPEF LVQRAVHANPPTLLLPAVHMLEVGDPPDLAFLRKALPRLHAWFSW LHQSQAQPLPLSYRWRGRDPALPTLLNPKTLPSGLDDYPRASHPSVT ERHLDLRCWVALGARVLTRLAEHLGEAEVAAELGPLAASLEAAES LDELHWAPELGVFADFGNHTKAVQLKPRPPQGLVRVVGRPQPQLQ YVDALGYVSLFPLLLRLDPTSSRLGPLLDILADSRHLWSPFGLRSL AASSSFYQQRNSEHDPYWRGAVWLVNLYALGALHHYGHLEGP HQARAACKLHGLRANVVGNVWRQYQATGFLWEQYSDRDGRGMG CRPFHGWTSLVLLAMAEDY	MOGS
227	MAAEADGPKRLLVLPILLPEKCYDQFLVQWDLHVPCLKILLSKGL GLGIVAGSLLVKLPQVFKILGAKSAEGLSLQSVMLELVALTGTMVY SITNFPFSSWGEALFLMLQTITICFLVMHYRGQTVKGVAFACYGL	MPDU1

	VLLVLLSPLTPLTVVTLLQASNPVAVVVGRLLQAATNYHNGHTGQL SAITVFLFFGGSLARIFTSIQETGDPLMAGTFVVSSLCNGLIAAQLLF YWNAKPPHKQKKAQ	
228	MAAPRVFPLSCAVQQYAWGKMGSNSEVARLLASSDPLAQIAEDKP YAEIWMGTHPRGDAKILDNRISQKTLQWIAENQDSLGSKVKDTFN GNLPFLFKVLSVETPLSIQAHNPKELAEKLHLQAPQHYPDANHKPE MAIALTPFQGLCGFRPVVEIVTFLKKVPEFQFLIGDEAATHLKQTMS HDSQAVASSLQSCFSLMKSEKKVVVEQLNLLVKRISQQAAGNN MEDIFGELLQLHQYYPGDIGCFIYFLNLLTLKPGEAMFLEANVPH AYLKGDCVECMACSDNTVRAGLTP KFIDVPTLCEMPLSYTPSSSKDRLFLPTRSQEDPYLSIYDPPVPDFTIMK TEVPGSVTEYKVLALDSASILLMVQGTVIASPTTQTPIPLQRGGVLF IGANESVSLKLTPEPKDLLIFRACCLL	MPI
229	MAAAALGSSSGSASPAVAELCQNTPETFLEASKLLLLTYADNLRNPN DEKYRSIRIGNTAFSTRLLPVRGAVECLFEMGFEEGETHLIFPKKASV EQLQKIRDLIAIERSSRLDGSNKSCHKVSSQQAASSTQLPTTPSSNPS GLNQHTRNRQGSDDPPSASTVAADSAILEVLQSNIQHVLVYENPAL QEKALACIPVQELKRKSQEKLSRARKLDDKGINISDEDFLLLELLHWF KEE FFHWVNNVLCSKCGGQTRSRDRSLLPSDDELKWGAKEVEDHYCDA CQFSNRFPYNNPEKLLLETRCGRCGEWANCFTLCCRAVGFEARYV WDYTDHVWTEVYSPSQQRWLHCDACEDVCDKPLLIEIGWGKKLS YVIAFSKDEVVDVTWRYSCHEEVIARRTKVKEALLRDTINGLNKQ RQLFLSENRRKELLQRIIVELVEFISPKTPKPGELGGRISGSVAWRVA RGEMGLQRKETLFIPCENKISKQLHLCYNIVKDRYVRVSNNNQITIS GWENGVMKESIFRKVETDWHMVYLARKEGSSFAIYISWKFECGS VGLKVDISIRTSSTQTFQTGTVEWKLRSDDTAQVELTGDNSLHSAADF SGATEVILEAELSRGDGDVAWQHTQLFRQSLNDHEENCLEIIKFSDDL	NGLY1
230	MVKIVTVKTQAYQDQKPGTSGLRKRKVVFQSSANYAENFIQSIISTV EPAQRQEATLVVGGDGRFYMKEAIQLIARIAAANGIGRLVIGQNGIL STPAVSCIIRKIKAIIGGIILTASHNPGGPNNGDFGIKFNISNGGPAPEAIT DKIFQISKTIIEYAVCPDLKVDLGLVGLKQQFDLENKFKPFTVEIVDS VEAYATMLRSIFDFSALKELLSGPNRLKIRIDAMHGVVGPYVKKILC EELGAPANSVAVNCVPLEDFGGHHPDPNLTYAADLVETMKSGEHDF GAAFDGDGDRNMILGKHGFFVNPSDSVAVIAANIFSIPYFQQTGVRG FARSMPTSGALDRVASATKIALYETPTGWKFFGNLMDASKLSLCGE ESFGTGSDDHIREKDGLWAVLAWLSILATRKQSVEDILKDHQWKYGR NFFTRYDYEEVEAEGANKMMKDLEALMFDRSFVGKQFSANDKVY TVEKADNFEYSDPVDGSISRNLQGLRLIFTDGSRIVFRLSGTGSAGATI RLYIDSYEKDVAKINQDPQVMLAPLISIALKVSQQLQERTGRTAPTIVIT	PGM1
231	MDLGAITKYSALHAKPNGLILQYGTAGFRTKAHLDHVMFRMGLL AVLRSKQTKSTIGVMVTASHNPEEDNGVKLVDPPLGEMLAPSWEH ATCLANAEQDMQVRVLIDISEKEAVNLQQDAFVVIGRDTRPSSEKLS QSVIDGVTVLGGQFHDYGLLTPQLHYMVYCRNTGGRYGKATIEG YYQKLSKAFVELTKQASCSDGEYRSLKVDCANGIGALKLREMEHY FSQGLSVQLFNDGSKGKLNHLGADGVKSHQKPPQGMEIKSNERCC SFDGDADRIVYYYHDADGHFHLIDGDKIATLISSFLKELLVEIGESLN IGVVQTAYANGSSSTRYLEEVMKVPVYCTKTGVKHLHHKAQEFDIG VYFEANGHGTAFTSTAVEMKIKQSAEQLEDKKRKAAMLENIIDLF NQAAGDAISDMLVIEAILALKGLTVQQWDALYTDLPNRQLKVQVA DRRVISTTDAERQAVTPPGLQEAINDLVKKYKLSRAFVRPSGTEDV VRVYAEADSQESADHLAHEVSLAVFQLAGGIGERPQPGF	PGM3

232	MGSQEVLGHAARLASSGLLLQVLFRLITFVLNFIIRFLSKEIVGVV NVRLTLLYSTTLFLAREAFRRACLSGGTQRDWSQTLNLLWLTVP VFWSLFLGWIWLQLLEVDPNPVPHYATGVVLFGLSAVVELLGEPF WVLAQAHMFVKLVIAESLSVILKSVLTAFVLVWLPWGLYIFSLA QLFYTTVLVLCYVIYFTKLLGSPESTKLQTLVPSRITDILLPNITRNGA FINWKEAKLTWSFFKQSFLKQILTEGERYVMTFLNVLNFGDQGVYD IVNNLGSVLARLIFQPIEESFYIFFAKVLERGKDATLQKQEDVAVAA AVLESLLKLALLAGLTITVFGFAYSQALDIYGGTMLS SSGSGPVLLR SYCLYVLLLAINGVTECF TFAAMSKEEVDRYNFVMLALSSSFLVLS YLLTRWCGSVGFILANCFNMGIRITQSLCFIHRYYRRSPHRPLAGLH LSPVLLGTFALSGGVTA VSEVFLCCEQGWPARLAHIAVGAFCLGAT LGTAFLTETKLIHFLRTQLGVPRRTDKMT	RFT1
233	MATYLEFIQQNEERDGVRFWSNVWVPSRLEATRMVVPLACLLTPLK ERPDLPPVQYEPVLC SRPTCKAVLNPLCQVDYRAKLWACNFCFQRN QFPAYGGISEVNQPAELMPQFSTIEYVIQRGAQSPLIFLYVVDTCLE EDDLQALKESLQMSLSLLPPDALVGLITFGRMVQVHELSCGISKSY VFRGTDLTAKQIQDMLGLTKPAMPMQQAQPAQPQEHFPAASSRFL QPVHKIDMNLTDLLGELQRDPWPVTQGKRPLRSTGVALSIAVGLLE GTFPNTGARIMLFTGGPPTQGPVMVVGDELKPIRSWHDIEKD NARF MKKATKHYEMLANRTAANGHCIDIYACALDQTGLLEMKCCANLT GGYMVMGDSFNLSLFKQTFQRIFTKDFNGDFRMAFGATLDVKT SR ELKIAGAIGPCVSLNVKGPCVSENELGVGGTSQWKICGLDPTSTLGI YFEVVNQHNTPIPQGGGGAIQFVTHYQHSSTQRRIRVTTIARNWAD VQSQLRHIEAAFDQEA AAVLMARLGVFRAESEEGPDVLRWLD RQLI RLCQKFGQYNKEDPTSFRLSDFSFLYPQFMFHLRRSPFLQVFNNSPD ESSYRHHFARQDLTQSLIMIPILYSYSFHGPPEPVLLDSSSILADRI LLMDTFFQIVYILGETIAQWRKAGYQDMPEYENFKHLLQAPLDDAQ EILQARFPMPRYINTEHGGSQARFLLSKVNPSQTHNNLYAWGQETG APILTDDVSLQVFM DHLKKLAVSSAC	SEC23B
234	MAAPRDNVTLLFKLYCLAVMTLMAAVYTIALRYTRTSDKELYFST TAVCITEVIKLLLSVGILAKETGSLGRFKASLRENVLGSPKELLKLSV PSLVYAVQNNMAFLALSNDAAVYQVTYQLKIPCTALCTVLM LNR TSLKLQWVSVMFCAGVTLVQWKPAQATKVVEQNPLLGFGAIAI AVLCSGFAGVYFEKVLKSSDTS LWVRNIQMYLSGHI VTLAGVYLS GAEIKEKGFFYGYTYVWFVIFLASVGGLYTSVVVKYTDNIMKGFS AAAIVLSTIASVMLFGLQITLTFALGTLLVCVSIYLYGLPRQDTTSI QQGETASKERVIGV	SLC35A1
235	MAAVGAGGSTAAPGPGAVSAGALEPGTASAAHRLKYISLAVLVV QNASLILSIRYARTLPGDRFFATTAVVMAEVLKGLTCLLLLFAQKRG NVKHLVFLHEAVLVQYVDTLKLAVPSLIYTLQNNLQYVAISNLPA ATFQVTYQLKILTTALFSVLM LNRSLRSLQWASLLLLFTGVAIVQAQ QAGGGGPRPLDQNPAGLA AVVASCLSSGFAGVYFEKILKGSSGSV WLRNLQLGLFGTALGLVGLWWAEGTAVATRGGFFGYTPAVWGVV LNQAFGGLLVAVVVKYADNILKGFATSLSIVLSTVASIRLFGFHVDP LFALGAGLVIGAVYLYSLPRGA AKAIASASASASGPCVHQPPGQPP PPQLSSHRGDLITEPFLPKLLTKVKGS	SLC35A2
236	MNRAPLKR SRILHMALTGASDPSAEAEANGEKPFLLRALQIALVVS LYWVTSISMVFLNKYLLDSPSLRLDTPIFVTIFYQLVTTLLCKGLSA LAACCPGAVDFPSRLDLRVARSVLPLSVVFIGMITFNNLCLKYVGV AFYNVGRSLTTVFNVLLSYLLLKQTTSFYALLTCGIIIIGGFWLGVDQ EGAEGTLSWLGT VFGVLASLCVSLNAIYTTKVLP AVDGSIWRLTFY	SLC35C1

	NNVNACILFLPLLLLLGELQALRDFAQLGSAHFWGMMTLGGLFGFA IGYVTGLQIKFTSPLTHNVSG TAKACAQTVLAVLYYEETKSFLWWTSNMMVLGGSSAYTWVRGW EMKKTPEEPSPKDSEKSAMGV	
237	MAAMASLGALALLLSSLRCSAEACLEPQITPSYYTTSDAVISTET VFIVEISLTCKNRVQNMALYADVGGKQFPVTRGQDVGRYQVSWSL DHKSAHAGTYEVRFFDEESYLLRKAQRNEDISIIPPLFTVSVDHR GTWNGPWVSTEVLAAAIGLVIYYLAFSAKSHIQA	SSR4
238	MAPWAEAEHSALNPLRAVWLTLTAAFLLLTLLQLLPGLLPGCAIF QDLIRYGKTKCGEPSRPAACRAFDVPKRYFSHFYIISVLWNGFLLWC LTQSLFLGAPFPSWLHGLLRILGAAQFQGELALSFLVLVFLWLHS LRRLFECLYVSVFSNVMIHVVQYCFGLVYYVLVGLTVLSQVPMDG RNAYITGKNLLMQARWFHILGMMMFIWSSAHQYKCHVILGNLRKN KAGVVIHCNHRIPFGDWFEYVSSPNYLAELMIYVSMVTFGFHNL WWLVVTNVFFNQALSAFLSHQFYKSKFVSYPKHRKAFLPFLF	SRD5A3
239	MAAAPGNGRASAPRLLLLFLVPLWAPAAVRAGPDEDLSHRNKE PPAPAQQQLQPQPVAVQGPEPARVEKIFTPAAPVHTNKEDPATQTNL GFIHAFVAAISVIIVSELGDKTFFIAAIMAMRYNRLTVLAGAMLALG LMTCLSVLFGYATTVIPRVYTYVSTVLFAIFGIRMLREGLKMSPDE GQEELEEVQAEKKKDEEFQRTKLLNGPGDVETGTSITVPQKKWLH FISPIFVQALTTLFLAEWGDRSQTITVLAAREDPYGVAVGGTVGHC LCTGLAVIGGRMIAQKISVRTVTIIGGIVFLAFAFSALFISPDGFG	TMEM165
240	MSSWLGGLGSLGQSLGQVGGSLASLTGQISNFTKDMLMEGTEEV EALPDSRTKEIEAIHAILRSENERLKKLCTDLEEKHEASEIQIKQOST SYRNQLQQKEVEISHLKAQIALQDQLKLQSAQSVPSGAGVPAT TASSSFAYGISHPSAFHDDDMDFGDISSSQEINRLSNEVSRLESEV GHWRIHAQTSKAQGTDNDSQSEICKLQNIKELKQNRSQEIDDHQHE MSVLQNAHQKQLTEISRRHREELSDYEERIEELENLLQGGSGVIET DLSKIYEMQKTIQVLQIEKVESTKKMEQLEDKIKDINKKLSSAENDR DILRREQEQLNVEKRQIMEECENLKLECSKLQPSAVKQSDTMTEKE RILASQASVEEVFRLQQALSDAENEIMRLSSLNQDNSLAEDNLKLK MRIEVLEKEKSLLSQEKEELQMSLLKLNNEYEVIKSTATRDISLSEL HDLRLNLEAKEQELNQSISEKETLIAEIEELDRQNQEATKHMILIKDQ LSKQQNEGDSIISKLKQDLNDEKKRVHQLEDDKMDITKELDVQKEK LIQSEVALNDLHLTKQKLEDKVENLVDQLNKSQESNVSQKENLEL KEHIRQNEEELSRIENELMQSLNQDSNSNFKDTLLKEREAEVRNLKQ NLSELEQLNENLKKVAFDVKMENEKLVLACEDVRHQLEECLAGNN QLSLEKNTIVETLKMEKGEIEAELCWAKKRLLEEANKYEKTIEELSN ARNLNTSALQLEHEHLIKLNQKKDMEIAELKKNIEQMDTDHKETKD VLSSSLEEKQQLTQLINKKEIFIEKLKERSSKLQEELDKYSQALRKNE ILRQTIEEKDRSLGSMKEENNLHLEELERLREEQSRTAPVADPKTLD SVTELASEVSQLNTIKEHLEEEIKHHQKIIDQDNQSKMQLLQSLQE KKEMDEFQRYQHEQMNAHTHTQLFLEKDEEIKSLQKTIEQIKTQLHEER QDIQTDNSDIFQETKVQSLNIENGSEKHDLSKAETERLVKGIKERELE IKLLNEKNISLTKQIDQLSKDEVGKLTQIIQKQDLEIQALHARISSTSH TQDVVYLQQLQAYAMEREKVFAVLNEKTRENSHLKTEYHKMMD IVAAKEAALIKLQDENKKLSTRFESSQDMFRETIQNLRIIREKDIEI DALSQKCQTLTAVLQTSSTGNEAGGVNSNQFEELLQERDKLKQQV KKMEEWKQQVMTTVQNMQHESAQLQEELHQLQAQVLVDSNNS KLQVDYTGLIQSYEQNETKLKNFGQELAQVQHSIGQLCNTKDLLL KLDIISPQLSSASLLTPQSAECLRASKSEVLSESELQLEELRKS	TRIP11

	QEKDATIRTLQENNHRLSDSIAATSELERKEHEQTDSEIKQLKEKQD VLQKLLKEKDLLIKAKSDQLLSSNENFTNKVNENELLRQAVTNLKE RILILEMDIGKLKGENEKIVETYRGKETHEYQALQETNMKFSMMLRE KEFECHSMKEKALAFEQLLKEKEQGKTGELNQLLNAVKSQMKEKTV VFQQERDQVMLALKQKQMENTALQNEVQRLRDKEFRSNQELERLR NHLLESDSYTREALAAEDREAKLRKKVTVLEEKLVSSSNAMENAS HQASVQVESLQEQNLNVSKQRDETALQLSVSQEQVKQYALSLANL QMVLEHFQEEKAMYSAELEKQKQLIAEWKKNAENLEGKVISLQE CLDEANAALDSASRLTEQLDVKEEQIEELKRQNELRQEMLDDVQK KLMSLANSSSEGKVDKVLMRNLFIGHFHTPKNQRHEVLRLMGSILGV RREEMEQLFHDDQGGVTRWMTGWLGGGSKSVNTPLRPNQQSSV NSSFSELFVKFLETESHPSIPPKLSVHDMKPLDSPGRRKRDTNAPES FKDTAESRSGRRTDVNPFLAPRSAAVPLINPAGLPGGGPHLLLKPIS DVLPTFTPLPALPDNSAGVVLKDLLKQ	
241	MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLLCIQLGGGQKKKE NLLAEKVEQLMEWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTA LQPQRQCSVCRQANEYQILANSWRYSSAFCKLFFSMVDYDEGT DVVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAAEQLAKWIA DRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTG WAMVSLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVSYIHGSSQA QFVAESHIILVLNAITMGMVLLNE AATSKGDVGKRRIICLVGLGLVVFSSFLLSIFRSKYHGYPSDLDFE	TUSC3
242	MVCVLVLAAAAGAVAVFLILRIWVVLRSMDVTPRESLSILVVAGSG GHTTEILRLLGSLSNAYSPRHYVIADTDEMSANKINSFELDRADRP SNMYTKYIHRIPRSREVQQSWPSTVFTTLHSMWLSFPLIHRVKPDL VLCNGPGTCVPICVSALLGILGIKKVIIYVESICRVETLSMSGKILF HLSDYFIVQWPALKEKYPKSVYLGRIV	ALG14
243	MRLREPLLSGSAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGR DLSRLPQLVGSTPLQGGSNSAAAIGQSSGELRTGGARPPPLGASS QPRPGGDSSPVVDSGPGPASNLTSVPVPHTTALSLPACPEESPLLVP MLIEFNMPVDLELVAKQNPVVKMGGRYAPRDCVSPHKVAIIIPFRN RQEHLKYWLYLHPVLQRQQLDYGIYVINQAGDTIFNRAKLLNVGF QEALKDYDYTCFVFSVDLIPMNDHNAYRCFSQPRHISVAMDKFGF SLPYVQYFGGVSALSQKQFLTNGFPNNYWGWWGGEDDDIFNRLVFR GMSISRPNVVGRCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDG LNSLTYQVLDVQRYPLYTQITVDIGTPS	B4GALT1
244	MGYFRCARAGSFGRRRKMEPSTAARAWALFWLLLPLLGAVCASGP RTLVLLDNLNVRETHSLFFRSLKDRGFELTFKTADDPSSLIKYGEFL YDNLIIFPSVEDFGGNINVETISAFIDGGGSVLVAASSDIGDPLREL SECGIEFDEEKTAVIDHHNYDISDLGQHTLIVADTENLLKAPTIVGKS SLNPILFRGVGMVADPDNPLVLDILTGSSTSYFFPDKPITQYPHAVG KNTLLIAGLQARNNARVIFSGSLDFFSDSFFNSAVQKAAPGSQRYVSQ TGNIELAVALSRLWVFKEEGLRVGPVSHHRVGETAPPNAYTVTDL VEYSIVIQQLSNGKWVPFDGDDIQLEFVRIDPFVRTFLKKKGGKYSV QFKLPDVYGVFQFKVDYNRLGYTHLYSSTQVSVRPLQHTQYERFIP SAYPPYASAFSMMLGLFIFISIVFLHMKEKEKSD	DDOST
245	MTGLYELVWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAA SAAVLAPLGFTRLRKPPAVGRNRRHHRHPRGGSCLA AAHHRMRWR ADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVVWCMAVGISYISVY DHQGIFKRNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQV LNCHLAVKVLSPEDGKADIVRAAQDFCQLVAQKQKRPTDLVDVTL	NUS1

	ASLLSSNGCPDPLVLKFGPVDSTLGLFPWHIRL TEIVSLPSHLNISYE DFFSALRQYAAACEQRLGK	
246	MAPPGSSTVFLALTIHASTWALTPTHYLT KHDVERLKASLDRPFTN LESIFYSIVGLSSLGAQVPDAKKACTYIRSNLDPSNVDSLIFYAAQAS QALSGCEISISNETKDLLLLAAVSEDSSVTQIYHAVAALSGFGLPLASQ EALSALTARLSKEETVLATVQALQTASHLSQQADLRSIVEEIEDLVA RLDELGGVYLQFEEGLETTALFVAATYKLMDHVGTEPSIKEDQVIQ LMNAIFS KKNFESLSEAFSVASAAAVLSHNRYPVVPVVVPEGSASD THEQAILRLQVTNVLSQLPTQATVKLEHAKSVASRATVLQKTSFTP VGDVFELNFMNVKFSSGYDFLVEVEGDNR YIANTVELRVKISTEV GITNVLDLSTVDKDQSIAPKTTRVTYPAKAKGTFIADSHQNFALFFQL VDVNTGAELTPHQTFRVLHNQKTGQEVVFVAEPDNKNVYKFELDT SERKIEFDSASGT YTLYLIGDATLKNPILWNVADVVIKFPEEEAPST VLSQNLFTPKQEIQHLFREPEKRPPTV VSNTFTALILSPLLLLFALWIRIGANVSNTFAPSTIIFHLGHAAMLGL MYVYFELQLNMFQTLKYLAILGSVTF LAGNRMLAQQA V KRTAH	RPN2
247	MTTYLEFIQQNEERDGVRF SWNVWPSSRLEATRMVVPVAALFTPLK ERPDLPIQYEPVLC SRTTCRAVLNPLCQVDYRAKLWACNFCYQRN QFPSPYAGISELNQPAELLQFSSIEYVVL RGPQMPLIFLYVVDTCME DEDLQALKESMQMSLSLLPPTALVGLITFGRMVQVHELGC EGISKS YVFRGTKDLSAKQLQEMGLSKVPLTQATRGPVQVQPPPSNRFLQP VQKIDMNLTDLLGELQRDPWPVPQGKRPLRSSGV ALSIAVGLLECT FPNTGARIMMFIGGPATQGGPM VVGDELKTPIRSWHDIDKDNAKYVKKG TKHFEALANRAATTGHVI DIYACALDQTGLLEMKCCPNLTGGY MVMGDSFNTSLFKQTFQRVF TKDMHGQFKMGFGGTLEIKTSREIKISGAIGPCVSLNSKGPCVSENEI GTGGTCQWKICGLSPTTTLAIYFEVVNQHNAPIPQGGRGAIQFVTQY QHSSGQRRIRVTTIARNWADAQTQIQNIAASFDQEAAILMARLAIY RAETEEGPDVLRWLDRQLRLCQKFGEYHKDDPSSFRFSETFSLYPQ FMFHLRPPSSFLQVFNNSPDESSYRHHFMRQDLTQSLIMIQPILYAY SFSGPPEPRLDSSILADRILLMDTFFQILYHGETIAQWRKSGYQD MPEYENFRHLLQAPVDDAQEILHSRFPMPRYIDTEHGGSQARFLLSK VNPSQTHNNMYAWGQESGAPILTDDVSLQVFMDHLKKLAVSSAA	SEC23A
248	MFANLKYVSLGILVFQTTSLVLTMRYSRTLKEEGPRYLSSTAVVVA ELLKIMACILLVYKDSKCSLRALNRVLHDEILNKPMETLKLAIPSGIY TLQNNLLYVALSNLDAATYQVTYQLKILTTALFSVSMLSKKLGVYQ WLSLVILMTGVAFVQWPSDSQLDSKELSAGSQFVGLMAVLTACFSS GFAGVYFEKILKETKQSVWIRNIQLGFFGSIFGLMGVYIYDGELVSK NGFFQGYNRLTWIVVVLQALGGLVIAAVIKYADNILKGFATSLSIILS TLISYFWLQDFVPTSVFFLGAILVITATFLYGYDPKPAGNPTKA	SLC35A3
249	MGLLVFVRNLLLALCLFLVLGFLYYSAWKLHLLQWEEDSNSVVL S FDSAGQTLGSEYDRLGFLNLD SKLPAELATKYANFSEGACKPGYA SALMTAIFPRFSKPAPMFLDD SFRKWARIREFVPPFGIKGQDNLIKAI LSVTKEYRLTPALDSLRCRRCIIVGNNGVLANKSLGSRIDDYDIVVR LNSAPVKGFEDVGSKTTLRITYPEGAMQRPEQYERDSLFLVLAGFK WQDFKWLKYIVYKERSASDGFWKS VATRPKEPPEIRILNPYFIQE AAFTLIGLPFNNGLMGRGNIPTLGSVAVTMALHGCDEVAVAGFGY DMSTPNAPLHYETVRMAAIKESWTHNIQREKEFLRKLVKARVITD LSSGI	ST3GAL3
250	MTKFGFLRLSYEKQDTLLKLLLSMAAVLSFSTRLFAVLRFESVIHEF DPYFNYRTRFLAEEGFYKFHNWFD DRAWYPLGRIIGGTIYPGLMIT SAAIYHVLHFFHITIDIRNVCVFLAPLFSSFTTIVTYHLTKELKDAGA	STT3A

	GLLAAAMIAVVPGYISRSVAGSYDNEGIAIFCMLLTYYMWIKAVKT GSICWAAKCALAYFYMVSSWGGYVFLINLIPLHVLVLMLTGRFSHR IYVAYCTVYCLGTILSMQISFVGFQPVLSSEHMAAFGVFGLCQIHAF VDYLRSKLNPPQQFEVLFRSVISLVGFVLLTVGALLMLTGKISPWTGR FYSLLDPSYAKNNIPIASVSEHQPTTWSSYYFDLQLLVFMFPVGLYY CFSNLSGARIFIIMYGVTSMYFSAVMVRLMLVLAPVMCILSGIGVSQ VLSTYMKNLDIRPDKSKKKQDSTYPIKNEVASGMILVMAFFLITY TFHSTWVTSEAYSSPSIVLSARGGDGSRIFDDFREAYYWL RHNTPE DAKVMSWWDYGYQITAMANRTILVDNNTWNNTHISRVGQAMAST EEKAYEIMRELDVSYVLVIFGGTGYSSDDINKFLWMVRIGGSTDT GKHIKENDYYTPTGEFRVDREGSPVLLNCLMYKMCYYRFGQVYTE AKRPPGFDRVRNAEIGNKDFELDVLEEAYTTEHWLVRIYKVKDLDN RGLSRT	
251	MAEPSAPESKHKSSLNSSPWSGLMALGNSRHGHGPGAQCAHKA GGAAPPKPAPAGLSGGLSQPAGWQSLLSFTILFLAWLAGFSSRLFAV IRFESIIHEFDPWFNYRSTHHLASHGFYEFLNWFDERAWYPLGRIVG GTVYPGLMITAGLIHWILNTLNTVHIRDVCVFLAPTFSGLTSISTFLL TRELWNQGAGLLAACFIAIVPGYISRSVAGSFDNEGIAIFALQFTYYL WVKSVKTGSVFWTMCCCLSYFYMVSAWGGYVFIINLIPLHVFVLLL MQRYSKRVIAYSTFYIVGLILSMQIPFVGFQPIRTSEHMAAAGVFA LLQAYAFLLQYLRDRLTKQEFQTLFLLGVSLAAGAVFLSVIYLTGTG YIAPWSGRFYSWDTGYAKIHIPIASVSEHQPTTWVSFFFDLHILVC TFPAGLWFCIKNINDERVFVALYASAVYFAGVMVRLMLTLTPVVC MLSAIAFSNVFEHYLGDDMKRENPPVEDSSDEDDKRNQGNLYDKA GKVRKHATEQEKTTEEGLGPNIKSIVTMLMLMLMMFAVHCTWVTS NAYSSPSVVLASYNHDGTRNILDFFREAYFWLRQNTDEHARVMSW WDYGQYIAGMANRTTLVDNNTWNNSHIALVGKAMSSNETAAYKI MRTLDVDYVLVIFGGVIGYSGDDINKFLWMVRIAGEHPKDIRES YFTPQGEFRVDKAGSPTLLNCLMYKMSYYRFGEMQLDFRTPPGFD RTRNAEIGNKDIKFKHLEEAFTSEHWLVRIYKVKAPDNRETLDHKP RVTNIFPKQKYLKSKTTKRKRGYIKNLVFKKGKKKSKKTV	STT3B
252	MARKSNLPVLLVPFLLCQALVRCSSPLPLVVNTWPFKNATEAAWR ALASGGSALDAVESGCAMCEREQCDSVGFSGSPDELGETTLDAMI MDGTTMDVGAVGDLRRIKNAIGVARKVLEHTHTLLVGESATTFA QSMGFINEDLSTTASQALHSDWLARNCPNYWRNVIPDSKYCGPY KPPGILKQDPIHKETEDDRGHDTIGMVVIHKTGHIAAGTSTNGIKFK IHGRVGDSPGAGAYADDTAGAAAATGNGDILMRFLPSYQAVEY MRRGEDPTIACQKVISRIQKHFPF FGAVICANVTGSYGAAKNLSTFTQFSFMVYNSEKNQPTTEKVDCI	AGA
253	MGAPRSLLLALAAGLAVARPPNIVLIFADDLGYGDLGCYGHPSSTTP NLDQLAAGGLRFTDFYVPVSLCTPSRAALLTGRLPVRMGMYPGVL VPSSRGGLPLEEVTVAEVLAAARGYLTMAGKWHLGVGPEGAFLPP HQGFHRFLGIPYSHDQGPCQNLTCFPPATPCDGGCDQGLVPIPLAN LSVEAQPPWLPGLEARYMAFAHDLMAAQRQDRPFFLYYASHHTH YPQFSGQSFAERSGRGPFGDSLMELDAAVGTLMTAIGDLGLLEETL VIFTADNGPETMRMSRGGCSGLLRC GKGTTYEGGVREPALAFWPGHIAPGVTHELASSLDLLPTLAALAGA PLPNVTLDDGFDLSPLLLGTGKSPRQSLFFYPSYPDEVGRGVFAVRTGK YKAHFFTQGSASDTTADPACHASSSLTAHEPPLLYDLSKDPGENY NLLGGVAGATPEVLQALKQLQLLKAQLDAAVTFGPSQVARGEDPA LQICCHPGCTPRPACCHCPDPA	ARSA

254	MGPRGAASLPRGPGPRRLLLPVVLPLLLLLLLAPPGSGAGASRPPHL VFLLADDLGWNDVGFHGSRI RTPHL DALAAGGVLLDNYYTQPLCT PSRSQLLTGRYQIRTGLQHIIWPCQPSCVPLDEKLLPQLLKEAGYTT HMGVKGWHLGMYRKECLPTRRGFDYFGYLLGSEDYYSHERCTLID ALNVTRCALDFRDGEEVATGYKNMYSTNIFTKRAIALITNHPPEKPL FLYLALQSVHEPLQVPEEYLKPYDFIQDKNRHHYAGMVSLMDEAV GNVTAALKSSGLWNNTVFIFSTDNGGQTLAGGNNWPLRGRKWSL WEGGVRGVGFVASPLLKQKGVKNRELIHISDWLPTLVKLARGHTN GTKPLDGFVDVWKTISEGSPSPRIELLHNIDPNFVDSSPCPRNSMAPAK DDSSLPEYSAFNSTVHAAIRHGNWKLLTGYPGCGYWFPPSQYNVS EIPSSDPPTKTLWLFIDIRDPEERHDL SREYPHIVTKLLSRLQFYHKKH SVPVYFPAQDPRCDPKATGVWGPWM	ARSB
255	MPGRSCVALVLLAAAVSCAVAQHAPPWTEDCRKSTYPPSGPTYRG AVPWYTINLDLPPYKRWHEMLMDKAPVLKVIVNSLKNMINTFVPSG KIMQVVDEKLPGLLGNFPGPFEEEMKGIAAVTDIPLGEIISFNIFYELF TICTSIVAEDKKGHLIHGRNMDFGVFLGWNINNDTWVITEQLKPLTV NLDFQRNKT VFKASSFAGYVGMLTGFKPGLFSLTLNERFSINGGY LGILEWILGKKDVMWIGFLTRTVLENSTSYEEAKNLLTKTKILAPAY FILGGNQSSEGCVITRDRKESLDVYELDAKQGRWYVVTNYDRWK HPFFLD DRRTPAKMCLNRTSQENISFETMYDVLSTKPVLNKLT VYT TLIDVTKGQFETYLRDCPDPCIGW	ASAH1
256	MSADSSPLVGSTPTGYGTLTIGTSIDPLSSSVSSVRLSGYCGSPWRVI GYHVVVWMMAGIPLLLFRWKPLWGVRLRLRPCNLAAHETLVIEIR DKEDSSWQLFTVQVQTEAIGEGSLEPSQSQ AEDGRSQA AVGAVPE GAWKDTAQLHKSEEAVSVGQKRVLRYLFGQGRYIWIETQQAFYQ VSLLDHGRSCDDVHRSRHGLSLQDQMVRAIYGPNVISIPVKSYPQ LLVDEALNPYYGFQAFSIALWLADHYWYALCIFISSISICLSLYKT RKQSQT LRDMVKLSMRVCVCRPGGEEEWVDSSELVPGDCLVLPQE GGLMPCDAALVAGECMVNESSLTGESIPVLKTALPEGLGPYCAETH RRHTLFCGTLILQARAYVGPVHLAVVTRTGCTAKGGLVSSILHPRP INFKFYKHS MKFVAALSVLALLGTIYSIFILYRNRVPLNEIVIRALDL VTVVVPPALPAAMTVCTLYAQSRLLRQGFICHIPLRLNLGGKLQLVC FDKTGTLTEDGLDVMGVVPLKQQAFLPLV PEPRRLPVGPLLRALATCHALSRLQDTPVGDPMDLKMVESTGWVL EEEPAAADSAGTQVLAVMRPPLWEPQLQAMEEPPVPVSVLHRFPFS SALQRMSVVVAWPGATQPEAYVKGSPELVAGLCNPETVPTDFAQM LQSYTAAGYRVVALASKPLPTVPSLEAAQQLTRDTVEGDLSLLGLL VMRNLLKPQTTPVIQALRRTRIRAVMVTGDNLQTAVTVARGCGMV APQEHLIIVHATHPERGQPASLEFLPMESPTAVNGVKDPDQAASYTV EPDPRSRHLALS GPTFGIIVKHFPKL LPKVLVQGTVFARMAPEQKTEL VCELQKLQYCVGMCGDGANDCG ALKAADVGISLSQAEASVSPFTSSMASIECVPMVIREGRCSLDTSFS VFKYMALYSLTQFISVLILYTINTNLGDLQFLAIDL VITTTVAVLMSR TGPALVLGRVRPPGALLSVPVLSLLQLQMVLTGVQLGGYFLTLAQ PWFPVPLNRTVAAPDNLPNYENTVVFSLSFQYLILAAAVSKGAPFRR PLYTNVPFLVALALLSSVLVGLVLPGLLQGPLALRNITDTGFKLLL LGLVTLNFGAFMLESVLDQCLPACLRRLRPKRASKKRFKQLEREL AEQPWPPLPAGPLR	ATP13A2
257	MGGCAGSRRRFS DSEGEETVPEPRLPLLDHQGAHWKNAVGFWLLG LCNNFSYVVMLSAAHDILSHKRTSGNQSHVDPGPTPIPHNSSSRFDC NSVSTAAVLLADILPTLVIKLLAPLGLHLLPYSPRVLVSGICAAGSFV LVAFSHSVGTSLCGVVFASISSGLGEVTFLSLTAFYPRAVISWWSSG TGGAGLLGALS YLGLTQAGLSPQQTLLSMLGIPALLASYFLLLTSP	CLN3

