

PPT1 GENE THERAPY**RELATED APPLICATIONS**

[0001] The present application claims priority to U.S. Provisional Application No. 63/491,205, filed March 20, 2023, U.S. Provisional Application No. 63/584,000, filed September 20, 2023, and U.S. Provisional Application No. 63/600,153, filed November 17, 2023, each of which are hereby incorporated by reference herein in their entirety.

REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

[0002] The contents of the electronic sequence listing (065830.23WO1.xml; Size: 315,901 bytes; and Date of Creation: March 8, 2024) is herein incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

[0003] Palmitoyl Protein Thioesterase 1 (PPT1) is a glycoprotein involved in removing thioester-linked fatty acyl groups, such as palmitate from protein. Full-length PPT1 contains 306 amino acids and includes a signal sequence of 27 amino acids. Removal of the signal sequence produces a mature PPT1. (Bellizzi *et al.*, (2000) PNAS 97:9 4573-4578.)

[0004] PPT1 is found in lysosomes where it aids in lipid-modified protein catabolism. PPT1 is also found in other locations, such as synaptosomes, synaptic vesicles, blood and cerebrospinal fluid (CSF). PPT1 is encoded by the *CLN1* gene.

[0005] Neuronal Ceroid Lipofuscinosis 1 (NCL1) is a progressive neurodegenerative disease resulting from reduced PPT1 activity. NCL1 can be caused by mutations leading to abolished or reduced PPT1 activity or expression, resulting in a deficiency in lysosomal PPT1 activity. As the disease progresses symptoms can include epilepsy, seizures, motor and cognitive deterioration, visual impairment, behavioral disturbances, sleep disturbances, and death. NCL1 is also referred to as CLN1, Batten disease or Infantile Neuronal Ceroid Lipofuscinosis (INCL), and symptoms have an onset of about 12-18 months. In some cases, mutations in PPT1 can lead to a later onset of symptoms, such as late infantile (about 2-4 year), juvenile or adult onset. NCL1 pathology can occur in different locations including the brain and spinal cord. (Gorenberg *et al.*, (2022) PLoS Biol. 20(3): e3001590; Simonati and Williams (2022) Front. Neurol. Mar 11;13:811686; Shyng *et al.*, (2017) PNAS 114 (29) E5920-E5929; and Bellizzi *et al.*, (2000) PNAS 97(9):4573-4578.)

[0006] References mentioning potential enzyme replacement and gene therapy treatments for NCL1 include Nelvagil *et al.*, (2022) J. Clin. Invest. 132(20);e163107; Griffey *et al.*, (2004) Neurobiology of Disease 16:360-369; Griffey *et al.*, (2006) Molecular Therapy 13(3):538-547;

Shyng *et al.*, (2017) PNAS 114 (29) E5920-E5929; International Patent Publication No. WO2017/219450; and International Patent Publication No. WO2020/223322.

BRIEF SUMMARY OF THE INVENTION

[0007] The present invention features PPT1 polypeptides and encoding nucleic acid constructs. PPT1 polypeptides described herein include polypeptides comprising a PPT1 amino acid sequence and, in different embodiments, further comprise a signal sequence, where either or both the mature PPT1 or signal sequence is modified from that present in the full-length naturally occurring human PPT1. PPT1 encoding constructs include nucleic acid sequences encoding PPT1 polypeptides having one or more differences from the natural occurring human PPT1 mature or full-length sequence and/or comprise a CpG reduced sequence. Uses of the polypeptides and encoding nucleic acid constructs include producing PPT1 polypeptides; increasing PPT1 activity in a subject; and treating a PPT1 related disorder, such as CLN1 disease in a subject.

[0008] Reference to a PPT1 polypeptide indicates the presence of sequence related to the mature sequence and includes either or both the full-length sequence comprising a signal sequence or a mature sequence not containing a signal sequence. The mature sequence can vary in size depending on signal sequence and may be subject to further processing. The PPT1 polypeptide sequence may be naturally occurring or a modification of the naturally occurring sequence. Reference to the full-length naturally occurring human PPT1 indicates SEQ ID NO: 29. The mature naturally occurring human PPT1 sequence is provided by SEQ ID NO: 1.

[0009] Thus, a first aspect of the present invention describes a polynucleotide comprising a nucleic acid sequence encoding a palmitoyl-protein thioesterase-1 (PPT1) polypeptide, wherein the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1, wherein:

(a) the PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or

(b) the PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or

(c) the PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus; and/or

(d) the nucleic acid sequence comprises a PPT1 encoding sequence having at least 85% identity to any of SEQ ID NOs: 61-94.

[0010] The signal sequence of SEQ ID NOs: 16-27 provides signal sequences not present in naturally occurring full-length human PPT1.

[0011] Reference to a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus, indicates glycine (G), valine (V) or leucine (L) is present in the location corresponding to aspartic acid (D) in the native mature PPT1 sequence (SEQ ID NO: 1).

[0012] Another aspect of the present invention describes a PPT1 polypeptide comprising a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1, wherein:

(a) the PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or

(b) the PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or

(c) the PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus.

[0013] Another aspect of the present invention is directed to an expression cassette comprising a nucleic acid sequence encoding a PPT1 polypeptide and one or more expression control elements operably coupled to the encoding nucleic acid sequence.

[0014] Reference to one or more expression control elements “operably linked” or “operably coupled” to nucleic acid encoding a PPT1 polypeptide indicates that the expression control element(s) impacts PPT1 polypeptide expression. PPT1 polypeptide expression can be impacted in different ways, such as increased production of PPT1 polypeptide mRNA transcripts, increased nuclear transport and stability of mRNA transcripts, and increased mRNA translation.

[0015] Another aspect of the present invention is directed to a recombinant viral vector nucleic acid comprising (a) an expression cassette comprising a nucleic acid sequence encoding a PPT1 polypeptide and one or more expression control elements operably linked to the encoding nucleic acid sequence and (b) 5' and/or 3' viral elements providing for viral packaging and/or replication.

[0016] Another aspect of the present invention is directed to a delivery vehicle comprising a viral or a non-viral vector and (a) a PPT1 polypeptide or (b) a polynucleotide, expression cassette, or recombinant viral vector nucleic acid comprising a sequence encoding for a PPT1 polypeptide.

[0017] Another aspect of the present invention is directed to a pharmaceutical composition comprising (a) a PPT1 polypeptide; (b) a polynucleotide, expression cassette, or recombinant viral vector nucleic acid comprising a sequence encoding for a PPT1 polypeptide; or (c) a delivery vehicle comprising (a) or (b); and a pharmaceutically acceptable carrier.

[0018] Another aspect of the present invention is directed to a method of increasing PPT1 activity, treating a PPT1 disease or disorder, or treating CLN1, in a subject comprising administration of: (a) a PPT1 polypeptide; (b) a polynucleotide, expression cassette, or

recombinant viral vector nucleic acid comprising a sequence encoding for a PPT1 polypeptide; (c) a delivery vehicle comprising (a) or (b); or (d) a pharmaceutical composition comprising (a), (b), or (c) and a pharmaceutically acceptable carrier.

[0019] Additional aspects of the present invention include: (a) a PPT1 polypeptide; (b) a polynucleotide, expression cassette, or recombinant viral vector nucleic acid comprising a sequence encoding for a PPT1 polypeptide; (c) a delivery vehicle comprising (a) or (b); or (d) a pharmaceutical composition comprising (a), (b), or (c) and a pharmaceutically acceptable carrier; (a), (b), (c), or (d) for use in medicine, increasing PPT1 activity, treating a PPT1 disease of disorder or treating CLN1 in a subject; and use of (a), (b), (c), or (d) in the preparation of a medicament (*e.g.*, for use in medicine, increasing PPT1 activity, treating a PPT1 related disorder, or treating CLN1 in a subject).

[0020] Additional aspects of the present invention include vector genome plasmids comprising recombinant viral vector nucleic acid encoding a PPT1 polypeptide, methods of producing recombinant viral vector nucleic acid encoding a PPT1 polypeptide, and methods of obtaining a PPT1 polypeptide.

[0021] Other features and advantages of the present invention are apparent from additional descriptions provided herein, including different examples. The provided examples illustrate different components and methodology useful in practicing the present invention. Such examples do not limit the claimed invention. Based on the present disclosure, the skilled artisan can identify and employ other components and methodology useful for practicing the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] FIG. 1 is a schematic of a rAAV PPT1 expression construct (or cassette), comprising different nucleic acid regions; 5'-inverted terminal repeat (ITR), elongation factor-1 alpha (EF-1 α) promoter, Kozak sequence (Kozak), signal sequence (SS also referred to as signal peptide), mature PPT1 sequence (human PPT1), bovine growth hormone polyadenylation sequence (bGH-pA), stuffer (stuffer sequence), synthetic polyadenylation sequence (synthetic pA) and 3'-ITR.

[0023] FIG. 2A and FIG. 2B are bar graphs illustrating PPT1 enzymatic activity levels in PPT1 knock out HeLa cells transfected with rAAV plasmids carrying different engineered human PPT1 constructs. PPT1 activity is expressed as a percent of PPT1 activity from PPT1 having its native signal peptide in the total cell lysates (FIG. 2A) and in cell culture supernatants (secreted PPT1) (FIG. 2B). Reference to "Dead" indicates a mutated PPT1 lacking enzymatic activity.

[0024] FIGs. 3A-3C illustrate a rAAV PPT1 nucleic acid/tdTomato reporter plasmid, and results using the plasmid to evaluate PPT1 secretion from transfected cells. FIG. 3A shows different rAAV nucleic acid regions and tdTomato regions. FIG. 3B provides a bar graph showing the ratio of the number of cells having PPT1 to the number of cells transfected with the PPT1/tdTomato reporter plasmid. FIG. 3C provides a bar graph showing results from FIG. 3B, normalized for the native construct. For both FIG. 3B and 3C, each point (clear circle; average of 30 fields) is an independent transfection (from 2-3 experiments), and the values are mean $\pm$ SD. P values by one way ANOVA, **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$.

[0025] FIG. 4 illustrates dose and time dependent increased PPT1 activity detected in the media of rat primary cortical neurons transduced with rAAV comprising Sp7-F.PPT1. Cultured neurons were transduced with three multiplicities of infection (MOI) (low, $1E+5$; mid, $5E+5$ and high, $1E+6$) and PPT1 activity in the culture media was analyzed at 3-, 4- or 6-days post transduction. "Dead" PPT1 refers to catalytically inactive PPT1. Untreated or cells treated with diluent served as negative control. Purified recombinant PPT1 was used as a positive control for assay. Each circle represents the values from independently transduced cells in a single well. Data are mean $\pm$ SD.

[0026] FIG. 5A, FIG. 5B and FIG. 5C illustrate serum PPT1 expression and activity at different time points from mice dosed with rAAV comprising Sp7-F.PPT1 viral vector nucleic acid. FIG. 5A illustrates production of glycosylated PPT1 activity in the serum of mice dosed intravenously (IV). FIG. 5B shows serum PPT1 activity of mice dosed IV. FIG. 5C shows serum activity of mice dosed IPa (delivered intraparenchymally to the hippocampus through stereotactic injection). ND, not detected. Data are mean $\pm$ SD. P values by ANOVA * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$.

[0027] FIG. 6 illustrates liver PPT1 activity from mice dosed IV with rAAV comprising Sp7-F.PPT1 viral vector nucleic acid. Mice injected with diluent served as negative controls.

[0028] FIG. 7 illustrates an immunohistochemical analysis showing increased PPT1 staining (indicated by black asterisk) in a mouse hippocampus upon hippocampal administration of rAAV comprising Sp7-F.PPT1 viral vector nucleic acid.

[0029] FIG. 8 provides an automated capillary-based immunoassay (JESS, ProteinSimple) analysis demonstrating glycosylation of overexpressed PPT1 in mice brain. Recombinant AAV comprising Sp7-F.PPT1 viral vector nucleic acid or diluent was administered to the hippocampus (four mice per group).

[0030] FIG 9 provides an automated capillary-based immunoassay (JESS, ProteinSimple) analysis of hippocampal protein extract, untreated (-) or treated (+) with deglycosylase, from two different mice, following administration of rAAV comprising Sp7-F.PPT1 viral vector to

the hippocampus. Reduced molecular weight upon deglycosylase treatment (lanes marked with +) suggested PPT1 was glycosylated.

[0031] FIG. 10 illustrates PPT1 activity assay in mice hippocampal lysate showing PPT1 expression in mice brain. Recombinant AAV comprising Sp7-F.PPT1 viral vector nucleic acid was administered to the hippocampus.

[0032] FIG. 11 are bar diagrams illustrating vector genome copy number (VGCN) as an indicator of viral transduction in different brain regions. Total DNA from frozen tissue was isolated and the VGCN was quantified by quantitative PCR (qPCR) using a standard curve. Each circle is data from one mouse. The height of each bar shows the average. Hpc, hippocampus; Br. Stem, brain stem; crblm, cerebellum; Cerv, cervical spinal cord; Thor, thoracic spinal cord, lumb; lumbar spinal cord. Mice dosed by diluent served as a negative control.