	EAQDPGEEEEAESAAARQPLIRTEAPESKPGSSSSLSLRERWTVFKGL LWYIVPLVVVYFAEYFINQGLFELLFFWNTSLSHAQQYRWYQMLY QAGVFASRSSLRCCRIRFTWALALLQCLNLVLLADVWFGFLPSIYL VFLIILYEGLLGGAAYVNTFHNIALETSDEHREFAMAATCISDTLGIS LSGLLALPLHDFLCQLS	
258	MAQEVDTAQGAEMRRGAGAARGRASWCWALALLWLAVVPGWS RVSGIPSRRHWPVPYKRFDFRKPDPYCQAKYTCPTGSPVMEGD DDIEVFRLQAPVWEFKYGDLLGHLKIMHDAIGFRSTLTGKNYTME WYELFQLGNCTFPHLRPEMDAPFWCNQGAACFFEGIDDVHWKENG TLVQVATISGNMFNQMAKWKQDNETGIYYETWNVKASPEKGAE TWFDSDYCSKFVLRFTFNKLAEFGAEFKNIETNYTRIFLYSGEPTYLG NETSVFGPTGNKTLGLAIKRFYYPFKPHLPTKEFLLSLLQIFDAVIVH KQFYLFYNFEYWFLPMKFPFIKITYYEIIPLPIRNKTLSSGL	CLN5
259	MEATRRRQHLGATGGPGAQLGASFLQARHGSVSADEAARTAPFHL DLWFYFTLQNWVLDGFRPIAMLVFPLEWFPLNKPVSVDYFHMAYN VITPFLLLKLIERSPRTLPRSITYVSIIFIMGASIHVGDSDVNHRLLSFG YQHLSVRENPIIKNLKPETLIDSFELLYYYDEYLGHCMWYIPFLIL FMYFSGCFTASKAESLIPGALLLVAPSGLYYWYLVTEGQIFILFIFT FAMLALVLHQKRKRLFLDSNGLFLFSSFALTLLLVALWVAWLWND PVLRRKKYPGVIYVPEPWAFYTLHVSSRH	CLN6
260	MNPASDGGTSESIFDLDYASWGIRSTLMVAGFVFYLGVFVCHQLS SSLNATYRSLVAREKVFWDLAATRAVFGVQSTAAGLWALLGDPVL HADKARGQQNWCWFHITTATGFFCFENVAVHLSNLIFRTFDLFLVI HHLFAFLGFLGCLVNLQAGHYLAMTLLLEMTPTFCVSWMLLKA GWSESLFWKLNQWLMIHMFHCRMVLTYHMWWVCFWHWDGLVS SLYLPHTLTLFLVGLALLTLINPYWTHKKTQQLNPDWNFAQPEA KSRPEGNGQLLRKKRP	CLN8
261	MIRNWLTIIFLPLKLVEKCESSVSLTVPPVVKLENGSSSTNVSLTLRP PLNATLVITFEITFRSKNITILELPDEVVVPVGTNSSFQVTSQNVGQL TVYLHGNHNSQTGPRIRFLVIRSSAISINQVIGWIYFVAWSISFYQV IMNWRKRKSVIGLSFDFVALNLTGFVAYSVFNIGLLWVPYIKEQFLK YPNGVNPVNSNDVFFSLHAVVLTIIIVQCCLYERGGQRVSWPAIGF LVLAWLFAFVTMIVAAGVTTWLQFLFCFSYIKLAVTLVKYFPQAY MNFYKSTEGWSIGNVLLDFTGGSFSLQMFLQSYNNQWTLIFGD PTKFGGLGVFSIVFDVFFIHFCLYRKRPGYDQLN	CTNS
262	MIRAAPPPLFLLLLLLLLLLVSWASRGEAAPDQDEIQRPLGLAKQPSF RQYSGYLKSGSGSKHLHYWFVESQKDPENSPVVLWLNGGPGCSSLD GLLTHEGPFLVQPDGVTLEYNPYSWNLIANVLYLESPAGVGFSYSD DKFYATNDTEVAQSNFEALQDFFRLFPEYKNNKLFLTGESYAGIYIP TLAVLVMQDPSMNLQGLAVGNGLSSYEQNDNSLVYFAYYHGLLG NRLWSSLQTHCCSQNKCNFYDNKDLECVTNLQEVARIVGNSGLNIY NLYAPCAGGVPSHFREYKDTVVVQD LGNIFTRLPLKRMWHQALLRSGDKVRMDPPCTNTTAASTYLNPNY VRKALNIPEQLPQWDMCNFLVNLQYRRLYRSMNSQYLKLLSSQKY QILLYNGDVMACNFMGDEWFVDSLQKMEVQRRPWLVKYGDS GEQIAGFVKEFSHIAFLTIKGAGHMVPTDKPLAAFTMFSRFLNKQPY	CTSA
263	MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIA KGPVSKYSQAVPAVTEGPIPEVLKNYMDAQYYGEIGIGTPPQCFTV VFDTGSSNLWVPSIHCKLLDIACWIHHKYNSDKSSTYVKNGTSDIH YSGSLSGYLSQDTVSVPCQSASSASALGGVKVERQVFG EATKQPGITFIAAKFDGILGMAYPRISVNNVLPVFDNLMQQKLVDQ NIFSFLSRDPDAQPGGELMLGGTDSKYKGSLSYLVNTRKAYWQ	CTSD

	VHLDQVEVASGLTLCKEGCEAIVDTGTSLMVGVPVDEVRELQKAIGA VPLIQGEYMIPCEKVSTLPAITLKLGGKGYKLSPEYTLKVSQAGKT LCLSGFMGMDIPPPSGPLWILGDVFIGRYYTVFDRDNNRVGF AEAA RL	
264	MAPWLQLLSLLGLLPGAVAAPAQPRAASFQAWGPPSPELLAPTRFA LEMFNRGRAAGTRAVLGLVRGRVRRAGQGSLSLEATLEPPCND PMVCRLPVSKKTLCSFQVLDELGRHVLLRKDCGPVDTKVPGAGEP KSAFTQGSAMISLSQNHDPNRNETFSSVISLLNEDPLSQDLVPKMA SIFKNFVITYNRTYESKEEARWRLSVFVNNMVRAQKIQALDRGTAQ YGVTKFSDLTEEEFRTIYLNLTLLRKEPGNKMKAQSVGDLAPPEWD WRSGAVTKVKDQGMCGSCWAFSVTGNVEGQWFLNQGTLLSLSE QELLDCKMDKACMGGPLSNAYSIAKNLGGLETEDDYSYQGHMQ SCNFSAEKAKVYINDSVELSQNEQKLAAWLAKRGPISVAINAFGMQ FYRHGISRPLRPLCSPWLIDHAVLLVGYGNRSDVPFWAIKNSWGT D WGEKGYYYLHRGSGACGVNTMASSAVVD	CTSF
265	MWGLKVLVLLPVVSFALYPEEILDTHWELWKKTHRKQYNNKVDEIS RRLIWEKNLKYISIHNEASLGVHTYELAMNHLGDMTSEEVVQKMT GLKVPLSHSRSDNTLYIPEWEGRAPDSVDYRKKGYVTPVKNQGC GSCWAFSSVGALEGLKKTGKLLNLSPQNLVDCVSENDGCGGGY MTNAFYVQKNRGIDSEDAYPYVGQEESEMYNPTGKAAKCRGYR EIEPEGNEKALKRAVARVGPVSVDAIDSLTSFQFYSGVYYDESCNSD NLNHAVALVGYGIQKGNKHWHIKNWGENWGNKGILMARNKNN ACGIANLASFPKM	CTSK
266	MADQRQRSLSTSGESLYHVLGLDKNATSDDIKKSYRKLALKYHPD KNPDNPEAADKFKEINNAHAILTDATKRNIYDKYGSGLYVAEQFG EENVNTYFVLSSWWAKALFVFCGLLTCCYCCCCCLCCCFNCCCGKC KPKAPEGEETEFYVSPEDLEAQLQSDEREATDTPIVIQPASATETTQL TADSHPSYHTDGFN	DNAJC5
267	MRAPGMRSRPAGPALLLLLLFLGAAESVRRAPPRRYTPDWPSLDS RPLPAWFDEAKFGVFIHWGVFSVPAWGSEFWFHHWQGEGRPQYQ RFMRDNYPPGFSYADFGPQFTARFFHPEEWADLFQAAGAKYVVL TKHHEGFTNWSPVSWNWSKDVGPHRDLVGEGLTALRKRNIYR LYHSLLEWFHPLYLLDKKNGFKTQHFVSAKTMPELYDLVNSYKPD LIWSDGEWECPDITYWNSTNFLSWLYNDSPVKDEVVVNDRWGQNC SCHHGGYYNCEDKFKPQSLPDHKWEMCTSIDKFSWGYRRDMALSD VTEESEIISELVQTVSLGGNYLLNIGPTKDGLIVPIFQERLLAVGK WLSINGEAIYASKPWRVQWEKNTTSVWYTSKGSAYYAIFLHWPEN GVLNLESPITTSTTKITMLGIQGDLLKWSTDPDKGLFISLPQLPPSAVP AEFAWTIKLTGVK	FUCA1
268	MGVRHPPCSHRLAVCALVSLATAALLGHILLHDFLLVPRELSGSSP VLEETHPAHQQGASRPGPRDAQAHPGRPRAVPTQCDVPPNSRFDCA PDKAITQEQCEARGCCYIPAKQGLQGAQMGPWCFFPPSYPSYKLE NLSSEMGYTATLTRTPTFFPKDILTLRLDVMETENRLHFTIKDP ANRRYEVPLETPHVHSRAPSPLYSVFSEEPFGVIVRRQLDGRVLLN TTVAPLFFADQFLQLSTSLPSQYITGLAEHLSPLMLSTSWTRITLWNR DLAPTPGANLYGSHPFYLALEDGGSAGHVLLNSNAMDVVLQPSPA LSWRSTGGILDVYIFLGPPEKSVVQQYLDVVGYPFMPYPYWGLGFHL CRWGYSSAITRQVVENMTRAHFPLDVQWNLDDYMDSRRDFTFN KDGRDFPAMVQELHQGGRRYMMIVDPAISSSGPAGSYRPPYDEGLR RGVFITNETGQPLIGKVWPGSTAFPDTNPTALAWWEDMVAEFHD QVPFDGMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVVGGLQA ATICASSHQFLSTHYNLHNLGYLGLTEAIAASHRALVKARGTRPFVISRST	GAA

	FAGHGRYAGHWTGDVWSSWEQLASSVPEILQFNLLGVPLVGADVC GFLGNTSEELCVRWTQLGAFYPMRNHNSLLSLPQEPYSFSEPAQQ AMRKALTRYALLPHLYTLFHQAHVAGETVARPLFLEFPKDSSTWT VDHQLLWGEALLITPVLQAGKAEVTGYFPLGTWYDLQTVPVEALG SLPPPPAAPREPAIHSEGQWVTLPAPLDTINVHLRAGYIIPLOQGPGLT TTESRQQPMALAVALTGGGEARGELFWDDGESLEVLERGAYTQVIF LARNNTIVNELVRVTSEGAGLQLQKVTVLGVATAPQQVLSNGVPVS NFTYSPDTKVLDCVSLLMGEQFLVSWC	
269	MAEWLLSASWQRRAKAMTAAAGSAGRAAVPLLLCALLAPGGAYV LDDSDGLGREFDGIGAVSGGGATSRLLVNYPEPYRSQILDYLFKPNF GASLHILKVEIGGDGQTTDGTESPMHYALDENYFRGYEWWMKE AKKRNPNTILIGLPWSFPGWLKGFDWPYVNLQLTAYYVVTWIVG AKRYHDLIDIDYIGIWNERSYNANYIKILRKMLNYQGLQVRKIIASDN LWESISASMLLDAELFKVVDVIGAHYPGTHSAKDAKLTGKKLWSSE DFSTLNSDMGAGCWGRILNQNYINGYMTSTIAWNLVASYYEQLPY GRCGLMTAQEPWSGHYVVEVPVWVSAHTTQFTQPGWYYLKTVGH LEKGSYVALTDGLGNLTIIETMSHKHSCIRPFLPYFNVSQQFATF VLKGSFSEIPELQVWYTKLGKTSERFLFKQLDSLWLLSDSGSFTLSL HEDELTTLTTLTGRKGSYPLPPKSQFPSTYKDDFNVDYPPFSEAPN FADQTGVFEYFTNIEDPGEHHFTLRQVLNQRPIITWAADASNTISIIGD YNWTNLTIKCDVYIETPDTGGVFIAGRVNKGILIRSARGIFFWIFAN GSYRVTDLAGWIIYALGRVEVTAKKWTTLTLTIKGHFTSGMLND KSLWTDIPVNFKNGWAAIGTHSFEFAQFDNFLVEATR	GALC
270	MAAVVAATRWWQLLLVLSAAGMGASGAPQPPNILLMLDDMGW GDLGVYGEPSRETPNLDMAAEGLLPNFYSANPLCSPSRAALLTG RLPIRNGFYTTNAHARNAYTPQEIYVGGIPDSEQLLPELLKKAGYVSKI VGKWHLGHRPQFHPLKHGFDEWFGSPNCHFGPYDNKARNIPVYR DWEMVGRYYEEFPINLKTGEANLTQIYLQEALDFIKRQARHHPFL YWAVDATHAPVYASKPFLGTSQRGRYGDVREIDDSIGKILELLQD LHVADNTFVFFTSDNAAALISAPEQGGSNPFLCGKQTTFEGGMRE PALAWWPGHVTAGQVSHQLGSIMDLFTTSLALAGLTPPSDRAIDGL NLLPTLLQGRLMDRPIFYRGTDLMAATLGQHKAHFWTWTNSWE NFRQGIDFCPGQNVSGVTTHNLEDHTKLPLIFHLGRDPGERFPLSFAS AEYQEALSRTSVVQQHQEALVPAQPQLNVCNWAVMNWAPPGCE KLGKCLTPPESIPKKCLWSH	GALNS
271	MQLRNPHELHGCALALRFLALVSWDIPGARALDNLARTPTMGWL HWERFMCNLDCQEEPDCISEKLFMEMAELMVSEGWKDAGYEYL CIDDCWMAQPDSEGRQLQADPQRFPHGIRQLANYVHSGGLKLGIIYA DVGNKTCAGFPGSFGYYDIDAQTFADWGVDDLKFDGCYCDLENL ADGYKHMSLALNRTGRSIVYSCEWPLYMWPFQKPNYTEIRQYCNH WRNFADIDDSWKSISILDWTSFNQERIVDVAGPGGWNDPDMVLIG NFGLSWNQVQTQMALWAIMAAPLFMSNDLRHISPOAKALLQDKD VIAINQDPLGKQGYQLRQGDNFEVWERPLSGLAWAVAMINRQEIG GPRSYTIAVASLGKGVACNPACFITQLLPVKRKLGFYEWTSRLRSHI NPTGTVLLQLENTMQMSLKDLL	GLA
272	MPGFLVRILPLLLVLLLLGPTRGLRNATQRMFEIDYSRDSFLKDGQP FRYISGSIHYSRVPRFYWKDRLLKMKMAGLNIAIQTYPWNFHEPW PGQYQFSEDHDEYFLRLAHELGLLVILRPGPYICAEWEMGGLPAW LLEKESILLRSSDPDYLAADVWKVLGVLPLKMKPLLYQNGGPVITVQ VENEYGSYFACDFDYLRFLQKRFRHHLGDDVVLFTTDGAHKTFLK CGALQGLYTTVDFTGSNITDAFLSQRKCEPKGPLINSEFYTGWLHDH WGQPHSTIKTEAVASSLYDILARG	GLB1

	ASVNLVYMFIGGTNFAYWNGANSFYAAQPTSYDYDAPLSEAGDLTE KYFALRNIIQKFEKVPEGPIPPSTPKFAYGKVTLKLTVGAAALDILC PSGPIKSLYPLTFIQVKQHYGFVLYRTTLPQDCSNAPLSSPLNGVHD RAYVAVDGPQGVLERNNVITLNTGKAGATLDLLVENMGRVNYG AYINDFKGLVSNLTLSSNLTDTWTFPLDTEDAVRSHLGGWGHRS HHDEAWAHNSSNYTLPAFYMGNFSPSGIPDLQDTFIQFPGWTKGQ VWINGFNLGRYPARGPQLTLFVPQHILMTSAPNTITVLELEWAPC SSDDPELCAVTFVDRPVIGSSVTYDHPKPKVEKRLMPPPPQKNKDS WLDHV	
273	MQSLMQAPLLIALGLLLAAPAQAHLKKPSQLSSFSWDNCDEGKDP VIRSLTLEPDPIVPGNVTLVSMGSTSVPLSSPLKVDLVLEKEVAGLW IKIPCTDYIGSCTFEHFCDVLDMLIPTGEPCEPLRTYGLPCHCPFKEG TYSLPKSEFVVPDLELPSWLTGTNYRIESVLSSSGKRLGCIKIAASLK GI	GM2A
274	MLFKLLQRQTYTCLSHRYGLYVCFLGVVVTIVSAFQFGEVVLEWSR DQYHVLFDSDYRDNIAKGSFQNRCLPMPIDVVYTWVNGTDLELLKE LQQVREQMEEEEKAMREILGKNTTEPTKKSEKQLECLLTHCIKVP LVLDPALPANITLKDLPSTLYPSFHSASDIFNVAKPKNPSTNVSVVVD STKDVEDAHSGLLKGNRQTVWRGYLTDDKEVPGLVLMQDLAFLS GFPPTFKETNQLKTKLPENLSSKVLLQLYSEASVALLKLNNPKDFQ ELNKQTKKNMTIDGKELTISPA YLLWDLAISQSKQDEDISASRFEDNEELRYSLSIERHAPWVRNIFI VTNGQIPSWLNLNPNRVTIVTHQDVFRNLSHLPTFSSPAIESHIHRIEG LSQKFIYLNDDVMFGKDVWPDDFYSHSKGQKVYLTWPVPNCAEGC PGSWIKDGYCDKACNNSACDWDGDCSGNSGGSRYIAGGGGTGSI GVGQPWFQGGGINSVSYCNQGCANSWLADKFCDQACNVLSCGFD AGDCGQDHFHELYKVILLPNQTHYIIPKGECLPYFSFAEVAKRGVEG AYSNDPIIRHASIANKWKTIHLIMHSGMNATTIHFNLTFQNTNDEEF KMQITVEVDTRREGPKLNSTAQKGYENLVSPITLLPEAEILFEDIPKEK RFPKFKRHDVNSTRRAQEEVKIPLVNISLLPKDAQLSLNTLDLQLEH GDITLKGYNLSKSALLRSFLMNSQHAHAKIKNAIITDETNDLSVAPQE KQVHKSLPNSLGVSERLQRLTFPAVSVKVNNGHDQGNPPDLLETT ARFRVETHQKTIGGNVTKEKPPSLIVPLESQMTKEKKITGKEKENS RMEENAENHIGVTEVLLGRKLQHYTDSYLGFLPWEKKKYFQDLLD EEESLKTQLAYFTDSKNTGRQLKDTFADSLRYVNKILNSKFGFTSRK VPAHMPHMIDRIVMQELQDMFPEEFDKTSFHKVRHSEDMQFAFSYF YYLMSAVQPLNISQVFDEVDTDQSGVLSREIRTLATRIHELPLSLQ DLTGLEHMLINCSKMLPADITQLNNIPTQESYYDPNLPVTKSLVT NCKPVTDKIHKAYKDKNKYRFEIMGEEIIAFKMIRTNVSHVVGQLD DIRKNPRKFVCLNDNIDHNHKAQTVKAVLRDFYESMFPIPSQFELP REYRNRFLHMHQEWRAVRDKLKFVTHCVLATLIMFTIFSFFAEQ LIALKRKIFPRRRIHKEASPNRIRV	GNPTAB
275	MAAGLARLLLLLGLSAGGPAPAGAAKMKVVEEPNAGVNNPFLPQ ASRLQAKRDPSPVSGPVHLFRLSGKCFSLVESTYKYEFPCFHNVTQH EQTFRWNAYSGLGIWHEWEIANNTFTGMWMRDGDACRSRQRQSK VELACGKSNRLAHVSEPSTCVYALTFETPLVCHPHALLVYPTLPEAL QRQWDQVEQDLADELITPQGHEKLLRTLFEDAGYLKTPEENEPTQL EGGPDSLGFTLENCRAHKELSKEIKRLKGLLTQHGIPTTRPTETS NLEHLGHETPRAKSPEQLRGDPG LRGSL	GNPTG
276	MRLPLAPGRLRRGSPRHLPSCPALLLVGGCLGVFGVAAGTRR PNVVLTLTDQDEVLGGMTPKKTALIGEMGMTFSSAYVPSALC	GNS

	<p>CPSRASILTGKYPHNHHVNNLTLEGNCSSKSWQKIQEPNTFPAILRS MCGYQTFAGKYLNEYGAPDAGGLEHVPLGWSYWYALEKNSKYY NYTLSINGKARKHGENYSVDYLTDLANVSLDFLDYKSNFEPFFM MIATPAPHSPWTAAPQYQKAFQNVFAPRNKNFNIHGTNKHWRIRQ AKTPMTNSSIQFLDNAFRKRWQTLLSVD DLVEKLVKRLEFTGELNNTYIFYTSDNGYHTGQFSLPIDKRQLYEFD IKVPLLVRGPGIKPNQTSKMLVANIDLGPITLDIAGYDLNKTQMDG MSLLPILRGASNLTWRSVDLVEYQGEGRNVDPTCPSLSPGVSQCFP DCVCEDAYNNTYACVRTMSALWNLQYCEFDQEVFVEVYNLTAD PDQITNIAKTIDPELLGKMNYRLMMLQSCSGPTCRTPGVFDPGYRFD PRLMFSNRGSVRTRRRFSKHL</p>	
277	<p>MWTLVSWVALTAGLVAGTRCPDQGFCPVACCLDPGGASYSCCRPL LDKWPTTLRHLGGPCQVDAHCSAGHSCIFTVSGTSSCCPFPEAVAC GDGHHCCPRGFHCSADGRSCFQSRGNSVGAIQCPDSQFECPDFST CCVMVDGSGGCCPMPQASCCEDRVHCCPHGAFCDLVHTRCITPTG THPLAKKLPAQRTNRAVALSSVMCPDARSRCPDGSTCCELPSTGY GCCPMPNATCCSDHLHCCPQDTVCDLIQSKLSKENATTDLTKLP AHTVGDVKCDMEVSCPDGYTCCRLQSGAWGCCPFTQAVCCEDHIH CCPAGFTCDTQKGTCEQGPQVPWMEKAPAHLSLPDPQALKRDVP CDNVSSCPSSDTCCQLTSGEWGCCPIPEAVCCSDHQHCCPQGYTCV AEGQCQRGSEIVAGLEKMPARRASLSHPRDIGCDQHTSCPVGQTCC PSLGGSWACCQLPHAVCCEDRQHCCPAGYTCNVKARSCEKEVVSA QPATFLARSPHVGKDVCEGEGHFCHDNQTCCRDNRQGWACCPY RQGVCCADRRHCCPAGFRCAARGTKCLRREAPRWDAPLRDPALRQ LL</p>	GRN
278	<p>MARGSAVAWAALGPLLWGCALGLQGGMLYPQESPSRECKELDGL WSFRADFSDNRRRGFEEQWYRRPLWESGPTVDMVPSSFNDSQD WRLRHFGVWVWYEREVILPERWTQDLRTRVLRIGSAHSYAIWVW NGVDTLEHEGGYLPFEADISNLVQVGPLPSRLRITIAINNTLTPTLPP GTIQYLTDTSKYPKGYFVQNTYFDFNYAGLQRSVLLYTTPTTYIDD ITVTTSEQDSGLVNYQISVKGSNLFKLEVRLDLENKVVANGTGT QQQLKVPVGSVSLWWPYLMHERPAYL YSLEVQLTAQTSGLPVSDFYTLPGVIRTVAVTKSQFLINGKPFYFHG VKNHEDADIRGKGFDWPLLVKDFNLLRWLGANAFRTSHYPYAEV MQMCDRYGIVVIDECPGVLALPQFFNNVSLHHHMQVMEEVVR DKNHPAVVMWSVANEPASHLESAGYYLKMVIAHTKSLDPSRPVTF VSNSNYAADKGAPYVDVICLNSYYSWYHDYGHLELIQLQLATQFE NWYKKYQKPIIQSEYGAETIAGFHQDPPLMFTEYQKSLEQYHLG LDQKRRKYVVGELIWNFADFMTESPTRVLGKKGIFTRQRQPKSA AFLLRERYWKIANETRYPHSVAKSQCLENSLFT</p>	GUSB
279	<p>MTSSRLWFSLLLAAAFAGRATALWPWPQNFQTSQRYVLYPNNFQ FQYDVSSAAQPGCSVLDEAFQRYRDLLFGSGSWPRPYLTGKRHTLE KNVLVSVVTPGCNQLPTLESVENYTLTINDDQCLLLSETVWGALR GLETFSQLVWKSAGTFFINKTEIEDFPRFPHRGLLDTSRHYLPLSSI LDTLDVMAYNKLNVFHWHLVDDPSFPYESFTFPELMRKGSYNPVT HIYTAQDVKEVIEYARLRGIRVLAEFDTPGHTLSWGPPIPGLLTPCY SGSEPSGTFGPVNPSLNNTYEFMSTFFLEVSSVFPDFYLHLGGDEV FTCWKSNPEIQDFMRKKGFGEFQKLESFYIQTLLDIVSSYGKGYVV WQEVFDNKVKIQPDTHIQVWREDIPVNYMKELELVTKAGFRALLSA PWYLNRIYSGPDWKDFYIVEPLAFEGTPEQKALVIGGEACMWGEY VDNTNLVPRLWPRAGAVAERLWSNKLTSDLTFAYERLSHFRCCELLR RGVQAQPLNVGFCEQEFEQT</p>	HEXA

280	MELCGLGLPRPPMLLALLLATLLAAMLALLTQVALVVQVAEAARA PSVSAKPGPALWPLPLSVKMTPNLLHLAPENFYISHSPNSTAGPSCTL LEEAFRRYHGYIFGFYKWHHEPAEFQAKTQVQQLLVSITLQSECD FPNISSDESYTLLVKEPVAVLKANRVWGALRGLETFSQLVYQDSYG TFTINESTIIDSPRFSHRGILIDTSRHYLPVKIILKTLDAMAFNKFNVLH WHIVDDQSFYQSFITPELSNKGSYSLSHVYTPNDVRMVIEYARLRG IRVLPEFDTPGHTLSWGKGQKDLLTPCYSRQNKLDSFGPINPTLNTT YSFLTTFKEISEVFPDQFIHLGGDEVEFKCWESNPKIQDFMRQKGF GTDFKKLESFYIQKVLDIATINKGSIVWQEVFDDKAKLAPGTIVEV WKDSAYPEELSRVTASGFPVILSAPWYLDLISYGQDWRKYYKVEPL DFGGTQKQKQLFIGGEACLWGEYVDATNLTPRLWPRASAVGERLW SSKDVRDMDDAYDRLTRHRCRMVERG IAAQPLYAGYCNHENM	HEXB
281	MTGARASAAEQRRAGRSGQARAAERAAGMSGAGRALAALLLAAS VLSAALLAPGGSSGRDAQAAPPRDLDDKKRAELKMDQALLLIHNE LLWTNLTVYWKSECCYHCLFQVLVNVQPSPKAGKPSAAAASVSTQ HGSILQLNDTLEEKEVCRLEYRFGFEGNYSLLVKNIHNGVSEIACDL AVNEDPVDSNLPVSI AFLIGLAVIIVISFLRLLLSLDDFNNWISKAISR ETDRLINSELGSPSRDPLDGDVQPATWRLSALPPRLRSVDTFRGIAL ILMVFNNGGGKYWYFKHASWNGLTVALDVFPWFVFIMGSSIFLS MTSILQRGCSKFRLLGKIAWRSFLLICIGIIVNPNYCLGPLSWDKVRI PGVLQRLGVTYFVVAVLELLFAKPVPEHCASERSCLSLRDITSSWPQ WLLILVLEGLWLGLTFLLPVPGCPTGYLGPGGIGDFGKYPNCTGGA AGYIDRLLLGDDHLYQHPSSAVLYHTEVAYDPEGILGTINSIVMAFL GVQAGKILLYKARTKDILIRFTAWCC ILGLISVALTKVSENEGFI PVNKNLWLSYVTTLSFAFFILLVLYPVV DVKGLWTGTFFYPGMNSILVYVGHEVFENYFPFQWKLKDNQSHK EHLTQNIVATALWVLIAYILYRKKIFWKI	HGSNAT
282	MAAHLPLICALFTLLDMAQGFGRGPLLPNRPFTTVWNANTQWCLE RHGVDVDVSVFDDV ANPGQTFRGPDMTIFYSSQLGTYPYYTPTGEP VFGGLPQNASLIAHLARTFQDILAAIPAPDFSGLAVIDWEAWRPRW AFNWDTKDIYRQRSRALVQAQHPDWPAPQVEAVAQDQFQGAARA WMAGTLQLGRALRPRGLWGFYGFPCDYNDFLSPNYTGQCPGIR AQNDQLGWLWGQSRALYPSIYMPAVLEGTGKSQMYVQHRVAEAF RVAVAAGDPNLPVLPYVQIFYDTTNHFLPLDELEHSLGESAAQGAA GVVLWVSWENTRTKESCQAIKEYMDTTLGPFILNVTSGALLCSQ ALCSGHGRCVVRTSHPKALLLNPAFSIQLTPGGGPLSLRGALSLE DQAQMAVEFKCRCYPGWQAPWCERKSMW	HYAL1
283	MPPPRRTGRGLLWLGLVLSSVCVALGSETQANSTTDALNVLLIIVDDL RPSLGCYGDKLVRSPNIDQLASHSLLFQNAFAQQAVCAPSRVSFLTG RRPDTRRLYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHP GISSNHTDDSPYSWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPV DVLDPVPEGLTLPDKQSTEQAIQLLEKMKTSASPFFLAVGYHKPHIPFR YPKEFQKLPLENITLAPDPEVPDGLPPVAYNPWMDIRQREDVQAL NISVPYGPVDFQIRKIRQSYFASVSYLDTQVGRLLSALDDLQLANS TIIAFTSDHGWALGEHGEWAKYSNFDVATHVPLIFYVPGRTASLPEA GEKLFYPYLDPFDSASQLMEPGRQSMDELVELVSLFPTLAGLAGLQVP PRCPVPSFHVLCREGKNLLKHFRFRDLEEDPYLPGNPRELIAYSQY PRPSDIPQWNSDKPSLKD IKIMGYSIRTIDYRYTVWVGFPNDEFAN FSDIHAGELYFVDSPLQDHNMYNDSQGGDLFQLLMP	IDS
284	MRPLRPRAALLALLASLLAAPPVAPAEAPHLVHVDAARALWPLRRF WRSTGFCPLPHSQADQYVLSWDQQLNLAYVGAVPHRGIKQVRTH	IDUA

	WLELVTTTRGSTGRGLSYNFTHLDDGYLDLLRENQLLPGFELMGSAS GHFTDFEDKQQVFEWKDLVSSLARRYIGRYGLAHVSKWNFETWNE PDHHDFFDNVSMTMQGFNLNYDACSEGLRAASPALRLGGPGDSFHT PPRSPLSWGLLRHCHDGTNFFTGEAGVRLDYISLHRKGARSSISILEQ EKVVAQQIRQLFPKFADTPIYNDEADPLVGWSLPQPWRADVITYAA MVVKVIAQHQNLLANTTSAPFYALLSNDNAFLSYHHPFAQRTLT ARFQVNNTRPPHVQLLRKPVLTAMGLLALLDEEQLWAEVSQAGTV LDSNHTVGVLASAHRPQGPADAWRAAVLIYASDDTRAHPNRSVAV TLRLRGVPPGGLVYVTRYLDNGLCSPDGEWRRLGRPVPFTAQEQFR RMRAAEDPVAAAPRPLPAGGRLTLRPALRLPSLLLHVHVCARPEKPP GQVTRLRALPLTQGGQLVLVWSDEHVGSKCLWTYEIQFSQDGKAYT PVSRRKPSTFNLVFPSPDTGAVSGSYRVRALDYWARPGPFSDPVPYLE VPVPRGPPSPGNP	
285	MVVVTGREPDSRRQDGAMSSSDAEDDFLEPATPTATQAGHALPLLP QEFPEVVPLNIGGAHFTTRLSTLRCYEDTMLAAMFSGRHYIPTDSEG RYFIDRDGTHFGDVLNFLRSGDLPPRERVRAVYKEAQYIAIGPLLE QLENMQPLKGEKVRQAFLGLMPYYKDHLEIRIVEIARLRAVQRKAR FAKLKVCVFKEEMPITPYECPLLSLRFERSESDGQLFEHHCEVDVS FGPWEAVADVYDLLHCLVTDLSAQGLTVDHQCIGVCDKHLVNHY YCKRPIYEFKITWW	KCTD7
286	MVCFRLFPVPGSLVLVCLVLGAVRSYALELNLTDSENATCLYAK WQMNFTVRYETTNTKYKTVTISDHGTVTYNGSICGDDQNGPKIAV QFGPGFSWIANFTKAASTYSIDSVSFSYNTGDNTTFPDAEDKGILTV DELLAIRIPLNDFRCNSLSTLEKNDVVQHYWDVLVQAFVQNGTVS TNEFLCDKDKTSTVAPTHTTVPSPTTPTPKEKPEAGTYSVNNGND TCLLATMGLQLNITQDKVASVININPNTTHSTGSCRSHTALLRLNSS TIKYLDVFAVKNENRFYLKEVNISMVYLVNGSVFSIANNLSYWDA PLGSSYMCNKEQTVSVSGAFQINTFDLRVQPFNVQTQGYSTAQDCS ADDDNFLVPIAVGAALAGVLILVLLAYFIGLKHHHAGYEQF	LAMP2
287	MGAYARASGVCARGCLDSAGPWTMSRALRPLPLPCLFLLLLAAA GARAGGYETCPTVQPNMLNVHLLPHTHDDVGWLKTVQYFYGIK NDIQHAGVQYILDSVISALLADPTRRFIYVEIAFFSRWWHQQTNAQ EVVRDLVRQGRLEFANGGWVMNDEAATHYGAIVDQMTLGLRFL DTFGNDGRPRVAWHIDPFGHSREQASLFAQMFGDGGFFGRLDYQD KWVRMQKLEMEQVWRASSTLKPPTADLFTGVLPGYNPPRNLWC DVLCDVQPLVEDPRSPEYNAKELVDYFLNVATAQGRYYRTNHTVM TMGSDFFQYENANMWFKNLDKLIRLVNAQQAKGSSVHVLYSTPAC YLWELNKANLTWSVKHDDFFPYADGPHQFWTGYFSSRPALKRYER LSYNFLQVCNQLEALVGLAANVGPGYSGDSAPLNEAMAVLQHHD AVSGTSRQHVANDYARQLAAGWGPCEVLLSNALARLRGFKDHFTF CQQLNISICPLSQTAARFQVIVYNPLGRKVNWMVRLPVSEGVFVK DPNGRTVPSDVVIFPSSDSQAHPPELLFSASLPALGFSTYSVAQVPR WKPQARAPQPIPRRSWSPALTIENEHIRATFDPDTGLLMEIMNMNQ QLLLPVRQTFFWYNASIGDNESDQASGAYIFRPNQKPLPVSRAWQI HLVKTPLVQEVHQNFSAWCSQVVRLYPGQRHLELEWSVGPIPVGD TWGKEVISRFDTPLETKGRFYTDSNGREILERRRDYRPTWKLNQTEP VAGNYYPVNTRIYITDGNMQLTVLTDTSQGGSSLRDGSLELMVHRR LLKDDGRGVSEPLMENGSGAWVRGRHLVLLDTAQAAAAGHRLLA EQEVLAPQVVLAPGGGAAYNLGAPPRTQFSGLRDLPPSVHLLTLA SWGPEMVLLRLEHQFAVGEDSGRNLAPVTLNLRDLFSTFTITRLQE TTLVANQLREAASRLKWTNTGTGTPHQTPYQLDPANITLEPMEIRTF LASVQWKEVDG	MAN2B1

288	MRLHLLLLLALCGAGTTAAELSYSLRGNWSICNGNGSLELPGAVPG CVHSALFQQGLIQDSYYRFNDLNYRWVSLDNWWTYSKEFKIPFEISK WQKVNILEGVDTVSKILFNEVTIGETDNMFNRYSFIDITNVVRDVNS IELRFQSAVLYAAQQSKAHTRYQVPPDCPPLVQKGECHVNFVRKEQ CSFSWDWGPSFPTQGIWKDVRIEAYNICHNLNYFTFSPIYDKSAQEWN LEIESTFDVSSKPVGGQVIVAIPKLQTQQTYSIELQPGKRIVELFVNI SKNITVETWWPHGHGNQTGYNMTVLFELDGGLNIEKSAKVYFRTV ELIEEPIKGSPLSFYFKINGFPIFLKGSNWIPADSFQDRVTSELLRLL QSVVDANMNTLRVWGGGIYEQDEFYELCDELGIMVWQDFMFACA LYPTDQGF LDSVTAEVAYQIKRLKSHPSIIIWSGNNENEEALMMNW YHISFTDRPIYIKDYVTLYVKNIRELVLADGKSRPFITSSPTNGAETV AEAWVSQNPNSNYFGDVHFYDYISDC WNWKVFPKARFASEYGYQSWPSFSTLEKVSSTEDWSFNSKFSLHRQ HHEGGNKQMLYQAGLHFKLPQSTDPLRTFKDTIYLTQVMQAQCVK TETEFYRRSRSEIVDQQGHTMGALYWQLNDIWQAPSWASLEYGGK WKMLHYFAQNFFAPLLPVGFENENTFYIYGVSDLHSDYSMTLSVRV HTWSSLEPVC SRVTERFVMKGGEAVCLYEPPVSELLRRCGNCTRES CVVSFYLSADHELLSPTNYHFLSSPKEAVGLCKAQITAIISQQGDIFV FDLETSAPFVWLDVGSIPGRFSDNGFLMTEKTRTILFYPWEPTSK NELEQSFHVTSLTDIY	MANBA
289	MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKY FFMSPCDKFRAKGRKPKCLMLQVVKILVVTVQLILFGLSNQLAVTF REENTIAFRHLFLLGYSDGADDTFAAYTREQLYQAIFHAVDQYLAL PDVSLGRYAYVRGGGDPWTNGSGLALCQRYYHRGHVDPANDTFDI DPMVVTDCIQVDPPERPPPPSDDLTLLESSSSYKNLTLKFHKLNVN TIHFRKLTINLQSLINNEIPDCYTFSVLITFDNKAHSGRIPISLETQAH QECHKPSVFQHGDNFRLLFDVVVILTCSLSFLLCARSLRGRFLLQN EFVGFMMWRQRGRVISLWERLEFVNGWYILLVTSVDLTISGTIMKIGI EAKNLASYDVCSILLGTSTLLVWVGVIYLYTFFHNYNIIATLRVALP SVMRFCCCVAVIYLYGFCGWIWLGYPYHVKFRSLSMVSECLFSLNG DDMFVTFAAMQAQQGRSSLVWLFSQLYLYSFISLFIYMVLSLFIALI TGAYDTIKHPGGAGAESESELQAYIAQCQDSPTSGKFRRGSGSACSLL CCCCGRDPSEEHSLLVN	MCOLN1
290	MAGLRNESEQEPLLGDTGPSREWDILETEEHYKSRWRSIRILYLTMF LSSVGFSVVMMSIWPLYLQKIDPTADTSFLGWVIASYSLGQMVASPIF GLWSNYRPRKEPLIVSILISVAANCLYAYLHIPASHNKYYMLVARGL LGIGAGNVAVRSYTAGATSLQERTSSMANISMCMQALGFILGPVFQ TCFTFLGEKGVTDVVIKLQINMYTTPVLLSAFLGILNIIILAILREHR VDDS GRQCKSINFEEASTDEAQVPQGNIDQVAVVAINVLFVTLFIFALFET IITPLTMDMYAWTQEQA VLYNGHILAALGVEAVVIFLGVKLLSKKIG ERAILLGLLIVVWVGFFILLPWGNQFPKIQWEDLHNSIPNTTFGEIII GLWKSPMEDDNERPTGCSIEQAWCLYTPVIHLAQFLTS AVLIGLGY VCNLMSYTLYSKILGPKPQGVYMGWLTASGSGARILGPMFISQVYA HWGPRWAFSLVCGIIVLTITLLGVVYKRLIALSVRYGRIQE	MFSD8
291	MLLKTVLLGHVAQVLMLDNGLLQTPPMGWLAWERFRCNINCDE DPKNCISEQLFMEMADRMAQDGWRDMGYTYLNIDDCWIGGRDAS GRLMPDPKRFPHGIPFLADYVHSLGLKLGIIYADMGNFTCMGYPGTT LDKVVQDAQTFAEWKVDMLKLDGCFSTPEERAQGYPKMAAALNA TGRPIAFSCSWPAYEGGLPPRVNYSLLADICNLWRNYDDIQDSWWS VLSILNWFVEHQDILQPVAGPGHWNDPDMLLIGNFGLSLEQSRAQM ALWTVLAAPLLMSTDLRTISAQNMDILQNPLMIKINQDPLGIQGRRI HKEKSLIEVYMRPLSNKASALVFFSCRTDMPYRYHSSLGQLNFTGS	NAGA