[0033] FIG. 12 are bar diagrams showing fold change (FC) in PPT1 activity in different brain and spinal cord regions of rAAV-injected animals relative to diluent-injected mice. PPT1 activity in tissue lysate was quantified using 4-methylumbelliferyl-6-thiopalmityl- β -D-glucopyranoside (MUTG) as a substrate. A standard curve (prepared by using known concentrations of 4 MU) was used to determine PPT1 activity. FC activity was calculated relative to the mean activity of diluent-injected animals. Each circle is the result of one mouse. The data is Mean $\pm$ SD. One way ANOVA, Tukey's test, * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$.

[0034] FIG. 13 are bar diagrams illustrating FC in PPT1 activity in the cerebrospinal fluid (CSF) of mice subjected to rAAV injection compared to those injected with the diluent. The FC is relative to the average activity level of the diluent group. Each data point on the graph represents data from a single mouse. The data is expressed as Mean $\pm$ SD. The statistical analysis involved a one-way ANOVA followed by Tukey's test. * $P < 0.05$.

[0035] FIG. 14 is a scatter plot illustrating the correlation between VGCN and PPT1 enzymatic activity in the brain. The Spearman correlation coefficient was $r = 0.066$, $p < 0.0001$. Each circle is data from a brain region from mice dosed with rAAV.

[0036] FIG. 15 shows detection of glycosylated and deglycosylated PPT1 protein in brain lysate analyzed by JESS. Protein extract from brain lysate from the cortex of two independent mice (mouse one: lanes 1 and 3; mouse 2: lanes 2 and 4) administered rAAV comprising Sp7-F.PPT1 was treated with deglycosylase and analyzed by JESS assay. The doublet seen in lanes 1-2 indicates two glycosylated forms of PPT1. The reduced molecular weight upon deglycosylase treatment (lanes 3-4) suggests PPT1 was glycosylated.

[0037] FIGs. 16A-16H are bar diagrams depicting vector genome copy number (VGCN) as a measure of viral transduction, across various region: cortical areas (FIG. 16A); thalamus (FIG.

16B); cerebellar cortex (FIG. 16C); hippocampus (FIG. 16D); corpus callosum (FIG. 16E); additional indicated brain areas (FIG. 16F); spinal cord (FIG. 16G); and liver (FIG. 16H). Quantitative PCR (qPCR) utilizing a standard curve was employed to quantify the VGCN from frozen tissue DNA. Each data point represents a single sheep. Vertical arrows indicate sample from the contralateral brain region (compared to injection side). Open circles correspond to sheep dosed with rAAV comprising Sp7-F.PPT1, while filled circles depict results from sheep dosed with a vector expressing GFP. Data is provided for the following regions: Frontal Cortex (FC), Motor Cortex (MC), Somato sensory cortex (SSC), Piriform Cortex (PC), Suprasylvian Gyrus (SSG), Ectolateral Gyrus (EcG), Entolateral Gyrus (EnG), Caudate (Cau), Choroid plexus (Ch Ple), Optic chiasm (Opt chi), Fornix (For), Periaqueductal gray (Periaq G), Olfactory bulb (Ol), Optic nerve (Op nerve), Hippocampus (HPC), Thalamus (Tha), Corpus callosum (Cca), Cerebellar cortex (Cer ctx), Spinal cord_Cervical (SC_Cer), Spinal cord_Thoracic (SC_Tho), and Spinal cord_Lumbar (SC_Lum).

[0038] FIGs. 17A-17D are bar diagrams showing functional PPT1 expression and secretion in brain and CSF of sheep injected with rAAV comprising Sp7-F.PPT1 (n=4 sheep) or GFP (n=2 sheep) transgenes. FIG. 17A is a bar diagram illustrating PPT1 activity in the cortex. FIG. 17B is a bar diagram illustrating PPT1 activity the thalamus. FIG. 17C is a bar diagram illustrating PPT1 activity in the cerebellar cortex. FIG. 17D is a bar diagram illustrating PPT1 activity in the caudate. PPT1 activity in tissue lysate was quantified using 4-methylumbelliferyl-6-thiopalmityl- β -D-glucopyranoside (MUTG) as a substrate. Each circle is the result of one tissue punch from the brain region. N denotes number of regions from 2 GFP dosed animals and 4 from PPT1 dosed animals. The data is Mean $\pm$ SEM. Statistical analysis by Mann-Whitney U test, *P<0.05, *** P<0.001.

[0039] FIG. 18 provides 95% confidence intervals showing mean PPT1 activity by overall treatment group. After taking the log-transformed activity results of all data in FIGs. 17A-17D and accounting for numbers of punches, and brain region differences in average value, a hypothesis test for average Treatment Type difference was conducted at the 0.05 alpha level. The average estimated fold change in PPT1 over GFP was 3.9, and with medians of 73 nmoles/mg/hr for the PPT1 vs. 19 nmoles/mg/hr for the GFP. **** P<0.0001, Weighted 2-way ANOVA.

[0040] FIG. 19 is a bar diagram showing percent change in PPT1 activity in the cerebrospinal fluid of (CSF) of sheep dosed with rAAV carrying Sp7-F.PPT1 (n=4 sheep) or GFP (n=2 sheep). The percent changes is relative to mean activity of control (GFP animals). Each circle is one animal. N denotes number of animals. The data is Mean $\pm$ SEM.

[0041] FIG. 20 shows the result of a JESS assay detecting the expression of recombinant PPT1 in the tissue lysate of the spinal cord of sheep injected with an rAAV vector carrying either PPT1 (sheep 1-4) or GFP (sheep 1-2). The first lane shows the position of the molecular weight standard (std). While protein bands likely representing sheep PPT1 were observed in all subjects, bands for human PPT1 were solely identifiable in animals treated with a vector encoding human PPT1. The abbreviations C, T, and L denote the cervical, thoracic, and lumbar segments of the spinal cord, respectively. KDa, Kilo Dalton.

[0042] FIG. 21 illustrates results showing an elevated level of mean PPT1 activity in the thoracic and lumbar segment of the spinal cord of sheep dosed with PPT1 carrying vector. Each circle represents the result from the spinal cord segment of an animal. The data is Mean $\pm$ SEM.

[0043] FIG. 22 is a bar diagram illustrating a rotarod assessment of PPT1 knockout (KO) mice dosed with rAAV (1E+11 vg/animal) comprising Sp7-F.PPT1 (1) or SPARC.PPT1 (2). KO mice were dosed with rAAV through bilateral ICV injection at postnatal day 1 and assessed at 7 months of age. Untreated (Un) or vehicle (Veh) treated KO mice served as negative controls. Native refers to rAAV vector encoding native human PPT1. Each circle represent one mouse. The bars are mean $\pm$ SEM. One-way ANOVA, Tukey's post hoc test, *P<0.05, *** P<0.001, ****P<0.0001.

[0044] FIG. 23 illustrates the ability of rAAV-Sp7-F.PPT1 and rAAV-SPARC.PPT1 to improve motor coordination and balance as assessed by latency to fall from an accelerating rod in Ppt1^{-/-} mice (KO). The values are mean+SEM. One-way ANOVA, Tukey's post hoc test. # # # #p<0.0001, WT, 7 mo vs KO, Un, 7 mo; KO, Veh, 7mo. ****p<0.0001, KO, Veh, 7 mo vs all dosed groups except rAAV-Sp7-F.PPT1, hi at 9mo. **p<0.005, KO, Veh, 7 mo vs rAAV-Sp7-F.PPT1, hi at 9 months. AAV=adeno-associated virus; ANOVA=analysis of variance; CNS=central nervous system; lo=low dose (1x10¹¹ vg/animal); hi= high dose (3.82x10¹¹ vg/animal); KO= Ppt1^{-/-}; mo=months; PND=post-natal day sec=seconds; SEM=standard error of mean; Un=untreated; Veh=vehicle; WT=wild-type.

[0045] FIG. 24 illustrates the effect of rAAV-Sp7-F.PPT1 and rAAV- SPARC.PPT1 on grip strength in Ppt1^{-/-} mice (KO). The values are mean+SEM. One-way ANOVA, Tukey's post hoc test. # # # #p<0.0001, WT, 7 mo vs KO, Un, 7 mo; KO, Veh, 7 mo. ****p<0.0001, KO, Veh, 7 mo vs all treated groups. AAV=adeno-associated virus; ANOVA=analysis of variance; CNS=central nervous system; lo=low dose (1x10¹¹ vg/animal); hi= high dose (3.82x10¹¹ vg/animal); KO= Ppt1^{-/-}; mo=months; PND=post-natal day sec=seconds; SEM=standard error of mean; Un=untreated; Veh=vehicle; WT=wild-type.

[0046] FIGs. 25A and 25B illustrates serum PPT1 activity in mice administered rAAV-Sp7-F.PPT1 or rAAV-SPARC.PPT1. FIG. 25A illustrates activity at different time points. FIG. 25B

illustrates activity at 8 months. In each group, the number of mice varied from 11 to 18. Gender distribution was approximately balanced within each group. Each circle represent a result from one mouse. Each point in FIG. 25A and each bar in FIG. 25B shows mean+SEM. One-way ANOVA, Tukey's post hoc test done on Log10 transformed valued. # # # # $p < 0.0001$, WT vs KO, Un; WT vs KO Veh; ** $p < 0.01$, KO, Veh vs Native, lo; *** $p < 0.001$, KO, Veh vs rAAV-SPARC.PPT1, lo; *** $p < 0.0001$, KO, Veh vs rAAV- Sp7-F.PPT1, lo; Native, hi; rAAV- Sp7-F.PPT1, hi; rAAV-SPARC.PPT1, hi. AAV=adeno-associated virus; ANOVA=analysis of variance; lo=low dose (1×10^{11} vg/animal); hi= high dose (3.82×10^{11} vg/animal); KO= *Ppt1*^{-/-}; mo=months; LLOQ=lower limit of quantification; PND=post-natal day sec=seconds; SD=standard deviation; Un=untreated; Veh=vehicle; WT=wild-type.

[0047] FIGs. 26A-C provide bar diagrams showing PPT1 activity in the cortex (FIG. 26A), brainstem (FIG. 26B) and cerebellum (FIG. 26C) of *Ppt1*^{-/-} mice injected with rAAV-Sp7-F.PPT1 or rAAV-SPARC.PPT1. Mice were dosed through bilateral intra-cerebro ventricular injections at PND 1. 'Native' refers to AAV vectors containing the unmodified human *PPT1* gene. The bars are mean±SEM. Each circle represent a mouse.

[0048] FIG. 27 is a schematic showing the design of an rAAV nucleic acid, present in a plasmid. Expression of the human PPT1 sequence is driven by a longer version of the EF1a promoter (SEQ ID NO: 173). SS=Signal sequence.

[0049] FIGs. 28A and 28B are bar graphs depicting PPT1 expressed in PPT1 knockout HeLa cells transfected with different AAV plasmids carrying codon-optimized PPT1 sequences. Different codon-optimized PPT1 cDNA constructs (excluding the signal sequence), labeled as CO followed by the number on the x-axis, were paired with signal sequences. FIG. 28A depicts Sp7F or SP7F (codon optimized). FIG. 28B depicts SpSPARC, or SpSPARC (codon optimized).

[0050] FIG. 29 provides survival curves illustrating the ability of CNS-targeted AAV gene therapy delivering functional human PPT1 in *Ppt1*^{-/-} mice to extend survival.

[0051] FIG. 30 is a bar diagram illustrating the effect of rAAV comprising nucleic acid encoding Sp7F.PPT1 or SPARC.PPT1 on brain weight in *Ppt1*^{-/-} mice. Mean+SD.

*** $P < 0.0001$, One-way ANOVA, Tukey's post hoc. AAV=adeno-associated virus; CNS-central nervous system; lo=low dose (1×10^{11} vg/animal); hi= high dose (3.82×10^{11} vg/animal); KO= *Ppt1*^{-/-}; PND=post-natal day; Un=untreated; Veh=vehicle; WT=wild-type.

DETAILED DESCRIPTION OF THE INVENTION

[0052] The present invention features PPT1 polypeptides and nucleic acid constructs encoding for PPT1 polypeptides. The polypeptides and encoding nucleic acid constructs can be used, for

example, to produce PPT1 polypeptides, to increase PPT1 activity in a subject and/or to treat a PPT1 related disease or disorder, such as CLN1.

[0053] Polynucleotides encoding a PPT1 polypeptide can be delivered to a subject by non-viral or viral delivery. Viral vectors that can be used include retroviral vectors, adenovirus vectors, AAV vectors, and herpes simplex viral vectors. Non-viral delivery includes naked DNA and the use of nanoparticles.

[0054] Reference to “subject” indicates a mammal, such as a human; non-human primate such as ape, gibbon, gorilla, chimpanzee, orangutan, and macaque; domestic animal such as a dog and cat; farm animal such as poultry, duck, horse, cow, goat, sheep and pig; and experimental animal such as mice, rat, rabbit, sheep, and guinea pig. A preferred subject is a human.

[0055] In certain embodiments, PPT1 polypeptides or nucleic acid constructs encoding for PPT1 polypeptide are evaluated for PPT1 expression and efficacy using a sheep model. (See, for example, Nelvagal *et al.*, (2022) J. Clin. Invest. 132(20);e163107, hereby incorporated by reference herein in its entirety; and the Example section provided below.)

[0056] Reference to an indicated percent identity to one or more reference sequences, and similar language throughout the specification providing for an indicated percent identity to one or more reference sequences, provides the indicated percent identity or percent identity range independently to each of the referenced sequences. In determining percent identity for a polynucleotide, RNA and the corresponding DNA are considered the same unless provided otherwise by the employed context, for example, providing reference to the polynucleotide being RNA or DNA. Corresponding RNA and DNA include uracil for thymine and replacement of the ribose backbone for the deoxyribose backbone.

[0057] Reference to a percent “identical”, “identity” and similar terminology are with respect to two sequences having maximal alignment in a particular area. The provided area is with respect to the indicated reference sequence. For example, sequence “identical” or “identity” to a PPT1 polypeptide of SEQ ID NO: 1 can be calculated by determining the number of identical amino acids in aligned sequences, dividing by the total number of amino acids in SEQ ID NO: 1 (279 amino acids) and multiplying by 100. Percent “identical” or “identity” for nucleic acid sequences can be determined in an analogous manner where nucleotides to the reference sequence are aligned to achieve maximal alignment taking into account nucleotide differences and gaps, dividing by the total number of nucleotides in the reference sequence and multiplying by 100.

[0058] Percent identical or identity for a PPT1 encoding sequence indicated to comprise two or more exons is determined independently of any intron. For calculating purposes, intron(s) are removed prior to the alignment.

[0059] In determining sequence identity to a reference sequence having one or more indicated variation, the particular variation chosen for determining sequence identity is that providing for the greatest sequence identity. For example, SEQ ID NO: 2 provides for X being G, V, or L; and for purposes of determining sequence identity to SEQ ID NO: 2, the X selected for determining sequence identity is that providing for the greatest sequence identity.

[0060] The terms “nucleic acid” and “polynucleotide” are used interchangeably herein to refer to all forms of nucleic acid, oligonucleotides, including deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). In discussing nucleic acids, a sequence or structure of a particular polynucleotide can be described herein according to the convention providing the sequence in the 5' to 3' direction.

[0061] In certain embodiments, nucleic acids include genomic DNA, cDNA, antisense DNA/RNA, plasmid DNA, linear DNA, (poly- and oligo-nucleotide), chromosomal DNA, spliced or unspliced mRNA, rRNA, tRNA, inhibitory DNA or RNA (RNAi, *e.g.*, small or short hairpin (sh)RNA, microRNA (miRNA), small or short interfering (si)RNA, trans-splicing RNA, or antisense RNA), locked nucleic acid analogue (LNA), oligonucleotide DNA (ODN) single and double stranded, immunostimulating sequence (ISS), riboswitches and ribozymes.

[0062] In certain embodiments, nucleic acids include naturally occurring, synthetic, and intentionally modified or altered polynucleotides. Nucleic acids can be single, double, or triplex, linear or circular, and can be of different lengths.

[0063] According to certain embodiments, the polynucleotide is a single-stranded (ssDNA) or a double-stranded DNA (dsDNA) molecule. According to certain embodiments, the dsDNA molecule is a minicircle, a nanoplasmid, open linear duplex DNA or a closed-ended linear duplex DNA (CELiD/ceDNA/doggybone DNA). According to certain embodiments, the ssDNA molecule is a closed circular or an open linear DNA.

[0064] A “transgene” refers to a nucleic acid that is intended or has been introduced into a cell and operably linked to a promoter. Transgenes include a heterologous polynucleotide sequence, such as nucleic acid encoding a PPT1 polypeptide and a heterologous promoter.

[0065] Certain embodiments are directed to “CpG reduced” or “CpG depleted”. “CpG reduced” or “CpG depleted” refer to (i) a nucleotide sequence wherein one or more of the CpG dinucleotides (or motifs) are removed from a reference nucleic acid sequence; and/or (ii) the percentage of CpGs in a referred to polynucleotide is 0% to 15%. In different embodiments, the CpG percentage is 0%, about 0.5%, about 1.0%, about 2.0%, about 3.0%, about 4.0%, about 5.0%, about 6%, about 7%, about 8%, about 9%, about 10%, about 11%, about 12%, about 13%, about 14%, or about 15% CpGs; and/or up to about 0.5%, up to about 1.0%, up to about 2.0%, up to about 3.0%, up to about 4.0%, up to about 5.0%, up to about 6%, up to about 7%, up to

about 8%, up to about 9%, up to about 10%, up to about 11%, up to about 12%, up to about 13%, up to about 14%, or up to about 15% CpGs.

[0066] CpG motifs can be suitably reduced or eliminated in a nucleotide sequence encoding a PPT1 protein and in other sequences that may be present in particular constructs (*e.g.*, expression cassettes and viral vectors). Other sequences that may be present include non-coding sequences such as 5' and 3' untranslated regions (UTRs), stuffer sequences, promoter, enhancer, polyadenylation signal, ITRs, and introns.

[0067] The singular forms “a,” “an,” and “the” include plural reference unless the context clearly dictates otherwise.

[0068] The conjunctive term “and/or” between multiple recited elements encompasses both individual and combined options. For instance, where two elements are conjoined by “and/or”, a first option refers to the applicability of the first option without the second, a second option refers to the applicability of the second option without the first, and a third option refers to the applicability of the first and second options together. Any one of the options is understood to fall within the meaning, and therefore satisfy the requirement of the term “and/or”. Concurrent applicability of more than one of the options is also understood to fall within the meaning of the term “and/or.”

[0069] Unless clearly indicated otherwise by the context employed the terms “or” and “and” have the same meaning as “and/or”.

[0070] Reference to terms such as “including”, “for example”, “*e.g.*,” “such as” followed by different members or examples, are open-ended descriptions where the listed members or examples are illustrative and other member or examples can be provided or used.

[0071] The terms “polypeptide,” “protein” and “peptide” can be used interchangeably to refer to an amino acid sequence without regard to function. Polypeptides and peptides contain at least two amino acids, while proteins contain at least about 10 amino acid acids. Amino acids include naturally occurring amino acids and amino acids provided by cellular modification.

[0072] Reference to “comprise”, and variations such as “comprises” and “comprising”, used with respect to an element or group of elements is open-ended and does not exclude additional unrecited elements or method steps. Terms such as “including”, “containing” and “characterized by” are synonymous with comprising. In the different aspects and embodiments described herein reference to an open-ended term such as “comprising” can be replaced by “consisting” or “consisting essentially of”.

[0073] Reference to “consisting of” excludes any element, step, or ingredient not specified in the listed claim elements, where such element, step or ingredient is related to the claimed invention.

[0074] Reference to “consisting essentially of” limits the scope of a claim to the specified materials or steps and those that do not materially affect the basic and novel characteristic(s) of the claimed invention.

[0075] The term “about” refers to a value within 10% of the underlying parameter (*i.e.*, plus or minus 10%). For example, “about 1:10” includes 1.1:10.1 or 0.9:9.9, and “about 5 hours” includes 4.5 hours or 5.5 hours. The term “about” at the beginning of a string of values modifies each of the values by 10%.

[0076] All numerical values or numerical ranges include integers within such ranges and fractions of the values or the integers within ranges unless the context clearly indicates otherwise. Thus, to illustrate, reference to reduction of 95% or more includes 95%, 96%, 97%, 98%, 99%, 100%, as well as 95.1%, 95.2%, 95.3%, 95.4%, 95.5%, etc., 96.1%, 96.2%, 96.3%, 96.4%, 96.5% and so forth and reference to a numerical range, such as “1-4” includes 1, 2, 3, 4 as well as 1.1, 1.2, 1.3, 1.4 and so forth. As a further illustration, “1 to 4 weeks” includes 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, or 28 days.

[0077] Further, reference to a numerical range, such as “0.01 to 10” includes 0.011, 0.012, 0.013 etc., as well as 9.5, 9.6, 9.7, 9.8, 9.9 and so forth. For example, a dosage of about “0.01 mg/kg to about 10 mg/kg” body weight of a subject includes 0.011 mg/kg, 0.012 mg/kg, 0.013 mg/kg, 0.014 mg/kg, 0.015 mg/kg etc., as well as 9.5 mg/kg, 9.6 mg/kg, 9.7 mg/kg, 9.8 mg/kg, 9.9 mg/kg and so forth.

[0078] Reference to an integer with more (greater) or less than includes numbers greater or less than the reference number, respectively. Thus, for example, reference to more than 2 includes 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more; and administration “two or more” times includes 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more times.

[0079] Various references including articles and patent publications are cited or described in the background and throughout the specification. Each of these references is herein incorporated by reference in their entirety. None of the references are admitted to be prior art with respect to any inventions disclosed or claimed. In some cases, particular references are indicated to be incorporated by reference herein to highlight the incorporation.

[0080] The definitions provided herein, including those in the present section and other sections of the application, apply throughout the present application.

[0081] Unless defined otherwise, all technical and scientific terms used herein have the same meaning commonly understood to one of ordinary skill in the art to which this invention pertains.

[0082] The description has been separated into various sections and paragraphs, and provides examples of various embodiments. These separations should not be considered as disconnecting

the substance of a paragraph or section or embodiments from the substance of another paragraph or section or embodiment. The provided descriptions have broad application and encompasses all the combinations of the various sections, paragraphs and sentences that can be contemplated. The discussion of any embodiment is meant only to be exemplary and is not intended to suggest the scope of the disclosure, including the claims (unless otherwise provided in the claims), is limited to these examples.

[0083] The instant invention is generally disclosed herein using affirmative language to describe the numerous embodiments of the instant invention. The instant invention also specifically includes embodiments in which particular subject matter is excluded, in full or in part, such as substances or materials, method steps and conditions, protocols, or procedures. For example, in certain embodiments of the instant invention, materials and/or method steps are excluded. Thus, even though the instant invention is generally not expressed herein in terms of what the instant invention does not include, embodiments that are not expressly excluded in the instant invention are nevertheless disclosed herein.

I. PPT1 Polypeptide

[0084] PPT1 polypeptides comprise a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1. In certain embodiments the polypeptide comprises a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1, wherein: (a) the PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or (b) the PPT1 amino acid sequence comprises a G, V or L substitution for D at its amino terminus; and/or (c) the PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus. SEQ ID NO: 38 is an example of a full-length polypeptide where the region corresponding to the human mature sequence has a leucine-glutamine-histidine-leucine added to the N-terminus of the naturally occurring sequence.

[0085] In certain embodiments a mature sequence contains a deletion at its N-terminus. SEQ ID NO: 37 is an example of a full-length where the mature sequence has an aspartic acid-proline-proline-alanine deletion at N-terminus of the naturally occurring sequence.

[0086] In certain embodiments, the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96%, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence to SEQ ID NO: 1. In further embodiments the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 2, and X is glycine; the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96%, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the

sequence of SEQ ID NO: 2, and X is valine; the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96%, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 2, and X is leucine; the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 3; or the PPT1 polypeptide comprises a PPT1 amino acid comprising the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus sequence and having at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence to SEQ ID NO: 4.

[0087] PPT1 is a well characterized enzyme, where different amino acids responsible for activity and different mutations resulting in decreased activity are well-known. (See, e.g., Kumar *et al.*, *Advances in Protein Chemistry and Structural Biology* (2022) 132:89-109; Bellizzi *et al.*, *PNAS* (2000) 97(9):4573-4578; and [hyperlink://www.uniprot.org/uniprotkb/P50897/entry](https://www.uniprot.org/uniprotkb/P50897/entry) (December 3, 2022) each of which of are hereby incorporated by reference herein in their entirety.)

[0088] A signal sequence provides the amino sequence of a signal peptide, where signal peptides are short N-terminal amino acid sequences providing for protein secretion. The terms signal sequence and sequence peptide are used interchangeably herein. Signal sequences direct proteins to or through the endoplasmic reticulum secretory pathway and are generally cleaved within the endoplasmic reticulum prior to secretion. Thus, a signal peptide enhances secretion of a polypeptide from the cell as compared to the secretion level of the corresponding polypeptide lacking a signal peptide.

[0089] The presence of a signal sequence in a PPT1 polypeptide facilitates extracellular secretion of the mature PPT1 polypeptide. The secreted polypeptide can be taken up by another cell providing for cross-correction.

[0090] In certain embodiments the signal peptide comprises the amino acid sequence of any of SEQ ID NOs: 16-27, or comprises an amino acid sequence differing by any of SEQ ID NOs: 16-27 by one amino acid. In further embodiments, the signal peptide comprises the amino acid sequence of SEQ ID NO: 16; comprises the amino acid sequence of SEQ ID NO: 19; or comprises the amino acid sequence of SEQ ID NO: 23.