	VIYEAQDVYSGDIISGLRDETNTFTVIINPSGVVMWYLYPIKNLEMSQ Q	
292	MEAVAVAAAAGVLLLAGAGGAAGDEAREAAAVRALVARLLGPGP AADFSVSVERALAAKPGLDYSLSGGGAARVRVRGSTGVAAAAGL HRYLRDFCGCHVAWSGSQLRLPRPLPAVPGELTEATPNRYRYQ VCTQSYSFVWWDWARWEREIDWMALNGINLALAWSGQEAIWQR VYLALGLTQAEINEFFTGAFLAWGRMGNLHTWDGPLPPSWHIKQL YLQHRVLDQMRSFGMTPVLPFAFAGHVPEAVTRVFPQVNVTKMGS WGHFNCSYSCSFLAPEDPIFPIIGSLFLRELIKEFGTDHIYGADTFNE MQPPSSEPSYLAATTAVYEAMTAVDTEAVWLLQGWLQHQHQPQF WGPAQIRAVLGAVPRGRLLVLDLFAESQPVYTRTASFQGGQPFWCM LHNFGGNHGLFGALEAVNGGPEAARLFPNSTMVGTGMAPEGISQN EVVYSLMAELGWRKDPVPDLAAWVTSFAARRYGVSHPDAGAAWR LLRSVYNCSGEACRGNRSPLVRRPSLQMNTSIWYNRSDVFEAWR LLLTSAPLATSPAFRYDLDLTRQAVQELVSLYEEARSAYLSKEL ASLLRAGGVLAYELLPALDEVLASDSRFLGWSWLEQARAAAVSEAE ADFYEQNSRYQLTLWGPEGNILDYANKQLAGLVANYTTPRWRLFL EALVDSVAQGIPIQQHQFDKNVFQLEQAFVLQSKQRYPSQPRGDTVD LAKKIFLKYYPRWVAGSW	NAGLU
293	MTGERPSTALPDRRWGPRILGFWGGCRVWVFAAIFLLSLAASWSK AENDFGLVQPLVTMEQLLWVSGRQIGSVDTFRIPITATPRGTLLAF AEARKMSSSDEGAKFIALRRSMDQGSTWSPTAFIVNDGDVDPGLNL GAVVSDVETGVVFLFYSLCAHKAGCQVASTMLVWSKDDGVSWST PRNLSLDIGTEVFAPGPGSGIQKQREPRKGRLIVCGHGTLERDGVFC LLSDDHGASWRYGSGVSGIPYQPKQENDFNDECQPYELPDGSVV INARNQNNYHCHCRIVLRSYDACDTLRPRDVTDFPELVDPVVAAGA VVTSSGIVFFSNPAHPEFRVNLTLRWSFSNGTSWRKET VQLWPGPSGYSSLATLEGSMGDGEEQAPQLYVLYEKGRNHYTESISV AKISVYGT	NEU1
294	MTARGLALGLLLLLLCPAQVFSQSCVWYGECEGIAYGDKRYNCEYS GPPKPLPKDGYDLVQELCPGFFFFGNVSLCCDVRQLQTLKDNLQPL QFLSRCPSCFYNNLNLFCELTCSPRQSQFLNVTATEDYVDPVTNQTK TNVKELQYYVQGQSFANAMYNACRDVEAPSSNDKALGLLCGKDAD ACNATNWIEYMFNKDNGQAPFTITPVFSDFPVHGMEPMNNATKGC DESVDEVTAPCSQCDCSIVCGPKQPPPPAPWTILGLDAMYVIMWI TYMAFLLVFFGAFFAVWCYRKRYFVSEYTPIDSNIASFVNASDKGE ASCCDPVSAAFEGCLRRLFTRWGSFCVRNPGCVIFFSLVFITACSSGL VFVRVTNPVDLWSAPSSQARLEKEYFDQHFQPFRTTEQLIIRAPLT DKHIYQPYPSGADVFPFGPPLDIQILHQVLDLQIAIENITASYNETVT LQDICLAPLSPYNTNCTILSVLNYFQNSHVSVDHKKGDDFFVYADY HTHFLYCVRAPASLNDTSLLDHDPCLGTGGPVFPWLVLGGYDDQN YNNATALVITFPVNNYYNDTEKLQRAQAWKEFINFVKNYKNPNL TISFTAERSIEDELNRESDSVFTVVISYAIMFLYISLALGHMKSCRRL LVDSKVSGLIAGILIVLSSVACSLGVFSYIGLPLTLIVIEVIPFLVLA VDNIFILVQAYQRDERLQGETLDQQLGRVLGEVAPSMFLSSSFSETVA FFLGALSVMFAVHTFSLFAGLAVFIDFLLQITCFV SLLGLDIKQEKNRDLIFCCVRGAEDGTSVQASESCLFRFFKNSYSPL LLKDWMRPIVIAIFVGVLSFSIAVLNKVDIGLDQSLSMPPDDSYMVDY FKSISQYLHAGPPVYFVLEEGLDYSKQNMVCGGMGCNNDLSLV QQIFNAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRVDNITDQFC NASVVDPAVCRCRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCG KGGHAAAYSSAVNILLGHGTRVGATYFMTYHTVLQTSADFIDALKK ARLIASNVTTETMGINGSAYRVFPYSVFYVFYEQYLTIIDDTIFNLGVS	NPC1

	LGAIFLVTMVLLGCELWSAVIMCATIAMVLVNMFGVMWLWGISLN AVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRVERAEEALAHMGS SVFSGITLTKFGGIVVLAFAKSQLIFQIFYFRMYLAMVLLGATHGLIFL PVLLSYIGPSVKNKAKSCATEERYKGTERRERLLNF	
295	MRFLAATFLLLALSTAAQAEPVQFKDCGSVDGVIKEVNVSPCPTQP CQLSKGQSYSVNVTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPDGC KSGINCPIQDKTYSYLNKLPVKSEYPSIKLVVEWQLQDDKNQSLFC WEIPVQIVSHL	NPC2
296	MSCPVPACCALLLVGLCRARPRNALLLADDGGFESGAYNNSAIA TPHLDALARRSLLFRNAFTSVSSCSPSRASLLTGLPQHONGMYGLH QDVHFFNSFDKVRSLPLLSQAGVRTGIIGKKHVGPETVYPDFAYT EENGSVLQVGRNITRIKLLVRKFLQTQDDRPFFLYVAFHDPHRCGHS QPQYGTFCFKFGNGESGMGRIPDWTPQAYDPLDVLVPYFVPNTPAA RADLAAQYTTVGRMDQGVGLVLQELRDAGVLNDTLVIFTSDNGIPF PSGRTNLYWPGTAEPLLVSSPE HPKRWGGVSEAYVSLDDLTPILDWFSIPYPSYAIFGSKTIHLTGRSL LPALAEPLWATVFGSQSHHEVTMSYPMRSVQHRHFRLVHNLNFK MPFPIDQDFYVSPTFQDLLNRTTAGOPTGWYKDLRHYYRARWEL YDRSRDPHETQNLATDPRFAQLLEMLRDQLAKWQWETHDPWVCA PDGVLEEKLSPOCQPLHNEL	SGSH
297	MASPGCLWLLAVALLPWTCASRALQHLDPPAPLPLVIWHMGMDSC CNPLSMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFFLNVNSQV TTVCQALAKDPKLQGGYNAMGFSQGGQFLRAVAQRCPSPPMINLIS VGGQHQQVFGLPKPCGESSHICDFIRKTLNAGAYSKVQERLVQAE YWHDPKEDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKF LNDSDVDPVDSEWFGFYRSGQAKETIPLQETSLYTQDRLGLKEMDN AGQLVFLATEGDHLQLSEEFYAHIPFLG	PPT1
298	MYALFLLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGA VKHCLQTVWNKPTVKSLPCDICKDVVTAAGDMLKDNATEEEILVY LEKTCDWLKPKNMSASCKEIVDSYLPVILDIIKGEMSRPGEVCSALN LCESLQKHLAELNHQKQLESNKIPELDMTEVVAPFMANIPLLLYPQ DGPRSKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHV EECDRLGPGMADICKNYISQYSEIAIQMMMHHMQPKEICALVGFCDE VKEMPMQTLVPAKVASKNVIPALELVEPIKKHEVPAKSDVYCEVCE FLVKEVTKLIDNNKTEKEILDAFDKMC SKLPKSLSEECQEVVDYGS SILSILLEEVSPELVCSMLHLCSGTRLPALT VHVTQPKDGGFCEVCK KLVGYLDRNLEKNSTKQEILAALEKGC SFLPDYQKQCDQFVAEYE PVLIEILVEVMDPSFVCLKIGACPSAHKPLLGTCKIWGPSYWCQNT ETAAQCNAVEHCKRHVWN	PSAP
299	MRSPPVRLARNDGEESTDRTPLLPGAPRAEAAPVCCSARYNLAILA FFGFFIVYALRVNLSVALVDMVDSNTTLEDNRTSKACPEHSAPIKVH HNQTGKKYQWDAETQGWLGSFFYGYIITQIPGGYVASKIGGKMLL GFGILGTAVLTLFTPIAADLGVGPLIVLRALEGLGEGVTFPAMHAM WSSWAPPLERSKLLSISYAGAQLGTVISLPLSGIICYMNWTVFYF FGTIGIFWFLWIWLVSDTPQKHKRISHYEKEYILSSLRNQLSSQKSV PWVPILKSLPLWAIVVAHFSYNWTFYTLTLTLLPTYMKEILRFNVQEN GFLSSLPYLGSWLCMILSGQAADNLRKWNFSTLCVRRIFSLIGMIG PAVFLVAAGFIGCDYSLAVAFLTISTTLGGFCSSGFSINHLDIAPSYA GILLGITNTFATIPGMVGPVIAKSLTPDNTVGEWQTVFYIAAAINVFG AIFFTLFAKGEVQNWALNDHHGHRH	SLC17A5
300	MPRYGASLRQSCPRSGREQQDGTAGAPGLLWMGLVLALALALAL ALALSDSRVLWAPAEAHPLSPQGHPARLHRIVPRLRDVFGWGNLTC	SMPD1

	PICKGLFTAINLGLKKEPNVARVGSVAIKLCNLLKIAPPAVCQSIVHL FEDDMVEVWRRSVLSPSEACGLLLGSTCGHWDIFSSWNISLPTVPKP PPKPPSPAPGAPVSRILFLTDLHWDHLDYLEGTDPCADPLCCRRGS GLPPASRPGAGYWGEYSKCDLPLRTLLESLLSGLGPAGPFDMVYWT GDIPAHADVWHQTRQDQLRALTTVTALVRKFLGPVPVYPVAVGNHES TPVNSFPPPFIEGNHSSRWLYEAMAKAWEPWLPALRTLRIGGFY ALSPYPGLRLISLNMNFCSRENFLLINSTDPAGQLQWLVGELQAA EDRGDKVHIIGHIPPGHCLKSWSWNYRIVARYENTLAAQFFGHTH VDEFEVFYDEETLSRPLAVAFAPSATTYIGLNPGRVYQIDGNYS SSHVVLDHETIYLNLTQANIPGAIPHWQLLYRARETYGLPNTLPTAW HNLVYRMRGDMQLFQTFWFLYHKGHPPSEPCGTPCRLATLCAQLS ARADSPALCRHLMPLDGLPEAQLWPRPLFC	
301	MAAPALGLVCGRCPELGLVLLLLLLSLLCGAAGSQEAGTGAGAGSL AGSCGCGTPQRPAGHSSAAAHRYSSREANAPGPVPPERQLAHSKM VPIPAGVFTMGTDDPQIKQDGEAPARRVTIDAFYMDAYEVSNTFE KFVNSTGYLTEAEKFGDSFVFEGMLSEQVKTNIQQA VAAAPWWLP VKGANWRHPEGPDSTILHRPDHPVLHVSWNDAYCTWAGKRLP TEAEWEYSCRGLHNRLFPWGNKLQPKGQHYANIWQGEFPVTNTG EDGFQGTAPVDAFPNGYGLYNIVGNAWEWTSDDWWTVHHSVEET LNPKGPPSGKDRVKKGGSYMCHRSYCYRYRCAARSQNTPDSSASN LGFRCAADRLPTMD	SUMF1
302	MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEEL SLTFALRQQNVERLSELVQAVSDPSSPQYGYLTLENVADLVRPSPL TLHTVQKWLLAAGAQAQCHSVITQDFTCWLSIRQAELLPLGAEFHH YVGGPTETHVVRSPHPYQLPQALAPHVDFVGGGLHRFPPTSSLRQRPE PQVTGTVGLHLGVTPSVIRKRYNLTSQDVSGSGTSNNSQACAQFLEQ YFHDSDLAQFMRLFGGNFAHQASVARVVGQQGRGRAGIEASLDVQ YLMSAGANISTWVYSSPGRHEG QEPFLQWLMLLSNESALPHVHTVSYGDDDESLSSAYIQRVNTELMK AAAGFLTLFASGDSGAGCWSVSGRHQFRPTFPASSPYVTTVGGTS FQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSPHLPSSYF NASGRAYPDVAALSDGYWVVSNRVPIPVWSGTSASTPVFGGILSLIN EHRILSGRPPLGFLNRLYQQHGAGLFDVTRGCHESCLDEEVEGQGF CSGPGWDPVTGWGTPNFPALLKTLLNP	TPP1
303	MSDKLPYKVADIGLAAWGRKALDIAENEMPGLMRMRERYASASKPL KGARIAGLHMTVETAVLIETLVTLGAEVQWSSCNIFSTQDHAAAAI AKAGIPVYAWKGETDEEYLWCIEQTLYFKDGPLNMILDDGGDLTN LIHTKYPQLLPGIRGISEETTTGVHNLYKMMANGILKVPAINVNDV TKSKFDNLYGCRESLIDGIKRA TDVMIAGKVAVVAGYGDVGKGA QALRGFGARVIITEIDPINALQAAMEGYEVTMTDEACQEGNIFVTTT GCIDIILGRHFEQMKDDAIVCNIG HFDVEIDVKWLNENAVEKVNIPQVDRLKNGRRILLAEGR LVN LGCAMGHPSFVMSNSFTNQVMAQIELWTHPDKYPVGVHFLPKKLD EAVAEHLGKLVNKLTKLTEKQAQYLGMSCDGPFKPDHYRY	AHCY
304	MVDSVYRTRSLGVAAEGLPDQYADGEAARVWQLYIGDTRSRTAEY KAWLLGLLRQHGCQRVLDVACGTGVDSIMLVEEGFSVTSVDASDK MLKYALKERWNRREHAPFDKWVIEEANWMTLDDKDVPSAEGGFD AVICLGNSFAHLPDCKGDQSEHRLALKNIASMRAGGLLVIDHRNY DHILSTGCAPPKNIYYKSDLTKDVTTSVLIVNNKAHMTLDTYTVQ VPGAGQDGSPGLSKFRLSYYPHCLASFTELLQAAFGGKCQHSVLGD FKPYKPGQTYIPCYFIHVLKRTD	GNMT

305	MNGPVDGLCDHSLSEGVFMFTSESVGEGHPDKICDQISDAVLDAHL KQDPNAKVACETVCKTGMVLLCCEITSMAMVDYQRVVRDTIKHIG YDDSAKGFDFKTCNVLVALEQQSPDIAQCVHLDRNEEDVGAGDQG LMFGYATDETEECMPLTIILAHKLNARMADLRRSGLLPWLRPDSKT QVTVQYMQDNGAVIPVRIHTIVISVQHNEEDITLEEMRRALKEQVIRA VVPKAYLDEDTVYHLQPSGRFVIGGPQGDAGVTGRKIIVDTYGGW GAHGGGAFSGKDYTKVDRSAAYAARWVAKSLVKAGLCRRVLVQ VSYAIGVAEPLSISIFTYGTSTQKTERELLDVVHKNFDLRPGVIVRDL LKKPIYQKTACYGHFGRSEFPWEVPRKLVF	MAT1A
306	MEKGPVRAPAEKPRGARCSNGFPERDPPRPGPSRPAEKPPRPEAKSA QPADGWKGERPRSEEDNELNLPNLAAAYSSILSSLGENPQRQGLLK TPWRAASAMQFFTKGYQETISDVLNDAIFDEHDHDEMVIKIDIMFS MCEHHLVPFVGKVHIGYLPNKQVLGLSKLARIVEIYSRRLQVQERL TKQIAVAITEALRPAGVGVVEATHMCMVMRGVQKMNSKTVTST MLGVFREDPKTREEFLLIRS	GCH1
307	MAGKAHRLSAEERDQLLPNLRAVGVWNELEGRDAIFKQFHFKDFNR AFGFMTRVALQAEKLDHHPWFNVYNKVHITLSTHECAGLSERDIN LASFIEQVAVSMT	PCBD1
308	MSTEGGRRRCQAQVSRISFSASHRLYSKFLSDEENLKLFGKCNNP NGHGHNYKVVTVHGEIDPATGMVMNLADLKKYMEEAIMQPLDH KNLDMDPYFADVSTTENVAVYIWDNLQKVLPVGVLYKVKVYE TDNNIVVYKGE	PTS
309	MAAAAAAGEARRVLVYGGRGALGSRVQAFRARNWWWASVDVV ENEEASASIIKMTDSFTEQADQVTAEVGKLLGEEKVDAILCVAGG WAGGNAKSKSLFKNCMLMWKQSIWTSTISSHLATKHLKEGGLTL AGAKAALDGTGPMIGYGMAGAVHQLCQSLAGKNSGMPPGAAAI AVLPVTLDTPMNRKSMPEADFSSWTPLEFLVETFDHWITGKNRPSS GSLIQVVTTEGRTELTPAYF	QDPR
310	MEGGLGRAVCLLTGASRGFGRTLAPLLASLLSPGSVLVLSARND LRQLEAELGAERSGLRVVRPADLGAELQQLLGAELRELPRPKGL QRLLLINNAGSLGDVSKGFVDLSSTQVNNY WALNLTSMLCLTSSV LKAFPDSPGLNRTVNNISLCLQPFKGWALY CAGKAARDMLFQV LALPEPNVRVLNYAPGLDTDMQQLARETSVDPDMRKGLQELKAK GKLVDCKVSAQKLLSLEKDEFKSGAHVDFYDK	SPR
311	MDAILNYRSEDTEYTYLLGCDELSSVEQILAEFKVRALECHPDKHP ENPKAVETFQKLQKAKEILTNEESRARYDHWRRSQMSMPFQQWEA LNDSVKTSMHVVRGKKDLMLEESDKTHTTKMENEECNEQRERK KEELASTAEKTEQKEPKPLEKSVSPQNSDSSGFADVNGWHLRFRWS KDAPSELLRKFRNYEI	DNAJC12
312	MLLPAPALRRALLSRPWTGAGLRWKHTSSLKVANEPVLAFTQGSPE RDALQKALKDLKGRMEAIPC VVGDEEVWTS DVQYQVSPFNHGHK VAKFCYADKSLNKAIEAALAARKEWDLKPIADRAQIFLKAADMLS GPRRAEILAKTMVGQGTVIQAEIDAAELIDFFRFNAKYAVELEG QQPISVPPSTNSTVYRGLEGFVAAISPFNFTAIGGNLAGAPALMGNV VLWKPSDTAMLASYAVYRILREAGLPNNIIQFVPADGPLFGDTVTS EHLCCGINFTGSVPTFKHLWKQVAQ NLDRFHTFPRLAGECGGKNFHFVHRSADVSVVSGTLRSafeYGGQ KCSACSRLYVPHSLWPQIKGRLL EEHSRIKVGDP AEDFGTFFSAVID AKSFARIKKWLEHARSSPSLTILAGGKCDDSVGYFVEPCIVESKDPQ EPIMKEEIFGPVLSVYVYPDDKYKETLQLVDSTTSYGLTGAVFSQDK	ALDH4A1

	DVVQEATKVLRNAAGNFYINDKSTGSIVGQQPFGGARASGTNDKP GGPHYLRWTSPQVIKETHKPLGDWSYAYMQ	
313	MALRRALPALRPCIPRFVQLSTAPASREQPAAGPAAVPGGGSATAV RPPVPAVDGNAQEAAYRSRRTWELARSLVLRLCAWPALLARHEQ LLYVSRKLLGQRLFNKLMKMTFYGHFVAGEDQESIQLLRHYRAFG VSAILDYGVEEDLSPEEAHEKEMESCTSAERDGS GTNKRDKQYQA HRAFGDRRNGVISARTYFYANEAKCDSHMETFLRCIEASGRVSDDG FIAIKLTALGRPQFLLQFSEVLAKWRCFFHQMAVEQQAGLAAMDT KLEVAVLQESVAKLGIASRAEIEDW FTAETLGVSGTMDLLDWSSLIDSRTKLSKHLVVPNAQTGQLEPLLSR FTEEEELQMTRMLQRMVDLAKKATEMGMVRLMVD AEQTYFQPAISR LTLEMQRKFNVEKPLIFNTYQCYLKDAYDNVTLDVELARREGWCF GAKLVRGAYLAQERARAAEIGYEDPINPTYEATNAMYHRCLDYVL EELKHNAKAKVMVASHNEDTVRFALRRMEELGLHPADHQVYFGQ LLGMCDQISFPLGQAGYPVYKYVPYGPVMEVLPYLSRRALNSSLM KGTHRERQLLWLELLRRLRTGNLFHRPA	PRODH
314	MTTYSKDGAKPERGRFLHFHSVTFWVGNNAKQAASFYCSKMGFEPL AYRGLETGSREVSVSHVIKQKIVFVLSSALNPWNKEMGDHLVKHG DGVKDIAFEVEDCDYIVQKARERGA KIMREPWVEQDKFGKVKFAV LQTYGDTTHTLVEKMNYIGQFLPGYEAPAFMDPLL PKLPKCSLEMI DHIVGNQPDQEMVSASEWYLKNLQFHRFWSVDDTQVHTEYSSLSI VVANYEESIKMPINEPAPGKKKSQIQEYVDYNGGAGVQHIALKTEDI ITAIRHLRERGLEFLSPSTYYKQLREKLKTAKIKVKENIDALEELKI LVVDYDEKGYLLQIFTKPVQDRPTLFLEVIQRH NHQGFGAGNFNSLF KAFEEEEQNLRGNLTNMETNGVVPGM	HPD
315	MEFSSPSREECPKPLSRV SIMAGSLTG LLLLQAVSWASGARPCIPKSF GYSSVVCVCNATYCD SFDPP TFPALGTFSRYESTRSGRRMELSMGPI QANHTGTG LLLTLQPEQKFQKVKGF GGAMTDAAALNILALSPPAQ NLLLKS YFSEEGIGYNIIRVPMASCD FSIRTYTYADTPDDFQLHNFSL PEEDTKLKIPLIHRALQLAQRPV SLLASPWTSP T WLKTNGAVNGKGS LKGQP GDIYHQTWARYFVKFLDAYAEHKLQFWAVTAENEPSAGLLSGYPF QCLGFTPEHQ RDFIARDLGPTLANSTHHNVRLMLDDQRLLLP HWA KVVLTDPEAAKYVHGIAVHWYLD FLAPAKATLGETHRLFPNTMLF ASEACVGSKFWEQSVRLGSWDRGMQYSHSITNLLYHVVGWTDW NLALNPEGPNWVRNFVDSPIVDITKDTFYKQPMFYHLGHFSKFIP EGSQRVGLVASQKNDLDAVALMHPDGS AVVVVLNRSSKDVPLTIK DPAVG FLETISPGYSIHTYLWRRQ	GBA
316	MAELKYISGFGNECSSEDP RCPGSLPEGQNNPQVCPYNLYAEQLSGS AFTCPRSTNKR SWLYRILPSVSHKPFESIDEGQVTHNWDEVDPDPNQ LRWKPF EIPKASQKKVDFVSGLHTLCGAGDIKSNNGLAIHIFLCNTS MENRCFYNSDGD FLIVPQKGNLLIYTEFGKMLVQPNEICVIQRGMRF SIDVFEETR GYILEVYG VHFELPD LGPIGANLANPRDFLIPIAWYED RQVPGGYTVINKYQGKLFAAKQDVSPFN VVAWHGNYTPYKYNLK NFMVINSVAFDHADPSIFTVLTAKSVRPGVAIADVFIFPPRWGVADK TFRPPYYHRNCMSEFMGLIRGHYEA KQGGFLPGGSLHSTMTPHGP DADC FEKASKVKLAPERIADGTMAFMFESSLSLAVTKWGLKASRCL DENYHKCWEPLKSHFTPN SRNPAEPN	HGD
317	MGVLGRVLLWLQLCALTQAVSKLVWPNTDFDV AANWSQNRTPCA GGAVEFPADKMVS VLVQEGHAVSDMLLPLD GELVLASGAGFGVSD VGSHLDCGAGEP AVFRDSDRFSWHD PHLWRS GDEAPGLFFVDAER VPCRHDDVFFPPSASFRVGLGPGASPV RVR SIALGRTFTRDEDLAV	AMN

	FLASRAGRRLRFHGPAGLSVGPEDCADPSGCVCGNAEAQPWICAALL QPLGGRCPPQAACHSALRPQGQCCDLGCAVLLTHGPAFDLERYRA RILDFTLGLPQYHGLQVAVSKVPRSSRLREADTEIQVVLVENGPELG GAGRLARALLADVAENGEALGVLEATMRESGAHVWGSSAAGLAG GVAAAVLLALLVLLVAPPLLRAGRLRWRHEAAAAPAGAPLGFRN PVFVDVTASEELPLPRRLSLVPKAAADSTSHSYFVNPLFAGAEAEA	
318	MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQ AAGPSSSGSCPPTKFQCRSTGLCVPLTWRCRDLDCSDGSDEEECRIE PCTQKQGCPPPPGLPCPCTGVSDCSGGTDKLRNCSRLACLAGELR CTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPV TLESVTSRLNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVI AAAVALSASLVTATLLLLSWLRAQERLRPLGLLVAMKESLLLSEQK TSLP	CD320
319	MMNMSLPFLWSLLTLLIFAEVNGEAGELELQRQKRSINLQQPRMAT ERGNLVFLTGSAQNIEFRTGSLGKIKLNDEDLSECLHQIQKNKEDIIE LKGSAGLPPQNISSQIYQLNSKLVDLERKFQGLQQTVDKVCSSNPC QNGGTCLNLHDSFFCICPPQWKGPLCSADVNECEIYSGTPLSCQNGG TCVNTMGSYSCHCPPEITYGPQCASKYDDCEGGSVARCVRHIGICEDL MREQAGEPKYSCVCDAGWMFSPNSPACTLDRDECSFQPGPCSTLV QCFNTQGSFYCGACPTGWQNGYICEDINECEINNGGCSVAPPVEC VNTPGSSHCQACPPGYQGDGRVCTLTDCSVSNGGCHPDASCSSTL GSLPLCTCLPGYTGNGYGPNGCVQLSNICLSHPCLNGQCIDTVSGYF CKCDSGWTGVNCTENINECLSNPCLNGGTCVDGVDSFSCECTRLWT GALCQVPQQVCGESLSGINGSFSYRSPDVGYVHDVNCFWVIKTEMG KVLRTFTFFRLESMDNCPHEFLQVYDGDSSSAFQLGRFCGSSLPHE LLSSDNALYFHLHYSEHLRNGRGFTVRWETQQPECGGILTGPYGSIKS PGYPGNYPGRDCVWIVVTSPDLLVTFTFGTSLSEHHDDCNKDYLEI RDGPLYQDPLLKGFCTTFSVPPLQTTGPFARIHFHSDSQISDQGFHIT YLTSPSDLRCGGNYTDPEGELFLPELSGPFTHTRQCVYMMKQPQGE QIQINFTHVELQCQSDSSQNYIEVRDGETLLGKVCNGTISHIKSITN SVWIRFKIDASVEKASFRAVYQVACGDELTEGEGVIRSPFFPNVYPGE RTCRTWTHQPQSQVILLNFTVFEIGSSAHCETDYVEIGSSSILGSPENK KYCGTDIPSFITSVYNFLYVTFVKSSSTENHGFMAKFSAEIDLACGEIL TESTGTIQSPGHPNVYPHGINCTWHILVQPNHLIHLMFETFHLEFHY NCTNDYLEVYDSTDSETSLGRYCGKSIPPSLTSSGNSL MLVFVTDSDLAYEGFLINYEISAATAACLQDYTDLGTFTSPNFPNN YPNNWECIYRITVRTGQLIAVHFTNFSLEEAGNYYTDFLEIRDGGYE KSPLLGIIFYGNSLPPTIISHSNKLWLKFKSDQIDTRSGFSAYWDGSST GCGGNLTSSGTFISPNYMPYYHSSECYWWLKSSHGSAFELEFKDF HLEHHPNCTLDYLAVYDGPSSNSHLLTQLCGDEKPLIRSSGDSMFI KLR TDEGQQGRGFKAERYQTCENVVIVNQTYGILESIGYPNPYSENQHC NWTIRATTGNTVNYTFLAFDLEHHINCSTDYLELYDGPRQMGRYCG VDLPPPGSTTSSKLQVLLTLDGVGRREKGFQMQWFVYCGGELSG ATGSFSSPGFPNRYPPNKECIWYIRTDPGSSIQLTIHDFDVEYHSRCN FDVLEIYGGPDFHSPRIAQLCTQRSPENPMQVSSTGNELAIRFKTDLS INGRGFNASWQAVTGGCGGIFQAPSGEIHSPNYPSPYRSNTDCSWVI RVDNRHRVLLNFTDFDLEPDQSCIMAYDGLSSTMSRLARTCGREQL ANPIVSSGNSLFLRFQSGPSRQNRGFRAQFRQACGGHILTSSFDTVSS PRFPANYPNQNCWSIIQAQPPLNHITLSFTHFELERSTTCARDFVEIL DGGHEDAPLRGRYCGTDMHPITSFSSALTLRFVSDSSISAGGFHTT VTASVSACGGTFYMAEGIFNSPGYPDIYPPNVECVWNIVSSPGNRLQ LSFISFQLEDSQDCSRDFVEIREGNATGHLVGRYCGNSFPLNYSSIVG	CUBN

	<p>HTLWVRFISDGS SGTGFQATFMKIFGNDNIVGTHGKVASPFWPEN YPHNSNYQWTVNVNASHVVHGRILEMDIEEIQNCYYDKLRIYDGPS IHARLIGAYCGTQTESFSSTGNSLTFHFYSDSSISGKGFLLEWFAVDA PDGVLPTIAPGACGGFLRTGDAPVFLFSPGWPDYSNRVDCTWLIQ APDSTVELNILSLDIESHRTCAYDSLVRDGDNNLAQQLAVLGCREIP GPIRSTGEYMFIRFTSDSSVTRAGFNASFHKSCGGYLHADRGIIITSPK YPETYPSNLNCSWHVLVQSGLTIAVHFEQPFQIPNGDSSCNQGDYLV LRNGPDICSPPLGPPGGNGHFCGSHASSTLFTSDNQMFVQFISDHSNE GQGFKIKYEAKSLACGGNVYIHDADSAGYVTSNHPHNYPPHADCI WILAAPPETRIQLQFEDRFDIEVTPNCTSNYLELRDGVDSADAPILSKF CGTSLPSSQWSSGEVMYLRFRSDNSPTHVGFKAKYSIAQCGRVPG QSGVVESIGHPTLPYRDNLFCWEHLQGLSGHYLTISFEDFNLQNSSG CEKDFVEIWDNHTSGNILGRYCGNTIPDSIDTSSNTAVVRFVTDGSV TASGFRRLFESSMEECGDLQGSIGTFTSPNYPNPNPHGRICEWRITA PEGRRITLMFNRLRLATHPSCNNEHVIVFNGIRSNSPQLEKLCSSVNV SNEIKSSGNTMKVIFFTDGSRPYGGFTASYTSSSEDAVCGGSLPNTPE GNFTSPGYDGVNRNYSRNLNCEWTLNPNQGNSSISIHFDIFYLESHQ DCQFDVLEFRVGDADGPLMWRLCGPSKPTLPLVIPYSQVWIHFVTN ERVEHIGFHAKYSFTDCGGIIGD SGVITSPNYPNAYDSLTHCSSLLE APQGHTITLTFSDFDIEPHTTCAWDSVTVRNGGSPESPIIGQYCGNSN PRTIQSGSNQLVVTFNSDHSLQGGGFYATWNTQTLGCGGIFHSDNG TIRSPHWPQNFPENSRCSWTAITHKSKHLEISFDNNFLIPSGDGQCQN SFVKVWAGTEEVDKALLATGCGNVAPGPVITPSNTFTAVFQSQEAP AQGFSASFVSRCGSNFTGPSGYIISPYPKQYDNNMNCTYVIEANPL SVVLLTFVSFHLERSAVTGSCVNDGVHIIRGYSVMSTPFATVCG DEMPAPLTIAGPVLLNFYSNEQITDFGFKFSYRIISCGGVFNFFSSGITS PAYSYADYPNDMHCLYTITVSDDKVIELKFSDFDVVPSTSCSHDYL AIYDGANTSDPLLKFCGSKRPPNVKSSNNSMLLVFKTDSFQTAGK WKMSFRQTLGPQQGCGGYLTGSNNTFASPDSDSNGMYDKNLNCV WIIAPVNVKVIHLTFNTFALEAASTRQRCLYDYVKLYDGDSENANLA GTFCGSTVPAPFISSGNFLTQFISDLTLEREGFNATYTIMDMPCGGT YNATWTPQNISSPNSSDPDVPFSICTWVIDSPPHQVQKITVWALQLT SQDCTQNYLQLQDSPQGHGNSRFQFCGRNASAVPVFYSSMSTAMVI FKSGVVNRNSRMSFTYQIADCNRDHYHKAFGNLRSPGWPNDYDNDK DCTVTLTAPQNHTISLFFHSLGIENSVECRNDFLEVRNGSNSNSPLL KYCGTLLPNPVFSQNNELYLRFKSDSVTSDRGYEIIWTSSPSGCGGT LYGDRGSFTSPGYPGTYPNNTYCEWVLVAPAGRLVTINFYFISIDDP GDCVQNYLTLYDGPNAASSPSGPYCGGDTSIAPFVASSNQVFIKFHA DYARRPSAFRLTWDS</p>	
320	<p>MAWFALYLLSLLWATAGTSTQTQSSCSVPSAQEPLVNGIQVLMENS VTSSAYPNPSILIAMNLAGAYNLKAQKLLTYQLMSSDNDLTIGQL GLTIMALTSSCRDPGDKVSILQRQMENWAPSSPNAEASAFYGPSLAI LALCQKNEATLPIAVRFAKTLLANSSPFNVDTGAMATLALTCMYN KIPVGSEEGYRSLFGQVLKDIVEKISMKIKDNGIIGDIYSTGLAMQAL SVTPEPSKKEWNCKKTTDMILNEIKQGKFHNPMISIAQILPSLKGTKY LDVPQVTCSPDHEVQPTLPSNPGPGPTSASNITVIYTINNQLRGVELL FNETINVS VKSGSVLLVLEEAAQRKNPMFKFETMTSWGLVVSSIN NIAENVNHKTYWQFLSGVTPLNEGVADYIPFNHEHITANFTQY</p>	GIF
321	<p>MRQSHQLPLVGLLLFSFIPSQLCEICEVSEENYIRLKPLLNTMIQSNY NRGTS AVNVVLSLKLVGIIQITLMQKMIQKIYNVKSRLSDVSSGE LALILALGVCRNAEENLIYDYHLIDKLENKFQAEIENMEAHNGTPL TNYYQLSLDVLALCLFNNGNYSTAEEVNVHFTPENKNYYFGSQFSVD</p>	TCN1