[0091] In certain embodiments, the PPT1 polypeptide comprises a signal sequence and a PPT1 sequence wherein:

a) the signal sequence comprises the amino acid sequence of any of SEQ ID NOs: 16-21 and 24-27, or differing from any of SEQ ID NOs: 16-21 and 24-27 by one amino acid addition, deletion, or substitution; and the PPT1 polypeptide sequence comprises a sequence

with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 1;

b) the signal sequence comprises the amino acid sequence of any of SEQ ID NOs: 16-21 and 24-27, the PPT1 polypeptide sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 2 and X is G, V or L;

c) the signal sequence comprises the amino acid sequence of SEQ ID NO: 16 or differing from SEQ ID NO: 16 by one amino acid addition, deletion, or substitution; and the PPT1 polypeptide sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 1;

d) the signal sequence comprises the amino acid sequence of SEQ ID NO: 16 and the PPT1 polypeptide sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 2 and X is G, V, or L; and

e) the signal sequence comprises the amino acid sequence of SEQ ID NO: 16 and the PPT1 polypeptide sequence comprises the sequence of SEQ ID NO: 2, wherein X is G.

f) the signal sequence comprises the amino acid sequence of SEQ ID NO: 19 or differing from SEQ ID NO: 19 by one amino acid addition, deletion, or substitution; and the PPT1 polypeptide sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 1;

h) the signal sequence comprises the amino acid sequence of SEQ ID NO: 19 and the PPT1 polypeptide sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 2 and X is G, V or L; and

i) the signal sequence comprises the amino acid sequence of SEQ ID NO: 19 and the PPT1 polypeptide sequence comprises the sequence of SEQ ID NO: 2, wherein X is G.

[0092] In certain embodiments, the PPT1 polypeptide comprises a signal sequence and a PPT1 sequence wherein the signal sequence comprises the amino acid sequence of SEQ ID NO: 23 or differing from SEQ ID NO: 23 by one amino acid addition, deletion, or substitution; and the PPT1 polypeptide sequence comprises the amino acid sequence leucine-glutamine histidine-leucine at its N-terminus and has a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 4.

[0093] In certain embodiments concerning a PPT1 polypeptide comprising a signal sequence and a PPT1 sequence, the polypeptide comprises an amino acid sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to any of

SEQ ID NOs: 31-42, or differs from any of SEQ ID NOs: 31-42 by 1-10 amino acids differences, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 amino acids differences. Reference to amino acid differences indicates any combination of addition, substitution, and/or deletion.

[0094] In further embodiments the polypeptide comprises an amino acid sequence having a sequence identity at least 99% or 100% to SEQ ID NO: 31 or differs from SEQ ID NO: 31 by 1-5 amino acids; the polypeptide comprises the amino acid of SEQ ID NO: 31, where X is D or G; the polypeptide comprises an amino acid sequence having a sequence identity of at least 99% or 100% to SEQ ID NO: 34 or differs from SEQ ID NO: 34 by 1-5 amino acids; the polypeptide comprises the amino acid of SEQ ID NO: 34 where X is D or G; the polypeptide comprises the amino acid sequence leucine-glutamine histidine-leucine at its N-terminus and has a sequence identity at least 99% or 100% to SEQ ID NO: 38 or differs from SEQ ID NO: 38 by 1-5 amino acids; or the polypeptide comprises SEQ ID NO: 30.

II. Polynucleotide Comprising a PPT1 Encoding Sequence

[0095] A polynucleotide comprising a nucleic acid sequence encoding for a PPT1 polypeptide can be used to facilitate production and intracellular delivery of a PPT1 polypeptide. In certain embodiments, the polynucleotide comprises a nucleic acid sequence encoding a PPT1 polypeptide, wherein the PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1, wherein:

- (a) the PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or
- (b) the PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or
- (c) the PPT1 sequence comprises the amino acid sequence leucine-glutamine- histidine-leucine at its N-terminus; and/or
- (d) the nucleic acid sequence comprises a PPT1 encoding sequence having at least 85% identity to any of SEQ ID NOs: 61-94.

[0096] In certain embodiments, (a) the polynucleotide encodes a PPT1 polypeptide comprising a PPT1 amino acid sequence having at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 1; (b) the encoded PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 2 and X is glycine; (c) the encoded PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 2, and X is valine; or (d) the encoded PPT1 polypeptide comprises a PPT1 amino acid sequence having at

least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence of SEQ ID NO: 2, and X is leucine. In further embodiments, the nucleotide acid sequence encoding for any of (a), (b), (c), or (d) comprises a nucleic sequence having a sequence identity of at least 85%, at least 90%, at least 95%, at least 97%, at least 99% or 100% to any of SEQ ID Nos: 61-94 or a variant thereof wherein the first three nucleotides encode for G, D, V or L, or encodes for G. Nucleotides encoding for G, V, D or L are provided in Table 1.

[0097] Table 1

	Glycine	Valine	Leucine	Aspartic Acid
Codons	GGT GGC GGA GGG	GTT GTC GTA GTG	TTA TTG CTT CTC CTA CTG	GAT GAC

[0098] In certain embodiments nucleic acid encoding a mature PPT1 amino acid comprise the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus sequence and have at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, or 100% identity to the sequence to SEQ ID NO: 4.

[0099] In certain embodiments the nucleic acid encoding for a mature PPT1 polypeptide further comprises a signal sequence comprising the amino acid sequence of any of SEQ ID NOs: 16-27, or comprises an amino acid sequence differing by any of SEQ ID NOs: 16-27 by one amino acid. In further embodiments, the signal encoding sequence is the sequence of any of SEQ ID NOs: 43-58. Reference to “signal encoding sequence” indicates a nucleotide sequence encoding for the signal sequence or signal peptide.

[0100] In certain embodiments, the polynucleotide comprises the signal encoding sequence of any of SEQ ID NOs: 16-21 and 24-27; SEQ ID NO: 16; or SEQ ID NO: 19.

[0101] In certain embodiments, the signal encoding sequence comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 43-47, and the signal sequence comprises the sequence of SEQ ID NO: 16, or differing from SEQ ID NO: 16 by one amino acid. In further embodiments, the nucleic acid sequence comprises the signal encoding sequence of any of SEQ ID NOs: 43, 44, 45, 46, or 47.

[0102] In certain embodiments, the signal encoding sequence comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 43 and the signal sequence comprises the sequence of SEQ ID NO: 16, or differing from

SEQ ID NO: 16 by one amino acid. In a further embodiment, the signal encoding sequence comprises the sequence of SEQ ID NO: 43.

[0103] In certain embodiments, the signal encoding sequence comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 50, and the signal encoding sequence comprises the sequence of SEQ ID NO: 19, or differing from SEQ ID NO: 19 by one amino acid. In different embodiments, the nucleic acid sequence comprises the sequence of SEQ ID NO: 50.

[0104] In certain embodiments, the signal encoding sequence comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 54 and the signal encoding sequence comprises the sequence of SEQ ID NO: 23, or differing from SEQ ID NO: 23 by one amino acid. In different embodiments, the nucleic acid sequence comprises the sequence of SEQ ID NO: 54.

[0105] In certain embodiments, the polynucleotide encodes for a PPT1 polypeptide comprising a signal sequence and a PPT1 sequence wherein:

a) the signal sequence comprises SEQ ID NO: 16, the signal encoding sequence comprises a sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 43 and the PPT1 sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 1 or SEQ ID NO 2;

b) the signal encoding sequence comprises the sequence of SEQ ID NO: 43 and the PPT1 sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NOs: 1 or 2;

c) the signal encoding sequence comprises the sequence of SEQ ID NO: 43 and the PPT1 sequence comprises a sequence with a sequence identity of at least 97% to SEQ ID NO: 1;

d) the signal encoding sequence comprises the sequence of SEQ ID NO: 43 and the PPT1 sequence comprises a sequence with a sequence identity of at least 97% to SEQ ID NO: 2 and X is G, V or L;

e) the signal encoding sequence comprises the sequence of SEQ ID NO: 43 and the PPT1 sequence comprises a sequence with a sequence identity of at least 99% to SEQ ID NO: 2 and X is G;

f) the signal encoding sequence comprises the sequence of SEQ ID NO: 43; and the PPT1 sequence comprises the sequence of SEQ ID NO: 1;

g) the signal encoding sequence comprises the sequence of SEQ ID NO: 43; and the PPT1 sequence comprises the sequence SEQ ID NO: 2, wherein in X is G; and

h) in different embodiments the sequence encoding the polypeptide sequence for any of (a) to (g) comprises a sequence with a sequence identity of at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 61-79 and 80-94, or to any of SEQ ID NOs: 61-79 and 80-94 wherein the first three nucleotides are a codon from Table 1. In further embodiments, the sequence encoding the polypeptide sequence for any of (a) to (g) comprises a sequence with a sequence identity of at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 62-64, 71, 74, 78, 79 and 83 wherein the first three nucleotides are a codon from Table 1; the sequence encoding the polypeptide sequence for any of (a) to (g) comprises a sequence with a sequence identity of at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 64 wherein the first three nucleotides are a codon from Table 1; or the sequence encoding the polypeptide sequence for any of (a) to (g) comprises a sequence with a sequence identity of at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 79 wherein the first three nucleotides are a codon from Table 1

[0106] In certain embodiments, the polynucleotide encodes for a PPT1 polypeptide comprising a signal sequence and a PPT1 sequence wherein:

a) the signal sequence comprises SEQ ID NO: 19 and the signal encoding sequence comprises a sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 50; and the PPT1 sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 1 or SEQ ID NO 2;

b) the signal encoding sequence comprises the sequence of SEQ ID NO: 50 and the PPT1 sequence comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NOs: 1 or 2;

c) the signal encoding sequence comprises the sequence of SEQ ID NO: 50 and the PPT1 sequence comprises a sequence with a sequence identity of at least 97% to SEQ ID NO: 1;

d) the signal encoding sequence comprises the sequence of SEQ ID NO: 50 and the PPT1 sequence comprises a sequence with a sequence identity of at least 97% to SEQ ID NO: 2 and X is G, V, or L;

e) the signal encoding sequence comprises the sequence of SEQ ID NO: 50 and the PPT1 sequence comprises a sequence with a sequence identity of at least 99% to SEQ ID NO: 2 and X is G;

f) the signal encoding sequence comprises the sequence of SEQ ID NO: 50 and the PPT1 sequence comprises the sequence of SEQ ID NO: 1;

g) the signal encoding sequence comprises the sequence of SEQ ID NO: 50 and the PPT1 sequence comprises the sequence of SEQ ID NO: 2, wherein in X is G; and

h) in different embodiments the nucleic acid encoding the polypeptide sequence for any of (a) to (g) comprises a sequence with a sequence identity of at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 61-79 and 80-94, or to any of SEQ ID NOs: 61-79 and 80-94 wherein the first three nucleotides are a codon from Table 1.

[0107] In certain embodiments, the polynucleotide encodes for a PPT1 polypeptide comprising a signal sequence and a PPT1 sequence wherein:

a) the signal sequence comprises SEQ ID NO: 23, the signal encoding sequence comprises a sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 54, and the PPT1 sequence comprises a leucine-glutamine-histidine-leucine at its N-terminus and a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 4;

b) the signal encoding sequence comprises the sequence of SEQ ID NO: 54 and the PPT1 sequence comprises a leucine-glutamine histidine-leucine at its N-terminus and a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 4;

c) the signal encoding sequence comprises the sequence of SEQ ID NO: 54 and the PPT1 sequence comprises a leucine-glutamine histidine-leucine at its N-terminus and a sequence identity of at least 97% to SEQ ID NO: 1;

d) the signal encoding sequence comprises the sequence of SEQ ID NO: 54 and the PPT1 sequence comprises a leucine-glutamine-histidine-leucine at its N-terminus and a sequence identity of at least 98% to SEQ ID NO: 4;

e) the signal encoding sequence comprises the sequence of SEQ ID NO: 54 and the PPT1 sequence comprises a leucine-glutamine-histidine-leucine at its N-terminus and a sequence identity of at least 99% to SEQ ID NO: 4; and

f) the signal encoding sequence comprises the sequence of SEQ ID NO: 54 and the PPT1 sequence comprises SEQ ID NO: 4.

[0108] In certain embodiments, the nucleic acid (a) encodes a PPT1 polypeptide comprising a sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to any of SEQ ID NOs: 29-42, or differs from any of SEQ ID NOs: 29-42 by 1-10 amino acids, 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 amino acids; (b) encodes a PPT1 polypeptide comprising an amino acid sequence having a sequence identity at least 99% or 100% to SEQ ID NO: 31 or differs from SEQ ID NO: 31 by 1-5 amino acids; (c) encodes a PPT1 polypeptide comprising the amino acid of SEQ ID NO: 31, where X is D or G; (d) encodes a PPT1 polypeptide comprising an amino acid sequence having a sequence identity of at least 99% or 100% to SEQ ID NO: 34 or differs from SEQ ID NO: 34 by 1-5 amino acids; (e) encodes a

PPT1 polypeptide comprising the amino acid of SEQ ID NO: 34 where X is D or G; (f) encodes a PPT1 polypeptide comprising the amino acid sequence of SEQ ID NO: 38; (g) encodes for (b) wherein the nucleic acid comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 43; (h) encodes for (b) wherein the nucleic acid comprises a sequence having a sequence identity of at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 107-125 and 168; (i) encodes for (c) wherein the nucleic acid comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 99; (h) encodes for (c) wherein the nucleic acid comprises a sequence having a sequence identity of at least 85%, at least 90%, at least 95% or 100% to any of SEQ ID NOs: 126-140 and 161-167; or (i) encodes for (f) wherein the nucleic acid comprises a sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to SEQ ID NO: 103.

[0109] In certain embodiments, the polynucleotide comprises a nucleic acid sequence encoding for a PPT1 polypeptide described in Section I. *supra*.

[0110] In certain embodiments, the polynucleotide sequence encoding the PPT1 polypeptide comprises two or more exons coding for the PPT1 polypeptide and one or more introns.

[0111] In certain embodiments, the polynucleotide comprises a PPT1 encoding sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95% or 100% to the sequence of any of SEQ ID NOs: 95-106,

[0112] Reference to a sequence provided in the present application, such as in Table 2 *infra* that includes a stop codon, includes embodiments where the stop codon is not present, multiply stop codons are present, and different stop codons are present.

[0113] Reference to a sequence provided in the present application, such as in Table 2 *infra* encoding a protein that does provide a stop codon includes embodiments where the stop codon is not present immediately after the provided sequence, a stop is present, multiply stop codons are present, and different stop codons are present.

[0114] In certain embodiments, PPT1 encoding nucleotide sequences contains 0-5, 0-10, or 0-15 CpGs; 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 CpGs; 0%, about 0.5%, about 1.0%, about 2.0%, about 3.0%, about 4.0%, or about 5.0% CpGs; and/or up to about 0.5%, up to about 1.0%, up to about 2.0%, up to about 3.0%, up to about 4.0%, or up to about 5.0% CpGs.

III. Expression Cassettes

[0115] Polynucleotide expression cassettes contain a nucleic acid encoding for a PPT1 polypeptide operably linked to one or more expression control elements. Expression control can be affected, for example, at the level of transcription, translation, splicing, and message stability. Expression control elements are typically located 5' ("upstream") or 3' ("downstream") of a

transcribed nucleic acid. Expression control elements can also be located within the transcript (*e.g.*, in an intron), adjacent to or at a distance away from the transcribed sequence. One or more expression control elements of the same or different type may be present. Examples of expression control elements include a promoter, enhancer, an intron, polyadenylation signal, a Kozak sequence, post-transcriptional regulator elements and a termination sequence.

[0116] A promoter is a DNA region where transcription is initiated. In general, transcribed nucleic acid is located 3' of a promoter sequence. In certain embodiments, a promoter sequence is coupled to an enhancer. Enhancers are DNA regions that increase promoter transcription. Enhancers can be adjacent or inside a promoter or can be distal. Typically, enhancers are located upstream of a promoter, but can be located downstream or within a promoter sequence.

[0117] Expression control elements such a promoter and an enhancer can be chosen to preferentially drive expression in a particular cell or tissue type. Expression control elements are typically active in particular cells, tissues or organs because they are recognized by transcriptional activator proteins, or other regulators of transcription, that are unique to a specific cell, tissue or organ type. (*See, e.g.*, Green, M. and Sambrook, J. (2012) *Molecular Cloning: A Laboratory Manual*. 4th Edition, Vol. II, Cold Spring Harbor Laboratory Press, New York; and Ausubel *et al.*, (2010) *Current protocols in molecular biology*, John Wiley & Sons, New York).

[0118] The incorporation of tissue specific regulatory elements in the expression constructs provides for at least partial tissue tropism for the expression of PPT1 protein. Reference to a promoter or enhancer specific for a particular cell type of tissue, indicates the promoter or enhancer provides higher levels of expression and/or secretion in the indicated cell or tissue type. Examples of promoters specific for liver are the transthyretin (TTR) gene promoter; human alpha 1-antitrypsin (hAAT) promoter; the apolipoprotein A-I promoter; albumin, Miyatake *et al.*, *J. Virol.*, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.* 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot, et al., *Hum. Gene Ther.*, 7:1503-14 (1996); human Factor IX promoter; thyroxine binding globulin (TBG) promoter; TTR minimal enhancer/promoter; alpha-antitrypsin promoter; LSP (845 nt) (requires intronless scAAV); and LSP1 promoter. An example of an enhancer active in liver is apolipoprotein E (apoE) HCR-1 and HCR-2 (Allan *et al.*, *J. Biol. Chem.*, 272:29113-19 (1997)).

[0119] Expression control elements also include ubiquitous or promiscuous promoters and promoters/enhancers capable of driving polynucleotide expression in many different cell types. Such elements include the EF1-alpha promoter, the cytomegalovirus (CMV) immediate early promoter/enhancer sequences, the Rous sarcoma virus (RSV) promoter/enhancer sequences, phosphoglycerate kinase (PKG) promoter, CAG (composite of the CMV enhancer, the chicken beta actin promoter (CBA) and the rabbit beta globin intron) (*see, e.g.*, Boshart *et al.*, (1985)

Cell, 41:521-530), the SV40 promoter, the dihydrofolate reductase promoter, and the cytoplasmic b-actin promoter.

[0120] Examples of CNS specific promoters include: neuron specific promoters such as the NSE (neuronal specific enolase), synapsin or NeuN, platelet-derived growth factor (PDGF), platelet-derived growth factor B-chain (PDGF- β), methyl-CpG binding protein 2 (MeCP2), Ca^{2+} /calmodulin-dependent protein kinase II (CaMKII), metabotropic glutamate receptor 2 (mGluR2), neurofilament light (NFL) or heavy (NFH), β -globin minigene n β 2, preproenkephalin (PPE), enkephalin (Enk), and excitatory amino acid transporter 2 (EAAT2) promoters; astrocyte specific promoters such as the glial fibrillar acidic protein (GFAP) and EAAT2 promoters; oligodendrocyte specific promoters such as the myelin basic protein (MBP)/myelin-associated glycoprotein and oligodendrocyte transcription factor 2 promoter; neurons/hypothalamus specific promoters such as the proopiomelanocortin (POMC) promoter; and neuron/spinal cord specific promoter such as superoxide dismutase 1 (SOD1). (See, *e.g.*, U.S. Patent Publication No. 2021/214749 and Adeno-Associated Virus Vectors (2019), Ed. Castle., 1st Edition, Springer New York, New York, NY.; both of which are hereby incorporated by reference herein in their entirety.)

[0121] Additional promoters include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), herpes simplex virus (HSV) promoter, SFFV promoter, rat insulin promoter, TBG promoter, the desmin promoter and similar muscle-specific promoters, synthetic promoters, hybrid promoters, and promoters with multi-tissue specificity.

[0122] Expression control elements also can impact expression in a manner that is regulatable by a signal or stimuli increasing or decreasing expression. A regulatable element increasing expression of transcribed nucleic acid in response to a signal or stimuli is also referred to as an “inducible element” (*i.e.*, is induced by a signal). Typically, the amount of increase or decrease conferred by such elements is proportional to the amount of signal or stimuli present. Particular examples include zinc-inducible sheep metallothioneine (MT) promoter; the steroid hormone-inducible mouse mammary tumor virus (MMTV) promoter; the tetracycline-repressible system (Gossen, *et al.*, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)); the tetracycline-inducible system (Gossen *et al.*, Science 268: 1766-1769 (1995); see also Harvey *et al.*, Curr. Opin. Chem. Biol. 2:512-518 (1998)); the RU486-inducible system (Wang *et al.*, Nat. Biotech. 15:239-243 (1997) and Wang *et al.*, Gene Ther. 4:432-441 (1997); and the rapamycin-inducible system (Magari *et al.*, J. Clin. Invest. 100:2865-2872 (1997); and Rivera *et al.*, Nat. Medicine. 2:1028-1032 (1996)). Other examples of regulatable control elements include those regulated by a specific physiological state such as temperature, acute phase, or development.

[0123] In certain embodiments the expression cassette further comprises one or more introns independent of PPT1 encoding nucleotide acid. A variety of different introns can be used to enhance gene expression. Examples of introns that may be used include the rabbit β -globin intron with splice donor/splice acceptor, SV40 intron with splice donor/splice acceptor, human β -globin introns, intron 2 of the human hemoglobin beta gene, hFIX int1 (intron 1 of the human coagulation factor IX gene), CBA-rHHB (synthetic intron derived from the fusion of the intron 1 of the chicken beta actin gene and intron 2 of the rabbit hemoglobin beta), CBA (intron 1 of the chicken beta actin gene), hGH (intron 1 of the human growth hormone gene), hFIX synth (synthetic intron derived from different portions of the human coagulation factor IX gene and present in the pLIVE vector, Mirus Bio, Madison, WI); human hemoglobin subunit beta (HBB2) synthetic intron, and optimized HBB2; and chimeric introns such as introns made up of the 5'-splice donor of the first human β -globin intron and the branch and 3'-acceptor site from the intron that is between the leader and the body of the immunoglobulin gene heavy chain variable region. (Buck *et al.*, Int. J. Mol. Sci. (2020), 21, 4197; Ronzitti *et al.*, Mol. Ther. Methods Clin Dev. (2016) Jul 20;3:16049; and the HBB-IGG intron provided by the pCMVNT™ vector.)

[0124] In certain embodiments the expression cassette comprises a post-transcriptional regulatory element. Post-translational regulatory elements such as Woodchuck post-transcriptional regulatory element (WPRE) and Hepatitis B regulatory element can increase gene expression. (Buck *et al.*, Int. J. Mol. Sci. (2020), 21, 4197.)

[0125] Polyadenylation signal sequences provide for the formation of a polyA tail, which facilitates nuclear export, translation and/or mRNA stability, and may also be involved in transcription termination. Examples of polyadenylation signal sequences include SV40 late polyadenylation signal, bovine growth hormone polyA (bGHpA) signal sequence, synthetic polyA, mouse β -globin pA, rabbit β -globin pA, and H4-based pA. (Buck *et al.*, Int. J. Mol. Sci. (2020), 21, 4197.)

[0126] In certain embodiments, the expression cassette comprises a Kozak consensus sequence or a variation thereof. Kozak consensus sequences play a role in translation initiation. The Kozak consensus sequence and variations are provided in, for example, McClements *et al.*, (2021) Molecular vision, 27, 233–242, hereby incorporated by reference herein.

[0127] In certain embodiments the expression cassette comprises from 5' to 3' operatively coupled to the PPT1 encoding sequence: a promoter or promoter/enhancer, an intron, a Kozak sequence, the PPT1 encoding sequence and a polyadenylation signal. In certain embodiments, the intron comprises the amino acid sequence of SEQ ID NO: 11.

[0128] In certain embodiments the expression cassettes further comprises a miRNA target sequences, which in further embodiments is incorporated into the 3' UTR of the expression

cassette. A miRNA target sequence is recognized by miRNA present in particular cells or tissues leading to degradation of mRNA transcripts. Based on the presence of certain miRNA in particular cells, incorporating a miRNA target sequence(s) can be used to reduce expression in certain cells or tissue types. Multiple tandem repeats of miRNA target sequences can be used to increase degradation. (Geisle *et al.*, (2016) World Journal of Experimental Medicine 6(2): 37–54.)

[0129] In certain embodiments, the expression cassette encodes for a PPT1 polypeptide as provided in Section I. *supra*., comprises a nucleic acid sequence encoding for a PPT1 polynucleotide as provided in Section II *supra*; and/or comprises a nucleic acid sequence having a sequence identity of at least 80%, at least 85%, at least 90%, at least 95%, at least 99% or 100% to any of SEQ ID NOs: 141-143, 169 and 170.

[0130] In certain embodiments, the expression cassette nucleotide sequence contains any of 0-5, 0-10, 0-15, 0-50, or 0-100 CpGs; 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 CpGs; 0%, about 0.5%, about 1.0%, about 2.0%, about 3.0%, about 4.0%, about 5.0%, about 6%, about 7%, about 8%, about 9%, about 10%, about 11%, about 12%, about 13%, about 14%, or about 15% CpGs; and/or up to about 0.5%, up to about 1.0%, up to about 2.0%, up to about 3.0%, up to about 4.0%, up to about 5.0%, up to about 6%, up to about 7%, up to about 8%, up to about 9%, up to about 10%, up to about 11%, up to about 12%, up to about 13%, up to about 14%, or up to about 15% CpGs.

IV. Recombinant Viral Vector Nucleic Acid

[0131] Polynucleotide recombinant viral vector nucleic acid contain 5' and/or 3' viral elements providing for viral packaging and may provide for additional activities such as self-priming, DNA replication, promoter activity, genome integration, or episomal concatemerization. The 5' and 3' elements are generally located at or near the 5' and 3' terminal end of the recombinant viral vector nucleic acid and can be naturally occurring or modified versions of naturally occurring sequences. Examples of 5' and 3' elements include adenovirus ITRs, adeno-associated virus ITRs and packaging sequence; and retrovirus 5' and 3' long terminal repeats (LTRs) and packaging sequences. (Naso *et al.*, (2017) BioDrugs, 31(4), 317–334; Bulcha *et al.*, (2021) Sig. Transduct. Target Ther. 6:53 (2021); and Liu and Seol (2020) BMB Reports; 53(11):565-575.)

[0132] The term “recombinant,” as a modifier of nucleic acid or a vector indicates a combination of elements that does not occur in nature. For example, a recombinant viral vector nucleic acid provides 5' and/or 3' viral elements along with an expression cassette containing one or more elements not naturally associated with the 5' and/or 3' elements. Similarly, a viral vector, such as an rAAV vector may contain a naturally occurring or modified capsid, encapsidating recombinant viral vector nucleic acid.