	GAMAVLALTCVKKSLINGQIKADEGSLKNISIYTKSLVEKILSEKKE NGLIGN TFSTGEAMQALFVSSDYNNENDWNCQQTLLNTVLTEISQGAFSPNPA AAQVLPALMGKTFDINKDSSCVSASGNFNISADEPITVTPPDSQSYI SVNYSVRINETYFTNVTVLNGSVFLSVMEKAQKMNDTIFGFTMEER SWGPIYITCIQGLCANNNDRTYWELLSSGGEPLSQGAGSYVVRNGENL EVRWSKY	
322	MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRL SLEHLNPSIYVGLRLSSLQAGTKEDLYLHSLKLGYYQCLLGSAFSED DGDCQKGKPSMGQLALYLLALRANCEFVRGHKGDRLVSQLKWFLE DEKRAIGHDHKGHPHTSYYYQYGLGILALCLHQKRVHDSVVDKLLY AVEPFHQGHHSVDTAAMAGLAFTCLKRSNFPNGRRQRITMAIRTVR EEILKAQTPEGHFGNVYSTPLALQFLMTSPMRGAELGTACLKARVA LLASLQDGAQNALMISQLLPVLNHKTYIDLIFPDCLAPRVMLEPAA ETIPQTQEHSVTLQVLSLLPPYRQSVLSTVEDVLKKAHELGGFT YETQASLSGPYLTSMVGKAAGEREFWQLLRDPNTPLLQGIADYRPK DGETIELRLVSW	TCN2
323	MQQKTKLFLQALKYSIPHLGKCMQKQHLNHYNFADHCYNRIKLKK YHLTKCLQNKPKISELARNIPSRFSCKDLQPVKQENKPLPENMDA FEKVRTKLETQPQEEYEIINVEVKHGGFVYYQEGCCLVRSKDEEAD NDNYEVLNLEELKLDQPFIDCIRVAPDEKYVAAKIRTEDSEASTCVI IKLSDQPVMEASFPNVSSFEWVKDEEDEDVLFYTFQRNLRCHDVYR ATFGDNKRNERFYTEKDPSYFVFLYLTKDSRFLTNIMNKTTSEVWL IDGLSPWDPPVLIQKRIHGVLYYVEHRDDEL YILTNVGEPTFEKLMR TAADTPAIMNWDLFFTMKRNTKVIDLDMFKDHCVLFLKHSNLLYV NVIGLADDSVRSLKLPPWACGFIMDTNSDPKNCPFQLCSPIRPPKYY TYKFAEGKLFEETGHEDPITKTSRVLRLLEAKSKDGKLVPMTVFHK DSEDLQKKPLLHVYGYGMDLKMNFRRPERRVLVDDGWILAYCH VRGGGELGLQWHADGRLTKKLNLGLADLEACIKTLHGQGFSPSLT TLTAFSAGGVLAGALCNSNPVELVRAVTLLEAPFLDVLNTMMDTTLPL T LEELEEWNPPSSDEKHKNYIKRYCPYQNIKPQHYPHSIHITAYENDER VPLKGIVSYTEKLKEAIAEHAKDTGEGYQTPNILDIQPGGNHVIEDS HKKITAQIKFLYEELGLDSTSVFEDLKKYLKF	PREPL
324	MAFANLRKVLISDSLDPCCRKILQDGGGLQVVEKQNLKEELIAELQD CEGLIVRSATKVTADVINA AEKLQVVGRAAGTGVNDVLEAATRKG LVMNTPNGNSLSAAELTCGMIMCLARQIPQATASMKDGKWERKKF MGTELNGKTLGILGLGRIGREVATRMQSFGMKTIGYDPIISPEVSASF GVQQPLPLEEIWPLCDFITVHTPLLSTTGLLNDNTFAQCKKGVRVVN CARGGIVDEGALLRALQSGQCAGAALDVFTTEPPRDRALVDHENVI SCPHLGASTKEAQSRCGEEIA VQFVDMVKGKSLTG VVNAQALTSAFSPHTKPWIGLAEALGTMRA WAGSPKGTIQVITQGTSLKNAGNCLSPAVIVGLLKEASKQADVN NAKLLVKEAGLNVTTSHSPAAPGEQGFGECLLAVALAGAPYQAVG LVQGTTPVLQGLNGAVFRPEVPLRRDLPLLLFRTQTSDPAMLPTMIG LLAEAGVRLLSYQTSLSVSDGETWHVMGISSLLPSLEAWKQHVTEAF QFHF	PHGDH
325	MDAPRQVVNFPGPAKLPHSVLLEIQKELLDYKGVGISVLEMSHRS SDFAKIINNTENLVRELLAVPDNYKVIFLQGGGCGQFSVPLNLIGL KAGRCADYVVTGAWSAKAAEEAKKFGTINIVHPKLGSYTKIPDPST WNLNPDASYVYYCANETVHGVEFDFIPDVKGAVLVCDMSSNFLSK PVDVSKFGVIFAGA QKNVGSAGVTVVIVRDDLLGFALRECPSVLEY	PSAT1

	KVQAGNSSLYNTPPCFSIYVMGLVLEWIKNNGGAAAMEKLSSIKSQ TIYEIIDNSQGFYVCPVEPQNRSKMNIPFRIGNAKGDDALEKRFLDK ALELNMLSLKGHRVGGIRASLYNAV TIEDVQKLA AFMKKFLEMH QL	
326	MVSHSELRLKFYSADAVCFD V DSTVIREEGIDELAKICGVEDAVSE MTRRAMGGAVPFKAALTERLALI QPSREQVQRLIAEQPHLTPGIRE LVSRLQERNVQVFLISGGFRSIVEHVASKLNIPATNVFANRLKFYFN GEYAGFDETQPTAESGGKGKVIKLLKEKFHFKKIIMIGDGATDMEA CPPADAFIGFGGNVIRQQVKDNAK WYITDFVELLGELEE	PSPH
327	MQRAVS VVARLGFRLQAFPPALCRPLSCAQEVLRRTPLYDFHLAHG GKMVAFA GWSLPVQYRDSHTD SHLHTRQHCSLFDVSHMLQTKILG SDRVKLMESLVVG DIAELRPNQGTLSLFTNEAGGILDDLIVTNTSEG HLYVVS NAGCWEKDLALMQDKVRELQNQGRDVGLEVL DNALLAL QGPTAAQVLQAGVADDLRKLPFMTSAVMEVFGVSGCRVTRCGYT GEDGVEISVPVAGAVHLATAILKNPEVKLAGLAARDSLRLEAGLCL YGN DIDEHTTPVEGSLSWTLGKRRRAAMD FPGAKVIVPQLKGRVQ RRRVGLMCEGAPMRAHSPILNMEGTKIGTVTSGCPSPSLKKNVAMG YVPCEYSRPGTMLLVEVRRKQQMAV VSKMPFVPTNYITLK	AMT
328	MALRVVRSVRALLCTLRAPVSPAAPCPPRPWQLGVGAVRTLRTGP ALLSVRK FTEKHEWVT TENGIGTVGISNFAQEALGDV VYCSLPEVG TKLNKQDEFGALESVKAASELYSPLSGEVTEINEALAENPGLV NKSC YEDGWLIKMTLSNPSELDELMSEEAYEKYIKSIEE	GCSH
329	MQSCARAWGLRLGRGVGGRRLAGGSGPCWAPRSRDSSSGGDS AAAGASRLLERLLPRHDDFARRHIGPGDKDQREMLQTLGLASIDELI EKTVPANIRLKRPLK MEDPVCENEILATLHAISSKNQI WRSYIGMGY YNCSVPQTILRN LLENSGWITQYTPYQPEVSQGRLESLLNYQTMVC DITGLDMANASLLDEGTAAAEALQLCYRHNKRRKFLVDPRCHPQTI AVVQTRAKYTGVLTELKLPCEMDFSGKDVSGVLFQYPDTEGKVED FTELVERAHQSGSLACCATDLLALC ILRPPGEFGVDIALGSSQRFGVPLGYGGPHAAFFAVRESLVRMMPGR MVGVTTRDATGKEVYRLALQ TREQHRRDKATSNICTAQALLANMA AMFAIYHGSHGLEHIARRVHNATLILSEGLKRAGHQLQHDLFFDTL KIQC GCSVKEVLGRAAQRI NFRLFEDGTLGISLDET VNEKDLDDL WIFGCESSAELVAESMGEECRGIPGSVFKRTSPFLTHQVFN SYHSET NIVRYMKKLENKDISLVHSMIPLG SCTMKLNSSELAPITWKEFANI HPFVPLDQAQGYQQLFRELEKDLCELTGYDQVCFQPNSGAQGEYA GLATIRAYLNQKGEGHRTVCLIPKSAHG TNPASAHMAGMKIQPVEV DKYGNIDAVHLKAMVDKHKENLA AIMITYPSTNGVFEENISDVCDL IHQHGGQVYLDGANMNAQVGICRPGDFGSDVSHLNLHKTFCIPHG GGGPGMGPVGKHLAPFLPNHPVISLKR NEDACPVGTVSAAPWGS SSILPISWAYIKMMGGKGLKQATETAILNANYMAKRLETHYRILFR GARGYVGHEFILDTRPFKKSANIEAVDVAKRLQDYGFHAPTMSWP VAGTLMVEPTESDKAELDRFCDAMISIRQEIADIEEGRIDPRVNPLK MSPHSLTCVTSSHWD RPYSREVA AFPLPFVKPENKFWP T IARIDDIY GDQHLVCTCPPMEVYESPFSEQKRASS	GLDC
330	MSLR CGDAARTLGPRVFGRYFCSPVRPLSSLPDKKKELLQNGPD LQ DFVSGDLADRSTWDEYKGNLKRQKGERLRLPPWLKTEIPMGKNYN KLKNTLRNLNLHTVCEEARCPNIGECWGGGEYATATATIMLMGDT CTRGCRCFSVKTARNPPPLDASEPYNTAKAIAEWGLDYVVLTSVDR DDMPDGGAEHIAKT VSYLKERNPKILVECLTPDFRGDLKAIEKVALS GLDVYAHNVETVPELQSKVRDPRANFDQSLRVLKHAKKVQPDVIS	LIAS

	KTSIMLGLGENDEQVYATMKALREADVDCLTLGQYMQPTRRHLK VEEYITPEKFKYWEKVGNELGFHYTASGPLVRSSYKAGEFFL KNLVAKRKTKDL	
331	MAATARRGWGAAAVAAGLRRRFCHMLKNPYTIKKQPLHQFVQRP LFPLPAAFYHPVRYMFIQTQDTPNPNSLKFIKPKVLETRTMDFTP AAFRSPLARQLFRIEGVKSFFGPDFITVTKEEELDWNLLKPDYAT IMDFASGLPLVTEETPSGEAGSEEDDEVVAMIKELLDTRIRPTVQE DGGDVIIYKGFEDGIVQLKLQGSCTSCPSSIITLKNGIQNMLQFYIPEV EGVEQVMDDESDEKEANSP	NFU1
332	MSGGDTRAAIARPRMAAAHGPVAPSSPEQVTLLPVQRSFLLPPFSGA TPSTSLAESVLKVWHGAYNSGLLPQLMAQHSLAMAQNGAVPSEAT KRDQNLKRGNGWGNQIEFVLTSVGYAVGLGNVWRFYLCYRNGGG AFMFYFIMLIFCGIPLFFMELSFQQFASQGLGVWRISPMFKGVGY GMMVVSTYIGIYNNVICIAFYFFSSMTHVLPWAYCNPWNTHD CAGVLDASNLNNGSRPAALPSNLSHLLNLSLQRTSPSEYWRLYVL KLSDDIGNFGEVRLPLLGLGVSWLVVFLCLIRGVKSSGKVYFTA TFPYVVLTLFVRGVTLEGAFDGMYYLTPQWDKILEAKVWGDAAS QIFYSLGCAWGLITMASYNKFHNNCYRDSVSIITNCATSVYAGFV IFSILGFMANHLGVDVSRVADHGPGLAFVAYPEALTLLPISPLWSLL FFFMLILLGLGTQFCLLETLVTAIVDEVGNEWILQKKTYVTLGVA GFLGLIPLTSQAGIYWLLMDNYAASFSLVVISCMCVAIMYIYGHR NYFQDIQMMLGFPPPLFFQICWRFVSPAIFFILVFTVIQYQPITYNHY QYPGWAVAIGFLMALSSVLCIPLYAMFRLCRTDGDTLQLKKNATK PSRDWGPALLEHRTGRYAPTIA SPEDGFEVQPLHPDKAQIPVGSN GSSRLQDSRI	SLC6A9
333	MEPSSKKLTGRLMLAVGGAVLGSQFGYNTGVINAPQKVIEEFYNQ TWVHRYGESILPTTLTWSLSVAIFSVMIGSFSVGLFVNRFGR NSMLMMNLLAFVSAVLMGFSLKGSFEMILGRFIIYVYCGLTTF VPMYVGEVSPTALRGALGTLHQLGIVVGILIAQVFLGDSIMGNKDL WPLLLSIIFIPALLQCIVLPFCPEPRFLLINRNEENRAKSVLKKLRGT ADVTHDLQEMKEESRQMMREKKVTILELFRSPAYRQPILIAVVLQ SQQLSGINAVFYYSTSIFEKAGVQQPVYATIGSGIVNTAFTVVSLFVV ERAGRRTLHLIGLAGMAGCAILMTIALALLEQLPWMSYLSIVAFGF VAFFEVGPPIPWFIVAEFSQGPRPAAIAVAGFSNWTSNFIVGMC FYVEQLCGPYVFIIFTVLLVLFIFTYFKVPETKGRTFDEIASGFRQG GASQSDKTPE ELFHPLGADSQV	SLC2A1
334	MDPSMGVNSVTISVEGMTCSNVWTIEQQIGKVNGVHHIKVSLEEK NATHYDPKLQTPKTLQEAIDDMGFDAVIHNPDPVLTDTLFLTVA SLTLPWDHIQSTLLKTKGVTDIKIYPQKRTVAVTIIPSVNANQIKELV PELSLDTGTLEKKSGACEDHSMAQAGEVVLKMKVEGMTCHSCTST IEGKIGKLQGVQRIKVSLDNQEATIVYQPHLISVEEMKKQIEAMGFP AFVKKQPKYLKLGAIIDVERLKNTPVKSSEGSQQRSPSYTNDSTATFII DGMHCKSCVSNIESTLSALQYVSSIVVSLNRSIVKYNAASSVTPESL RKAIEAVSPGLYRVISITSEVESTSNPSSSSSLQKIPLNVVSQPLTQETV INIDGMTCSNVQSIEGVISKKPGVKSVIRVSLANSNGTVEYDPLLTSP ETLRGAIEDMGFDATLSDTNEPLVVIAQPSSEMPLLTSTNEFYTKGM TPVQD KEEGKNSSKCYIQVTGMTCA SCVANIERNLRREEGIYSILVALMAG KAEVRYNPAVIQPPMIAEFIRELGFGATVIENADEGDGVLELVVRG MTCASCVHKIESSLTKHRGILYCSVALATNKAHIKYDPEIIGPRDIIHT IESLGFEASLVKKDRSASHLDHKREIRQWRSSFLVSLFFCIPVMGLMI	ATP7A

	<p>YMMVMDHHFATLHHNQNSKEEMINLHSSMFLERQILPGLSVMNL LSFLLC VPVQFFGGWYFYIQAAYKALKHKTANMDVLIVLATTIAFAYSLIILLV AMYERAKVNPITFFDTPPMLFVFIALGRWLEHIAKGKTSEALAKLIS LQATEATIVTLDSNILLSEEQVDVELVQRGDIIKVVPGGKFPVDGR VIEGHSMVDESLITGEAMPVAKKPGSTVIAGSINQNGSLLICATHVG ADTTLSQIVKLVEEAQTSKAPIQQFADKLSGYFVPFIVFVSIATLLVW IVIG FLNFEIVETYFPGYNRSISRTEIIRFAFQASITVLCIACPCSLGLATPT AVMVGTVGAQNGILIKGGEPLMAHKVKVVVFDKTGTITHGTPV VNQVKVLTESNRISHHKILAIVGTAESNSEHPLGTAITKYCKQELDTE TLGTCIDFQVVPGCCISCKVTNIEGLLHKNNWNIEDNNIKNASLVQI DASNEQSSTSSSMIIDAQISNALNAQQYKVLIGNREWMIRNGLVINN DVN DFMTEHERKGRTAVLVAVDDELCLIAIADTVKPEAELAIHILKSMG LEVVLMTGDNSTARSIASQVGITKVFAEVLPSHKVAKVKQLQEEG KRVAMVGDGINDSPALAMANVGIAIGTGTDAIEAADVVLIRNDLL DVVASIDLSRKTVKRIRINFVFALIYNLVGPIAAGVFMPIGLVLQPW MGSAAMAASSVSVVLSSLFLKLYRKPTYESYELPARSQIGQKSPSEI SVHVGIDDTSRNSPKLGLLDRIVNYSRASINLLSDKRSLSNVVTSEP DKHSLLVGDFREDDDTAL</p>	
335	<p>MMRFMLLFSRQGLRLQKWYLATSDKERKKMVRELMQVVLARKP KMCSFLEWRDLKVYKRYASLYFCCAIEGQDNELITLELIHRYVEL LDKYFGSVCELDIIFNEKAYFILDEFLMGGDVQDTSKKSVLKAIEQ ADLLQEEDESPRSVLEEMGLA</p>	AP1S1
336	<p>MKILILGIFLFCSTPAWAKEKHYYIGIETTWDYASDHGEKKLISVD TEHSNIYLQNGPDRIGRLYKKALYLQYTDDETRTTIEKPVWLGLGPI IKAETGDKVYVHLKNLASRPYTFHSHGITYYKEHEGAIYPDNTTDF QRADDKVYPGEQYTYMLLATEEQSPGEGDGNVCVTRIIYHSHIDAPK DIASGLIGPLIICKKDSLDKEKEKHIDREFVVMFSSVDENFSWYLED NIKTYC SEPEKVDKDNEDFQESNRMYSVNGYTFGSLPGLSMCAEDRVKWYL FGMGNEVDVHAFFHQAALTNKNYRIDTINLFPATLFDAYMVAQN PGEWMLSCQNLNHLKAGLQAFFQVQECNKSSSKDNIRGKHVRHYY IAAEEIHWNYAPSGIDIFTKENLTAPGSDSAVFFEQGTTRIGGSYKKL VYREYTDASFTNRKERGPREEHLGILGPVIWAEVGDITRVTFHNKG AYPLSIEPIGVRFNKNNEGTYYSNPNPQSRVPPSASHVAPTETFTY EWTVPKEVGPTNADPVCLAKMY SAVDPTKDIFTGLIGPMKICKKGSLSHANGRQKDVDKEFYLFPTVFDE NESLLEDNIRMFTTAPDQVDKEDEDFQESNKMMSMNGFMYGNQP GLTMCKGDSVVWYLFSAAGNEADVHGIFSGNTYLWRGERRDAN LFPQTSLLHMPDTEGTFNVECLTDDHYTGGMKQKYTVNQCRRQ SEDSTFYLGERTYYIAAVEVEWDYSPQREWEKELHHLQEONVSNF LDKGEFYIGSKYKKVVYRQYTDSTFRVPVERKAEHLGILGPQLH ADVGDVKVIFKNMATRPYSIHAHGVQTESSTVPTLPGETLTYVW KIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCRRPYLK VFNPRRKLEFALLFLVFDENESWYLDNINIKTYSHPKVNKDDEEFI ESNKMHAINGRMFGNLQGLTMHVGDEVNWYLMGMGNEIDLHTV HFHGHFSFYKHGRGVYSSDVFDFPGTYQTLEMFPRTPGIWLLHCHV TDHIHAGMETTYTVLQNEDETKSG</p>	CP
337	<p>MSPTISHKDSSRRPGNFHSLDMKSGPLPPGGWDDSHLDSAGRE GDREALLGDTGTGDFLKAPQSFRALSSILLLLFLYVLQGIPLGLAGS IPLILQSKNVSYTDQAFFSFVFWPFSKLLWAPLVDAVYVKNFGRRK</p>	SLC33A1

	SWLVPTQYILGLFMIYLSQVDRLLGNTDDRTDPDIALTVAFFLFEF LAATQDIAVDGWALTMLSRENVGYASTCNSVGQTAGYFLGNVLF ALESADFCNKYLRFQPPRGIVTSLDFLFFWGTVFLITTTLVALLKK ENEVSVVKEETQGITDTYKL LFAIHKMPAVLTFCLLILTAKIGFSAADAVTGLKLVEEGVPKEHLALL AVPMVPLQIILPLIISKYTAGPQPLNTFYKAMPYRLLLGLEYALLVW WTPKVEHQGGFPIYYYIVVLLSYALHQVTVYSMYVSIMAFNAKVS DPLIGGTMYMTLLNTVSNLGGNWPSTVALWLVDPLTVKECVGASNQ NCRTPD AVELCKKLGGSCVTALDGYVYESIICVFIGFGWWFFLGPKF KKLQDEGSSSWKCKRNN	
338	MSAVCGGAARMRLRTPGRHGYAAEFSPYLPGRACATAQHYGIAGC GTLILLDPDEAGRLRFRSFDWNDGLFDVTWSENNEHVLITCSGDGSL QLWDTAKAAGPLQVYKEHAQEVYSVDWSQTRGEQLVVSWSWDQT VKLWDPTVGKSLCTFRGHESIYSTIWSHIPGCFASASGDQTLRIWD VKAAGVRIVIPAHQAEILSCDWCKYNENLLVTGAVDCSLRGWDLR NVRQPVFELLGHTYAIRRVKFSPPHASFVSLASCSYDFTVRFWNFSKPD SLLETVEHHTEFTCGLDFSLQSPQVADCSWDETIKIYDPACTIPA	PEX7
339	MEQLRAAARLQIVLGHGRPSAGAVVAHPTSGTISSASFHPQQFQY TLDNNVLTLEQRKFYEENGFLVIKNLVPDADIQRFRNEFEKICRKEV KPLGLTVMRDVTISKSEYAPSEKMITKVQDFQEDKELFRYCTLPEIL KYVECTGPNIMAMHTMLINKPPDSGKKTSRHPLHQDLHYFPFRPS DLIVCAWTAMEHISRNNGLVVLPGTHKGSLKPHDYPKWEAGGVNK MFHGIQDYEENKARVHLVMEKGDVTFHPLLIHGSGQNKQTQGRK AISCHFASADCHYIDVKGTSQENIEKEVVGIAHKFFGAENSVNLKDI WMFRARLVKGERTNL	PHYH
340	MAEAAAAAGGTGLGAGASYGSAADRDRDPDPDRAGRRLRVLSGH LLGRPREALSTNECKARRAASAATAAPTATPAAQESGTIPKKRQEV MKWNGWGYNDKFIENKKGQIELTGKRYPLSGMGLPTFKEWIQNT LGVNVEHKTTSKASLNPSDTPPSVVNEDFLHDLKETNISYSQEADDR VFRAHGHCLHEIFLLREGMFERIPDIVLWPTCHDDVVKIVNLACKY NLCIPIGGGTSVSYGLMCPADETRTHSLDTSQMNRLWVDENNLTA HVEAGITGQELERQLKESGYCTGH EPDSLEFSTVGGWVSTRASGMKKNIYGNIEDLVVHIKMTVPRGIEK SCQGPRMSTGPDIIHFMGSEGTGLGVITEATIKIRPVPEYQKYGSVAF PNFEQGVACLREIAKQRCAPASIRLMDNKQFQFGHALKPQVSSIFTS FLDGLKKFYITKFKGFDPNQLSVATLLFEGDREKVLQHEKQVYDIA AKFGGLAAGEDNGQRGYLLTYVIAIYRDLALEYYVLGESFETSAPW DRVVDLCRNVKERITRECKEKG VQFAPFSTCRVTQTYDAGACIYFY FAFNRYGISDPLTVFEQTEAAAREEILANGGSLSHHHGVGKLKQW LKESISDVGFGLKSVKEYVDPNNIFGNRNLL	AGPS
341	MESSSSNSYFSVGPTSPSAVVLLYSKELKKWDEFEDILEERRHVSD LKFAMKCYTPLVYKGITPCKPIDIKCSVLNSEEIHYVIKLSKESLQS VDVLREEVSEILDEMSEKLRRLGAIRFCAFTLSKVFKQIFSKVCVNEE GIQKLQRAIQEHPPVLLPSHRSYIDFLMLSFLLYNYDLPVPVIAAGM DFLGMKMGVCELLRMSGAFFMRRTFGGNKLYWAVFSEYVKTMLRN GYAPVEFFLEGTRSRSAKTLTPKFGLLNIVMEPFFKREVFDTYLVPI ISYDKILEETLYVYELLGVPKPKESTTGLLKARKILSENFSGSIHVYFG DPVSLRSLAAGRMSRSSYNLVPRYIPQKQSEDMHAFVTEVAYKMEL LQIENMVLSPWTLIVAVLLQNRPSMDFDALVEKTLWLKGLTQAFGG FLIWPDKNPAEEVVPASILLHSNIASLVKDQVILKVDSDGSEVVDGL MLQHITLLMCSAYRNQLLNIFVRPSLVAVLQMTPGFRKEDVYSCF RFLRDVFADEFIFLPGNTLKD FEEGCYLLCKSEAIQVTTKDILVTEKG	GNPAT

	NTVLEFLVGLFKPFVESYQIICKYLLSEEDHFSEEQYLA AVRKFTSQ LLDQGTSQCYDVLSSDVQKNALAACVRLGVVEKKKINNNCIFNVN EPATTKLEEMLGCKTPIGKPATAKL	
342	MPVLSRPRPWRGNTLKRTAVLLALAAAYGAHKVYPLVRQCLAPARG LQAPAGEPTQEASGVAAAKAGMNRVFLQRLLWLLRLLFPRVLCRE TGLLALHSAALVSRTFLSVYVARLDGRLARCIVRKDPRAFGWQLLQ WLLIALPATFVNSAIRYLEGQLALSFRSRLVAHA YRLYFSQQTYYRV SNMDGRLRNPQSLTEDVV AFAASVAHL YSNLTKPLLDVAVTSYT LLRAARSRGAGTAWPSAIAGLVVFLTANVLRAFSPKFGELVAEEAR RKGELRYMH SRVVANSEEIAFYGGHEVELALLQRSYQDLASQINLIL LERLWYVMLEQFLMKYVWSASGLLMVAVPIITATGYSESDAEAVK KAALEKKEEELVSETEAFTIARNLLTAAADAIERIMSSYKEVELA GYTARVHEMFQVFEDVQRCHFKRPRELEDAQAGSGTIGRSGVRVE GPLKIRGQVVDVEQGHICENIPIVTPSGEVVVASLNIRVEEGMHLLITG PNGCGKSSLFRLGGLWPTYGGVLYKPPQRMFYIPQRPYMSVGS LDQVIYPDSVEDMQRKGYSEQDLEAILDVVHLHHILQREGGWEAM CD WKDVLSGGEKQRIGMARMFYHRPKYALLDECTSAVSIDVEGKIFQ AAKDAGIALLSITHRPSLWKYHHTLLQFDGEGGWKFEKLDSAARLS LTEEKQRLEQQLAGIPKMQRRLQELCQILGEAVAPAHVPAPSPQGP GGLQGAST	ABCD1
343	MNPDLRRERDSASFNPPELLTHILDGSPEKTRRRREIENMILNDPDFQ HEDLNFLTRSQR YEVA VRKSAIMVKKMREFGIADPDEIMWFKKLHL VNFVEPVGLNYSMFIPTLLNQGTTAQKEKWLLSSKGLQIIGTYAQTE MGHGHTRLGLETTATYDPETQEFILNSPTVTSIKWWPGGLGKTSNH AIVLAQLITKGKCYGLHAFIVPIREIGTHKPLPGITVGDIGPKFGYDEI DNGYLKMDNHRIPRENMLMKYAQVKPDGTYYVKPLSNKLT YGTMV FVRSFLVGEAARALSKACTIAIRYSAVRHQSEIKPGEPEPQILDFQTQ QYKLFPLLATAYAFQFVGAYMKETYHRINEGIGQGDLSELPELHAL TAGLKFTSWTANTGIEACRMACGGHGYSHCSGLPNIYVNFPSCT FEGENTVMMLQATARFLMKSYDQVHSGKLVCGMVSYLNDLPSQRIQ PQQVAVWPTMVDINSPELSTEAYKLRAARLVEIAAKNLQKEVIHRK SKEVAWNLTSDVLVRASEAHCHYVVVKLFSEKLLKIQDKAIQAVLR SLCLLYSLYGISQNAGDFLQGSIMTEPQITQVNQRVKELLTLIRSDAV ALVDAFDFQDVTLGSVLGRYDGNVYENLFEWAKNSPLNKA EVHES YKHLKSLQSKL	ACOX1
344	MWGS DRLAGAGGGGA AVTV AFTNARD CFLHLPRRLVAQLHLLQN QAIEVVWSHQPAFLSWVEGRHFSDQGENVAEINRQVGQKLGLSNG GQVFLKPCSHV VSCQQVEVEPLSADDWEILELHAVSLEQHLLDQIRI VFPKAIFPVVDQQTYIFIQIVALIPAASYGRLETDTKLLIQPKTRRA KENTFSKADA EYKKLHSYGRDQKGMMKELQTKQLQSNTVGITESN ENESEIPVDSSSVASLWTMIGSIFSQSEKKQETSWGLTEINAFKNMQ SKVVPLDNIFRVCKSQPPSIYNASATSVFHKHCAIHVFPWDQEYFDV EPSFTVTYGKLVKLLSPKQQSKTKQNVLSPEKEKQMSEPLDQKKI RSDHNEEDEKACVLQVVWNGLEELNNAIKYTKNVEVLHLGKVWIP DDLKRRLNIEMH AVVRITPVEVTPKIPRSLKLQPRENLPKDISEEDIK TVFYSWLQQSTTTMLPLVISEEEFIKLETKDGLKEFSLSIVHSWEKEK DKNIFLLSPNLLQKTTIQVLLDPMVKEEN SEEIDFILPFLKLSSLGGVNSLGVSSLEHITHSLLGRPLSRQLMSLVAG LRNGALLLTGGKSGKSTLAKAICKEAFDKLDAHVERVDCKALRG KRLENIQKTLEVAFSEAVWMQPSVLLDDLDLIAGLPAPVEHEHSP DAVQSQR LAHALNDMIKEFISMGSLVALIATSQSQQSLHPLLVSAAQ GHIFQCVQHIQPPNQEQRC EILCNVKNKLD CDINKFTDLDLQHVAK	PEX1

	ETGGFVARDFTVLVDRAIHSRLSRQSISTREKLVLTTLDFQKALRGF LPASLRSVNLHKPRDLGWDKIGGLHEVRQILMDTIQLPAKYPELFA NLPIRQRTGILLYGPPGTGKTLLAGVIARESRMNFISVKGPELLSKYI GASEQAVRDIFIRAQAAKPCILFFDEFESIAPRRGHDNTGVTDRVFN QLLTQLDGVEGLQGVYVLAATSRPDLIDPALLRPGRLDKCVYCPPP DQVSRLEILNVLSDSLPLADDVDLQHVASVTDSTGADLKALLYNA QLEALHGMMLSSGLQDGSSSSSDSLSSMVFLNHSSGSDDSAGDG ECGLDQSLVSLEMSEILPDESKFNMRYLYFGSSYESELGNGTSSDLS SQCLSAPSSMTQDLPGVPGKDQLFSQPPVLRASQEGCQELTQEQR DQLRADISIIKGRYRSQSGEDES MNQPGPIKTRLAISQSHLMTALGHT RPSISEDWKNFAELYESFQNPKRKRNQSGTMFRPGQKVTLA	
345	MASRKENAKSANRVLRISQLDALELNKALEQLVWSQFTQCFHGFKP GLLARFEPEVKACLWVFLWRFTIYSKNATVGQSVLNKYKNDFSPN LRYQPPSKNQKIWYAVCTIGGRWLEERCYDLFRNHHLASFGKVQK CVNFVIGLLKLGG LINFLIFLQRGKFATLTERLLGIHSVFCKPQNICEV GF EYMNRELLWHGFAEFLIFLLPLINVQKLKAKLSSWCIPLTGAPNS DNTLATDISIKGECALCGEWPTMPHTIGCEHIFCYFCAKSSFLFDVYFTC PKCGTEVHSLQPLKSGIEMSEVNAL	PEX2
346	MLRSVWNFLKRHKKKCIFLGTVLGGVYILGKYGQKKIREIQEREA EYIAQARRQYHFESNQRTCNMVLSMLPTLREALMQQLNSESALTAL LKNRPSNKLIEWEDLKHSFTRSTVAVYSTCMLVLLRVQLNIIGGYI YLDNAAVGKNGTTILAPPDVQQQYLSSIQHLLGDGLTELITVIKQAV QKVLGVSLSKHSLSLLDLEQKLKEIRNLVEQHKSSSWINKDGSKPLL CHYMMPDEETPLAVQACGLSPRDITIKLLNETRDMLES PDFSTVLN TCLNRGFSRLLDNMAEFFRPTEQDLQHGNMNSLSSVSLPLAKIPIV NGQIHSVCSETPSHFVQDLLTMEQVKDFAANVYEAFASTPQQLEK	PEX3
347	MAMRELVEAECGGANPLMKLAGHFTQDKALRQEGLRPGWPWPPGA PASEAASKPLGVASEDELVAEFLQDQNAPLVSRAPQTFKMDDLAE MQQIEQSNFRQAPQRAPGVADLALSENWAQEFLAAGDAVDVTQD MYNETDWSQEFISEVTDPLSVSPARWAE EYLEQSEEKLWLGEPEGTA TDRWYDEYHPEEDLQHTASDFVAKVDDPKLANSEFLKFVRQIGEG QVSLESAGSGRAQAEQWAAEFIQQGTSDAWVDQFTRPVNTSAL DMEFERAKSAIESDVDFWDKLQAELEEMAKRDAEAHPWLSDYDDL TSATYDKGYQFEEENPLRDHPQPFEGLRRLQEGDLPNAVLLFEAA VQQDPKHMEAWQYLGTTQAENEQELLAISALRRCLELKPDNQTAL MALAVSFTNESLQRQACETLRDWLRYTPAY AHLVTPAEEGAGGAG LGPSKRILGSLSDSLFLEVKEFLAAVRLDPTSIDPDVQCGLGVLFN LSGEYDKAVDCFTAALSVRPN DYLLWNKLGATLANGNQSEEAVAA YRRALELQPGYIRSRYNLGISCINLGAHREAVEHFLEALNMQRKSRG PRGEGGAMSENIWSTLR LALSMLGQSDAYGAADARDLSTLLTMFG LPQ	PEX5
348	MALAVLRVLEPFPPTETPPLAVLLPPGGPWPAELGLVLALRPAGESP AGPALLVAALEGPDAGTEEQGPGPPQLLVS RALLRLLALGSGAWVR ARAVRRPPALGWALLGTS LGPGLGPRVGP LLVRRGETLPVPGPRVL ETRPALQGLLGP GTRLAVTEL RGRARLC PESGDSSRPPPPPVSSFA VSGTVRRLQGV LGGTGDSLGVSR SCLRGLGLFQGEVWVVAQARES SNTSQPHLARVQVLEPRWDLS DRLGPGSGPLGEPLADGLALVPATL AFNLGCDPLEMGELRIQRYLEGS IAPEDKGSCSLLPGPPFARELHIEIVSSPHYSTNGNYDGVLYRHFQIPR VVQEGDVLCVPTIGQVEILEGSPEKLPRWREMFFKVKKTVGEAPDG PASAYLADTTHTSLYMGSTLSVPWLPSEESTLWSSLSPPGLEALV	PEX6

	<p>SELCAVLKPRQLQPGGALLTGTSSVLLRGPPGCGKTTVVAAACSHLG LHLLKVPCSSSLCAESSGAVETKLQAFSRARRCRPAVLLLTAVDLLG RDRDGLGEDARVMAVLRHLLLNEDPLNSCPPLMVVATTSTRAQDLP ADVQTAFPHLEVPALSEGQRLSILRALTAHLPLGQEVNLAQLARR CAGFVVGDLYALLTHSSRAACTRIKNSGLAGGLTEEDEGELCAAGF PLLAEDFGQALEQLQTAHSQAVGAPKIPSVSWHDVGGGLQEVKKEIL ETIQLPLEHPELLSLGLRRSGLLLHGPPGTGKTLLAKAVATECSLTFL SVKGPELINMYVGQSEENVREVFARARAAAPCIIFFDELDSLAPSRG RSGDSGGVMDRVVSQLLAELDGLHSTQ DVFVIGATNRPDLLDPALLRPGRFDKLVFVGANEDRASQLRVLSAIT RKFKLEPSVSLVNVLDCCPPQLTGADLYSLCSDAMTAALKRRVHDL EEGLEPGSSALMLTMEDLLQAAARLQPSVSEQELLRYKRIQRKFAA C</p>	
349	<p>MAPAAASPPEVIRAAQKDEYYRGGLRSAAGGALHSLAGARKWLE WRKEVELLSDVAYFGLTTLAGYQTLGEEYVSIIQVDPRIHVPSSLR RGVVLTLHAVLPYLLDKALLPLEQELQADPDSGRPLQGS LGPGGRG CSGARRWMRHHTATLTEQQRALLRAVFLVRQGLACLQRLHVAW FYIHGVFYHLAKRLTGITYLRVRS L PGEDLRARVSYRLLGVISLLHL VLSMGLQLYGFRQQRARKEWRLHRGLSHRRASLEERAVSRNPLC TLCLEERRHPTATPCGHLFCWECITAW CSSKAECPLCREKFPPQKLIYLRHYR</p>	PEX10
350	<p>MAEHGAHFTAASVADDQPSIFEVVAQDSLMTAVRPALQHVVKVLA ESNPTHYGF LWRWFDEIFTLLDLLLQQHYLSRTSASFSENFYGLKRI VMGDTHKSQRLASAGLPKQQLWKSIMFLVLLPYLKVKLEKLVSSL REEDEYSIHPPSSRWKR FYRAFLAAYPFVNMAWEGWFLVQQLRYIL GKAQHHSPLRLAGVQLGRLTVQDIQALEHKPAKASMMQPARSV SEKINSALKKAVGGVALSLSTGLSVGVFFLQFLDWWYSSSENQETIKS LTALPTPPPVHLDYNSDSPLLPKMKTVCPLCRKTRVNDTVLATSG YVFCYRCVFHYVRSHQACPITGYPTVQHILKLYSPEN</p>	PEX12
351	<p>MASQPPPPPKPWETRRIPGAGPGPGPTFQSADLGPTLMTRPGQPA LTRVPPPIPRPSQQTGSSSVNTFRPAYSSFSFGYAYGNSFYGGYSP YSYGYNGLYGNRLRVDDLPPSRFVQQAEESSRGAFQSIESIVHAFAS VSMMDATFSAVYNSFRAVL DV ANHFSRLKIHF TKVFS AFALVRTI RYLYRRLQRM LGLRRGSENE DLWAESEGTVACLGAEDRAATS AKS WPIFLFFAVILGGPYLIWKLLSTHSDEVTD SINWASGEDDHVVARAE YDFAAVSEEEISFRAGDMLNLALKEQQPKVRGWLLASLDGQTTGLI PANYVKILGKRKGRKTVES SKVSKQQQSFTNPTLT KGATVADSLDE QEAAFESVFVETNKVPVAPDSIGKDGEKQDL</p>	PEX13
352	<p>MASSEQAEQPSQPSSTPGSENVLPREPLIATAVKFLQNSRVRSPLAT RRAFLKKKGLTDEEIDMAFQQSGTAADEPSSLGPATQVVPVQPPHLI SQPYSPAGSRWRDY GALAIIMAGIAFGFHQLYKKYLLPLILGGREDR KQLERMEAGLSELSGSVAQTVTQLQTTLASVQELLIQQQKIQELA HELAAAKATTSTNWILESQNINELKSEINSLKGLLLNRRQFPSPSPAP KIPSWQIPVKSPSPSPAAVNHHSSSDISPVSNESTSSSPGKEGHSPEG STVTYHLLGPQEEGEGVVDVKGQVRMEVQGEEEKREDKEDEEDEE DDD VSHVDEEDCLGVQREDRRGGDGQINEQVEKLRRPEGASNESE RD</p>	PEX14
353	<p>MEKLRLGLRLYQEYVTRHPAATAQLETA VRGFSYLLAGRFADSHE LSELVYSASNLLVLLNDGILRKELRKKLPVSLSQKLLTWLSVLECV EVFMEMGA AKVWGEVGRWL VIALVQLAKAVLRMLLLLWFKAGL QTSPPIVPLDRETQAQPPDGDHSPGNHEQSYVGKRSNRVVRTLQNT PSLHSRHWGAPQQREGRQQQHHEELSATPTPLGLQETIAEFLYIARP</p>	PEX16