[0133] In certain embodiments, the viral vector nucleic acid sequence comprises a 5' UTR and a 3' UTR, or a 5' ITR and a 3' ITR and (1) comprises a sequence encoding for the polypeptide of Section I *supra*.; (2) comprises a nucleic acid sequence encoding for a PPT1 polynucleotide as provided in Section II *supra*.; and/or (3) comprises an expression cassette encoding for a PPT1 polynucleotide as provided in Section III *supra*.

[0134] In certain embodiments, the viral vector comprises a polyA signal operatively linked to the 3' ITR, where the polyA signal antagonizes potential transcription initiating from the 3'-ITR. The operatively linked polyA signal is upstream of the 3'-ITR.

[0135] In certain embodiments, the viral vector nucleic acid contains any of 0-5, 0-10, 0-15, 0-50, 0-100, or 0 to 150 CpGs; 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 44, 46, 47, 48, 49 or 50 CpGs; 0%, about 0.5%, about 1.0%, about 2.0%, about 3.0%, about 4.0%, about 5.0%, about 6%, about 7%, about 8%, about 9%, about 10%, about 11%, about 12%, about 13%, about 14%, or about 15% CpGs; and/or up to about 0.5%, up to about 1.0%, up to about 2.0%, up to about 3.0%, up to about 4.0%, up to about 5.0%, up to about 6%, up to about 7%, up to about 8%, up to about 9%, up to about 10%, up to about 11%, up to about 12%, up to about 13%, up to about 14% or up to about 15% CpGs.

[0136] In certain embodiment, the recombinant viral vector nucleic acid comprises a sequence having at least 95% identity, at least 97% identity, at least 99% identity, or 100% identity to the sequence of any of SEQ ID NOs: 144-154, 171 and 172.

V. Viral Vectors

[0137] In certain embodiments the gene delivery vehicle is a viral vector comprising a protein capsid encapsidating recombinant viral vector nucleic acid. The viral vector can deliver the viral vector nucleic acid to cells or tissues. Depending on the particular vector, the viral vector may further comprise a viral envelope. Examples of viral vectors that can be used for gene delivery include adenovirus vectors, rAAV, retrovirus vectors and herpes simplex vectors.

[0138] Different serotypes exist within different types of viruses. The different serotypes can provide for different activities, such as cell or tissue tropism and likelihood of generating a host immune response. The term "serotype" broadly refers to both serologically distinct viruses as well as viruses not serologically distinct that can be within a subgroup or a variant of a given serotype. Serologic distinctiveness can be determined based on the lack of cross-reactivity between antibodies to one capsid as compared to another capsid. Such cross-reactivity differences are usually due to differences in capsid protein sequences/antigenic determinants (*e.g.*, due to VP1, VP2, and/or VP3 sequence differences of AAV serotypes).

[0139] As more naturally occurring virus isolates are discovered or capsid mutants generated, there may or may not be serological differences with any of the currently existing serotypes. Thus, in cases where the new virus has no serological difference, this new virus would be a subgroup or variant of the corresponding serotype.

V.A. Adenovirus Vectors

[0140] Adenoviruses are non-enveloped double-stranded DNA viruses. Recombinant adenovirus vectors comprise recombinant adenovirus nucleic acid lacking one or more protein involved in viral replication, and further comprise an adenoviral capsid. Recombinant adenovirus vectors can be produced containing different amounts of adenoviral DNA. The adenovirus (Ad) genome is flanked by hairpin-like inverted terminal repeats (ITRs) varying in length from 30–371 bp at its termini. The ITRs serve as self-priming structures that promote primase-independent DNA replication. A packaging signal located at the left arm of the genome is required for viral genome packaging. (Liu and Seol (2020) BMB Reports, 53(11):565-575; and Bulcha *et al.*, (2021) Sig. Transduct. Target Ther. 6:53.)

[0141] In certain embodiments, the recombinant adenovirus vector is a third-generation vector, which are also referred to as “gutless” or “helper-dependent”. Gutless vectors can be produced from recombinant adenovirus nucleic acid where all, or substantially all viral sequences, except for the ITRs and the packaging signal, are not present. Gutless adenovirus vectors are high capacity vectors able to accommodate up to about 36 kb of DNA insert. Preferred recombinant adenovirus nucleic acid is about 27 kb to about 37 kb. Stuffer sequences can be added to recombinant adenovirus nucleic acid to increase nucleic acid size and capsid incorporation. Preferred stuffer sequences avoid coding sequences, repetitive sequences, recombination sequences, and immunogenic sequences. (Liu and Seol (2020) BMB Reports, 53(11):565-575; Bulcha *et al.*, (2021) Sig. Transduct. Target Ther. 6:53; and Sandig *et al.*, PNAS (2000) 97(3):1002-1007, each of which are hereby incorporated by reference herein in their entirety.)

[0142] In certain embodiments, recombinant adenovirus vectors can be produced based on rare human serotypes or chimpanzee serotypes. The use of chimpanzee and rare human serotypes may be helpful in reducing host immune response against recombinant adenovirus vectors due to preexisting immunity. (Guo *et al.*, (2018) Human vaccines & immunotherapeutics, 14(7):1679–1685 and Bulcha *et al.*, (2021) Sig. Transduct. Target Ther. 6:53.)

[0143] Adenovirus vectors can be produced by supplying viral proteins needed for vector production *in trans* using for example, appropriate helper viruses or plasmids and cell lines. (Liu and Seol (2020) BMB Reports, 53(11):565-575; and Bulcha *et al.*, (2021) Sig. Transduct. Target Ther. 6:53.)

V.B. Recombinant AAV Vectors

[0144] Recombinant adeno-associated viral (also referred to herein as “rAAV”) vector are based on the adeno-associated virus. The adeno-associated virus is a single-strand DNA virus containing a 4.7-kb genome flanked by 145-nt ITRs on both ends of the genome. ITR activity is important for self-priming and packaging, and may also provide additional activity such as promoter activity. AAV 5' and 3' ITRs can vary in size and the 5' and 3' inverted repeats need not be exact inverted repeats.

[0145] A rAAV vector contains AAV recombinant nucleic acid and a viral capsid. The rAAV recombinant nucleic acid lacks one or more AAV proteins involved in viral replication. In certain embodiments, the rAAV vector contains an AAV 5' and/or 3' ITR along with a DNA insert. In certain embodiments, rAAV nucleic acid comprise a 5' ITR and/or 3' ITR independently selected from 5' and 3' ITRs provided in AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.10, AAVrh.74 and AAV3B ITRs. In further embodiments 5' and 3' ITRs are present, and both ITRs are from the same serotype genome.

[0146] In certain embodiments the 5' ITR comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 8; and the 3' ITR comprises independently (a) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 9; (b) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 158; (c) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 159; or (d) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 160.

[0147] In certain embodiments the 3' ITR comprises a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 9; and the 5' ITR independently comprises (a) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 10; (b) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 156; or (c) a sequence with a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% to SEQ ID NO: 157.

[0148] Recombinant adeno-associated viral vectors typically accept inserts of DNA having a size range generally about 4 kb to about 5.2 kb. If needed, stuffer sequence can be used to increase rAAV nucleic acid size and packaging efficiency. In different embodiments, the rAAV nucleic acid including stuffer is 4-5.2 kb, 3.0-5.5 kb, 4.0-5.0 kb, 4.3-4.8 kb, about 4.2 kb, about 4.3 kb, about 4.4 kb, about 4.5 kb, about 4.6 kb, or about 4.7 kb. Preferred stuffer sequences

avoid coding sequences, repetitive sequences, recombination sequences, and immunogenic sequences.

[0149] In certain embodiment the rAAV is a self-complementary adeno-associated virus vector (scAAV) or short hairpin adeno-associated virus vector (shAAV). scAAV and shAAV provide for a double-stranded rAAV nucleic acid that can be incorporated into an AAV capsid. scAAV and shAAV comprise inverted dimeric repeats providing intramolecular double-stranded DNA. scAAV can be produced by mutating an ITR terminal resolution site so that rep fails to nick the terminal resolution site. shAAV can utilize a short hairpin to produce double-stranded AAV nucleic acid. scAAV and shAAV being double-stranded DNA provide an advantage in circumventing the DNA synthesis step required for single-stranded rAAV nucleic acid upon entry into a cell. A potential disadvantage of scAAV and shAAV is the size of DNA inserts that can be incorporated is reduced by about half compared to single-stranded rAAV nucleic acid. (U.S. Patent No. 10,457,940; Xie *et al.*, Mol Ther. (2017) 25(6):1363-1374; and McCarty Mol. Ther. (2008) 16(10):1648-1656; each of which are hereby incorporated by reference herein in their entirety.)

[0150] Naturally occurring AAV capsids contain viral proteins VP1, VP2 and VP3 in a ratio of about 1:1:10. AAV vectors can be produced where all three viral proteins are based upon a particular serotypes or where one, two or all three viral protein are based on different serotypes.

[0151] Recombinant AAV capsid and nucleic acid can be based on the same serotype (or subgroup or variant), or can be different from each other. In certain embodiments, a rAAV nucleic acid has the same serotype genome (*e.g.*, ITRs) as the encapsidating capsid protein.

[0152] In different embodiments, the rAAV capsid comprises a protein having a sequence identity of at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.1%, at least 99.2%, at least 99.3%, at least 99.4%, at least 99.5%, at least 99.9% or 100% identical to a VP1, VP2 or VP3 of any of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.74, AAV3B, AAV-2i8, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, or AAV1/rh.10; or VP1 of SEQ ID NO: 12 or SEQ ID NO: 15.

[0153] Recombinant AAV capsids comprising VP1 of SEQ ID NO: 12 is described, for example, in U.S. Patent No. 9840719; and rAAV capsids comprising VP1 of SEQ ID NO: 15 is described, for example, in US 9,169,299; both of these patents are incorporated herein by reference.

[0154] In certain embodiments, AAV capsids comprises VP1, VP2 and VP3 each independently having a sequence identity of at least 80%, at least 90%, at least 95% or 100% to a VP1, VP2 or

VP3 of any of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.74, AAV3B, AAV-2i8, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, AAV1/rh.10, SEQ ID NO: 12 or SEQ ID NO: 15; as well as variants (*e.g.*, capsid variants, such as amino acid insertions, additions, substitutions and deletions) thereof. (See, for example, U.S. Patent Nos. 9,909,142 and 9,840,719 disclosing RHM4-1, RHM15-1, RHM15-2, RHM15-3/RHM15-5, RHM15-4 and RHM15-6; U.S. Patent Publication No. 2013/0059732 and U.S. Patent No. 9,169,299, disclosing LK01, LK02, and LK03; and U.S. Patent No. 11,110,153; the disclosures of which are herein incorporated in their entirety.)

[0155] In certain embodiments, the capsid comprises VP1 having the sequence of SEQ ID NO: 12; VP2 having the sequence of SEQ ID NO: 13; and VP3 having the sequence of SEQ ID NO: 14.

[0156] In certain embodiments, the AAV capsid can cross the blood brain barrier and provide for CNS expression. Examples of such AAV capsids and the design of AAV capsids able to provide for CNS expression are provided in Chen *et al.*, (2021) J. Control. Release 333, 129-138 (*e.g.*, AAV9, AAV-PHP-B, AAV-PHP.eB, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, and AAV1/rh.10), U.S. Patent No. 9,585,971, and Goertsen *et al.*, (2022) Nat. Neurosci. 25, 106–115 (2022), each of which are incorporated by reference herein in its entirety.

[0157] The AAV genome contains two main genes: *rep* and *cap*. Transcription from the *rep* gene is initiated from two different promoters resulting in the production of nonstructural proteins designated Rep78, Rep68, Rep52, and Rep40. The rep proteins function in genome replication and/or encapsidation. The *cap* gene encodes for structural proteins making up the capsid (VP1, VP2 and VP3); a non-structural assembly-activating protein (AAP), which performs functions related to capsid assembly; and the membrane-associated accessory protein, which may be associated with production phases of the replication cycle. (Maurer and Weitzman (2020) Hum. Gene Ther. 31(9-10):499-511, hereby incorporated by reference herein in its entirety.)

[0158] AAV requires helper virus functions to complete its replication cycle. Helper virus functions can be supplied by different viruses in permissive cell lines. Permissive cell lines are cell lines able to support viral replication. Examples of helper viruses for AAV include adenovirus, HSV-1, HPV-16, and HBV1 which can be used in conjunction with, for example, permissive primate cells; and baculovirus which can be used in conjunction with, for example, permissive insect cells such as sf9. (Maurer and Weitzman (2020) Hum. Gene Ther. (2020) 31(9-10):499-511 and Meier *et al.*, (2020) Viruses 19;12(6):662, both of which are herein incorporated by reference herein in their entirety.)

[0159] Recombinant AAV can be produced by supplying viral proteins needed for vector production *in trans* using, for example, appropriate helper viruses or plasmids and cell lines. In certain embodiments, rAAV is produced using a rAAV vector genome plasmid. The plasmid comprises that portion of the rAAV nucleic acid ultimately packaged or encapsidated to form a viral (e.g., rAAV) vector. The “plasmid backbone,” contains elements important for propagation and recombinant virus production. Except for possible 3' ITR and/or 5' ITR cloning remnants the plasmid backbone is not itself packaged or encapsidated into virus particles.

[0160] The vector genome plasmid may contain regions such as an origin of replication and a selectable marker. Additional sites that may be present include cloning sites.