	LLHLLSLGLWGQRSWKPWLLAGVVDVTSLSLLSDRKGLTRRERRE LRRRTILLYYLLRSPFYDRFSEARIL FLLQLLADHVPGLVTRPLMDYLPTWQKIYFYSWG	
354	MAAAEEGCSVGAEADRELEELLESALDDFDKAKPSPAPPSTTTAPD ASGPQKRSPGDTAKDALFASQEKFFQELFDSELASQATAEFKAMK ELAAEEPHLVEQFQKLSEAAGRVGSDMTSQQEFTSCLKETLSGLAK NATDLQNSSMSEEELTKAMEGLGMDEGDGEGNILPIMQSIMQNLLS KDVLYPSLKEITEKYPEWLQSHRESLPPEQFEKYQEHSVMCKICEQ FEAETPTDSETTQKARFEMVLDLMMQQLQDLGHPPKELAGEMPPGLN FDLDALNLSGPPGASGEQCLIM	PEX19
355	MKSDSSTSAAPLRGLGGPLRSSEPVRAVPAVPAVDLLEEADLLV VHLDFAALETCEAWQSLANHAVAEPPAGTSLEVVKCSLCVVGIQ ALAEMDRWQEVLSWVLQYYQVPEKLPPKVLELCILLYSKMQEPGA VLDVVGAWLQDPANQNLPEYGALAEFHVQRVLLPLGCLSEAEELV VGSAAFGEERRLDVLQAIHTARQQQKQEHSGSEEAQKPNLEGSVSH KFLSLPMLVRQLWDSAVSHFFSLPFFKSSLLAALILCLLVVRFPASP SSLHFLYKLAQLFRWIRKAASFRLYQ LRIRD	PEX26
356	MALQGISVVELSGLAPGPFCAVLADFGARVVRVDRPGSRYDVSR LGRGKRSLVLDLKQPRGA AVLRRLLCKRSDVLEPFRRGVMEKLQL GPEILQRENPRLIYARLSGFGQSGSFCLAGHDINYALSGVLSKIGR SGENPYAPLNLADFAAGGLMCALGIIMALFDRTRTGKGQVIDANM VEGTAYLSSFLWKTQKLSLWEAPRGQNMLDGGAPFYTTYRTADGE FMAVGAIEPQFYELLIKGLGLKSDLPNQMSMDDWPMMKKKFADV FAEKTKAEWQCIFDGTDACVTPVLTFFEEVVHHDHNGERGSFITSEE QDVSPRPAPLLLNTPAIPSFKRDPFGEHTEEILEEFGFSREEIYQLNSD KIIESNKVKASL	AMACR
357	MAQTPAFDKPKVELHVHLDGSIKPETILYYGRRRGIALPANTAEGLL NVIGMDKPLTLPDFLAKFDYYMPAIAAGCREAIKRIAYEFVEMKAKE GVVYVEVRYSPHLLANSKVEPIPNWQAEGDLTPDEVVALVGGLQ EGERDFGVKARSILCCMRHQPNWSPKVVELCKKYQQQTVVAIDLA GDETIPGSSLLPGHVQAYQEAVKSGIHRTVHAGEVGSAEVVKEAVD ILKTERLGHGYHTLEDQALYNRLRQENMHFEICPWSSYLTGAWKPD TEHAVIRLKNQDQANYSLNTDDPLIF KSTLDTDYQMTKRDGMGFTEEEFKRLNINAAKSSFLPEDEKRELLDL LYKAYGMPPSASAGQNL	ADA
358	MAAGGDHGSPPDSYRSPLASRYASPEMCFVFSDRYKFRTWRQLWL WLAEAEQTLGLPITDEQIQEMKSNLENIDFKMAAEEEEKRLRHDVMA HVHTFGHCCPKAAGIIHLGATSCYVGDNTDLIILRNALDLLLPKLAR VISRLADFAKERASLPTLGFTHFQPAQLTTVGKRCCLWIQDLCMDL QNLKRVRDDLRFGRGVKGTGTQASFLQLFEGDDHKVEQLDKMVTE KAGFKRAFIITGQTYTRKVDIEVLSVLASLGASVHKICTDIRLLANLK EMEPEFEKQQIGSSAMPYKRNPMSERCCSLARHMLTLVMDPLQT ASVQWFERTLDDSANRRICLAEAFLTADTILNTLQNISEGLVVYPKV IERRIRQELPFMATENIIMAMVKAGGSRQDCHEKIRVLSQQAASVVK QEGGDNDLIERIQVDAYFSPHSQLDHLLDPSSFTGRASQQVQRFLEE EVYPLLKPYESVMKVKAELCL	ADSL
359	MNVRIFYSVSQSPHSLLSLLFYCAILESRSISATMPLFKLPAEEKQIDD AMRNFAEKVFASEVKDEGGRQEISPFVDICPISHHEMQAHIFHLE TLSTSTEARRKKRFQGRKTVNLSIPLSETSSSTKLSHIDEYISSPTYQT VPDFQRVQITGDYASGVTVEDFEIVCKGLYRALCIREKYMQKSFQR FPKTPSKYLRNIDGEAWVANESFYPVFTPPVKKGEDPFRTDNLPENL	AMPD1

	GYHLKMKDGVVYVYPNEAAVSKDEPKPLPYPNLDTFLDDMNFLLA LIAQGPVKTYTHRRLLKFLSSKFQVHQMLNEMDELKELKNNPHRDF YNCRKVDTHIHAAACMNQKHLLRFIKKSYQIDADRVVYSTKEKNL TLKELFAKLKMHYPYDLTVDSLVDHAGRQTFQRFDKFNDKYNPVGA SELRDLYLKTDNYINGEYFATIIKEVGADLVEAKYQHAEPRLSIYGR SPDEWSKLSSWFVCNRIHCPNMTWMIQVPRIYDVFRSKNFLPHFGK MLENIFMPVFEATINPQADPELSVFLKHIT GFDSVDDESKHSGHMFSSKSPKPQEWTLKKNPSYTYYYAYMYANI MVLNSLRKERGMNTFLFRPHCGEAGALTHLMTAFMIADDISHGLNL KKSPVLQYLFFLAQIPIAMSPLSNNSLFLEYAKNPFLDFLQKGLMISL STDDPMQFHFTKEPLMEEYAIAAQVFKLSTCDMCEVARNSVLQCGI SHEEKVKFLGDNYLEEGPAGNDIRRTNVAQIRMAYRYETWCYELN LIAEGLKSTE	
360	MATEGMILTNHDHQIRVGVLTVDSCFRNLAEDRSGINLKDLVQDP SLLGGTISAYKIVPDEIEEIKETLIDWCDEKELNLILTTGGTGFAPRDV TPEATKEVIEREAPGMALAMLMGSLNVTPLGMLSRPVCGRGKTLII NLPGSKKGSQECFQFILPALPHAIDLRLDAIVKVEVHDELEDLPSP PPLSPPTTSPHKQTEDKGVCCEEEEEEEKKDSGVASTEDSSSSHITAA AIAAKIPDSIISRGVQVLPRTASLSTTPSESPRAQATSRLSTASCPTP KVQSRCSSKENILRASHSAVDITKVARRHRMSPPFLTSMDFKAFITVL EMTPVLGTEIHNRYRDMGRVLAQDVYAKDNLPPFPASVKDGYAVR AADGPGDRFIIGESQAGEQPTQTVMPGQVMRVTGAPIPCGADAVV QVEDTELIRESDDGTEELEVRILVQARPGQDIRPIGHDIKRGECVLAK GTHMGPS EIGLLATVGVTEVEVNKFPVVAVMSTGNELLNPEDDLLPGKIRDSN RSTLLATIQEHEGYPTINLGIVGDNPDLLNALNEGISRADVIITSGGVS MGEKDYLKQVLDIDLHAQIHFGRVFMKPGLPPTTFATLIDIDGVRKIIF ALPGNPVSAVVTCLNFVVPALRKMVGILDPRTIHKARLSCDVKLD RPEYHRCILTWHHQEPLPWAQSTGNQMSSRLMSMRANGLMLLP KTEQYVELHKGGEVVDVMVIGRL	GPHN
361	MAGAAAESGRELWTFAGSRDPSAPRLAYGYGPGSLRELRAREFSRL AGTVYLDHAGATLFSQSQLESFTSDLMENTYGNPHSQNISKLTHD TVEQVRYRILAHFHTTAEDYTVIFTAGSTAALKLVAEAFPWVSQGP ESSGSRFCYLTDSHTSVVGMNRNVTMAINVISTPVRPEDLWSAEERSA SASNPDQQLPHLFCYPAQSNFSGVRYPLSWIEEVKSGRLHPVSTPGK WVFLLDAAASYVSTSPLDLSAHQADFVPISFYKIFGFPTGLGALLVHN RAAPLLRKTYFGGGTASAYLAGEDFYIPRQSV AQRFEDGTISFLDVI ALKHGFDTLERLTGGMENIKQHTFTLAQYTYVALSSLQYPNGAPVV RIYSDSEFSSPEVQGPINFNVLDDKGNIIGYSQVDKMASLYNIHLRT GCFCNTGACQRHLGISNEMVRKHFAQGHVCGDNMDLIDGQPTGSV RISFGYMSTLDDVQAFLRFIIDTRLHSSGDWPVPQAHADTGETGAPS ADSQADVIPA VMGRRLSPQEDALTGSRVWNNSSSTVNAPVAPPV CDVARTQPTPSEKAAGVLEGALGPHVVTNLLYPIKSCAAFEVTRW PVGNGQLLYDRSWMVVNHNHGVCLSQKQEPRLCLIQPFIDLRQRIMV IKAKGMEPIEVPLEENSERTQIRQSRVCADRVSTYDCGEKISSWLSTF FGRPCHLIKQSSNSQRNAKKKHGKDQLPGTMATLSLVNEAQYLLIN TSSILELHRQLNTSDENGKEELFSLKDLRLFRANIIINGKRAFEEK WDEISIGSLRFQVLGPCHRCQMIDQQTGQRNQHVFKLSESRETK VNFGMYLMHASLDLSSPCFLSVGSQVLPVLKENVEGHDLPASEKHQ DVTS	MOCOS
362	MAARPLSRMLRRLLRSSARSCSSGAPVTQPCPGESARAASEEVSRRL QFLREHAAPFSAFLTDSFGRQHSYLRISLTKCNLRCQYCMPEEGVP LTPKANLLTTEEILTLARLFVKEGIDKIRLTGGEPLIRPDVVDIVAQLQ	MOCS1

	<p>RLEGLRTIGVTTNGINLARLLPQLQKAGLSAINISLDTLVPAKFEFIVR RKGFHKVMEGIHKAIELGYNPVKVNCVVMRGLNEDELDDFAALTE GLP</p> <p>LDVRFIEYMPFDGNKWNFKKMVSYPEMLDTRVQQWPELEKVPEEE SSTAKAFKIPGFQGGQISFITSMEHFCGTCNRLRITADGNLKVCLFGN SEVSLRDHLRAGASEQELLRIIGAAGVGRKKRQHAGMFSISQMKNRP MILIELFLMFPNSPPANPSIFSWDPLHVQGLRPRMSFSSQVATLWKG CRVPQTPPLAQQLGSGSFQRHYTSRADSDANSKCLSPGSWASAAP SGPQLTSEQLTHVDSEGRAAMVDVGRKPDTERVAVASAVVLLGPV AFKLVQQNQLKKGDALVVAQLAG VQAAKVTSQLIPLCHHVVALSHIQVQLELDSTRHAVKIQASCRARGPT GVEMEALTSAAVAALTYDMCKAVSRDIVLEEIKLISKTTGGQRGDF HRA</p>	
363	<p>MENGYTYEDYKNTAEWLLSHTKHRPQVAIICGSGLGGLTDKLTQA QIFDYGEIPNFRSTVPGHAGRLVFGFLNGRACVMMQGRFHMIEG YPLWKVTFPVRVHLLGVDTLVVTNAAGGLNPKFEVGDIMLIRDHI NLPGFSGQNPLRGPNDERFGDRFPAMSDAYDRTMRQALSTWQOM GEQRELQEGTYVMVAGPSFETVAECRVLQKLGAAGVGMSTVPEVI VARHCGLRVFGSLITNKVIMDYESLEKANHEEVLAAGKQAAQKLE QFVSILMASIPLPKAS</p>	PNP
364	<p>MTADKLVFFVNGRKKVVEKNADPETTLAYLRRKLGSLGSKLGCGE GGCGACTVMLSKYDRLQNKIVHFSANACLAPICSLHHVAVTTVEGI GSTKTRLHPVQERIAKSHGSQCGFCTPGIVMSMYTLLRNQPEPTMEE IENAFQGNLCRCTGYRPILQGFRTFARDGGCCGGDGNPNCCMNQ KKDHSVSLSPSLFKPEEFTPLDPTQEPIFPPELLRLKDTPRKQLRFEGE RVTWIAQSTLKELLDLKAQHPDAKL VVGNTIEIGIEMKFKNMLFPMI VCPAWIPELNSVEHGPDGISFGAACPLSIVEKTLVDAVAKLPAQKTE VFRGVLEQLRWFAKQVKSASVGGNIITASPIDLNPVFMASGAK LTLVSRGTRRTVQMDHTFFPGYRKTLSPEEILLSIEIPYSREGEYFSA FKQASRREDDIAKVTSGMRVLFKPGTTEVQELALCYGGMANRTISA LKTTFKFKYLTVLQKLGQENLEDKCGKLDPTFASATLLFQKDPPADV QLFQEVPGKQSEEDMVGRPLPHLAADMQASGEAVYCDDIPRYENE LSLRLVTSTRAHAKIKSIDTSEAKKVPGFVCFISADDVPGSNITGICN DETVFADKDVTCVGHIIAGVAVADTPEHTQRAAQGVKITYEELPAIITI EDAIKNNSFYGPPELKIEKGDLLKGFSEADNVVSGEIIYGGQEHFYLE THCTIAVPKGEAGEMELFVSTQNTMKTQSFVAKMLGVPANRIVVR VKRMGGGFGGKETRSTVVSTAVALAAYKTGRPVRCMLDRDEDML ITGGR</p> <p>HPFLARYKVGFMKTGTVV ALEVDHFSNVGNTQDLSQSIMERALFH MDNCYKIPNIRGTGRLCKTNLPSNTAFRGFGGPQGMLIAECWMSEV AVTCGMPAEVRRKNLYKEGDLTHFNQKLEGFTLPRCWEECLASS QYHARKSEVDKFNKENCWKKRGLCIPTKFGISFTVPFLNQAGALLH VYTDGSVLLTHGGTEMGQGLHTKMVQVASRALKIPTSIIYISETST NTVPNTSPTAASVSADLNGQAVYACQTLKRLEPYKKKNPSGSWE DWVTAAYMDTVSLSATGFYRTPNLGYSFETNSGNPFHYFSYGVAC SEVEIDCLTGDHKNLRTDIVMDVGSSLNPAIDIGQVEGAFVQGLGLF TLEELHYSPEGLHTRGPSTYKIPAFGSIPIEFVSLLRDCPNKKAIYA SKAVGEPPLFLAASIFFAIKDAIRAARAQHTGNNVKELFRLDSPATPE KIRNACVDKFTTLCVTGVPENCKPWSVRV</p>	XDH
365	<p>MLLLHRAVVLRLQQACRLKSIPSRICIACSTNDSFQQRPSLTFSGD NSSTQGWRVMGTLLGLGAVLAYQDHCRAAQESTHIYTKEEVSSH TSPETGIWVTLGSEVFDVTEFVDLHPGGPSKMLLAAGGPLEPFWAL</p>	SUOX

	YAVHNQSHVRELLAQYKIGELNPEDKVAPTVETSDPYADDPVRHPA LKVNSQRPFNAEPPPELLTENYITPNPIFFTRNHLVPVNLDPDTYRLH VVGAPGGQSLSLDDHLNFPRIEITVTLQCAGNRRSEMTQVKEVK GLEWRTGAISTARWAGARLCDVLAQAGHQLCETEAHVCFEGLDSD PTGTAYGASIPLARAMDPEAEVLLAYEMNGQPLPRDHGFPVRVVP GVVGARHVKWLGRVSVQPEESYSHWQRRDYKGFSPVDWETVDF DSAPSIQELPVQSAITEPRDGETVESGEVTIKGYAWSGGGRAVIRVD VSLDGGTLTWQVAKLDGEEQRPRKAWAWRLWQLKAPVPAGQKEL NIVCKAVDDGYNVQPDVAPIWNLRGVLSNAWHRVHVYVSP	
366	MFHLRTCAAKLRPLTASQTVKTFSQNRPAARTFQQIRCYSAPVAA EPFLSGTSSNYVEEMYCAWLENPKSVHKSWDIFFRNTNAGAPPGTA YQSPPLSRGSLAAVAHAQSLVEAQPNVDKLVEDHLAVQSLIRAYQ IRGHVVAQLDPLGILDADLDSSVPADIISSTDKLGFYGLDESDLDKVF HLPTTTFIGGQESALPLREIIRLEMA YCQHIGVEFMFINDELCQWI RQKFETPGIMQFTNEEKRTLLARLVSTRFEEFLQRKWSSEKRFGL GCEVLIPALKTIIDKSENGVDYVIMGMPHRGRLNVLANVIRKELEQ IFCQFDSKLEAADESGSDVKYHLGMYHRRINRVTDNRITLSLVANP SHLEAADPVVMGKTKAEQFYCGDTEGKKVMSILLHGDAAFAGQGI VYETFHLSDLPSYTTHTGTVHVNNQIGFTTDPRMARSSPYPTDVA RVVNAPIFHVNSDDPEAVMYVCKVAAEWRSTFHKDVVDLV CYR RNGHNEMDEPMFTQPLMYKQIRKQKPV LQKYAELLVSQGVVNQPE YEEEISKYDKICEEAFARSKDEKILHIKHWLDSPWPGFTLDGQPRS MSCPSTGLTEDILTHIGNVASSVPVENFTIHGGLSRILKTRGEMVKNR TVDWALAEYMAFGSLLKEGIHIRLSGQDVERGTFSHRHHVLHDQN VDKRTICPMNHLWPNQAPYTVCNSSLSEYGV LGFELGFAMASPNAL VLWEAQFGDFHNTAQCIIDQFICPGQAKWVRQNGIVLLLPHGMEG MGPEHSSARPERFLQMCNDPDLVPLKEANFDINQLYDCNWVVV NCSTPGNFFHVLRRQILLPFRKPLIIFTPKSLLRHPEARSSFDEMLPGT HFQRVIPEDGPAAQNPENVKRLLFCTGKVYYDLTRERKARDMVGQ VAITRIEQLSPFPDLLLLKEVQKYPNAELAWCQEEHKNQGYDYVK PRLRTTISRAPVWYAGRDPAAAPATGNKKTHLTELRLLDTAFDL DVFKNFS	OGDH
367	MVGYDPKPDGRNNTKFQVAVAGSVSGLVTRALISPFQVIRFQLQ HERLSRSDPSAKYHGILQASRQILQEEGPTAFWKGHVPAQILSIGYG AVQFLSFEMLTTELVRHGSVYDAREFSVHFVCGGLAACMATLTVHP VDVLRTRFAAQGEKPKVYNTLRHAVGTMYRSEGPQVYKGLAPTLI AIFPYAGLQFSCYSSLKHLKYWAIPAEGKKNNENLQNLLCGSGAGVIS KTLTYPLDLFKRLQVGGFEHARA AFGQVRRYKGLMDCAKQVLQ KEGALGFFKGLSPSLLKAALSTGFMF FSYEFFCNV FHC MNRTASQR	SLC25A19
368	MASATAAAARRGLGRALPLFWRGYQTERGVYGYRPRKPESREPQG ALERPPVDHGLARLVTVYCEHGHA AKINPLFTGQALLENVPEIQA LVQTLQGPFFTAGLLNMGKEEASLEEV LVYLNQIYCGQISIETSQ LQ SQDEKDWFAKRFEELQKETFTTEERKHL SKLMLESQEFDHFLATKF STVKRYGGEAESMMGFFHELLKMSAYS GITDVIIGMPHRGRLNLL TGLLQFPPELMFRKMRLSEFPENFSATGDVLSHLTSSVDLYFGAHH PLHVTMLPNPSHLEAVNPVAVGK TRGRQQSRQDGDYSPDNSAQPGDRVICLQVHG DASFCGQGIVPETF TLNLP HFRIGGSVHLIVNNQLGYTTPAERGRSSLYCSDIGKLVGCAI IHVNGDSPEEVVRATRLAF EYQRQFRKDV IIDLLCYRQWGHNELDE PFYTNPIMYKIIIRARKSIPDTYAEHLIAGGLMTQEEVSEIKSSYYAKL NDHLNNMAHYRPPALNLQAHWQGLAQPEAQITTWSTGVPLDLLRF	DHTKD1

	VGMKSVEVPRELQMHSLLKTHVQSRMEKMMMDGIKLDWATAEAL ALGSLLAQGFNVRLSGQDVGRGT FSQRHAIVVCQETDDTYIPLNHMDPNQKGFLEVSNSPLSEEAVLGFE YGMIESPKLLPLWEAQFGDFNQAIIIFDTFISGGEAKWLLQSGIVI LLPHGYDAGPDHSSCRIERFLQMCDSAEEGVGDGTVNMFVVHPT TPAQYFHLLRRQMVRNFRKPLIVASPKMLLRPLAAVSTLQEMAPGT TFNPVIGDSSVDPKKVKTLVFCSGKHFYSLVKQRESLGAKKHDFAI RVEELCPFPLDSLQEQEMSKYKHVKDHIWSQEEPQNMGPWSFVSPRF EKQLACKLRLVGRPPPLPVPAV GIGTVHLHQHEDILAKTFA	
369	MASALSYVSKFSFVILFVTPLLLPLVILMPAKFVRCAYVILMAIY WCTEVIPLAVTSLMPVLLFPLFQILDSRQVCVQYMKDTNMLFLGGLI VAVAVERNLHKRIALRTLLWVGAKPARMLLGMGTALLSMWI SNTATTAMMVPIVEAILQQMEATSAATEAGLELVDKGKAKELPGSQ VIFEGPTLGQEDQERKRLCKAMTLCICYAASIGGTATLTGTGPNVV LLGQMNELFPDSKDLVNFASWFAFAFPNMLVMLLFAWLWLQFVY MRNFVKKSWGCGLESKKNEKAALKVLQEEYRKLGPLSFAEINVLC FFLLVILWFSRDPGFMPGWLTVAWVEGETKYVSDATVAIFVATLLFI VPSQKPKFNFRSQTEERKTPFYPPPLLDWKVTQEKVPWGIVLLGG GFALAKGSEASGLSVWVGKQMEPLHAVPPAAITLILSLLVAVFTEC TSNVATTTLFLPIFASMSRSIGLNPLYIMLPCTLSASFAMLPVATPPN AIVFTYGHKLVADMVKTGVIMNIIGVFCVFLAVNTWGRAIFDLDFH PDWANVTHIET	SLC13A5
370	MYRALRLLARSRLVRAPAAALASAPGLGGAAPVPSFWPPNAARMA SQNSFRIEYDTFGELKVPNDKYYGAQTVRSTMNFKIGGVTERMPTP VIKAFGILKRAAAEVNQDYGLDPKIANAIMKAADEVAEGKLNDFHP LVVWQTGSGTQTNMNVNEVISNRAIEMLGELGSKIPVHPNDHVN KSQSSNDTFPTAMHIAAAIEVHEVLLPGLQKLHDALDAKSKEFAQII KIGRTHQTQDAVPLTLGQEFSGYVQQVKYAMTRIKAAMPRIYELAAG GTAVGTGLNTRIGFAEKVAAKVAAALTGLPFVTAPNKFEEALAAHDA LVELSGAMNTTACSLMKIANDIRFLGSGPRSGLGELILPENEPGSSIM PGKVNPTQCEAMTMVAAQVMGNHVAVTVGGSNGHFELNVFKPM MIKNVLHSARLLGDASVSFTENCVVGIQANTERINKLMNESLMLVT ALNPHIGYDKAAKIAKTAHKNGSTLKETAIELGYLTAEQFDEWVKP KDMLGPK	FH
371	MWRVCARRAQNVAPWAGLEARWTALQEVPGTPRVTSRSGPAPAR RNSVTTGYGGVRALCGWTPSSGATPRNRLLLQLLGSPGRRYYSLPP HQKVPLPSLSPTMQAGTIARWEKKEGDKINEGDLIAEVETDKATVG FESLEECYMAKILVAEGTRDVPIGAIICITVGKPEDIEAFKNYTLDS APTQAAPAPTPAATASPPTPSAQAPGSSYPHMQVLLPALSPMTM GTVQRWEKKVGEKLSEGDLLAEIETDKATIGFEVQEEGYLAKILVPE GTRDVPGLTPLCIIVEKEADISAFADYRPTVTDLKPQVPPPTPPVA AVPPTPQPLAPTPSAPCPATPAGPKGRVFSPLAKKLAVEKGIDLTQ VKGTGPDGRITKKDIDSFVPSKVAPAPAAVVPPTGPGMAPVPTGVFT DIPISNIRRVIAQRLMQSKQTIPHYLSIDVNMGEVLLVRKELNKILE GRSKISVNDFIKASALACLKVPEANSSWMDTVIRQNHVVVDVSAV STPAGLITPIVFNHIAHKGVTIANDVVS LATKAREGKLQPHEFQGGTF TISNLGMFGIKNFSAIINPPQACILAIGASEDKLPADNEKGFDVASM MSVTLSCDHRVVDGAVGAQWLAEFRKYLEKPITMLL	DLAT
372	MAGALVRKAADYVRSKDFRDYLMSTHFWGPVANWGLPIAAINDM KKSPEIISGRMTFALCCYSLTFMRFA YKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA	MPC1

373	MRKMLAAVSRVLSGASQKPASRVLVASRNFANDATFEIKKCDLHR LEEGPPVTTVLTREDGLKYRMMQTVRRMELKADQLYKQKIIRGF CHLCDGQEACCVGLEAGINPTDHLITAYRAHGFTFTRGLSVREILAE LTGRKGGCAKKGKGGSMHMYAKNFYGGNGIVGAQVPLGAGIALAC KYNGKDEVCLTLYGDGAANQGQIFEAYNMAALWKLPCIFICENNR YGMGTSVERAAASTDYKRGDFIPGLRVDGMDILCVREATRFAAA YCRSGKGPILMELQTYRYHGHSMSPGVSYRTREEIQEVRSKSDPIM LLKDRMVNSNLASVEELKEIDVEVRKEIEDAAQFATADPEPPLEELG YHIYSSDPPFEVRGANQWIKFKSVS	PDHA1
374	MAAVSGLVRRPLREVSGLLKRRFHWTPAAPALQVTVRDAINQGMDE ELERDEKVFLGEEVAQYDGAYKVSRLWKKYGDKRIIDTPISEMG FAGIAVGAAMAGLRPICEFMTFNFSMQAIDQVINSAAKTYMSGGL QPVPVIFRGPNGASAGVAAQHSQCFAAWYGHCPGLKVSPWNSED AKGLIKSAIRDNNPVVLENELMYGVPFEPPEAQSKDFLIPIGKAKI ERQGTHTVSSHSPVGHCLEAAVLSKEGVECEVINMRTIRPMDM ETIEASVMKTNHLVTVEGGWPQFG VGAEICARIMEGPAFNFLDAPAVRVTGADVPMPLYAKILEDNSIPQVK DIIFAIKKTLLNI	PDHB
375	MAASWRLGCDPRLRLRYLVGFPGRRSVGLVKGALGWSVSRGANWR WFHSTQWLRGDPIKILMPSLSPTMEEGNIVKWLKKEGEAVSAGDAL CEIETDKAVVTLTASDDGILAKIVVEEGSKNIRLGSLIGLIVEEGEDW KHVEIPKDVGPPPPVSKPSEPRPSPEPQISIPVKKHIPGTLRFRLSPAA RNILEKHSLDASQGTATGPRGIFTKEDALKLVQLKQTGKITESRPTP APTATPTAPSPLQATAGPSYPRPVIPPVSTPGQPNVGTFTFTEIPASNIR RVIAKRLTESKSTVPHAYATADCDLGAVLKVRQDLVKDDIKVSVN DFIIKAAAVTLKQMPDVNVSWDGEGPKQLPFIDISVAVATDKGLLTP IHKDAAAKGIQEIADSVKALSKKARDGKLLPEEYQGGSFSSISNLGMF GIDEFTAVINPPQACILAVGRFRPVCLKTEDEEGNAKLQQRQLITVT MSSDSRVVDDELATRFLKSFKANLENPIRLA	PDHX
376	MPAPTQLFFPLIRNCELSRIYGTACYCHHKHLCCSSSYIPQSRLRYTP HPAYATFCRPKENWWQYTQGRRYASTPQKFYLTTPQVNSILKANE YSFKVPEFDGKNVSSILGFDSNQLPANAPIEDRRSAATCLQTRGMLL GVFDGHAGCACSQAVSERLFYIIAVSLLPHETLLEIENAVESGRALL PILQWHKHPNDYFSKEASKLYFNLSLRTYWQELIDLNTGESTDIDVKE ALINAFKRLDNDISLEAQVGDPSFLNYLVLRVAFSGATACVAHVD GVDLHVANTGDSRAMLGVEEDGWSAVTSLNDHNAQNERELER LKLEHPKSEAKSVVKQDRLLGLLMPFRAFGDVKFKWSIDLQKRVIE SGPDQLNDNEYTKFIPPNYHTPPYLTAEPVYHRLRPQDKFLVLAT DGLWETMHRQDVVRIVGEYLTGMHHQQPIAVGGYKVTLGQMHGL LTERRTKMSSVFEDQNAATHLIRHAVGNNEFGTVDHERLSKMLSPL EELARMYRDDITHIVVQFNSHVVGAYQNQE	PDP1
377	MLEKFCNSTFWNSSFLDSPEADLPLCFEQTVLVWIPLGYLWLLAPW QLLHVYKSRTKRSSTTKLYLAKQVFGVGLLILAAIELALVLTEDSGQ ATVPAVRYTNPSTLYLGTWLLVLLIQYSRQWCVCQKNSWFLSLFWILS ILCGTFQFQTLIRTLQGDNSNLAYSCLFFISYGFQILILIFSAFSENNE SSNPNSSIASFLSSITYSWYDSIILKGYKRPLTLEDVWEVDEEMKTKT LVS KFETHMKRELQKARRALQRRQEKSSQQNSGARLPGLNKNQSQSQD ALVLEDVEKKKKKSGTKKDVPSWLMKALFKTFYMVLLKSFLKL VNDIFTFVSPQLLKLISFASDRDTYLVIGYLCAILLFTAALIQSFCLQ CYFQLCFKLGKVRTAIMASVYKKALTLNLARKEYTVGETVNLMI	ABCC2

	<p>SVDAQKLM DVTNFMHMLWSSVLQIVLSIFFLWRELGPSVLAVGV MVLVIPINAILSTKSKTIQVKNMKNKDKRLKIMNEILSGIKILKYFAW EPSFRDQVQNLRRKKELKNLLAFS QLQCVVIFVFQLTPVLVSVVTFVSVYVLVDSNNILDAQKAFSTITLFNI LRFPLSMLPMMISSMLQASVSTERLEKYLGGDDLDTSAIRHDCNFD KAMQFSEASFTWEHDSEATVRDVNLDIMAGQLVAVIGPVGSGKSSL ISAMLGEMENVHGHITIKGTTAYVPQQSWIQNGTIKDNILFGTEFNE KRYQQVLEACALLPDLEMLPGGDLAEIGEKGINLSGGQKQRISLAR ATYQNLDIYLLDDPLSAVDAHVGKHIFNKVLGPNGLLKGGKTRLLVT HSMHFLPQVDEIVVLGNGTIV EKGSYSALLAKKGEFAKNLKTFLRHTGPEEEATVHDGSEEDDDYG LISSVEEIPEDAASITMRRENSFRRTLRSRSSRNGRHLKSLRNSLKTRN VNSLKEDEELVKGQKLIKKEFIETGKVKFSIYLEYLQAIGLFSIFFILA FVMNSVAFIGSNLWLSAWTSDSKIFNSTDYPASQRDMRVGVY GAL GLAQGIFVFIAHFWSAFGFVHASNILHKQLLNNILRAPMRFFDTTPT GRI VNRFAGDISTVDDTLPLQSLRSWITCFLGIISTLVMICMATPVFTIIVIPL GIYVSVQMFYVSTSRQLRRLDSVTRSPIYSHFSETVSGLPVIRAFEH QQRFLKHNEVRIDTNQKCVFSWITSNRWLAI RLELVGNLTVFFSAL MMVIYRDTLSGDTVGFVLSNALNITQTLNWLVRMTSEIETNIVAVE RITEYTKVENEAPWVTDKRPPDWPSKGKIQFN NYQVRYRPELDLV LRGI TCDIGSMEKIGVVGR TGAGKSSLTNCLFRILEAAGGQIIIDGV DIASIG LHDLREKLTIIPQDPILFSGSLRMNLDPFNNYSDEEIWKALELAHLKS FVASLQLGLSHEVTEAGGNLSIGQRQLLCLGRALLRKS KILVLDEAT AAVDLETDNLIQTITIQNEFAHCTVITIAHRLHTIMDS DKVMVLDNGK IIECGSPEELLQIPGPFYFMAKEAGIENVNSTKF</p>	
378	<p>MDQNQHNLNKTAEAPSENKKTRYCNG LKMFLAALSLSFIAKTLGAI IMKSSIIHIERRFEISSSLVGFIDGSFEIGNLLVIVFVS YFGSKLHRPKLI GIGCFIMGIGGVLTALPHFFMGYYRYSKETNINSENSTSTLSTCLIN QILSLNRASPEIVGKGCLKESGSYMWIYVFMGNMLRGIGETPIVPLG LSYIDDFAKEGHSSLYLGILNAIAMIGPIIGFTLGS LFSKMYVDIGYV DLSTIRITPTDSRWVGAWWLNFLVSGLFSIISSIPFFFLPQTPNKPQKE RKASLSLHVLETNDEKDQTANLTNQGKNITKNVTGFFQSFKSILTNP LYVMFVLLTLLQVSSYIGAFTYVFKYVEQQY GQPSSKANILLGVITIP IFASGMFLGGYIHKFKLNTVGI AKFSCFTAVMSLSFYLLYFFILCEN KSVAGLTMTYDGNNPVTSHRDVPLSYCNSDCNCDE SQWEPVCGNN GITYISPCLAGCKSSSGNKKPIVFYNCSCLEVTGLQNRNYS AHLGEC PRDDACTRKFFYFFVAIQVLNLFSSALGGTSHV MLIVKIVQPELKSLA LGFHSMVIRALGGILAPIYFGALIDTT CIKWSTNNCGTRGSCRTYNST SFSRVYLG LSSMLRVSSLVLYIILIYAMKKKYQEKDINASENGS VMD EANLES LNKNKH FVPSAGADSETHC</p>	SLCO1B1
379	<p>MDQHQLNKTAEASASSEKKKTRRCNGFKMFLAALSFSYIAKALGGI IMKISITQIERRFDISSLAGLIDGSFEIGNLLVIVFVS YFGSKLHRPKLI GIGCLLMGTGSILTS LPHFFMGYYRYSKETHINPENSTSSLSTCLINQ TLFNGTSPEIVEKDCVKESGSHMWIYVFMGNMLRGIGETPIVPLGIS YIDDFAKEGHSSLYLGSLNAIGMIGPVIGFALGSLFAKMYVDIGYV DLSTIRITPKDSRWVGAWWLGFLVSGLFSIISSIPFFFLPKNPNKPQKE RKISLSLHVLTNDDRNQTANLTNQGKNVTKNVTGFFQSLKSILTNP LYVIFLLLTLLQVSSFIGSFTYVFKYMEQQY GQSASHANFLLGIITPT VATGMFLGGFIKKFKLSLVGI AKFSFLTSMISFLQLLYFPLICESKS VAGLTLT YDGNNSVASHVDVPLSYCNSECNCDE SQWEPVCGNNGI TYLSPCLAGCKSSSGIKKHTVFYNCSCVEVTGLQNRNYS AHLGEC</p>	SLCO1B3

	RDNTCTRKFFIYVAIQVINSLSFATGGTTFILLTVKIVQPELKALAMG FQSMVIRTLGGILAPIYFGALIDKTCMKWSTNSCGAQQACRIYNSVF FGRVYLGSLIALRFPALVLYIVFIFAMKKKFQGKDTKASDNERKVM DEANLEFLNNGEHFVPSAGTDSKTCNLDMQDNAAAN	
380	MGEPEGQSPSPRSSHSGSPPTLSTLTLLLLCGHAHSQCKILRCNAEYVS STLSLRGGGSSGALRGGGGGGRGGGVGSGGLCRALRSYALCTRRT ARTCRGDLAFHSAVHGIEDLMIQHNC SRQGPTAPPPRGPALPGAGS GLPAPDPCDYEGFRSRLHGRPPGFLHCASFDPHVRSFHHHFHTCR VQGAWPLLDNDFLFVQATSSPMALGANATATRKLTHIFKNMQECID QKVYQAEVDNLPVAFEDGSINGGDRPGGSSLSIQTANPGNHVEIQA AYIGTTIIIRQTAGQLSFSIKVAEDVAMAFSAEQDLQLCVGGCPSQR LSRSERNRRGAITIDTARRLCKEGLPVEDAYFHSCVFDVLISGDPNFT VAAQALEDARAFPLDLEKLHLFSPDAGVPLSSATLLAPLLSGLFVL WLCIQ	HFE2
381	MHQRHPRARCPPLCVAGILACGFLGCGWGPSHFQQSCLQALEPQAV SSYLSPGAPLKGRPPSPGFQRQRQRRAAGGILHLELLVAVGPDVF QAHQEDTERYVLTNLNIGAELLRDP SLGAQFRVHLVKMVILTEPEG APNITANLTSSLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPD GNRQVRGVTQLGGACSPWCLITEDTGFDLGVTIAHEIGHSFGLEH DGAPGSGCGPSGHVMASDGAAPRAGLAWSPCSRRLSLLSAGRA RCVWDPPRPQPGSAGHPPDAQPLYYSANEQCRVAFGPKAVACTF AREHLDMCQALSCHTDPLDQSSCSRLLVPLLDGTECGVEKWCSKG RCRSLVELTPIAAVHGRWSSWGPRSPCSRSCGGGVVTRRRQCNNPR PAFGGRACVGADLQAEMCNTQACEKTQLEFMSQQCARTDQGQLRS SPGGASFYHWGAAPVPHSQGDALCRHMCRAIGESFIMKRGD SFLDG TRCMPSPGREDGTLSCVSGSCRTFGCDGRMDSQQVWDRQCVC GG DNSTCSPRKG SFTAGRAREYVTFLTVPNLTSVYIANHRPLFTHLAV RIGGRYVAVAGKMSISPNTTYP SLLEDGRVEYRVALTEDRLPRLEEIRI WGPLQEDADIQVYRRYGEEYGNLTRPDITFTYFQPKPRQAWVWAA VRGPCSVSCGAGLRWVNYSCLDQARKELVETVQCQGSQQPPAWPE ACVLEPCPPYWAVGDFGPCSASC GGGLRERPVRCVEAQGSLLKTL P PARCRAGAQQPAVALET CNPQPCPARWEVSEPSCTSAGGAGLAE NETCVPGADGLEAPVTEGPGSVDEKLPAPEPCVGMSCPPGWGHL D ATSAGEKAPSPWGSIRTGAQAAHVWTPAAGSCSVSCGRGLMELRF LCMDSALRVPVQEELCGLASKPGSRREVCQAVPCPARWQYKLAAC SVSCGRGVVRRILYCARAHGEDDGEEILLDTQCQGLPRPEPQEACSL EPCPPRWKVMSLGPCSASCGLTARRSVACVQLDQGDQDVEVDEAA CAALVRPEASVPLIADCTYRWHVGTWMECSVSCGDGIQRRRDTC LGPQAQAPVPADFCQHLPKPVTVRGCWAGPCVGGQTPSLVPHEEA AAPGRTTATPAGASLEWSQARGLLFSPAPQPRLLPGPQENS VQSSA CGRQHLEPTGTIDMRGPGQADCAVAIGRPLGEVVT LRVLESSLNC S AGDMLLLWGRLTWRKMCRKLLDMTFSSKTNTLVVRQRCGRPGGG VLLRYGSQ LAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGCRL F INVAPHARIAHALATNMGAGTEGANASYILIRDTHSLRTTAFHGQQ VLYWESESSQAEMEFSEGLKAQASLRGQYWT LQSWVPEMQDPQS WKGKEGT	ADAMTS13
382	MSRPLSDQEKRRKQISVRGLAGVENVT ELKKNFNRLHFTLVKDRN VATPRDYFALAHTVRDHLVGRWIRTQQHYEYKDPKRIYYLSLEFY MGRTLQNTMVNLALENACDEATYQLGLDMEELEEIEEDAGLGNGG LGRLAACFLDSMATLGLAAYGYGIRYEF GIFNQKISGGWQMEEAD DWLRYGNPWEKARPEFTLPVHFYGHVEHTSQGAKWVDTQVVLAM PYDTPVPGYRNNVVNTMRLWSAKAPNDFNLKDFNVGGYIQAVLD	PYGM

	<p>RNLAENISRVLYPNDNFFEGKELRLKQEFVVAATLQDIIRRFKSSK FGCRDPVRTNFDAFPDKVAIQLNDTHPSLAIPELMRILVDLERM DWDKAWDVTVRTCAVTNHTVLPEALERWVPHLLETLLPRHLQIIYE INQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINMAHLICAGSHA VNGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWLVLCNPG LAEVIAERIGEDFISDLQRLKLLSFVDDEAFIRDVAKVKQENKLF AAYLEREYKVHINPNSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKR EPNKFFVPRTVMIGGKAAPGYHMAKMIIRLVTAIGDVVNHDPAVG DRLRVIFLENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMKFMLN GALTIGTMDGANVEMAEAGEENFFIFGMRVEDVDKLDQRGYNAQ EYYDRIPELRQVIEQLSSGFFSPKQPDLFKDIVNMLMHHDRFKVFAD YEDYIKCQEKVSALYKNPREWTRMVRINIATSGKFSSDRITIAQYARE IWGVEPSRQRLPAPDEAI</p>	
383	<p>MLSFVDTRTLLLAVTLCLATCQSLQEETVRKGPAGDRGPRGERGP PGPPGRDGEDGPTGPPGPPGPPGPPGLGGNFAAQYDGKGVGLGPGP MGLMGPRGPPGAAGAPGPQGFQGPAGEPGEPGQTGPAGARGPAGP PGKAGEDGHPGKPRPGERGVVGPQGARGFPGTPGLPGFKGIRGHN GLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGP GPAGARGSDGSVGPVGPAGPIGSAGPPGFPAGPKGEIGAVGNAG PAGPAGPRGEVGLPGLSGPVGPPGNP GANGLTGAKGAAGLPVAGAPGLPGRGIPGPVGAAGATGARGLV GEPGPAGSKGESGNGKEPGSAGPQGPGPSGEEGKRGPNGEAGSAG PPGPPGLRGSPGSRGLPGADGRAGVMGPPGSRGASGPAGVRGPNGD AGRPGEPLMGPRGLPGSPGNIGPAGKEGPVGLPGIDGRPGPIGPAG ARGEPGNIGFPGPKGPTGDPGKNGDKGHAGLAGARGAPGPDGNNG AQGPPGPQGVQGGKGEQGP GPPGFQGLPGPSGPAGEVGKPGERGL HGEFGLPGPAGPRGERGPPGESGAA GPTGPIGSRGPSGPPGPDGNKGEPPGVVGA VGTAGPSGPSGLPGERGA AGIPGGKGEKGEPLRGEIGNPGRDGARGAPGAVGAPGPAGATGD RGEAGAAGPAGPGRGSPGERGEVGPAGPNGFAGPAGAAAGQPGA KGERGAKGPKGENGVVGPTGPVGAAGPAGPNGPAGSRGDGGP PGMTGFPGAAGRTGPPGPSGISGPPGPPGAGKEGLRGPRGDQGPV GRTGEVGA VGPFGAGEKGPSGEAGTAGPPGTPGPQGLLAPGILG LPGSRGERGLPGVAGAVGEPGLGIAGPPGARGPPGAVGSPGVNGA PGEAGRDNPGNDGPPGRDGQPGHKGERGYPGNIGPVGAAGAPGP HGPVGPAGKHGNRGETGPSGPVGPAGAVGPRGPSGPQGIRGDKGEP GEKGPRGLPGLKGHNGLQGLPGIAGHHGDQAGPSVGPAGPRGPA GPSGPAGKDGRTGHPGTVPAGIRGPQGHQGPAGPPGPPGPPGPPG VSGGGYDFGYDGDYRADQPRSAAPSLRPKDYEV DATLKS LNNQIET LLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGCTMDAIKVY CDFSTGETCIRAQPENIPAKNWYRSSKDKKHVWLGETINAGSQFEY NVEGVTSKEMATQLAFMRLANYASQNITYHCKNSIAYMDEETGN LKKAVILQGSNDVELVAEGNSRFTYTVLVDGCSKKTNEWGKTHIEY KTNKPSRLPFLDIAPLDIGGADQEFFVDIGPVCFK</p>	COL1A2
384	<p>MNNLLCCALVFLDISIKWTTQETFPKYLYHYDEETSHQLLCDKCPPG TYLKQHCTAKWKTVCAPCPDHYTDSWHTSDECLYCSPVCKELQY VKQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVVQAGTPER NTVCKRCPDGFFSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNI CSGNSESTQKCGIDVTLCEEAFRRFAVPTKFTPNWLSVLVDNLPGTK VNAESVERIKRQHSSQEQT FQLLKLWKHQKNDQDIVKKIHDIDLCE NSVQRHIGHANLTFEQLRSLMESLPGKKVGAEDIEKTIKACKPSDQI LKLLSLWRKNGDQDTLKGMLHALKHSKTYHFPKTVTQSLKKTIRF LHSFTMYKLYQKL FLEMIGNQVQSVKISCL</p>	TNFRSF11B

385	<p>MAQQANVGELLAMLDSMPLGVRDDVTAVFKENLNSDRGPMLVNT LVDYYLETSSQPALHILTTLQEPHDKHLLDRINEYVGKAATRLSILSL LGHVIRLQPSWKHKLSQAPLLPSLLKCLKMDTDVVVLTGVLVLIT MLPMIPQSGKQHLLDDFDIFGRLLSSWCLKKPGHVAEVYLVHLHASV YALFHRLYGYMPCNFVSFLRSHYSMKENLETFEVVVKPMMEHVRI HPELVTGSKDHELDPRRWKRLETHDVVIECAKISLDPTEASYEDGYS VSHQISARFPHRSADVTTSPYADT QNSYGCATSTPYSTRMLLNMPGQLPQTLSSPSTRLITEPPQATLW SPSMVCGMTTPPTSPGNVPPDLSPYPSKVFGTTAGGKGTPLGTPATS PPPAPLCHSDDYVHISLPQATVTPPRKEERMDSARPCLRQHLLND RGSEPPGSKGSVTLSDLPGLGDLASEEDSIEKDKEEAISRELSEIT TAAEAPVVRGGFDSPFYRDSLPGSQRKTHSAASSSQGASVNPEPLH SSL DKLGPDPKQAFTPIDLPCGSADESPAGDRECQTSLETSIFTPSPCKIP PPTRVGFGSGQPPPYDHLFEVALPKTAHHFVIRKTEELLKKAKGNT EDGVPSTSPMEVLDRLIQQGADAHSKELNKLPLPSKSVDWTHFGGS PPSDEIRLDRDQLLLHNQLLYERFKRQQHALNRRLLRKVIKAAAL EEHNAAMKDQLKLQEKDIQMWKVSQKEQARYNQLQEQRDTMVT KLHSQIRQLQHDREEFYNSQELQTKLEDCRNMAELRIELKKANN KVCHTELLLSQVSQKLSNSESQQQMEFLNRQLLVGEVNELYEQ LQNKHSDTTKEVEMMKAAAYRKELEKNRSHVLQQTQRLDTSQKRIL ELESHLAKKDHLLEQKKYLEDVKLQARGQLQAAESRYEAQKRIT QVFELEILDLYGRLEKDGLLKKLEEEKAEAAEAEEERLDCCNDGCS DSMVGHNEEASGHNGETKTTPRPSARGSSGSRGGGSSSSSELSTP EKPPHQRAGPFSSRWETTMGEASASIPTTVGSLPSSKSFLGMKAREL FRNKSESQCDEDGMTSSLSESLKTELGKDLGVEAKIPLNLDGPHSP PTPDSVGQLHIMDYNETHHEHS</p>	TSC1
386	<p>MAKPTSKDSSLKEKFKILLGLGTPRPNPRSAEGKQTEFIITAEILRELS MECGLNNRIRMIGQICEVAKTKKFEHHAVALWKAVADLLQPERPL EARHVALALLKAIVQGQGERLGVLRALFFKVIKDYPSNEDLHERLE VFKALTDNGRHITYLEEELADFLVQWMDVGLSSEFLVLVNLVKFN SCYLDEYIARMVQMICLLCVRTASSVDIEVSLQVLDVAVVCYNCLPA ESLPLFIVTLCRTINVKELCEPCWKLNRNLLGTHLGHSIYNMCHL MEDRAYMEDAPLLRGAVFFVGMALWGAHRLYSLRNSPTSVLPSFY QAMACPNVVSYEIVLSITRLIKKYRKELQVVAWDILLNIIERLLQQL QTLDSPELRTIVHDLLTTVEELCDQNEFHGSQERYFELVERCADQRP ESSLLNLISYRAQSIHPAKDGWIQNLQALMERFFRSESARGAVRIKVL DVLSFVLLINRQFYEEELINSVVISQLSHIPEDKDHQVRKLATQLLDV LAEGCHTHHFNSLDIIKVMARSLSPPELEERDVAAYSASLEDVK TAVLGLLVILQTKLYTLPASHATRVYEMLVSHIQLHYKHSYTLPIAS SIRLQAFDFLLLRADSLHRLGLPNKDGVRFRSPYCVCDYMEPERGS EKKTSGPLSPPTGPPGPAPAGPAVRLGSPVPSLLFRVLLQCLKQESD WKVLKLVGLRPESLRYKVLIFTSPCSVDQLCSALCSMLSGPKTLER LRGAPEGFSRTDLHLAVVPVLTALISYHNYL DKTKQREMVYCLEQGLIHRCASQCVVALSICSVMPPDIIKALPVLV VKLTHISATASMAVPLLEFLSTLARLPHLYRNFAAEQYASVFAISLP YTNPSTKFNQYIVCLAHHVIAWMFIRCRLPFRKDFVPFITKGLRSNVL LSFDDTPEKDSFRARSTSLNERPKSLRIARPPKQGLNNSPPVKEFKES SAAEAFCRSISVSEHVVRISRIQTSLSASLGADENSVAQADDSLK NLHL ELTETCLDMMARYVFSNFTAVPKRSPVGEFLLAGGRTKTWLVGNK LVTVTTSVGTGTRSLGLDSGELQSGPESSSPGVHVRQTKEAPAKL ESQAGQQVSRGARDRVRSMSSGGHGLRVGALDVPASQFLGSATSPG</p>	TSC2

	<p>PRTAPAAKPEKASAGTRVPVQEKTNLAAYVPLLTQGWAEILVRRPT GNTSWLMSLENPLSPFSSDINNMPLOELSNALMAAERFKEHRDTAL YKSLSVPAASTAKPPPLPRSNTVASFSSLYQSSCQQLHRSVSWADS AVVMEEGSPGEVPVLVEPPGLEDV EAALGMDRRTDAYSRSSSVSSQEEKSLHAEELVGRGPIERVVSSEG GRPSVDLSFQPSQPLSKSSSPELQTLQDILGDPGDKADVGRLSPEVK ARSQSGTLDGESAAWSASGEDSRGQPEGPLSSSPRSPGLRPRGYTI SDSAPSRRGKRVERDALKSRATASNAEKVPGINPSFVFLQLYHSPFF GDESINKPILLPNESQSFERSVQLLDQIPSYDTHKIAVLVVGEGQSNSE LA ILSNEHGSYRYTEFLTGLGRLELKDQCQDPKVYLGGLDVCGEDGQFT YCWHDIMQAVFHIAITLMPKTDVDKHRCDKCRHLGNDVFSIVYND SGEDFKLGTIKGQFNHVIVTPLDYECNLVSLQCRKDMEGLVDT VAKIVSDRNLPFVARQMALHANMASQVHHSRSNPTDIYPSKWIARL RHIKRLRQRICEEAAYSNPSLPLVHPPSHSKAPAQTPAEPTPGYEVG QRKRLISSVEDFTEFV</p>	
387	<p>MAAKSQPNIPKAKSLDGVTDNRASQGWGRAWEVDWFLASVIF LLLAFPIVYYFIMACDQYSCALTGPVVDIVTG HARLSDIWAKTPIT RKAQAQLYTLWVTFQVLLYTSPLDFCHKFLPGYVGGIQEAVTPAGV VNKYQINGLQAWLLTHLLWFANAHLLSWFSPTIIFDNWIPLWCAN ILGYAVSTFAMVKGYFFPTSARDCKFTGNFFNYMMGIEFNPRIGK WDFDKLFFNGRPGIVAWTLINLSFAAKQRELHSHVTNAMVLNVNL QAIYVIDFFWNETWYLKTIDICH HFGWYLGWGDVWLPYLYTLQGLYLVYHPVQLSTPHAVGVLLLG LVGYIIFRVANHQKDLFRRTDGRCLIWGRKPKVIECSYTSADGQRH HSKLLVSGFWGVARHFNYVGDLMGSLAYCLACGGGHLLPYFYIY MAILLTHRCLRDEHRCASKYGRDWERYTAAVPYRLLPGIF</p>	DHCR7
388	<p>MSLSNKLTLDKLDVKGRVVMRVDFNVPKNNQITNNQRIKAAVP SIKFLDNGAKSVVLMShLGRPDGVMPDPKYSLEPVAVELKSLGK DVLFLKDCVGPVEVEKACANPAAGSVILLENLRFHVEEEGKGKDasG NKVKAEPKIEAFRASLSKLGDVYVNDAFGTAHRAHSSMVGvNPL QKAGGFLMKKELNYFAKALESPEPFLAILGGAKVADKIQLINNML DKVNEMIIGGGMAFTFLKVLNNMEIGTSLFDEEGAKIVKDLMSKAE KNGVKITLPVDFVTADKFDENAKTGQATVASGIPAGWMGLDCGPE SSKKYAEAVTRAKQIVWNGPVGVFEWEAFARGTKALMDEVV KATSRGCITIIGGGDTATCCAkwNTEDKVSHVSTGGGASLELLEGK VLPgVDALSNI</p>	PGK1
389	<p>MGTSALWALWLLALCWAPRESGATGTGRKAKCEPSQFQCTNGR CITLLWKCDGDEDCVDGSDKNCVKKTCAESDFVCNNGQCVPSRW KCDGDPDCEDGSDSPEQCHMRTCRIHEISCGAHSTQCIPVSWRCD GENDCDSGEDEENCGNITCSPDEFTCSSGRCSIRNFVCNGQDDCSDG SDELDCAPPTCGAHEFQCSTSSCIPISWVCDDADCDSDQSDSLEQC GRQPVHTKCPASEIQCGSGECIHKKWRCDGDPDCDKGSDDEVNCP RTRCPDQFECEDGSCIHGSRQCNGI RDCVDGSDDEVNCKNVNQCLGPGKFKCRSGECIDISKVCNQEQDCR DWSDEPLKECHINECLVNNGGCSHICKDLVIGYECDAAGFELIDRK TCGDIDECQNPgICSQICINLKGgYKCECSRGYQMDLATGVCKAVG KEPSLIFTNRRDIRKIGLERKEYIQLVEQLRNTVALDADIAAQKLFW ADLSQKAIFSASIDDKVGRHVKMIDNVYNPAAIAVDWVYKTIYWT DAASKTISVATLDGTRKFLFNSDLREPASIAVDPLSGFVYWSDWG EPAKIEKAGMNGFDRRPLVTADIQ WPNGITLDLIKSRLYWLDSKLHMLSSVDLNGQDRRIVLKSLEFLAHP LALTIFEDRVYWIDGENEAVYGANKFTGSELATLVNNLNDAAQDIIV</p>	VLDLR

	YHEL VQPSGKNWCEEDMENGGEYLCLPAPQINDHSPKYTCSCPSG YNVEENGRDCQSTATTVTYSETKDTNTTEISATSGLVPGGINVTTAV SEVSVPKGTSAAWAILPLLLL VMAAVGGYLMWRNWQHKNMKS MNFDPNPVYLKTTEEDLSIDIGRHSASVGHTYPAISVVSTDDDLA	
390	MEPSSLELPADTVQRIAAELKCHPTDERVALHLDEEDKLRHFRECFY IPKIQDLPPVDLSLVNKDENAIYFLGNSLGLQPKMVKTYLEEELDKW AKIAAYGHEVGKRPWITGDESIVGLMKDIVGANEKEIALMNALTVN LHLLMLSFFKPTPKRYKILLEAKAFPSDHYAIESQLQLHGLNIEESMR MIKPREGEETLRIEDILEVIEKEGDSIAVILFSGVHFYTGQHFNPATIK AGQAKGCYVGFDLAHAVGNVELYLHDWGVDFACWCSYKYL NAG AGGIAGAFIHEKHAHTIKPALVGWFGHELSTRFKMDNKLQIPGVC GFRISNPILLVCSLHASLEIFKQATMKALRKKSVLLTGYLEYLIKHN YGKDKAATKKPVVNIITPSHVEERGQLTITFSVPNKDVFQELEKRG VVCDKRNPNGIRVAPVPLYN SFHDVYKFTNLLTSILDSAETKN	KYNU
391	MFPGCPRLWVLVVLGTSWVGWGSQGTEAAQLRQFYVAAQGISWS YRPEPTNSSLNLSVTSFKKIVYREYEPYFKKEKPQSTISGLLGPTLYA EVGDIIKVFHFNKADKPLSIHPQGIRYSKLESEGASYLDHTFPAEKMD DAVAPGREYTYEWSISED SGPTHDDPPCLTHIYYSHENLIEDFN SGLI GPLLICKKGTLEGGTQKTFDKQIVLLFAVFDESKSWSQSSSLMYTV NGYVNGTMPDITVCAHDHISWHLLGMSSGPELFSIHFNQVLEQNH HKVSAILTVSATSTTANMTVGPEGKWISSLTPKHLQAGMQAYIDIK NCPKKTRNLKKITREQRRHMKRWEYFIAAEEVIWDYAPVIPANMD KKYRSQHLDNFSNQIGKHYKKVMYTQYEDESFTKHTVNPNMKED GILGPIIRAQVRDTLKIVFKNMASRPYSIYPHGVTFSPYEDEVNSSFTS GRNNTMIRAVQPGETYTYKWNILEFDEPTENDAQCCLTRPYYSVDVI MRDIASGLIGLLICKSRSLDRRGIRAA DIEQQAVFAVF DENKSWYLEDNINKFCENPDEVKRDDPKFYESNIM STINGYVPESITTLGFCFDDTVQWHFCSVGTQNEILTIHFTGHSFIYG KRHEDTLTLFPMRGESVTVTMDNVGTWMLTSMNSSPRSKKLRLKF RDV KCI PDDDEDSYEIFEPPESTVMATRKMHDRLEPEDEESDADYD YQNRLAAALGIRSFNRSSLNQEEEEFNLTALALENGTEFVSSNTDIIV GSNYSSPSNISKFTVNNLAEPQKAPSHQQAATTAGSPLRHLIGKNSVL NSSTA EHSSPYSEDPIEDPLQPDVTGIRLLSLGAGEFKSQEHAKHKG KVERDQA AKHRFSWMKLLAHKVGRHLSQDTGSPSGMRPWEDLPS QDTGSPSRMRPWKDPPSDLLLLKQSNSSKILVGRWHLASEKGSYEII QDTDEDTA VNNWLISPQNASRAWGESTPLANKPGKQSGHPKFPRV RHKSLQVRQDGGKSRLKKSQFLIKTRKKKKEKHTHHAPLSPRTFHP LRSEAYNTFSERRLKHSLVLHKSNETSLPT DLNQTLPSMDFGWIASLPDHNQNSSNDTGQASCPPGLYQTVPPEEH YQTFPIQDPDQMHSTSDPSHRSSPELSEMLEYDRSHKSFPDTDISQMS PSSEHEVWQTVISPDLSQVTLSPELSQTNLSPDLSHTTTLSPELIQRNLS PALGQMPISPDLSHTTTLSPDLSHTTTLSDLSQTNLSPELSQTNLSPAL GQMPLSPDLSHTTTLSDFSQTNLSPELSHMTLSPELSQTNLSPALGQ MP ISPDLSHTTTLSDFSQTNLSPELSQTNLSPALGQMPLSPDPSHTTTLSD LSQTNLSPELSQTNLSPDLSEMP LFADLSQIPLTPDL DQMTLSPDLGE TDLSPNFGQMSLSPDLSQVTLSPDISDTLLPDLSQISPPDL DQIFYP SESSQSLLLQEFNESFPYDLGQMPS PSSPTLNDTFLSKEFNPLVIVGL SKDGTDYIEIIPKEEVQSEDDYAEIDYVPYDDPYKTDVVRTNINSSRD PDNIAAWYLRNNGNRRNYYIAAEEISWDYSEFVQRETDIEDSDDIP EDTTYKKVVFRKYLDSTFTKRDPGRGEYEEHLGILGPIIRAEVDDVIQ VRFKNLASRPYSLHAHGLSYEKSSEGKTYEDDSPEWFKEDNAVQPN SSYTYVWHATERSGPESPGSACRAWAYYS AVNPEKDIHSGLIGPLLI	F5

	<p>CQKGILHKDSNMPMDMREFVLLFMTFDEKKSWYYEKKSRSSWRLT SSEMK KSHEFHAINGMIYSLPGLKMYEQEWVRLHLLNIGGSQDIHVVFHFG QTLLENGNKQHQLGVWPLLPGSFKTLEMKASKPGWWLLNTEVGE NQRAGMQTPFLIMDRDCRMPMGLSTGHIISDSQIKASEFLGYWEPRL ARLNNGGSYNAWSVEKLAAEFASKPWIQVDMQKEVIITGIQTQGA HYLKSCYTTEFYVAYSSNQINWQIFKGNSTRNVMYFNGNSDASTIK ENQFDPPIVARYIRISPTRAYNRPTLRLELQGCEVNGCSTPLGMENG KIENKQITASSFKKSWWGDYWEPFR ARLNAQGRVNAWQAKANNKQWLEIDLLKIKKITAIIQTGCKSLSS EMYVKSYTIHYSEQGVWKPYPRLKSSMVDKIFEGNTNTKGHVKNF FNPIISRFRVIPKTWNQSIARLELFGCDIY</p>	
392	<p>MGPTSGPSLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHD AQGDVPVTVTVHDFPGKKLVLSSEKTVLTPATNHMGNVTFITIPANR EFKSEKGRNKFVTVQATFGTQVVEKVVLVSLQSGYLFITDKTIYTP GSTVLYRIFTVNHKLLPVGRVTVMNIENPEGIPVKQDSLSSQNQLGV LPLSWDIPELVNMGQWKIRAYYENSPQVVFSTEFVKEYVLPSPFEVI VEPTEKFYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQRIS LPESLKRIPIEDGSGEVVLSRKVLLDGVQNPRAEDLVGKSLYVSATV ILHSGSDMVQAERSGIPIVTSPIYQIHFTKTPKYFKPGMPFDLMVFVT NPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPLSITV RTKKQELSEAEQATRTMQALPYSTVGNSNNYLHLSVLRTLRPGET LNVNFLLRMDRAHEAKIRYYTYLIMNKGRLKAGRQVREPQDLV VLPLSITDFIPSFRLVAYYTLIGASGQREVVADSVWVDVKDSCVGS LVVKSQSEDQRQVPVPGQMTLKIIEGDHGARVVLVAVDKGVFVLNK KNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQ TAQRAELQCPQPAARRRRSVQLTEKRMDKVKGYPKELRKCCEDG MRENPMRFSCQRRTRFISLGEACKKVFLDCCNYITELRRQHARASH LGLARSNLDEIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLM NIFLKDSITTWEILAVSMSDKKGICVADPFVETVMQDFFIDLRLPYSV VRNEQVEIRAVLYNYRQNLKVRVELLHNPACSLATTKRRHQQ TVTIPPKSSLSVPYVIVPLKTGLQEVEVKAAYVHHFISDGVRLSLKV VPEGIRMNKTVAVRTLDPERLREGVQKEDIPPADLSDQVPDTESET RILLQGTTPVAQMTEDA VDAERLKHLLVTPSGCGEQNMIGMTPTVIA VHYLDETEQWEKFGLEKRQGALELIKGYTQQLAFRQPSSAFAAFV KRAPSTWLTA YVVKVFSLAVNLIIDSQVLCGAVKWLILEKQKPDGVFQEDAPVIH QEMIGGLRNNNEKDMALTAFLVLSLQEAKDICEEQVNSLPGSITKAG DFLEANYMNLQRSYTVIAIAGYALAQMGRLKGPLLNKFLTAKDKN RWEDPGKQLYNVEATSYALLALLQLKDFDFVPPVVRWLNEQRYYG GGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITH RIHWESASLLRSEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKD QLTCNKFDLKVTIKPAPETEKRPQDAKNTMILEICTRYRGDQDATM SILDISMMTGFAPDTDLLKQLANGVDRIYISKYELDKAFSDRNTLIY LDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFY HPEKEDGKLNKLCRDELRCRAEENCIFIQKSDDKVTLERLDKACEP GVVYVYKTRLVKVQLSNDFDEYIMAIEQTIKSGSDEVQVGGQRTFIS PIKCREALKLEKKHYLMWGLSSDFWGEKPNLSYIIGKDTWVEHWP EEDECQDEENQKQCQDLGAFTESMVVFGCPN</p>	C3
393	<p>MGPRLSVWLLLLPAALLLHEEHSRAAAKGGCAGSGCGKCDCHGV KGQKGERGLPGLQGVIGFPGMQGPEGPGQPPGQKGDGTGEPGLPGTK GTRGPPGASGYPGNPGLPGIPQDGPFGPPGIPGCNGTKGERGPLGP PGLPGFAGNPGPPGLPGMKGDPGEILGHVPGMLLKGERGFPGIPGTP</p>	COL4A1

	<p>GPPGLPGLQGPVGPFGTGPFGPPGPPGPPGEGKQMGLSFQGPKGDK GDQGVSGPPGVPGQAQVQEKGDFAKGEKGQKGEPPGFQGMPPGVG EKGEPPGKPPRGKPGKDGDKGEKSGPFPGEPPGYPLIGRQGPQGE KGEAGPPGPPGIVIGTGPLGEKGERGYPGTPGPRGEPGPKGFPGLPG QGPPLPVPQAGAPGFPGERGEKGDRGFPGTSLPGPSGRDGLPGP PGSPGPPGQPGYTNGIVECQPGPPGDQPPGIPGQPGFGEIGEKGQK GESCLICDIDGYRGPPGPQGPGEIGFPGQPGAKGDRGLPGRDGVAG VPGPQGTPLIGQPGAKGEPGEFYFDLRLKGDKGDPGFPQPGMPG RAGSPGRDGHPLPGPKGSPGSGVLKGERGPPGGVGFPGSRGDTGP PGPPGYGPAGPIGDKQAGFPGGPGSPGLPGPKGEPGKIVPLPGPPG AEGLPSPGFPGPQDGRGFPGTPGRPLPGEKGA VGPQPGIFGPPPG PKGVDGLPGDMGPPGTPGRPGFNGLPGNPGVQGGKGEPPVGLPGL KGLPGLPGIPGTPEKGSIGVPGVPGEHGAIGPPGLQGIRGEPGPPGL PGSVGSPGVPGIPPGARGPPGGQPPGLSGPPGIKGEKGFPFGPGLD MPGPKGDKGAQGLPGITGQSLPGLPGQQGAPGIPGFPGSKGEMGV MGTPGQPGSPGPGVAPGLPGEKGDHGFPGSSGPRGDPGLKGDKG VGLPGKPGSMDKVDMSGMKGQKGDQGEKGQIGPIGEKGSRGDPGT PGVPGKDGQAGQPGQPGPKGDPGISGTPGAPGLPGPKGSGVGMGLP GTPGEKGVPGIPGPGQSPGLPGDKGAKGEKGQAGPPGIGIPGLRGEK GDQGIAGFPSPGEKGEKGSIGIPGMPGSPGLKSPGSGVGYPGSPGLP GEKGDGLPLDGLPGVKGEAGLPGTPGTPGAGQKGEPPGSDGIPG SAGEKGEPPGLPGRGFPGFPAKGDGSKGEVGFPLAGSPGIPGSK GEQGFMPGPPGQGPGLPGSPGHATEGPKGDRGPQGPGLPGLPGP MGPPGLPGIDGVKGDGNPGWPGAPGVPGPKGDPGFQGMPPGIGGS PGITGSKGDMGPPGVPGFQGPGLPGLQGIKGDQGDQGVPGAKGLP GPPGPPGPYDIKGEPPGLPGEPPGLKGLQGLPGPKGQQGVTLVG IPGPPGIPGFDGAPGQKGEMGPAGPTGPRGFPGPPGPDGLPGSMGPP GTPSVDHGFVTRHSQTIDDPQCPSTKILYHGYSLLYVQGNRAH GQDLGTAGSCLRKFTMPFLFCNINNVCNFASTRNDYSYWLSTPEPM PMSMATIPGENIRPFISRCVCEAPAMVMAVHSQTIQIPPCPSGWSL WIGYSFVMTSAGAEAGSGQALASPGSCLEEFRSAPFIECHGRGTCNY YANAYSFWLATIERSEMFKKTPSTLKAGELRTHVSRQVCMRRT</p>	
394	<p>MRLAKIICLMLWAICVAEDCNELPPRRNTEILTGSWSDQTYPEGTQ AIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKCQKRPCGHPGDT FGTFTLTGGNVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDI PICEVVKCLPVTAPENGKIVSSAMEPDREYHFGQAVRFVCNSGYKIE GDEEMHCSDDGFWSKEPKCVEISCKSPDVINGSPISQKIIYKENERF QYKCNMGYEYSERGDAVCTESGWRPLPSCEEKSCDNPYIPNGDYSP LRIKHRTGDEITYQCRNGFYPATRGNTAKCTSTGWIPAPRCTLKPCD YPDIKHGGLYHENMRRPYFPVAVGKYYSYCYDEHFETPSGSYWDH IHCTQDGWSPA VPCLRKCYFPYLENGYNQNYGRKFVQGSIDVAC HPGYALPKAQTTVTCMENGWSPTPRCIRVKTCSSIDIENGFISESQ YTYALKEKAKYQCKLGYVTADGETSGSITCGKDGWSAQPTCIKSC DIPVFMNARTKNDFTWFKLNDTLDYECHDGYESNTGSTTGSIVCGY NGWSDLPICYERECLEPKIDVHLVPDRKKDQYKVGVLKFSCKPGF TIVGPNVSVQCYHFGLSPLPICKEQVQSCGPPPELLNGNVKEKTKEE YGHSEVVEYYCNPRFLMKGPNKIQCVDGEWTTLPVCIVEESTCGDI PELEHGW AQLSSPPYYYGDSVEFNCSSESFTMIGHRSITCIHGVWTQL PQCAIDKLKCKSSNLIIEHLKNKKEFDHNSNIRYRCRGKEGWI HTVCINGRWDPEVNCSMAQIQLCPPPQIPNSHNMTTTLNYRDGEK VSVLCQENYLIQEGEEITCKDGRWQSIPLCVEKIPCSQPPQIEHGTINS SRSSQESYAHGTLKSYTCEGGFRISEENETTCYMGKWSSPPQCEGLP CKSPPEISHGVVAHMSDSYQYGEVTKCFEGFGIDGPAIAKCLGEK</p>	CFH

	WSHPPSCIKTDCLSLPSFENAIPMGEKKDVYKAGEQVTTYTCATYYK MDGASNVTCINSRWTGRPTCRDTSCVNPPTVQNAIYIVSRQMSKYPS GERVRYQCRSP YEMFGDEEVMLNNGNWTEPPQCKDSTGKCGPPPIDNGDITSFPLSV YAPASSVEYQCQNLQLEGNKRITCRNGQWSEPPKCLHPCVISREIM ENYNIALRWTAQKQLYSRTGESVEFVCKRGYRLSSRSHLTRTTCWD GKLEYPTCAKR	
395	MEPRPTAPSSGAPGLAGVGETPSAAALAAARVELPGTAVPSVPEDA APASRDGGGVRDEGPAAAGDGLGRPLGPTPSQSRFQVDLVSENAG RAAAAAAAAAAAAAAAAAGAGAGAKQTPADGEASGESEPAKGSEEA KGRFRVNFVDPAASSSAEDSLSDAAGVGVDGPNVSFQNGGDTVLS GSSLHSGGGGGSGHHQHYYDYDTHNTYYLRTFGHNTMDAVPRIDH YRHTAAQLGEKLLRPSLAELHDELEKEPFEDGFANGEESTPTRDAV VTYTAESKGVVKFGWIKGVLVRCMLNIWGVMLFIRLSWIVGQAGI GLSVLVIMMATVVTITGLSTSAIATNGFVRGGGAYYLISRLGPEF GGAIGLIFAFANAVAVAMYVVGFAETVVELLKEHSILMIDEINDIRII GAITVVILLGISVAGMEWEAKAQIVLLVILLAIAGDFVIGTFIPLESKK PKGFFGYKSEIFNENFGPDFREEETFFSVFAIFFPAATGILAGANISGD LADPQSAIPKGTLLAILITTLVYVGVIAVSV GSCVVRDATGNVNDTIVTELNTCTSAACKLNDFSSCESSPCSYGL MNMFQVMSMVSFGFTPLISAGIFSATLSSALASLVSAKIFQALCKDNI YPAFQMFAGYKGNNEPLRGYILTFILALGFILIAELNVIAPISNFFL ASYALINFSVFHASLAKSPGWRPAFKYYNMWISLLGAILCCIVMFVI NWWAALLTYVIVLGLYIYVYKPKPDVNWGSSTQALTYLNAHQHSI RLSGVEDHVKNFRPQCLVMTGAPNSRPALLHLVHDFTKNVGLMIC GHVHMGPRRQAMKEMSIDQAKYQRWLKKNMKAFYAPVHADDL REGAQYLMQAAGLGRMKPNTLVLGFKKDWLQADMVDVDMYINL FHDAFDIQYGVVIRLKEGLDISHLQGGQEELLSSQEKSPGTDVVS VEYSKSDLDTSKPLSEKPTHKVEEEDGKTATQPLLKESKGPIVPL NVADQKLEASTQFQKKQGKNTIDVWWLFDDGGLTLLIPYLLTTK KKWKDCKIRVFIGGKINRIDHRRAMATLLSKFRIDFSDIMVLGDIN TKPKKENIIAFEEIIEPYRLHEDDKEQDIADKMKEDPWRITDNELEL YKTKTYRQIRLNELLKEHSSTANIIVMSLPVARKGAVSSALYMAWL EALSKDLPPILLVRGNHQSVLTFYS	SLC12A2
396	MAASKKAVLGPLVGAVDQGTSSSTRFLVFNSKTAELLSHHQVEIKQE FPREGWVEQDPKEILHSVYECIEKTCEKLGQLNIDISNIKAIGVSNQR ETTVVWDKITGEPLYNAVWLDLRTQSTVESLSKRIPGNNNFVKS TGLPLSTYFSAVKLRWLLDNVRKVQKAVEEKRALFGTIDSWLIWSL TGGVNGGVHCTDVTNASRTMLFNIHSLEWDKQLCEFFGIPMEILPN VRSSSEIYGLMKISHSVKAGALEGVPISGCLGDQSAALVGQMCQFQIG QAKNTYGTGCFLLCNTGHKCVFSDHGLLTTVAYKLGRDKPVYYAL EGSVAIAGAVIRWLRDNLGHIKTSEEIEKLAKEVGTSYGCYFVPAFSG LYAPYWEPSARGIICGLTQFTNKCHIAFAALEAVCFQTREILDAMNR DCGIPLSHLQVDGGMTSNKILMQLQADILYIPVVKPSMPETTALGAA MAAGAAEGVGWVWSLEPEDLSAVTMERFEPQINAESEIRYSTWKK AVMKSMGWVTTQSPESGDPSIFCSLPLGF FIVSSMVMILIGARYISGIP	GK
397	MDVGSKEVLMESPPDYSAAPRGRFGIPCCPVHLKRLIVVVVVVLIV VVIVGALLMGLHMSQKHEMVMLEMSIGAPEAQQLALSEHLVTTA TFSIGSTGLVVDYQQLLIAYPAPGTCCYIMKIAPEIPSLEALNRK VHNFMQECSLQAKPAVPTSKLQAEGRDAGSAPSGGDPFLGMAV NTLCEVPLYI	SFTPC