[0161] Recombinant AAV can be produced from different types of cell lines including HeLa, A549, BHK, Vero, and HEK293, or derivatives thereof. In certain embodiments, HEK293 cells are used (American Type Culture Collection Accession Number ATCC CRL1573). Other host cell lines appropriate for rAAV vector production are described in, for example, Robert *et al.*, (2017) *Biotechnol. J.* (2017) 12(3), 1600193; and International Application No. PCT/US2017/024951; the disclosures of which are herein incorporated in its entirety.

[0162] Recombinant AAV can be cultured under a variety of different conditions suitable for providing cell growth and gene expression. References describing rAAV manufacturing include Clément and Grieger (2016) *Mol. Ther. Methods Clin. Dev.* 16;3:16002; Robert *et al.*, (2017) *Biotechnol. J.* 12(3), 1600193; and *Adeno-Associated Virus Vectors* (2019), Ed. Castle., 1st Edition, Springer New York, New York, NY.; each of which are hereby incorporated by reference herein in their entirety.)

[0163] In certain embodiments, AAV helper functions are introduced into the host cell by transfecting the host cell with an AAV helper construct either prior to, or concurrently with, the transfection of an AAV expression vector. A host cell having AAV helper functions can be referred to as a “helper cell” or “packaging helper cell.” AAV helper constructs are thus sometimes used to provide at least transient expression of AAV *rep* and/or *cap* genes to complement missing AAV functions necessary for productive AAV transduction. AAV helper constructs often lack AAV ITRs and can neither replicate nor package themselves. These constructs can be, for example, in the form of a plasmid, phage, transposon, cosmid, virus, or virion. A number of AAV helper constructs have been described, such as the commonly used plasmids pAAV/Ad and pIM29+45 which encode both *rep* and *cap* expression products. A number of other vectors are known which encode *rep* and/or *cap* expression products. Recombinant AAV can be produced, for example, as described in US Patent 9,408,904; and International Application Nos. PCT/US2017/025396 and PCT/US2016/064414, the disclosures of which are herein incorporated in their entirety.

[0164] In certain embodiments, a rAAV vector is produced by a rAAV production cell comprising rAAV helper virus activity. The genome of the rAAV production cell comprises rAAV nucleic acid, the *rep* gene and the *cap* gene.

[0165] In certain embodiments, a rAAV vector is produced by culturing a rAAV permissive cell comprising an AAV genome plasmid, where the rAAV permissive cell further comprises *rep* and *cap* genes provided either as part of the cell genome and/or by one or more separate plasmids; and helper virus activity either as part of the cell genome and/or provided by one or more separate plasmids. In further embodiments, (a) the rAAV permissive cell line is a packaging cell, wherein the genome of the packaging cell comprises the *cap* gene and the *rep* gene; (b) the *rep* gene, *cap* gene, and helper activity are provided from the same plasmid; or (c) the *rep* gene and *cap* gene are provided by a *rep/cap* plasmid and helper activity is provided by a helper plasmid.

[0166] In certain embodiments involving the use of HSV helper functions, the helper functions are provided by genes encoding for at least UL5, UL8, UL52, and ICP8.

[0167] In certain embodiments involving the use of adenovirus helper functions, the helper function are provided by genes encoding for at least E1A, E1B19K, E1B55K, E2A, E4orf6 and VA RNA. In certain embodiments E1, E2A and VR RNA functions are provided by a helper plasmid, where additional helper functions are provided by a host strain.

[0168] In certain embodiments, rAAV vector is obtained by producing rAAV using methods described herein and purifying the rAAV. Purification of rAAV can performed using techniques such as gradient-based purification, column-based, and combined methods. (See, *e.g.*, Ayuso *et al.*, (2010), Curr Gene Ther. (2010) 10(6):423-36, hereby incorporated by reference herein in its entirety.)

V.C. Retrovirus Vectors

[0169] Retroviruses are enveloped, single-stranded RNA viruses comprising 5' and 3' LTRs, and a signal packaging sequence located just outside of the LTR. Different types of retrovirus vectors can contain different amounts of viral genome. In certain embodiments, the retrovirus vector is a lentiviral vector based on HIV, retaining all cis-acting sequences needed for viral RNA packaging, reverse transcription and proviral DNA integration, while removing all HIV protein-coding genes. Lentiviral vectors have a packaging capacity of up to about 9 kb. If needed, stuffer sequence can be used to increase rAAV nucleic acid size and packaging efficiency. Lentiviral vectors can be produced by supplying viral proteins needed for vector production *in trans* using appropriate plasmids and cell lines. (Bulcha *et al.*, (2021) Sig. Transduct. Target Ther. 6:53.)

VI. Non-Viral Vectors

[0170] In certain embodiments, the gene delivery vehicle is a non-viral vector. Non-viral vectors include nanoparticles and naked nucleic acid. Preferred non-viral vectors are nanoparticles. A variety of different nanoparticles can be employed including lipid nanoparticles (LNP), polymeric nanoparticles, lipid polymer nanoparticles (LPNP), protein and peptide-based nanoparticles, DNA dendrimers and DNA-based nanocarriers, carbon nanotubes, microparticles, microcapsules, inorganic nanoparticles, peptide cage nanoparticles, and exosomes. (See, e.g., Riley and Vermerris *Nanomaterials* (2017) 201, 7, 94; Thomas et al., *Molecules* (2019), 24, 3744; Bochicchio et al., (2021), 13, 198; Munagala et al., *Cancer Letters* (2021), 505, 58; Fu et al., (2020) *NanoImpact* 20, 100261; Neshat et al., (2020) *Current Opin. Biotechnol.* 66:1-10; Ouranidis et al., (2022) *Biomedicines*, 10, 50; and Qin et al., *Signal Transduct Target Ther.* (2022) May 21;7(1):166, each of which are hereby incorporated by reference herein in their entirety.)

[0171] If desired, a nanoparticle can target a cell type using, for example, targeting ligands recognizing a target cell receptor. Examples of targeting ligands include carbohydrates (e.g., galactose, mannose, glucose, and galactomannan), endogenous ligands (e.g., folic acid and transferrin), antibodies and protein/peptides (e.g., RGD, epidermal growth factor, and low density lipoprotein) and peptides. (For example, Teo *et al.*, *Advanced Drug Delivery Reviews* (2016), 98, 41.)

[0172] Nanoparticles can be used to deliver PPT1 encoding polynucleotide constructs to a cell. In different embodiments, nanoparticles can deliver additional therapeutic compounds; and one or more additional compounds is provided in different nanoparticles. Reference to compound includes small molecules and large molecules (e.g., therapeutic proteins and antibodies).

[0173] The production of different nanoparticles and incorporation of nucleic acid and other compounds is well known in the art. Examples of publications illustrating incorporation of nucleic acid in a particular nanoparticle such as an LPNP and a LNP include Teo *et al.*, *Advanced Drug Delivery Reviews* (2016) 98, 41; Bochicchio *et al.*, *Pharmaceutics* (2021) 13, 198; Mahzabin and Das, *IJPSR* (2021) 12(1), 65; and Teixeira *et al.*, (2017) *Prog. Lipid Res.* Oct;68:1-11 (each of which are hereby incorporated by reference herein in their entirety). Factors that may impact small molecule incorporation into a nanoparticle include hydrophobicity and the presence of an ionizable moiety. (See, e.g., Nii and Ishii, *Int. J. Pharm.* (2005) 298:198-205; and Chen *et al.*, *J. Control. Release* (2018) 286:46-54.)

VI.A. Lipid-Based Delivery Systems

[0174] Lipid-based delivery systems include the use of a lipid as a component. Examples of lipid-based delivery systems include liposomes, LNPs, micelles, and extracellular vesicles.

[0175] A “lipid nanoparticle” or “LNP” refers to a lipid-based vesicle useful for delivery of

nucleic acid molecules and having dimensions on the nanoscale. In different embodiments the nanoparticle is from about 10 nm to about 1000 nm, about 50 nm to about 500 nm, or about 50 nm to about 200 nm.

[0176] DNA is negatively charged. Thus, it can be beneficial for the LNP to comprise a cationic lipid such as, for example, an amino lipid. Exemplary amino lipids are described in U.S. Patent Nos. 9,352,042, 9,220,683, 9,186,325, 9,139,554, 9,126,966 9,018,187, 8,999,351, 8,722,082, 8,642,076, 8,569,256, 8,466,122, and 7,745,651 and U.S. Patent Publication Nos. 2016/0213785, 2016/0199485, 2015/0265708, 2014/0288146, 2013/0123338, 2013/0116307, 2013/0064894, 2012/0172411, and 2010/0117125, all of which are incorporated herein in their entirety. In certain embodiments, the LNP comprises amino lipids described in U.S. Patent No. 9,512,073, hereby incorporated herein in its entirety.

[0177] The terms “cationic lipid” and “amino lipid” are used interchangeably herein to include lipids and salts thereof having one, two, three, or more fatty acid or fatty alkyl chains and a pH-titratable amino group (*e.g.*, an alkylamino or dialkylamino group). The cationic lipid is typically protonated (*i.e.*, positively charged) at a pH below the pKa of the cationic lipid and is substantially neutral at a pH above the pKa. The cationic lipid can also be titratable cationic lipids. In certain embodiments, the cationic lipids comprise a protonatable tertiary amine (*e.g.*, pH-titratable) group; C18 alkyl chains, wherein each alkyl chain independently can have one or more double bonds, one or more triple bonds; and ether, ester, or ketal linkages between the head group and alkyl chains.

[0178] Cationic lipids include 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLinDMA), 1,2-dilinolenyloxy-N,N-dimethylaminopropane (DLenDMA), 1,2-di- γ -linolenyloxy-N,N-dimethylaminopropane (γ -DLenDMA), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)-[1,3]-dioxolane (DLin-K-C2-DMA, also known as DLin-C2K-DMA, XTC2, and C2K), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), dilinoleylmethyl-3-dimethylaminopropionate (DLin-M-C2-DMA, also known as MC2), (6Z,9Z,28Z,31 Z)-heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (DLin-M-C3-DMA, also known as MC3), salts thereof, and mixtures thereof. Other cationic lipids also include 1,2-distearoyloxy-N,N-dimethyl-3-aminopropane (DSDMA), 1,2-dioleyloxy-N,N-dimethyl-3-aminopropane (DODMA), 2,2-dilinoleyl-4-(3-dimethylaminopropyl)-[1,3]-dioxolane (DLin-K-C3-DMA), 2,2-dilinoleyl-4-(3-dimethylaminobutyl)-[1,3]-dioxolane (DLin-K-C4-DMA), DLen-C2K-DMA, γ -DLen-C2K-DMA, and (DLin-MP-DMA) (also known as 1-B11).

[0179] Still further cationic lipids include 2,2-dilinoleyl-5-dimethylaminomethyl-[1,3]-dioxane (DLin-K6-DMA), 2,2-dilinoleyl-4-N-methylpepiazzino-[1,3]-dioxolane (DLin-K-MPZ), 1,2-dilinoleylcarbamoyloxy-3-dimethylaminopropane (DLin-C-DAP), 1,2-dilinoleyloxy-3-

(dimethylamino)acetoxyp propane (DLin-DAC), 1,2-dilinoleyloxy-3-morpholinopropane (DLin-MA), 1,2-dilinoleyloxy-3-dimethylaminopropane (DLinDAP), 1,2-dilinoleythio-3-dimethylaminopropane (DLin-S-DMA), 1-linoleyloxy-2-linoleyloxy-3-dimethylaminopropane (DLin-2-DMAP), 1,2-dilinoleyloxy-3-trimethylaminopropane chloride salt (DLin-TMA.Cl), 1,2-dilinoleyloxy-3-trimethylaminopropane chloride salt (DLin-TAP.Cl), 1,2-dilinoleyloxy-3-(N-methylpiperazino)propane (DLin-MPZ), 3-(N,N-dilinoleylamino)-1,2-propanediol (DLinAP), 3-(N,N-dioleylamino)-1,2-propanediol (DOAP), 1,2-dilinoleyloxy-3-(2-N,N-dimethylamino)ethoxyp propane (DLin-EG-DMA), N,N-dioleyl-N,N-dimethylammonium chloride (DODAC), N-(1-(2,3-dioleyloxy)propyl)-N,N,N-trimethylammonium chloride (DOTMA), N,N-distearyl-N,N-dimethylammonium bromide (DDAB), N-(1-(2,3-dioleyloxy)propyl)-N,N,N-trimethylammonium chloride (DOTAP), 3-(N-(N',N'-dimethylaminoethane)-carbamoyl)cholesterol (DC-Chol), N-(1,2-dimyristyloxyprop-3-yl)-N,N-dimethyl-N-hydroxyethyl ammonium bromide (DMRIE), 2,3-dioleyloxy-N-[2(spermine-carboxamido)ethyl]-N,N-dimethyl-1-propanaminiumtrifluoroacetate (DOSPA), dioctadecylamidoglycyl spermine (DOGS), 3-dimethylamino-2-(cholest-5-en-3-beta-oxybutan-4-oxy)-1-(cis,cis-9,12-octadecadienoxy)propane (CLinDMA), 2-[5'-(cholest-5-en-3-beta-oxy)-3'-oxapentoxo]-3-dimethyl-1-(cis,cis-9',12'-octadecadienoxy)propane (CpLinDMA), N,N-dimethyl-3,4-dioleyloxybenzylamine (DMOBA), 1,2-N,N'-dioleylcarbamyl-3-dimethylaminopropane (DOcarbDAP), 1,2-N,N'-dilinoyleylcarbamyl-3-dimethylaminopropane (DLincarbDAP), dexamethasone-sperimine (DS) and disubstituted spermine (D2S) or mixtures thereof.