398	MEPGRRGAAALLALLCVACALRAGRAQYERYSFRRSFRDELMPLES AYRHALDKYSGEHWAESVGYLEISLRLHRLLRDSEAFCHRNCSAAP QPEPAAGLASYPELRLFGGLLRRRAHCLKRCKQGLPAFRQSQPSREV LADFQRREPYKFLQFAYFKANNLPKAIAAAHTFLLKHPDDEMMKR NMAYYKSLPGAEDYIKDLETKSYESLFIRAVRAYNGENWRTSITDM ELALPDFFKAFYECLAACEGSREIKDFKDFYLSIADHYVEVLECKIQ CEENLTPVIGGYPVEKFVATMYHY LQFAYYKLNLDLKNAAPCAVSYLLFDQNDKVMQQNLVYYQYHRDT WGLSDEHFQPRPEAVQFFNVTTLQKELYDFAKENIMDDDEGEVVE YVDDLLELEETS	CRTAP
399	MAVRALKLLTTLAVVAAASQAEVESEAGWGMVTPDLLFAEGTA AYARGDWPQGVVLSMERALRSRAALRALRLRCRTQCAADFPWELDP DWSPSPAQASGAAALRDLFFGGLLRRRAACLRRCLGPPAAHSLSEE MELEFRKRSPYNYLQVAYFKINKLEKAVAAAHTFFVGNPEHMEMQ QNLDYYQTMMSGVKEADFKDLETQPHMQEFRLGVRLYSEEQPQEA PHLEAALQEFVAYEECRALCEGPYDYDGYNYLEYNADLFQAITD HYIQVLNCKQNCVTELASHPREKPFEDFLPSHYNYLQFAYYNIGN YTQAVECAKTYLLFFPNDEVNMNQNLAYYAAMLGEEHTRSIGPRES AKEQRQSRILLEKELLFFAYDVFGIPFVDPDSWTPEEVIPKRLQEKQK SERETA VRISQEIGNLMKEIETLVEEKTKESLDVSRLTREGGPLLYEG ISLTMNSKLLNGSQRVVMDGVIDHECQELQRLTNVAATSGDGYR GQTSPTHPNEKFYGVTVFKALKLGQEGKVPLQSAHLYYNVTEKVR RIMESYFRLDTPLYFSYSHLVCRTAIEEVQAERKDDSHPVHVDNCIL NAETLVCVKEPPAYTFRDYSAILYLNGDFDGGNFYFTELDAKTVTA EVQPOCGRAVGFSSTENPHGVKAVTRGQRCAIALWFTLDPRHSE DRVQADDLVKMLFSPEEMDLSQEQPLDAQGPPEPAQESLSGSESK PKDEL	P3H1
400	MTLRLLLVAALCAGILAEAPRVRAQHRERVCTRLYAADIVFLLDGS SSIGRSNFREVRSFLEGLVLPFSGAASAQGVRFATVQYSDDPRTEFG LDALGSGGDVIRAIRESYKGGNTRTGAAILHVADHVFLPQLARPG VPKVCILITDGKSQDLVDTAQRLKGQGVKLFVAVGIKNADPEELKR VASQPTSDFFFFVNDFSILRTLPLVSRVCTTAGGVPVTRPPDDSTS APRDLVLSEPSQSLRVQWTAASGPVTGYKVQYTPLTGLGQPLPSE RQEVNVPAGETSVRLRGLRPLTEYQVTVIALYANSIGEA VSGTARTT ALEGPELTIQNTTAHSLLVAWRSVPGATGYRVTVRWVLSGGPTQQQE LGPQGQSVLLRDLEPGTDYEVTVSTLFGRSVGPATSLMARTDASVE QTLRPVILGPTSILLSWNLVPEARGYRLEWRRETGLEPPQKVVLPSD VTRYQLDGLQPGTEYRLTYTLLEGHEVATPATVVPTGPELPVSPVT DLQATELPGQVRVRSWSPVPGATQYRII VRSTQGVERTLVLPQSQTAFDLDVQAGLSYTVRV SARVGPREGSA SVLTVRREPETPLAVPGLRVVVS DATRVVAVGWPGVPGASGFRISWS TGSGPESSQTLPPDSTATDITGLQPGTTYQVAVSVLRGREGPAAVI VARTDPLGPVRTVHVTVQASSSVTITWTRVPGATGYRVSWHSAHGP EKSQVLVSGEATVAELDGLEPDTEYTVHVRHVAGVDGPPASVVVR TAPEPVGRVSRQLILNASSDVLRTWVGVTGATAYRLAWGRSEGGP MRHQILPGNTDSAIRGLEGGVSY SVRVTA LVGDREGTPVSIVVTTPEAPPALGTLHVVRGEHSLRLR WEPVPRAQGFLHWQPEGGEQSRVLGPELSSYHLDGLEPATQYR VRLSVLGPAGEGPSAEVTARTESPRVPSIELRVVDTSIDSVTLAWTP VSRASSYILSWRPLRGPGQEVPGSPQTLPGISSQRTGLEPGVSYIFS LTPVLDGVRGPEASVTQTPVCPRGLADVFLPHATQDNAHRAEATR RVLERLV LALGPLGPQAVQVGLLSYSHRPSPLFPLNGSHDLGILQRI RDMPYMDPSGNNLGTAVVTAHRYMLAPDAPGRRQHVPVGMVLLV	COL7A1

	DEPLRGDIFSPIREAQASGLNVVMLGMAGADPEQLRRLAPGMDSVQ TFFAVDDGPSLDQAVSGLATALCQASFTTQPRPEPCPVYCPKGQKG EPGEMGLRGQVGPDPGLPGRGTGAPGPQGPPTSATAKGERGFPGA DGRPGSPGRAGNPGTPGAPGLKGSPGLPGRGDPGERGPRGPKGEP GAPGQVIGGEGPGLPGRKGDGPGSPGPPGPRGPLGDPGPRGPPGLPT AMKGDKGDRGERGPPGPGEGGIAPGEPGLPGLPGSPGPQGPVGP KKGEKGDSSEDPGLPGQPGSPGEQGPGRPPGAIGPKGDRGFPGPL GEAGEKGERGPPGPAGSRGLPGVAGRPGAKGPEGPPGPTGRQGEKG EPGRPGDPAVVGPAVAGPKGEKGDVGPAGPRGATGVQGERGPPGL VLPDGPDPKGDGDRGPIGLTGRAGPPGDSGPPGEKGDGPRGPPGP VGPRGRDGEVGEKGDEGPPGDPGLPGKAGERGLRGAPGVRGPVGE KGDQGDGPDGDRNGSPGSSGPKGDRGEPGPPGPPGRLVDTGPGARE KGEPGDRGQEGPRGPKGDPGLPGAPGERGIEGFRGPPGPQGDGVR GPAGEKGDRGPPGLDGRSGLDGKPGAAGPSGPNAAAGKAGDPGRD GLPGLRGEQGLPGSPGPPGLPGKPGEDGKPLNGKNGEPGDPGEDG RKGEKGDGSGASGREGRDGPKEGGERGAPGILGPQGPGLPGVGP GFPGVPGGTGPKGDRGETGSKGEQGLPGERGLRGEPGSPVNVDRLL ETAGIKASALREIVETWDESSGFLPVPERRRGPKGDSGEQGPKE GPIGFPRGERGLKGDGDPGPQGPGLALGERGPPGPSGLAGEPGKPG IPGLPGRAGGVGEAGRPGERGERGEKGERGEQGRDGPGLPTGP PGPPGPKVSVDEPGPLSGEQGPGLKGAKGEPGSNGDQGPKGDRG VPGIKGDRGEPGPRGQDGNPGLPGERGMAGPEGKPLQGPRGPPGP VGGHGDGPPGAPLAGPAGPQGPSGLKGEGETGPPGRGLTGPTG AVGLPGPPGPSGLVGPQGPSGLPGQVGETGKPGAPGRDGASGKDG DRGSPGVPGSP GLPGPVGPKGEPGPTGAPGQAVVGLPGAKGEKGAPGGLAGDLVGE PGAAGDRGLPGRGEKGEAGRAGEPGDPGEDGQKGAPGPKGFKGD PGVGVPGSPGPPGPPGVKGDGLPGLPGAPGVVGFPGQTGPRGEMG QPGPSGERGLAGPPGREGIPGLGPPGPPGSVGPAGSLKGDKGDP GVGLPGRGERGEPGIRGEDGRPGQEGPRGLTGPPGSRGERGEKGD VGSAGLKGDKGDSAVILGPPGPRGAKGDMGERGPRGLDGDGKGRG DNGDPGDKGSKGEPGDKGSAGLPGLRGLLGPQGPAGIPGDPGS PGKDGVPGRGEKGDVGFMGPRGLKGERGVKGACGLDGEKGDGK EAGPPGRPGLAGHKGEMGEPGVPGQSGAPGKEGLIGPKGDRGFDG QPGPKGDQGEKGERGTPGIGGFPGSGNDGSAGPPGPPGSVGPGRPE GLQGQKGERGPPGERVVGAPGVPGAPGERGEQGRPGAPGRGEKG EAALTEDDIRGFVRQEMSQHCACQGQFIASGSRPLPSYAADTAGSQ LHAVPVLRVSHAEERVPEDDEYSEYSEYVEEYQDPEAPWDS DPCSLPLDEGSCTAYTLRWYHRAVTGSTEACHPFVYGGCGGNANR FGTREACERRCPPRVVQSQGTGTAQD	
401	MSIQENISSQLRSWVSKSQRD LAKSILIGAPGGPAGYLRRASVAQL TQELGTAFFQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIATIG PASRSVERLKEMIKAGMNIARLNFSHGSHEYHAESIANVREAVESFA GSPLSYRPVAILDTKGPEIRTGILQGGPESEVELVKSQVLVTVDPA FRTRGNANTVWVDYPNIVRVVPVGGRIYIDDLISLVVQKIGPEGLV TQVENGVLGSRKGVNLPQAQVDLPGLSEQDVRDLRFGVEHGVDI VFASFVRKASDVAAVRAALGPEGHGIKISKIENHEGVKRFDEILEVS DGIMVARGDLGIEIPAEEKVFLAQKMMIGRCNLAGKPVVCATQMLES MITKPRPTRAETSDVANAVLDGADCIMLSGETAKGNFPVEAVKMQ HAIAREAEAAVYHRQLFEELRRAAPLSRDPTEVTAIGAVEAAFKCC AAAIIVLTGTRSAQLLSRYRPRAAVIAVTRSAQAARQVHLRCRGVFP LLYREPPEAIWADDVDRRVQFGIESG KLRGFLRVGDLVIVVTGWRPGSGYTNIMRVLSIS	PKLR

402	MSSSPVKRQRMESALDQLKQFTTVVADTGDFHAIDEYKPQDATTNP SLILAAAQMPAYQELVEEAIAYGRKLGGSQEDQIKNAIDKLFVLFGA EILKKIPGRVSTEV DARLSFDKDAMVARARRLIELYKEAGISKDRILI KLSSTWEGIQAGKELEEQHGHCNMTLLFSFAQAVACAEAGVTLISP FVGRILDWHVANTDKKSYEPLEDPGVKSVTKIYNYKKFSYKTIVM GASFRNTGEIKALAGCDFLTISP KLLGELLQDNAKLVPVLSAKAAQA SDLEKIHLDEKSFRWLHNEDQMAVEKLSDGIRKFAADAVKLERML TERMFNAENGK	TALDO1
403	MRLAVGALLVCAVLGLCLAVPDKTVRWCAVSEHEATKCQSFRDH MKSVIPSDGPSVACVKKASYLDCIRAIANEADAVTL DAGLVYDAY LAPNNLKPVVAEFYGSKEDPQTFYYAVAVVKKDSGFQMNQLRGK KSCHTGLGRSAGWNIPIGLLYCDLPEPRKPLEKAVANFFSGSCAPCA DGTDFPQLCQLCPGCGCSTLNQYFGYSGAFKCLKDGAGDVA FVKH STIFENLANKADRDQYELLCLDNTRKPVDEYKDCHLAQVPSHTVVA RSMGGKEDLIWELLNQAQEHFGKDKSKEFQLFSSPHGKDLLFKDSA HGFLKVPPRMDAKMYLGYEYVTAIRNLREGTCPEAPTDECKP VKWCALSHHERLKCDEWSVNSVGKIECVSAETTEDCIAKIMNGEA DAMSLDGGFVYIAGKCGLVPVLAENYNKSDNCEDTPEAGYFAIAV VKKSASDLTWDNLKGKKSCHTAVGRTAGWNIPMGLLYNKINHCRF DEFFSEGCAPGSKKDSCLKCMGSGNLCEPNNKEGYGYTGAFR CLVEKGDVAFVKHQTPQNTGGKNPDPWAKNLNEKDYELLCLDG TRKPVEEYANCHLARAPNHAVVTRKDKACVHKILRQQQHLFGSN VTDCSGNFCLFRSETKDLLFRDDTVCLAKLHNRNTYEKYLGEYVVK AVGNLRKCSTSSLLEACTFRRP	TF
404	MAPQVLAFLGLLLAAATATFAAAQEECVCENYKLAVNCFVNNNRQ CQCTSVGAQNTVICSKLAAKCLVMKAEMNGSKLGRRAKPEGALQN NDGLYDPDCDESLFKAKQCNGTSMCWCVNTAGVRRTDKDEITC SERVRTYWIHIELKHKAREKPYDSKSLRTALQKEITTRYQLDPKFITSI LYENNVITIDL VQNSSQKTQNDVDIADVA YFEKDVKGESLFHSHK MDLTVNGEQLDLDPGQTLIYVDEKAPEFSMQGLKAGVIAVIVVVV IAVVAGIVVLVISRKKRMAKYEKA EIKEMGEMHRELNA	EPCAM
405	MPRAENWDEAEVGAEEAGVEEYGPPEEDGGEEESGAEEESGPESGPE ELGAEEMEAGRPRLRSVNSREPSQVIFCNRSRVLPLVWLNFD GEPQPYPTLPPGTGRRIHSYRGHLWLFRDAGTHDGLLVNQTELFVPS LNVGDGQPIFANITLPVYTLKERCLQVVRSLVKPENYRRLDIVRSLYE DLEDHPNVQKDLERLTQERIAHQRMGD	VHL
406	MKRVLVLLLAVAFGHALERGRDYENKVKCKEFSHLGKEDFTSLSL VLYSRKFPSGTFEQVSQLVKEVVSLEACCAEGADPDCYDTRTSAL SAKSCESNSPFPVHPGTAECCTKEGLERKLCMAALKHQPFPTTYV EPTNDEICEAFRKDPKEYANQFMWEYSTNYGQAPLSLLVSYTKSYL SMVGSCCTSASPTVCFLKERLQLKHLSTLTLNVRVCSQYAA YGEK KSRLSNLIKLAQKVPTADLEDVLPLAEDITNILSKCCESASEDCMAK ELPEHTVKLCDNLSTKNSKFEDCCQEKAMDV FVCTYFMPAAQLPE LPDVELPTNKDVCDPGNTKVMKDYTFELSRRTLPEVFLSKVLEPT LKSLGECCDVEDSTTCFNAKGPLLKELSSFIDKGQELCADYSENTF TEYKKKLAERLKA KL PDATPTELAKLVNKHSDFASNCCSINSPLYC DSEIDAELKNIL	GC
407	MPSSVSWGILLLAGLCLLPVSLAEDPQGDAQAQKTDTSHHQDQDHT FNKITPNLAFAFSLYRQLAHQSNTNIFFSVSIATAFAMLSLGTKA DTHDEILEGLNFNLTEIPEAQIHEGFQELLRTL NQPDSQLQLTTGNGL FLSEGLKLVDKFLEDVKKLYHSEAFVTNFGDTEEAKKQINDYVEKG	SERPINA1

	TQGKIVDLVKELDRDTVFALVNYIFFKGKWERPFVVKDTEEDFHV DQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFFLP DEGKLQHLENELTHDIITKFLNEDRRSASLHLPKLSITGTYDLKSVL GQLGITKVFNSGADLSGVTEEAPLKLSKAVHKAVLTIDEKGTEAAG AMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK	
408	MAAPAEP CAGQGVWNQTEPEPAATSLLSLCFLRTAGVWVPPMYL WVLGPYLLFIHHHGRGYLRMSPLFKAKMVLGFALIVLCTSSVAVA LWKIQQTPEAPEFLIHPTVWLTTMSFAVFLIHTERKKGVQSSGVLF GYWLLCFVLPATNAAQQASGAGFQSDPVRHLSTYLCLSLVVAQFV LSCLADQPPFFPEDPQQSNPCPETGAAPSKATFWWVSGLVWRGYR RPLRPKDLWSLGRENSSEELVSRLEKEWMRNRSAARRHNKAIAFKR KGGSGMKAPETEPFLRQEGSQWRPLL KAIWQVFHSTFLLGTLISLSDVFRFTVPKLLSLFLEFIGDPKPPAWKG YLLAVLMFLSACLQTLFEQQNMYRLKVLQMRLRSAITGLVYRKVL ALSSGSRKASAVGDVVNLVSVDVQRLTESVLYNLGLWLPLVWIVV CFVYLLQLLGPSALTAIAVFLSLLPLNFFISKRNHHQEEQMRQKDS RARLTSSILRNSKTIKFHGWEGAFDLRVLGIRGQELGALRTSGLFS VSLVSFQVSTFLVALVVFVAVHTLVAENAMNAEKAFTLTVLNINLK AQAFLPFSIHSVLQARVSFDRVLVTFCLLEVDPGVVDSSSSGSAAGK DCITIHSATFAWSQESPPCLHRINLTVPQGCLLAVVGPVGAGKSSLLS ALLGELSKVEGFVSIEGAVAYVPQEAUVQNTSVVENVCFGQELDPP WLERVLEACALQPDVDSFPEGIHTSIGEQGMNLSGGQKQRLSLARA VYRKAAYVLLDDPLAALDAHVGQHVFNQVIGPGGLLQGTTRILVT HALHILPQADWIIVLANGAIAEMGSYQELLQRKGALMCLLDQARQP GDRGEGETEPGTSTKDPGRGTSAGRRPELRRERSIKSVPEKDRTTSEA QTEVPLDDPDRAWPAGKDSIQYGRVKATVHLAYLRAVGTPLCLY ALFLFLCQQVASFCRGYWLWLWADDPVAGGQQTQAALRGGIFGLL GCLQAIGLFAASMAAVLLGGARASRLLFQRLLDVVRSPISFFERTPI GHLLNRFSKETDTVDVDIPDKLRSLMYAFGLLEVSLVVAVATPLA TVAILPLFLLYAGFQSLYVVSQCRLRLESASYSSVCSHMAETFQGS TVVRAF RTQAPFVAQNNARVDESQRISFPRLVADRWLAANVELLGNGLVFA AATCAVLSKAHLSAGLVGFSVSAALQVTQTLQWVVRNWDLENSI VSVERMQDYAWTPKEAPWRLPTCAAQPPWPQGGQIEFRDFGLRYR PELPLAVQGVSFKIHAGEKVGIVGRTGAGKSSLASGLLRLQEAAG GIWIDGVPIAHVGLHLTLRSRISIIPQDPILFPGSLRMNLDLLQEHSDEA IWAALETVQLKALVASLPGQLQYKCADRGEDLSVGQKQLLCLARA LLRKTQILILDEATAAVDPGTELQM QAMLGSWFAQCTVLLIAHRLRSVMDCARVLVMDKGQVAESGSPA QLLAQKGLFYRLAQESGLV	ABCC6
409	MQIELSTCFLLCLLRFCSATTRYYLGAVELSWDYMQSDLGELPVD ARFPPRVPKSFNFNTSVVYKKTFLFEFTDHLFNIAKPRPPWMGLLGP TIQAEVYDVTVITLKNMASHPVSLHAVGVSYWKASEGAHEYDDQTS QREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDL VKDLNSGLIGALLVCREGLAKEKTQTLHKFILLFAVFDEGKSWHSE TKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYW HVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDL GQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDD DLTDSEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWD YAPLV LAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTR EAIQHESGILGPLLYGEVGD TLLIIFKNQASRPYNIYPHGITDVRPLYS RRLPKGVKHLKDFILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYS FVNMERDLASGLIGPLLCYKESVDQQRGNQIMSDKRNVLFSVFDEN	F8

	<p>RSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSV CLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGE TVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYED SYEDISAYLLSKNNAIEPRSFQNSRHPSTRQKQFNATTIPENDIEKTD PWFAHRTMPKIQNVSSDLLMLLRQSPTPHGLSLSDLQEAKYETFS DDPS</p> <p>PGAIDSNNLSSEMTHFRPQLHHSMDMVFTPEGLQLRLNEKLGTTA ATELKKLDFKVSSTSNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDS QLDITLFGKKSSPLTESGGPLSLSEENNDSKLLSGLMNSQESSWGK NVSSTESGRLFKGKRAHGPAALLTKDNALFKVSISLLKTNKTSNNSAT NRKTHIDGPSLLIENSPSVWQNILESDTEFKKVTPLIHDRMLMDKNA TALRLNHMSNKTSSKNMEMVQQKKEGPIPPDAQNPDMSSFFKMLF LPESARWIQRTHGKNSLNSGQGPSKQLVSLGPEKSVEGQNFLSEKN KVVVGKGEFTKDVGLKEMVFPSSRNLFNLNLHENNTHNQEKK IQEEIEKKETLIQENVVLPQIHTVTGTKNFMKNLFLSTRQNVESYD GAYAPVLQDFRSLNDSTNRKHTAHFSKKGEEENLEGLGNQTKQI VEKYACTRISPNTSQQNFVTQSRKALKQFRLPLEETELEKRIIVDD TSTQWSKNMKHLTPSTLTQIDYNEKE</p> <p>KGAITQSPLSDCLTRSHSIPQANRSPPIAKVSSFPSIRPIYLTRVLFQD NSSHLPAASYRKKDSGVQESSHFLQGAKKNLNLAILTLEMTGDQR EVGSLGTSATNSVTYKKVENTVLPKPDLPKTSKGKVELLPKVHIYQK DLFPTETSNNGSPGHLDLVEGSLQGTGAIKWNEANRPGKVPFLRV ATESSAKTPSKLLDPLAWDNHYGTQIPKEEWSQEKSPKTAFAKKK DTILSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPPV LKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPR SFQKKTRHYFIAAVERLWDYGMSSSPHVLNRQAQSGSVPPQFKKVVF QEFTDGSFTQPLYRGELNEHLGLLPYIRAEVEDNIMVTFRNQASRP YSFYSSLISYEEDQRQGAEPKRFVKNPNETKTYFWKVQHMHMAPTKD EFDCKAWAYFSDVDLEKDVHSGLIGPLLCHTNTLNPAGHGRQVTV QEFALFTTIFDETKSWYFTENMERNRA</p> <p>PCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQIRRWYLLSM GSNENIHSIHFGHVFTVRKKKEEYKMALYNLYPGVFETVEMLPKA GIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITAS GQYQGWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIHGIKT QGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDS SGIKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGM ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNP KEWLQVDFQKTMKVTVTTQGVKSLLTSMYVKEFLISSSQDGHQW TLFFQNGKVVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIA LRMEVLGCEAQDLY</p>	
410	<p>MQRVNMIAMESPGLITICLLGYLLSAECTVFLDHENANKILNRPKRY NSGKLEEFVQGNLERECMEKCSFEEAREVFENTERTEFWKQYVD GDQCESNPCLNGGCKDDINSYECWCPFGFEGKNCELDVTENIKNG RCEQFCKNSADNKVVCSTEGYRLAENQKSCEPAVPFPCGRVSVSQ TSKLTRAETVFPDVEDYVNSTEAETILDNITQSTQSFNDFTRVVGGED AKPGQFPWQVVLNGKVDAFCGGSIVNEKWIVTAAHCVETGVKITV VAGEHNIEETEHEQKRNIRII</p> <p>PHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTNIFLKF GSGYVSGWGRVFKGRSALVLQYLRVPLVDRATCLRSTKFTIYN MFCAGFHEGGRDSCQGDGGPHVTEVEGTSFLTGIISWGEECAMKG KYGIYTKVSRYVNWIKETKLT</p>	F9
411	<p>MDPPRPALLALLPALLLLLLAGARAEEMLENVSLVCPKDATRF KHLRKYTYNYEAESSGVPGTADSRSATRINCKVELEVQQLCSFILK</p>	ApoB

<p>TSQCTLKEVYGFNPEGKALLKKTKNSEEFAAAMSRYELKLAIPGK QVFLYPEKDEPTYILNIRGHSALLVPPETEEAKQVLFLLDTVYGNCS THFTVKTTRKGNVATEISTERDLGQCDRFKPIRTGISPLALIKGMTRPL STLIS</p> <p>SSQSCQYTLDKRKHVAEAICKEQHLFLPFSYKNKYGMVAQVTQT LKLEDTPKINSRFFGEGTKKMGLAFESTKSTSPPKQAEAVLKTQLQEL KKLTISEQNIQRANLFNKLVELRGLSDEAVTSLLPQLIEVSSPITLQA LVQCGQPQCSTHILQWLKRVHANPLLDVVTYLVALIPEPSAQQLRE IFNMARDQRSRATLYALSHAVNNYHKTNPSTGTQELLDIANYLMEQI QDDCTGDEDYTYLILRVIGNMGQTMEQLTPELKSSILKCVQSTKPSL MIQKAAIQALRKMEPKDKD</p> <p>QEVLLQTFLLDDASPGDKRLAAYLMLMRSPSQAINKIVQILPWEQNE QVKNFVASHIANILNSEELDIQDLKKLVKEALKESQLPTVMDFRKFS RNYQLYKSVSLPSLDPASAKIEGNLIFDPNNYLPKESMLKTTLTAFG FASADLIEIGLEGKGFEPTLEALFGKQGFPPDSVNKALYWVNGQVP DGVSKVLVDHFGYTKDDKHEQDMVNGIMLSVEKLIKDLKSKEVPE ARAYLRILGEELGFASLHDLQLLGKLLLMGARTLQGIQPMIGEVIK GSKNDFFLHYIFMENAFELPTGAGLQLQISSSGVIAPGAKAGVKLEV ANMQAELVAKPSVSVEFVTNMGIIPDFARSGVQMNTNFFHESGLE AHVALKAGKLFIPSPKRPVKLLSGGNTLHLVSTTKTEVIPPLIENR QSWSVCKQVFPGLNYCTSGAYSNASSTDSASYPLTGDTRLELELR PTGEIEQYSVSATYELQREDRALVDTLKFTVQAEGAKQTEATMTFK YNRQSMTLSSSEVIPDFDVLGTILRVN</p> <p>DESTEGKTSYRLTLDIQNKKITEVALMGHLSCDTKEERKIKGVISIPR LQAEARSEILAHWSPAKLLLQMDSSATAYGSTVSKRVAWHYDEEKI EFEWNTGTNVDTKKMTSNFPVDLSYPKSLHMYANRLLDHRVPQT DMTFRHVGSKLIVAMSSWLQKASGSLPYTQTLQDHLNSLKEFNLQ NMGLPDFHIPENLFLKSDGRVKYTLNKNLSLKIEIPLPFGGKSSRDLK MLETVRTPALHFKSVGFHLPREFQVPTFTIPKLYQLQVPLLGVLDL STNVYSNLYNWSASYSGGNTST</p> <p>DHFSLRARYHMKADSVVDLLSYNVQGSGETTYDHKNTFTLSYDGS LRHKFLDSNIKFSHVEKLGNNPVSKGLLIFDASSSWGPMASVHLD SKKKQHLFVKEVKIDGQFRVSSFYAKGTYGLSCQRDPNTGRLNGES NLRFNSSYLQGTNQITGRYEDGTLSTSTDLQSGHIKNTASLKYENY ELTLKSDTNGKYKNFATSNKMDMTFSKQNALLRSEYQADYESLRF FSLLSGSLNSHGLELNADILGTDKINSGAHKATLRIGQDGISTSATTN LKCSLLVLENELNAELGLSGASMKLTTNGRFREHNAKFSLDGKAAL TELSLGSAYQAMILGVDSKNIFNFKVSQEGLKLSNDMMGSYAEMK FDHTNSLNIAGLSLDFSSKLDNIYSSDKFYKQTVNLQLQPYSLVTTL NSDLKYNALDLTNGKLRLEPLKLHVAGNLKGAYQNNIEIKHIY AIS SAALSASYKADTVAKVQGVFESHRLNTDIAGLASAIDMSTNYNSDS LHFSNVFRSVMAPFTMTIDAHTNGNGKLALWGEHTGQLYSKFLLK AEPLAFTFSDYKGSTSHHLVSRKSISAALEHKVSALLTPAEQTGTW KLKTQFNNNEYSQDLDAYNTKDKIGVELTGRTLADLTLLDSPIKVPL LLSEPINIIDALEMRDAVEKPQEFTIVAFVKYDKNQDVHSINLPFFET LQEYFERNRQTIIVVLENVQRNLKHINIDQFVRKYRAALGKLPQQA NDYLSNFSNWERQVSHAKEKLTAITKKYRITENDIQIALDDAKINFNE KLSQLQTYMIQFDQYIKDSYDLHDLKIAIANIIDEIIEKLKSLDEHYHI RVNLVKTIHDLHLFIENIDFNKSGSSTASWIQNVDTKYQIRIQIEKL QQLKRIHQNIDIQHLAGKQLQHIEAIDVRVLLDQLGTTISFERINDILE HVKHFFVINLIGDFEVAEKINAFRAKVHELIEREYVDQQIQVLMDKLV ELAHQYKLLKETIQKLSNVLQQVQIKDYFEKLVGFIDDAVKKLNELSF KTFIEDVNKFLDMLIKKLKSFYHQFVDETNDKIREVTQRLNGEIQ</p>	
---	--

	<p>LELPQKAEALKLFLEETKATVAVYLESLODTKITLIINWLQEALSSAS LAHMKAKFRETLEDTRDRMYQMDIQQELQRYLSLVGQVYSTLVTY ISDWWTAAKNLTDFAEQYSIQDWAKRMKALVEQGFTVPEIKTILG TMPAFEVSLQALQKATFQTPDFIVPLTDLRIPSVQINFKDLKNIKIPSR FSTPEFTILNTFHIPSFTIDFVEMKVKIIRTIDQMLNSELQWPVPDIYLR DLKVEDIPLARITLPDFRLPEIAIPEFIPTLNLNDFQVPDLHIPEFQLPH ISHTIEVPTFGKLYSILKIQSPLFTLDANADIGNGTTSANEAGIAASITA KGESKLEVLNDFQANALSNPKINPLALKESVKFSSSKYLRTTEHGSE MLFFGNAIEGKSNTVASLHTEKNTLELSNGVIVKINNQLTLDSENTKY FHKLNIPKLDSSQADLRNEIKTLLKAGHIAWTSSGKGSWKWACPR FSDEGTHESQISFTIEGPLTSFGLSNKINSKHLRVNQNLVYESGSLNFS KLEIQSQVDSQHVGHSVLTAKGMALFGEKAEFTGRHDAHLNGKV IGTLKNSLFFSAQPFEITASTNNEGNLKVRFPLRLTGKIDFLNNYALF LSPSAQQASWQVSARFNQYKYNQNFSAAGNNENIMEAHVGINGE ANLDFLNIPLTIPEMRLPYTHITPPLKDFSLWEKTGLKEFLKTTKQSF DLSVKAQYKKNKHRHSITNPLAVLCEFISQSIKSFDRHFENRNNAL DFVTKSYNETKIKFDKYKAEKSHDELPRTFQIPGYTVPVVNVEVSPF TIEMSAFGYVFPKAVSMPSFSILGSDVRVPSYTLILPSLELPVLHVPR NLKLSLPDFKELCTISHIFIPAMGNITYDFSFKSSVITLNTNAELFNQS DIVAHLSSSSSVIDALQYKLEGTTTLTRKRGLKLATASLSNKFVE GSHNSTVSLTTKNMEVSVATTTKAQIPILRMNFKQELNGNTKSKPT VSSSMEFKYDFNSSMLYSTAKGAVDHKLSLESLSYFSIESSTKGDV KGSVLSREYSGTIASEANTYLSKSTRSSVKLQGTSKIDDIWNLEVK ENFAGEATLQRIYSLWEHSTKNHLQLEGLFFTNGEHTSKATLELSPW QMSALV QVHASQPSSFHDFPDLGQEVALNANTKNQKIRWKNEVRIHSGSFQS QVELSNDQEKALHLDIAGSLEGHLRFLKNILPVYDKSLWDFLKLDVT TSIGRRQHLRVSTAFVYTKNPNNGYSFSIPVKVLADKFIIPGLKLNDLN SVLVMPTFHVPTDLQVPSCKLDFREIQIYKKLRTSSFALNLPPLPEV KFPEVDVLTKYSPEDSLIPFEITVPESQLTVSQFTLPKSVSDGIAAL DL NAVANKIADFELPTIIVPEQTIEIPSIKFSVPAGIVIPSFQALTARFEVDS PVYNATWSASLKNKADYVETVLDSTCSSTVQFLEYELNVLGTHKIE DGTLASKTKGTFAHRDFSAEYEEDGKYEGLQEWEGKAHLNIKSPAF TDLHLRYQDKKKGISTSAAASPAVGTVGMDMDDEDDDFSKWNFYYS QSSPDKKLTIKTEL RVRESDEETQIKVNWEEEAASGLLTSKDNVP KATGVLYDYVNKYHWEHTGLTLREVSSKLRRNLQNNAEWVYQGA IRQIDDIDVRFQKAASGTTGT YQEWKDKAQNLYQELLTQEGQASQGLKDNVFDGLVRVTQEFHM KVKHLIDSLIDFLNFRPFQFPGKPGIYTREELCTMFIREVGTVLSQVY SKVHNGSEILFSYFQDLVITLPFELRKHKLIDVISMYRELLKDLSKEA QEVFKAIQSLKTTEVLRNLQDLLQFIFQLIEDNIKQLKEMKFTYLINY IQDEINTIFSDYIPYVFKLLKENLCLNLHKFNEFIQNELQEASQELQQI HQY IMALREEYFDPSIVGWTVKYYELEEKIVSLIKNLLVALKDFHSEYIVS ASNFTSQLSSQVEQFLHRNIQEYLSILTDPDGKGKEKIAELSATAQEII KSQAIATKKIISDYHQFRYKLQDFSDQLSDYIEKFIAESKRIDLISI QNYHTFLIYITELLKKLQSTTVMNPYMKLAPGELTIIL</p>	
412	<p>MGTVSSRRSWWPLPLLLLLLLLLLPAGARAQEDDGDYEELVLAL RSEEDGLAEAPEHGTTATFHRCADPWRLPGTYVVVLKEETHLSQS ERTARRLQAQAARRGYLTKILHVFHGLLPGLVKMSGDLLELALKL PHVDYIEEDSSVFAQSIPWNLERITPPRYRADEYQPPDGGSLVEVYL LDTSIQSDHREIEGRVMVTD FENVPEEDGTRFHRQASKCDSHGTHL</p>	PCSK9

	AGVVSGRDAGVAKGASMRSLRVLNCQGKGTVSGTLIGLEFIRKSQ VQPVGPLVLLPLAGGYSRVLNAA CQRLARAGVVLVTAAGNFRDDACLYSPASAPEVITVGATNAQDQP VTLGTLGTNFGRCVDLFAPGEDIIIGASSDCSTCFVSQSGTSQAAAHV AGIAAMMLSAEPELTLAELRQRLIHFSKDVINEAWFPEDQRVLT LVAALPPSTHGAGWQLFCRTVWSAHS GPTRMATAVARCAPDEELL SCSSFSRSGKRRGERMEAQGGKLVCRAHNAFGGEGVYAIARCCLLP QANCSVHTAPPAEASMGTRVHCHQQGHVLTGCSSHWEVEDLGTH KPPVLRPRGQPNQCVGHREASIHASCCHAPGLECKVKEHGIPAPQE QVTVACEEGWTLTGCSALPGTSHVLGAYAVDNTCVVRSRDVSTTG STSEGAVTAVAICCRSRHLAQASQELQ	
413	MDALKSAGRALIRSPSLAKQSWGGRHRKLPENWTDRETLL MLFSLKYLGMTLVEQPKGEELSAAAIKRIVATAKASGKKLQKVTLK VSPRGILTDNLTNQLIENVSIYRISYCTADKMHDKVFAIYASQHNQ SLECHAFLCTKRKMAQAVTLTVAQAFKVAFEFWQVSKEEKEKRD ASQEGDVLGARQDCTPSLKS LVATGNLLDLEETAKAPLSTVSANT TNMDEVPRPQALSGSSVWELDDGLDEAFSRLAQSRTPNPQVLDTG LTAQDMHYAQCLSPVDWDKPDSSGTEQDDLF	LDLRAP1
414	MGDLSSLTPGGSMGLQVNRGSQSLEGAPATAPEPHSLGILHASYSV SHRVRPWWDTSCRQQWTRQILKDVSLYVESGQIMCILGSSGSGKT TLLDAMSGRLGRAGTFLGEVYVNGRALRREQFQDCFSYVLQSDTL LSSLTVRETLHYTALLAIRRGNPGSFQKKVEAVMAELSLSHVADR LIGNYSLGGISTGERRRVISAAQLLQDPKVMFLDEPTTGLDCMTANQ VVLVVELARRNRIVLTIHQPRSELFQLFDKIAILSFGELIFCGTPAEM LDFFNDGYPCEHSNPFDFYMDLTSVDTQSKEREIETSKRVQMIES AYKKSACHKTLKNIERMKHLKTLPMPVFKTKDSPGVFSKLGVLRR RVTRNLVRNKLAVITRLLQNLIMGLFLLFFVLRVRSNVLKGAIQDRV GLLYQFVGATPYTGMLNAVNLFPVLRVSDQESQDGLYQKWQMM LAYALHVLFPFSVATMIFSSVCYWTGLHPEVARFGYFSAALLAPH LIGEFLLTVLLGIVQNPINVSVALLSIAGVLVSGFLRNIQEMPIPF KIISYFTFQKYCEILVNEFYGLNFTCGSSNVSVTTNPMCAFTQGIQ FIEKTCPGATSRFTMNFLLIYSFIPALVILGIVVFKIRDHLISR	ABCG5
415	MAGKAAEERGLPKGATPQDTSGLQDRLFSSESNSLYFTYSGQPNT LEVRLDNYQVDLASQVPWFELAQFKMPWTSPSCQNSCELGIQNL FKVRSGQMLAIIGSSGCGRASLLDVITGRGHGGKIKSGQIWINQPS PQLVRKCVAHVRQHNQLLPNLTVRETLAFIAQMRLPRTFSQAQRDK RVEDVIAELRLRQCADTRVGNMYVRGLSGGERRRVSIGVQLLWNP GILILDEPTSGLDSTAHNLVKTL SRLAKGNRLVLISLHQPRSDIFRL DLVLLMTSGTPIYLGAQAQHMVQYFTAIGYPCPRYSNPADFYVDLTS DRRSREQELATREKAQSLAALFLEKVRDLDDFLWKAETKDLED CVESSTPLDTNCLPSPTKMPGAVQQFTTLIRRQISNDFRDLPTLLI HGAELCLMSMTIGFLYFGHSGIQLSFMDTAALLFMIGALIPFNVIL DVI SKCYSERAMLYYELEDGLYTTGPYFFAKILGELPEHCAYIIYGMPT YWLANLRPGLQPFLHFLLVWL VVFCRIMALAAAALLPTFHMASF FSNALYNSFYLAGGFMINLSSLWTPAWISKVSFLRWCFEGLMKIQ FSRRTYKMPPLGNLTIAVSGDKILSVMELDYPLIYIYLIVIGLSGGFM VLYYVSLRFIKQKPSQDW	ABCG8
416	MGPPGSPWQWVTLTLLGLLLPPAAPFWLLNVLPHTTPKAELSNHT RPVILVPGCLGNQLEAKLDKPDVVNWMCYRKTEFFTIWLDLNM F LPLGVDWIDNTRVVYNRSSGLVSNAPGVQIRVPGFGKTY SVEYLD SSKLAGYLHTLVQNLVNNGYVRDET VRAAPYDWRLEPGQQUEEY	LCAT

	RKLAGLVEEMHAAAYGKPVFLIGHSLGCLHLLYFLLRQPQAWKDRFI DGFISLGAPWGGSIKPMVLASGDNQGIPIIMSSIKLKEEQRITTSPW MFPSRMAWPEDHVFISTPSFNITGR DFQRRFADLHFEEGWYMWLQSRDLLAGLPAPGVEVYCLYGVGLPT PRTYIYDHGFPYTDVPVGLYEDGDDTVATRSTELCGLWQGRQPQPV HLLPLHGIQHLNMVFSNLTLEHINAILLGAYRQGPPASPTASPEPPPP E	
417	MKIATVSVLLPLALCLIQDAASKNEDQEMCHEFQAFMKNGKLFCEPQ DKKFFQSLDGMFINKCATCKMILEKEAKSQKRARHLARAPKATAP TELNCDDFKKGERDGFICPDYYEAVCGTDGKTYDNRCALCAENA KTGSQIGVKSEGECKSSNPEQDVCSAFRPFVRDGRLGCTRENDPVL GPDGKTHGNKCAMCAELFLKEAENAKREGETRIRRNAEKDFCKEY EKQVRNGRLFCTRESDPVRGPDGRMHGNKCALCAEIFKQRFSEENS KTDQNLGKAEEKTKVKREIVKLCSQYQNAKNGILFCTRENDPIRG PDGKMHGNLCSMCQAYFQAENEKKKAEARARNKRESGKA TSYAEELCSEYRKLVRNGKLACTRENDPIQGPDGKVHGNTCSMCEVF FQAEEEEKKKEGKSRNKRQSKSTASFEELCSEYRKS RNKNGRLFCT RENDPIQGPDGKMHGNNTCSMCEAFFQQEERARAKAKREAAKEICSE FRDQVRNGTLICTREHNPVRGPDGKMHGNKCAMCASVFKLEEEEEK KNDKEEKKGKVEAEKVKREAVQELCSEYRHYVRNGRLPCTRENDPI EGLDGKIHGNTCSMCEAFFQQAKEKERAEPRAKVKREAEKETCDE FRLLQNGKLFCTRENDPVRGPDGKTHGNKCAMCKAVFQKENEER KRKEEDQRNAAGHGSSGGGGGNTQDECAEYREQMKNRGLS CTRESDPVRDADGKSYNNQCTMCKAKLERAERKNEYSRSRNGT GSESGKDTCDFRSQMKNKGKLICTRESDPVRGPDGKTHGNKCTMC KEKLERAEEKKKKKEDEDRSNTGERSNTGERSNDKEDLCREFRSM QRNGKLICTRENPNVRGPYGYKMHINKCAMCQSIFDREANERKKKD EEKSSSKPSNNAKDECSEFRNYIRNNELICPRENDPVHGADGKFYTN KCYMCRAVFLTEALERAKLQEKPSHVRASQEEDSPDSFSSLDSEMC KDYRVLPRIGYLCPKDLKPVCGDDGQTYNNPCMLCHENLIRQTNTH IRSTGKCEESSTPGTTAASMPPSDE	SPINK5
418	MEKNGNNRKLVRVCATCNRADYSKLAPIMFGIKTEPEFFELDVVVL GSHLIDDYGNTYRMIEQDDFDINTRLHTIVRGEDEAAMVESVGLAL VKLPDVLNRLKPDIMIVHGDRFDALALATSAALMNIRILHIEGGEVS GTIDDSIRHAITKLAHYHVCCTRSAEQHLISMCDHDRILLAGCPSY DKLLSAKNKDYMSIIRMWLGDDVKSKDYIVALQHPVTTDIKHSIKM FELTLDALISFNKRTLVLFPNIDAGSKEMVRVMRKKGIEHHPNFRAV KHVPFDQFIQLVAHAGCMIGNSSCGVREVGAFGTPVINLGTRQIGRE TGENVLHVRDADTQDKILQALHLQFGKQYPCSKIYGDGNAVPRILK FLKSIDLQEPLQKKKFCFPVKENISQDIDHILETSLAVALDLGGTNLR VAIVSMKGEIVKKYTQFNPKTYEERINLILQMCVEAAAEAVKLNCRIL LGVGISTGGRVNPREGIVLHSTKLIQEWNSVDLRTPLSDTLHLPVWV DNDGNCAALAEKRFQGGKLENFVTL ITGTGIGGGIIHQHELIHGSSFCAELGHLVVS LDGPDSCSGSHGCIE AYASGMALQREAKKLHDEDLLLVEGMSVPKDEAVGALHLIQAACL GNAKAQSILRTAGTALGLGVVNILHTMNPSLVILSGVLASHYIHIVK DVIRQQALSSVQDQDVVVSDLVDPALLGAASMVLDTTTRRIY	GNE
419	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLWYQQKPDGTVKLL IYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDIATYFCQQGNTLP YTFGGGKLEITGSTSGSGKPGSGEGSTKGEVKLQESGPGLVAPSQS LSVTCTVSGVSLPDYGVSWIRQPPRKGLEWLGVWGETTYNSAL KSRLTIKDNSKSQVFLKMNSLQTDDAIYYCAKHYYYGGSYAMD YWGQGTSTVTVSS	Anti-CD19 scFv (FMC63)

420	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPDGTVKLL IYHTSRLHSGVPSRFSGSGSGTDYSLTISNLEQEDIATYFCQQGNTLP YTFGGGKLEITGGGSGGGGSGGGGSEVKLQESGPGLVAPSQSL VTCTVSGVSLPDYGVSWIRQPPRKGLEWLGVWIGSETTYNSALKS RLTIKDNSKSQVFLKMNSLQTD DTAIYYCAKHYYYGGSYAMDYWG GQGTSVTVSS	Anti-CD19 scFv (FMC63)
421	ESKYGPPCPPCP	IgG4 Hinge
422	TTTPAPRPPTPAPTIASQPLSLRPE	CD8 Hinge
423	IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPFPGPSKP	CD28
424	ACRPAAGGAVHTRGLDFACDIYIWAPLAGTCGVLLLSLVITLYC	CD8
425	FWVLVVVGGLACYSLLVTVAFIIFWV	CD28
426	FWVLVVVGGLACYSLLVTVAFIIFWV	CD28
427	RSKRSRLHSDYMNMTPRRPGPTRKHYPYAPPRDFAAYRS	CD28
428	KRGRKKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL	4-1BB
429	RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEM GGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGL YQGLSTATKDTYDALHMQALPPR	CD3zeta
430	RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPEM GGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGL YQGLSTATKDTYDALHMQALPPR	CD3zeta

CLAIMS

WHAT IS CLAIMED:

1. A targeted lipid particle, comprising:
 - (a) a lipid bilayer enclosing a lumen,
 - (b) a henipavirus F protein molecule or biologically active portion thereof; and
 - (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) single domain antibody (sdAb) variable domain, wherein the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof and/or wherein the sdAb is attached to the G protein or the biologically active portion thereof via a peptide linker, wherein the sdAb binds to a cell surface molecule of a target cell,
wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.
2. The targeted lipid particle of claim 1, wherein the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule.
3. The targeted lipid particle of claim 1 or claim 2, wherein the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells.
4. The targeted lipid particle of any of claims 1-3, wherein the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

5. The targeted lipid particle of any of claims 1-4, wherein the single domain antibody binds to an antigen or portion thereof present on a hepatocyte.
6. The targeted lipid particle of any of claims 1-5, wherein the cell surface molecule or antigen is selected from the group consisting of ASGR1, ASGR2 and TM4SF.
7. The targeted lipid particle of any of claims 1-4, wherein the single domain antibody binds to an antigen or portion thereof present on a T cell.
8. The targeted lipid particle of any of claims 1-4 and 7, wherein the cell surface molecule or antigen is CD8 or CD4.
9. The targeted lipid particle of any of claims 1-4, wherein the cell surface molecule or antigen is LDL-R.
10. A targeted lipid particle, comprising:
 - (a) a lipid bilayer enclosing a lumen,
 - (b) a henipavirus F protein molecule or biologically active portion thereof; and
 - (c) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, TM4SF5, CD8, CD4 and LDL-R, wherein the F protein molecule or the biologically active portion thereof and the targeted envelope protein are embedded in the lipid bilayer.
11. The targeted lipid particle of claim 10, wherein the binding domain is attached to the G protein via a linker.
12. The targeted lipid particle of claim 11, wherein the linker is a peptide linker.

13. The targeted lipid particle of any of claims 1-12, wherein the lipid particle is a lentiviral vector.

14. A lentiviral vector , comprising:

- (a) a henipavirus F protein molecule or biologically active portion thereof; and
- (b) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds CD4; and
- (c) a cargo comprising nucleic acid encoding a chimeric antigen receptor (CAR), wherein the CAR comprises (i) an extracellular antigen binding domain that binds CD19, optionally wherein the extracellular antigen binding domain is an scFv, (ii) a transmembrane domain and (iii) an intracellular signaling region comprising a CD3zeta signaling domain.

15. The lentiviral vector of claim 14, wherein the intracellular signaling region of the CAR further comprises a 4-1BB co-stimulatory signaling domain.

16. The lentiviral vector of claim 14 or claim 15, wherein the lentiviral vector is capable of delivering the nucleic acid encoding the CAR to T cells, optionally wherein the T cells are in vivo in a subject

17. A lentiviral vector , comprising:

- (a) a henipavirus F protein molecule or biologically active portion thereof; and
- (b) a targeted envelope protein comprising (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain, wherein the binding domain is attached to the C-terminus of the G protein or the biologically active portion thereof, and wherein the binding domain binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2 and TM4SF5.

18. The lentiviral vector of claim 17, wherein the lentiviral vector is capable of targeting hepatocytes.

19. The lentiviral vector of claim 17 and claim 18, further comprising an exogenous agent for delivery to hepatocytes and/or wherein the lentiviral vector is capable of delivering the exogenous agent to hepatocytes, optionally wherein the hepatocytes are in vivo in a subject.

20. The lentiviral vector of any of claims 14-19, wherein the binding domain is attached to the G protein via a linker, optionally wherein the linker is a peptide linker.

21. The targeted lipid particle of any of claims 10-13 or the lentiviral vector of any of claims 14-20, wherein the binding domain is a single domain antibody or is a single chain variable fragment (scFv).

22. The targeted lipid particle of any of claims 1-9, 12, and 13 or the lentiviral vector of claim 20 or claim 21, wherein the peptide linker comprises up to 65 amino acids in length, optionally 2 to 65 amino acids in length.

23. The targeted lipid particle of any of claims 1-9, 12, 13 and 22 or the lentiviral vector of any of claims 20-22, wherein the peptide linker is a flexible linker that comprises GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) or combinations thereof, or wherein the peptide linker comprises (GGS)_n, wherein n is 1 to 10; (GGGGS)_n (SEQ ID NO:42), wherein n is 1 to 10; or (GGGGGS)_n (SEQ ID NO:27), wherein n is 1 to 6.

24. The targeted lipid particle of any of claims 1-13, 22 and 23 or the lentiviral vector of any of claims 14-23, wherein the G protein or the biologically active portion thereof is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein, or is a functionally active variant or biologically active portion thereof.

25. The targeted lipid particle of any of claims 1-13 and 22-24 or the lentiviral vector of any of claims 14-24, wherein the G protein or the biologically active portion thereof is a wild-type NiV-G protein or a functionally active variant or biologically active portion thereof.

26. The targeted lipid particle or lentiviral vector of claim 24 or claim 25, wherein the NiV-G protein or functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO: 9, SEQ ID NO: 28 or SEQ ID NO:44, or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:9,SEQ ID NO:28 or SEQ ID NO:44.

27. The targeted lipid particle or lentiviral particle of any of claims 24-26, wherein the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9,SEQ ID NO:28 or SEQ ID NO:44).

28. The targeted lipid particle or lentiviral particle of any of claims 24-27, wherein the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and has the sequence set forth in any of SEQ ID NOS: 10-15, 35-40 or 45-50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at least at or about 84%, at least at or about 85%, at least at or about 86%, at least at or about 87%, at least at or about 88%, at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40 or 45-50.

29. The targeted lipid particle or lentiviral vector of any of claims 24-28, wherein the NiV-G protein is a biologically active portion that has a truncation at or near the N-terminus of the wild-type NiV-G selected from the group consisting of a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 15 amino acid

truncation at or near the N-terminus, a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, or a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, optionally wherein the wild-type NiV-G protein is set forth in SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44.

30. The targeted lipid particle or lentiviral vector of any of claims 24-29, wherein the NiV-G protein is a biologically active portion that has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

31. The targeted lipid particle of any of claims 1-13, and 22-30 or the lentiviral vector of any of claims 14-30, wherein the G-protein or the biologically active portion thereof is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3.

32. The targeted lipid particle or lentiviral vector of claim 31, wherein the mutant NiV-G protein comprises:

one or more amino acid substitutions corresponding to amino acid substitutions selected from the group consisting of E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

33. The targeted lipid particle or lentiviral vector of claim 31 or claim 32, wherein the mutant NiV-G protein comprises:

amino acid substitutions E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

34. The targeted lipid particle or lentiviral vector of any of claims 31-33, wherein the mutant NiV-G protein or the biologically active portion has the amino acid sequence set forth in SEQ ID NO: 16 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or

about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16.

35. The targeted lipid particle of any of claims 1-13, and 22-34 or the lentiviral vector of any of claims 14-34, wherein the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a Hendra virus F protein or is a functionally active variant or biologically active portion thereof.

36. The targeted lipid particle of any of claims 1-13, and 22-35 or the lentiviral vector of any of claims 14-35, wherein the F protein or the biologically active portion thereof is a wild-type NiV-F protein or a functionally active variant or a biologically active portion thereof.

37. The targeted lipid particle of any of claims 1-13, and 22-36 or the lentiviral vector of any of claims 14-36, wherein the NiV-F-protein or the functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO: 2, or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 2.

38. The targeted lipid particle of any of claims 1-13, and 22-37 or the lentiviral vector of any of claims 14-37, wherein the NiV-F protein is a biologically active portion thereof that comprises:

- i) a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2); and/or
- ii) a point mutation on an N-linked glycosylation site.

39. The targeted lipid particle or lentiviral vector of claim 38, wherein the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:5 or SEQ ID NO:7 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 5 or SEQ ID NO: 7.

40. The targeted lipid particle of any of claims 1-13, and 22-37 or the lentiviral vector of any of claims 14-37, wherein the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2).

41. The targeted lipid particle or lentiviral vector of claim 40, wherein the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:8 or SEQ ID NO:23 or an amino acid sequence that is encoded by a sequence of nucleotides encoding a sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 8 or SEQ ID NO:23.

42. The targeted lipid particle of any of claims 1-13, and 22-41 or the lentiviral vector of any of claims 14-41, wherein the F protein comprises the sequence set forth in SEQ ID NO:23 and the G protein comprises the sequence set forth in SEQ ID NO:16.

43. The targeted lipid particle of any of claims 1-13, and 22-42, wherein the lipid bilayer is derived from a membrane of a host cell used for producing a retrovirus or retrovirus-like particle.

44. The targeted lipid particle of claim 43, wherein the host cell is selected from the group consisting of 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.

45. The targeted lipid particle of any of claims 1-13, and 22-44, wherein the lipid bilayer is or comprises a viral envelope.

46. The targeted lipid particle of any of claims 1-13 and 22-45, wherein the targeted lipid particle comprises one or more viral components other than the F protein molecule and the G protein.

47. The targeted lipid particle of claim 46, wherein the one or more viral components are from a retrovirus.

48. The targeted lipid particle of any of claims 43-47, wherein the retrovirus is a lentivirus or a lentivirus like particle.

49. The targeted lipid particle of any of claims 1-13 and 22-48 or the lentiviral vector of any of claims 14-42, wherein the lipid particle or lentiviral vector further comprises an exogenous agent.

50. The targeted lipid particle or lentiviral vector of claim 49, wherein the exogenous agent is present in the lumen.

51. The targeted lipid particle or lentiviral vector of claim 49 or claim 50, wherein the exogenous agent is a protein or a nucleic acid, optionally wherein the nucleic acid is a DNA or RNA.

52. The targeted lipid particle or lentiviral vector of any of claims 49-51, wherein the exogenous agent is a nucleic acid encoding a cargo for delivery to the target cell.

53. The targeted lipid particle or lentiviral vector of any of claims 49-52, wherein the exogenous agent encodes a therapeutic agent or a diagnostic agent.

54. The targeted lipid particle or lentiviral vector of any of claims 49-53, wherein the exogenous agent encodes a membrane protein, optionally wherein the membrane protein is an antigen receptor for targeting cells expressed by or associated with a disease or condition.

55. The targeted lipid particle or lentiviral vector of claim 53, wherein the membrane protein is a chimeric antigen receptor (CAR).

56. The targeted lipid particle or lentiviral vector of any of claims 52-55, wherein the target cell is a T cell.

57. The targeted lipid particle or lentiviral vector of any of claims 49-53, wherein the exogenous agent is a nucleic acid comprising a payload gene for correcting a genetic deficiency, optionally a genetic deficiency in the target cell, optionally wherein the genetic deficiency is associated with a liver cell or a hepatocyte.

58. A polynucleotide comprising a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a single domain antibody (sdAb) variable domain, wherein the sdAb variable domain is attached to the C-terminus of the G protein or the biologically active portion thereof.

59. The polynucleotide of claim 58, wherein the single domain antibody binds a cell surface molecule present on a target cell.

60. The polynucleotide of claim 59, wherein the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule.

61. The polynucleotide of claim 59 or claim 60, wherein the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells.

62. The polynucleotide of any of claims 59-61, wherein the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

63. The polynucleotide of any claims 59-62, wherein the cell surface molecule or antigen is selected from the group consisting of ASGR1, ASGR2, TM4SF5, CD8, CD4, and low density lipoprotein receptor (LDL-R).

64. A polynucleotide comprising a nucleic acid sequence encoding (i) a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and (ii) a binding domain that binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, TM4SF5, CD4, CD8, and low density lipoprotein receptor (LDL-R).

65. The polynucleotide of claim 64, wherein the binding domain is a single domain antibody (sdAb) or is a single chain variable fragment (scFv).

66. The polynucleotide of any of claims 58-65, wherein the nucleic acid sequence is a first nucleic acid sequence and the polynucleotide further comprises a second nucleic acid sequence encoding a henipavirus F protein molecule or a biologically active portion thereof.

67. The polynucleotide of claim 66, wherein the polynucleotide comprises an IRES or a sequence encoding a linking peptide between the first and second nucleic acid sequences, optionally, wherein the linking peptide is a self-cleaving peptide or a peptide that causes ribosome skipping, optionally a T2A peptide.

68. The polynucleotide of any of claims 58-67, further comprising at least one promoter that is operatively linked to control expression of the nucleic acid, optionally expression of the first nucleic acid sequence and the second nucleic acid sequence.

69. The polynucleotide of any of claims 58-68, wherein the sdAb variable domain or the binding domain is attached to the G protein via an encoded peptide linker.

70. The polynucleotide of claim 69, wherein the encoded peptide linker comprises up to 65 amino acids in length, optionally 2 to 65 amino acids in length.

71. The polynucleotide of claim 69 or claim 70, wherein the encoded peptide linker comprises GS, GGS, GGGGS (SEQ ID NO:43), GGGGGS (SEQ ID NO:41) and combinations thereof, or . wherein the encoded peptide linker comprises (GGS) n , wherein n is 1 to 10; (GGGGS) n (SEQ ID NO:42), wherein n is 1 to 10; or (GGGGGS) n (SEQ ID NO:27), wherein n is 1 to 4.

72. The polynucleotide of any of claims 58-71, wherein the encoded G protein is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein or is a functionally active variant or a biologically active portion thereof, optionally wherein the variant is a variant thereof that exhibits reduced binding for the native binding partner.

73. The polynucleotide of any of claims 58-72, wherein the encoded G protein is a wild-type NiV-G protein or a functionally active variant or a biologically active portion thereof.

74. The polynucleotide of claim 72 or claim 73, wherein the NiV-G protein or functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO:9, SEQ ID NO: 28 or SEQ ID NO: 44 or comprises an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44.

75. The polynucleotide of any of claims 72-74, wherein the NiV-G protein is a biologically active portion that is truncated and lacks up to 40 contiguous amino acid residues at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9, SEQ ID NO:28 or SEQ ID NO:44).

76. The polynucleotide of any of claims 72-75, wherein the NiV-G protein is a biologically active portion that is truncated at the N-terminus of wild-type NiV-G and comprises the sequence set forth in any of SEQ ID NOS: 10-15, 35-40 or 45-50 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NOS: 10-15, 35-40 or 45-50.

77. The polynucleotide of any of claims 72-76, wherein the NiV-G protein is a biologically active portion that has a truncation at or near the N-terminus of the wild-type NiV-G selected from the group consisting of a 5 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 10 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 15 amino acid truncation at or near the N-terminus, a 20 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 25 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, a 30 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, or a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein, optionally wherein the wild-type NiV-G protein is set forth in SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO: 44.

78. The polynucleotide of any of claims 72-77, wherein the NiV-G protein is a biologically active portion that has a 34 amino acid truncation at or near the N-terminus of the wild-type NiV-G protein (SEQ ID NO:9,SEQ ID NO:28 or SEQ ID NO:44).

79. The polynucleotide of any of claims 58-78, wherein the G-protein is a mutant NiV-G protein that exhibits reduced binding to Ephrin B2 or Ephrin B3.

80. The polynucleotide of claim 79, wherein the mutant NiV-G protein comprises: one or more amino acid substitutions corresponding to amino acid substitutions selected from the group consisting of E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

81. The polynucleotide of claim 79 or claim 80, wherein the mutant NiV-G protein comprises amino acid substitutions E501A, W504A, Q530A and E533A with reference to numbering set forth in SEQ ID NO:28.

82. The polynucleotide of any of claims 79-81, wherein the mutant NiV-G protein comprises the amino acid sequence set forth in SEQ ID NO: 16 or an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at

least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:16.

83. The polynucleotide of any of claims 66-82, wherein the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a Hendra virus F protein or is a functionally active variant or biologically active portion thereof.

84. The polynucleotide of any of claims 66-83, wherein the F protein or the biologically active portion thereof is a wild-type NiV-F protein or a functionally active variant or a biologically active portion thereof.

85. The polynucleotide of claim 83 or claim 84, wherein the NiV-F-protein or the functionally active variant or biologically active portion thereof comprises the amino acid sequence set forth in SEQ ID NO: 2, or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 2.

86. The polynucleotide of claim 85 or claim 85, wherein the NiV-F protein is a biologically active portion thereof that comprises:

- i) a 20 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2); and/or
- ii) a point mutation on an N-linked glycosylation site.

87. The polynucleotide of claim 86, wherein the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:5 or SEQ ID NO:7 or an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about

82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:5 or SEQ ID NO: 7.

88. The polynucleotide of claim 85 or claim 86, wherein the NiV-F protein is a biologically active portion thereof that has a 22 amino acid truncation at or near the C-terminus of the wild-type NiV-F protein (SEQ ID NO:2).

89. The polynucleotide of claim 88, wherein the NiV-F protein or the biologically active portion has the sequence set forth in SEQ ID NO:8 or SEQ ID NO:23 or an amino acid sequence that is encoded by a sequence of nucleotides encoding a sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 8 or SEQ ID NO:23.

90. The polynucleotide of any of claims 66-89, wherein the F protein comprises the sequence set forth in SEQ ID NO:23 and the G protein comprises the sequence set forth in SEQ ID NO:16.

91. A vector, comprising the polynucleotide of any of claims 58-90.

92. The vector of claim 91, wherein the vector is a mammalian vector, viral vector or artificial chromosome, optionally wherein the artificial chromosome is a bacterial artificial chromosome (BAC).

93. A plasmid, comprising the polynucleotide of any of claims 58-90.

94. The plasmid of claim 93, further comprising one or more nucleic acids encoding proteins for lentivirus production.

95. A cell comprising the polynucleotide of any of claims 58-90 or the vector of claim 91 or claim 92, or the plasmid of claim 93 or 94.

96. A method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain, the method comprising:

a) providing a cell that comprises a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and a nucleic acid encoding a targeted envelope protein, the targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain;

b) culturing the cell under conditions that allow for production of a targeted lipid particle, and

c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

97. A method of making a pseudotyped lentiviral vector, the method comprising:

a) providing a producer cell that comprises a lentiviral viral nucleic acid(s), a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof, and a nucleic acid encoding a targeted envelope protein, said targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody;

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

c) separating, enriching, or purifying the lentiviral vector from the cell, thereby making the pseudotyped lentiviral vector.

98. A method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a binding domain, the method comprising:

a) providing a cell that comprises a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and a nucleic acid encoding a targeted envelope protein, the targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and binding domain, wherein the binding domain:

(i) binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2;

(ii) binds a cell surface molecule selected from the group consisting of CD4 or CD8, optionally human CD4 or human CD8; or

(iii) binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R;

b) culturing the cell under conditions that allow for production of a targeted lipid particle, and

c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

99. A method of making a pseudotyped lentiviral vector, the method comprising:

a) providing a producer cell that comprises a lentiviral viral nucleic acid(s), a nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof, and a nucleic acid encoding a targeted envelope protein, said targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and binding domain, wherein the binding domain:

(i) binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2;

(ii) binds a cell surface molecule selected from the group consisting of CD4 or CD8, optionally human CD4 or human CD8; or

(iii) binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R;

b) culturing the producer cell under conditions that allow for production of a lentiviral vector, and

c) separating, enriching, or purifying the lentiviral vector from the cell, thereby making the pseudotyped lentiviral vector.

100. The method of claim 98 or claim 99, wherein the binding domain is a single domain antibody or is a single chain variable fragment (scFv).

101. A method of making a targeted lipid particle comprising a henipavirus F protein molecule or biologically active portion thereof and a targeted envelope protein, comprising:

a) providing a cell that comprises the polynucleotide of any of claims 58-90 or the vector of claim 91 or claim 92, or the plasmid of claim 93 or claim 94;

b) culturing the cell under conditions that allow for production of a targeted lipid particle, and

c) separating, enriching, or purifying the targeted lipid particle from the cell, thereby making the targeted lipid particle.

102. A method of making a pseudotyped lentiviral vector, comprising:

a) providing a producer cell that comprises a lentiviral viral nucleic acid(s), and the polynucleotide of any of claims 58-90 or the vector of claim 91 or claim 92 or the plasmid of claim 93 or claim 94;

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

c) separating, enriching, or purifying the lentiviral vector from the cell, thereby making the pseudotyped lentiviral vector.

103. The method of claim 101 or claim 102, wherein prior to step (b) the method further comprises providing the cell a polynucleotide encoding a henipavirus F protein molecule or biologically active portion thereof.

104. The method of claim 96-103, wherein the cell is a mammalian cell.

105. The method of any of claims 96-104, wherein the cell is a producer cell comprising viral nucleic acid, optionally retroviral nucleic acid or lentiviral nucleic acid, and the targeted lipid particle is a viral particle or a viral-like particle, optionally a retroviral particle or a retroviral-like particle, optionally a lentiviral particle or lentiviral-like particle.

106. A producer cell comprising the polynucleotide of any of claims 58-90 or the vector of claim 91 or claim 92 or the plasmid of claim 93 or claim 94.

107. The producer cell of claim 106, further comprising nucleic acid encoding a henipavirus F protein or a biologically active portion thereof.

108. The producer cell of claim 106 or claim 107, wherein the cell further comprises a viral nucleic acid, optionally wherein the viral nucleic acid is a lentiviral nucleic acid.

109. A producer cell comprising (i) a viral nucleic acid(s) and (ii) nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and (iii) a nucleic acid encoding a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a single domain antibody (sdAb) variable domain, optionally wherein the viral nucleic acid(s) are lentiviral nucleic acids.

110. The producer cell of claim 109, wherein the single domain antibody binds a cell surface molecule present on a target cell.

111. The producer cell of claim 110, wherein the cell surface molecule is a protein, glycan, lipid or low molecular weight molecule.

112. The producer cell of claim 110 or claim 111, wherein the target cell is selected from the group consisting of tumor-infiltrating lymphocytes, T cells, neoplastic or tumor cells, virus-infected cells, stem cells, central nervous system (CNS) cells, hematopoietic stem cells (HSCs), liver cells or fully differentiated cells.

113. The producer cell of any of claims 110-112, wherein the target cell is selected from the group consisting of a CD3+ T cell, a CD4+ T cell, a CD8+ T cell, a hepatocyte, a haematopoietic stem cell, a CD34+ haematopoietic stem cell, a CD105+ haematopoietic stem cell, a CD117+ haematopoietic stem cell, a CD105+ endothelial cell, a B cell, a CD20+ B cell, a CD19+ B cell, a cancer cell, a CD133+ cancer cell, an EpCAM+ cancer cell, a CD19+ cancer cell, a Her2/Neu+ cancer cell, a GluA2+ neuron, a GluA4+ neuron, a NKG2D+ natural killer cell, a SLC1A3+ astrocyte, a SLC7A10+ adipocyte, or a CD30+ lung epithelial cell.

114. A producer cell comprising (i) a viral nucleic acid(s) and (ii) nucleic acid encoding a henipavirus F protein molecule or biologically active portion thereof and (iii) a nucleic acid encoding a targeted envelope protein comprising a henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof and a binding domain, wherein the binding domain:

(i) binds a cell surface molecule selected from the group consisting of ASGR1, ASGR2, and TM4SF5, optionally human ASGR1, human ASGR2 and human ASGR2;

(ii) binds a cell surface molecule selected from the group consisting of CD4 or CD8, optionally human CD4 or human CD8; or

(iii) binds a cell surface molecule that is low density lipoprotein receptor (LDL-R), optionally human LDL-R;

optionally wherein the viral nucleic acid(s) are lentiviral nucleic acid.

115. The producer cell of claim 114, wherein the binding domain is a single domain antibody or is a single chain variable fragment.

116. The producer cell of any of claims 109-115, wherein the F protein or the biologically active portion thereof is a wild-type Nipah virus F (NiV-F) protein or a Hendra virus F protein or is a functionally active variant or biologically active portion thereof.

117. The producer of any of claims 109-116, wherein: the F protein or the biologically active portion thereof is a wild-type NiV-F protein or a functionally active variant or a biologically active portion thereof.

118. The producer cell of any of claims 109-117, wherein the henipavirus F protein molecule or biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 2, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO: 8 or SEQ ID NO:23;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO:2, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO: 8 or SEQ ID NO:23.

119. The producer cell of any of claims 109-118, wherein the encoded G protein is a wild-type Nipah virus G (NiV-G) protein or a Hendra virus G protein or is a functionally active variant or a biologically active portion thereof, optionally wherein the variant is a variant thereof that exhibits reduced binding for the native binding partner.

120. The producer cell of any of claims 109-119, wherein the encoded G protein is a wild-type NiV-G protein or a functionally active variant or a biologically active portion thereof.

121. The producer cell of any of claims 109-120, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO:44;

(ii) an amino acid sequence having at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at least at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 9, SEQ ID NO:28 or SEQ ID NO:44.

122. The producer cell of any of claims 109-121, wherein the henipavirus envelope attachment glycoprotein G (G protein) or a biologically active portion thereof comprises:

(i) the sequence set forth in SEQ ID NO: 10-16, 35-40 or 45-51;

(ii) an amino acid sequence having at least at or about 80%, at least at or about 81%, at least at or about 82%, at least at or about 83%, at or about 84%, at least at or about 85%, at least at or about 86%, or at least at or about 87%, at least at or about 88%, or at least at or about 89%, at or about 90%, at least at or about 91%, at least at or about 92%, at least at or about 93%, at least at or about 94%, at least at or about 95%, at least at or about 96%, at least at or about 97%, at least at or about 98%, or at least at or about 99% sequence identity to SEQ ID NO: 10-16, 35-40 or 45-51.

123. The producer cell of any one of claims 109-122, wherein:

the producer cell has greater membrane (e.g., plasma membrane) expression of the targeted envelope protein compared to a reference producer cell that has incorporated into its membrane (e.g. plasma membrane) the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting moiety is a single chain variable fragment (scFv); and/or

the producer cell has the expression of the targeted envelope protein on a membrane (e.g., plasma membrane) of the producer cell is at least 20 proteins (e.g., at least 50, 100, 200, 500, 1000, 2000, 5000, or 10,000 proteins) per square micron; and/or the targeted envelope protein comprises at least 0.1% (e.g., at least 0.2%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, or 10%) of the total membrane (e.g., plasma membrane) proteins of the producer cell (e.g., by total protein weight).

124. A targeted lipid particle or pseudotyped lentiviral vector produced by the method of any of claims 96-105 or from the producer cell of any of claims 106-122.

125. The targeted lipid particle of any of claims 1-13, 21-57 and 124, wherein the targeted lipid particle has greater expression of the targeted envelope protein compared to a reference lipid particle that has incorporated into a similar lipid bilayer the same envelope protein but that is fused to an alternative targeting moiety, optionally wherein the alternative targeting moiety is a single chain variable fragment (scFv).

126. The targeted lipid particle of any of claims 1-13, 21-57, 124 and 125, the lentiviral vector of any of claims 13-42, 49-57 and 124, wherein:

the titer in target cells following transduction is at or greater than 1×10^6 transduction units (TU)/mL, at or greater than 2×10^6 TU/mL, at or greater than 3×10^6 TU/mL, at or greater than 4×10^6 TU/mL, at or greater than 5×10^6 TU/mL, at or greater than 6×10^6 TU/mL, at or greater than 7×10^6 TU/mL, at or greater than 8×10^6 TU/mL, at or greater than 9×10^6 TU/mL, or at or greater than 1×10^7 TU/mL; and/or

the targeted envelope protein is present on the surface of the targeted lipid particle at a density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

127. A composition comprising a plurality of targeted lipid particles or a plurality of lentiviral vectors of any of claims 1-57 and 124-126.

128. The composition of claim 127 further comprising a pharmaceutically acceptable carrier.

129. The composition of claim 127 or claim 128, wherein:

among the population of lipid particles or lentiviral vectors in the composition, greater than at or about 50%, greater than at or about 55%, greater than at or about 60%, greater than at or about 65%, greater than at or about 70%, or greater than at or about 75% are surface positive for the targeted envelope protein; and/or

the targeted envelope protein is present on the surface of the targeted lipid particles at an average density of at least about (0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2 or 0.5) targeted envelope proteins/nm².

130. A method of transducing a cell comprising transducing a cell with a lentiviral vector of any of claims 13-42, 49-57 and 124 or a composition comprising a lentiviral vector or plurality of lentiviral vectors of any of claims 13-42, 49-57 and 124.

131. The method of claim 130, wherein:

the targeted envelope protein of the lentiviral vector or targeted lipid particle targets CD4 and the cell is a CD4+ cell; or

the targeted envelope protein of the lentiviral vector targets CD8 and the cell is a CD8+ cell; or

the targeted envelope protein of the lentiviral vector targets ASGR1, ASGR2 or TM4SF5 and the cell is a hepatocyte.

132. A method of delivering an exogenous agent to a subject (e.g., a human subject), the method comprising administering to the subject the targeted lipid particle or lentiviral vector of any of claims 49-57, 125 and 126, wherein targeted lipid particle or lentiviral vector comprise the exogenous agent.

133. A method of delivering an exogenous agent to a subject (e.g., a human subject), the method comprising administering to the subject the composition of any of claims 127-129, wherein targeted lipid particle or lentiviral vectors of the plurality comprise the exogenous agent.

134. A method of delivering a chimeric antigen receptor (CAR) to a cell, comprising contacting a cell with the lentiviral vector of any of claims 14-16 and 19-42 or a targeted lipid particle of claim 55 or claim 56, wherein the lentiviral vector or targeted lipid particle comprise nucleic acid encoding the CAR.

135. A method of delivering a chimeric antigen receptor (CAR) to a cell, comprising contacting a cell with the composition of any of claims 127-129 wherein lentiviral vectors or targeted lipid particles of the plurality comprise nucleic acid encoding the CAR.

136. A method of delivering an exogenous agent to a hepatocyte, comprising contacting a cell with the lentiviral vector of any of claims 17-42 or a targeted lipid particle or lentiviral vector of claim 57.

137. A method of delivering an exogenous agent to a hepatocyte, comprising contacting a cell with the composition of any of claims 127-129, wherein lentiviral vectors or targeted lipid particles of the plurality comprise an exogenous agent for delivery to the hepatocyte.

138. The method of any of claims 134-137, wherein the contacting transduces the cell with lentiviral vector or the targeted lipid particle, optionally wherein the contacting is in vivo in a subject.

139. A method of treating a disease or disorder in a subject (e.g., a human subject), the method comprising administering to the subject a targeted lipid particle of any of claims 1-13, 21-57, and 124-126 or the lentiviral vector of any of claims 13-42, 49-57, 124 and 126 or the composition of any of claims 127-129.

140. A method of fusing a mammalian cell to a targeted lipid particle, the method comprising administering to the subject a targeted lipid particle of any of claims 1-13, 21-57, and 124-126, the lentiviral vector of any of claims 13-42, 49-57, 124 and 126 or the composition of any of claims 127-129.

141. The method of claim 140, wherein the fusing of the mammalian cell to the targeted lipid particle delivers an exogenous agent to a subject (e.g., a human subject).