[0180] A number of commercial preparations of cationic lipids can be used, such as, LIPOFECTIN® (including DOTMA and DOPE, available from GIBCO/BRL), and LIPOFECTAMINE® (comprising DOSPA and DOPE, available from GIBCO/BRL).

[0181] Additional ionizable lipids that can be used include C12-200, 3060i10, MC3, cKK-E12, bCKK-E12, Lipid 5, Lipid 9, ATX-002, ATX-003, and Merck-32. U.S. Patent Application Publication No. 2017/0367988, describes Merck-32.

[0182] In further embodiments, cationic lipid can be present in an amount from about 10% by molar ratio of the LNP to about 85% by molar ratio of the LNP, or from about 50% by molar ratio of the LNP to about 75% by molar ratio of the LNP.

[0183] LNP can comprise a neutral lipid. Neutral lipids can comprise a lipid species existing either in an uncharged or neutral zwitterionic form at physiological pH. Such lipids include diacylphosphatidylcholine, diacylphosphatidylethanolamine, ceramide, sphingomyelin, dihydrosphingomyelin, cephalin, and cerebroside. The selection of neutral lipids is generally guided by considerations including particle size and stability. In certain embodiments, the

neutral lipid component can be a lipid having two acyl groups (e.g., diacylphosphatidylcholine and diacylphosphatidylethanolamine).

[0184] Lipids having a variety of acyl chain groups of varying chain length and degree of saturation are available or can be isolated or synthesized. In certain embodiments, lipids containing saturated fatty acids with carbon chain lengths in the range of C14 to C22 can be used. In certain embodiments lipids with mono or di-unsaturated fatty acids with carbon chain lengths in the range of C14 to C22 are used. Additionally, lipids having mixtures of saturated and unsaturated fatty acid chains can be used. Exemplary neutral lipids include 1,2-dioleoyl-sn-glycero-3-phosphatidyl-ethanolamine (DOPE), 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), or a phosphatidylcholine. The neutral lipids can also be composed of sphingomyelin, dihydrosphingomyelin, or phospholipids with other head groups, such as serine and inositol.

[0185] In further embodiments, providing for neutral lipids, the neutral lipid can be present in an amount from about 0.1% by weight of the LNP to about 99% by weight of the LNP, or from about 5% by weight of the LNP to about 15% by weight of the LNP, e.g., about 1%, about 2%, about 3%, about 4%, about 5%, about 6%, about 7%, about 8%, about 9%, about 10%, about 15%, about 20%, about 25%, about 30%, about 35%, about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, or about 99%.

[0186] LNP can contain additional components such as sterols and polyethylene glycol. Sterols can confer fluidity to the LNP. As used herein “sterol” refers to a naturally occurring sterol of plant (phytosterols) or animal (zoosterols) origin as well as non-naturally occurring synthetic sterols, all of which are characterized by the presence of a hydroxyl group at the 3-position of the steroid A-ring. Suitable sterols include those conventionally used in the field of liposome, lipid vesicle or lipid particle preparation, most commonly cholesterol. Phytosterols include campesterol, sitosterol, and stigmasterol. Sterols also include sterol-modified lipids, such as those described in U.S. Patent Application Publication No. 2011/0177156. In different embodiments providing for a sterol, the sterol is present in an amount from about 1% by weight of the LNP to about 80% by weight of the LNP or from about 10% by weight of the LNP to about 25% by weight of the LNP.

[0187] Polyethylene glycol (PEG) is a water-soluble polymer of ethylene PEG repeating units with terminal hydroxyl groups. PEGs are classified by their molecular weights, for example, PEG 2000 has an average molecular weight of about 2,000 daltons, and PEG 5000 has an average molecular weight of about 5,000 daltons. PEGs commercially available from Sigma Chemical Co. and other companies include monomethoxypolyethylene glycol (MePEG-OH),

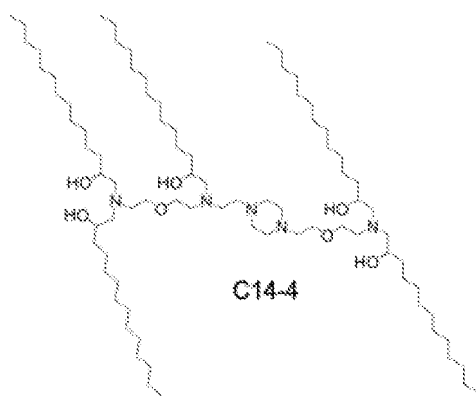
monomethoxypolyethylene glycol-succinate (MePEG-S), monomethoxypolyethylene glycol-succinimidyl succinate (MePEG-S-NHS), monomethoxypolyethylene glycol-amine (MePEG-NH₂), monomethoxypolyethylene glycol-tresylate (MePEG-TRES), and monomethoxypolyethylene glycol-imidazolyl-carbonyl (MePEG-IM).

[0188] In certain embodiments concerning PEG, PEG has an average molecular weight of about 550 to about 10,000 daltons and is optionally substituted by alkyl, alkoxy, acyl or aryl. In further embodiments, the PEG is substituted with methyl at the terminal hydroxyl position. In further embodiments, the PEG has an average molecular weight from about 750 to about 5,000 daltons, or from about 1,000 to about 5,000 daltons, or from about 1,500 to about 3,000 daltons, or from about 2,000 daltons, or from about 750 daltons.

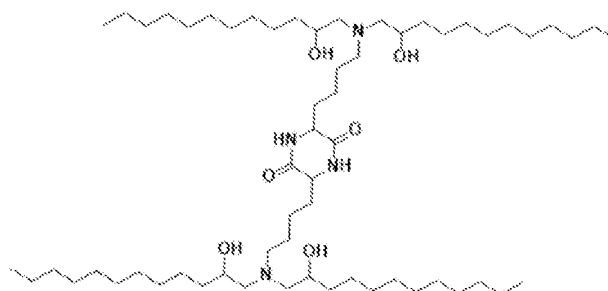
[0189] PEG-modified lipids include the PEG-dialkyloxypopyl conjugates (PEG-DAA) described in U.S. Patent Nos. 8,936,942 and 7,803,397. PEG-modified lipids (or lipid-polyoxyethylene conjugates) can have a variety of “anchoring” lipid portions to secure the PEG portion to the surface of the lipid vesicle. Examples of suitable PEG-modified lipids include PEG-modified phosphatidylethanolamine and phosphatidic acid, PEG-ceramide conjugates (*e.g.*, PEG-CerC14 or PEG-CerC20) which are described in U.S. Patent No. 5,820,873, PEG-modified dialkylamines and PEG-modified 1,2-diacyloxypropan-3-amines. In certain embodiments, the PEG-modified lipid can be PEG-modified diacylglycerols and dialkylglycerols. In certain embodiments, the PEG can be in an amount from about 0.1% by weight of the LNP to about 50% by weight of the LNP, or from about 5% by weight of the LNP to about 15% by weight of the LNP.

[0190] In further embodiments concerning LNP size, prior to encapsulating nucleic acid, LNPs have a size range from about 10 nm to 500 nm, or from about 50 nm to about 200 nm, or from 75 nm to about 125 nm.

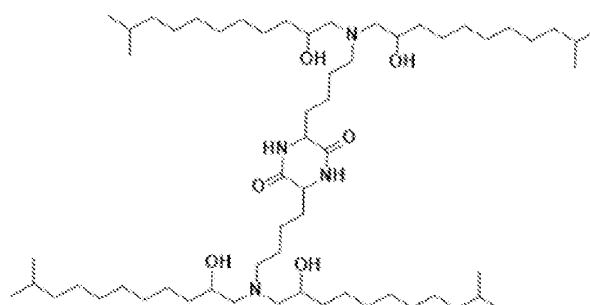
[0191] In certain embodiments concerning LNP, the LNP is described by Billingsley *et al.*, Nano Lett. 2020, 20, 1578 or Billingsley *et al.*, International Patent Publication No. WO 2021/077066 (both of which are hereby incorporated by reference herein in their entirety). Billingsley *et al.*, and WO2021/077066 describe LNPs containing lipid-anchored PEG, cholesterol, phospholipid and ionizable lipids. In certain embodiments, the LNP contains a C14-4 polyamine core and/or has a particle size of about 70 nm. C14-4 has the following structure.



[0192] In certain embodiments the LNP is made up of a cationic lipid or lipopeptide described by U.S. Patent No. 10,493,031, U.S. Patent No. 10,682,374 or International Patent Publication No. WO2021/077066 (each of which is hereby incorporated by reference herein in their entirety). In certain embodiments, the LNP contains a cationic lipid, a cholesterol-based lipid, and/or one or more PEG-modified lipids. In certain embodiments the LNP contains cKK-E12 (Dong *et al.*, PNAS (2014) 111(11), 3955):

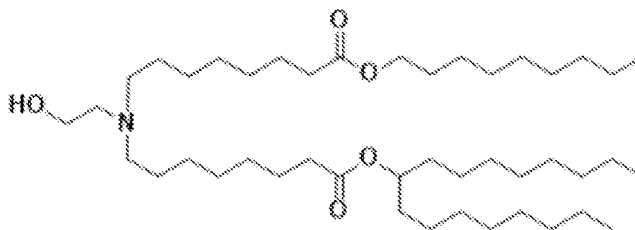


[0193] In certain embodiments the LNP comprises a modified form of cKK-E12 referred to herein as “bCKK-E12,” having the following structure:

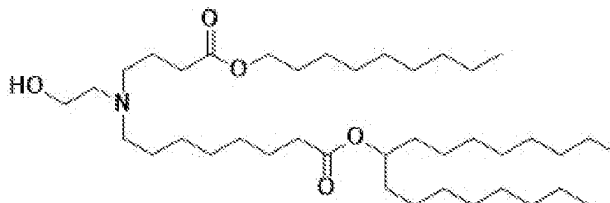


[0194] In certain embodiments the LNP comprises Lipid 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 as described by Sabnis *et al.*, Molecular Therapy 2018, 26:6, 1509-1519 (hereby incorporated by reference herein in its entirety). In certain embodiments the LNP comprises Lipid 5, 8, 9, 10, or 11 described in Sabnis *et al.*

[0195] Lipid 5 of Sabnis *et al.* has the structure:



[0196] Lipid 9 of Sabnis *et al.* has the structure:



[0197] Additional lipids that may be utilized include those described by Roces *et al.*, *Pharmaceutics*, 2020, 12, 1095; Jayaraman *et al.*, *Angew. Chem. Int. Ed.*, 2012, 51, 8529-8533; Maier *et al.*, www.moleculartherapy.org, 2013, Vol.21, No. 8, 1570-1578; Liu *et al.*, *Adv. Mater.* 2019, 31, 1902575, *e.g.*, BAMEA-O16B; Cheng *et al.*, *Adv. Mater.*, 2018, 30, 1805308, *e.g.*, 5A2-SC8; Hajj *et al.*, *Small*, 2019 15, 1805097, *e.g.*, 306Oi10; Du *et al.*, U.S. Patent Application Publication No. 20160376224; and Tanaka *et al.*, *Adv. Funct. Mater.*, 2020, 30, 1910575; each of which are hereby incorporated by reference herein in their entirety.

[0198] In further embodiments, the nanoparticle is an LNP. In further embodiments the LNP in mol% comprises, consists essentially, or consists, of the following components: (1) one or more cationic lipids from about 20% to 65%, one or more phospholipid lipids from about 1% to about 50%, one or more PEG-conjugated lipid from about 0.1 % to 10%, and cholesterol from about 0% to about 70%; or (2) one or more cationic lipids from about 20% to 50%, one or more phospholipid lipids from about 5% to about 20%, one or more PEG-conjugated lipids from about 0.1 % to 5%, and cholesterol from about 20% to about 60%. In further embodiments, the phospholipid lipid is a neutral lipid; and the phospholipid lipid is DOPE or DSPC.

[0199] In certain embodiments the LNP, in mole %, comprises, consists essentially, or consists of the following components: (1) cKK-E12 about 35%; C14-PEG2000, about 2.5%; cholesterol, about 46.5%; and DOPE, about 16%; (2) bCKK-E12 about 35%; C14-PEG2000, about 2.5%; cholesterol, about 46.5%; and DOPE, about 16%; (3) Lipid 9 about 50%; C14-PEG2000, about 1.5%; cholesterol, about 38.5%; and DSPC about 10%; (4) Lipid 5 about 50%; C14-PEG2000 about 1.5%; cholesterol about 38.5%; and DSPC about 10%; (5) ionizable lipid, about 50%; DSPC, about 10%; cholesterol, about 37.5%; and stabilizer (PEG-Lipid), about 2.5%; or (6) is GenVoy-ILM™ LNP (Precision NanoSystems).

VI.B. Polymer-Based Nanoparticles

[0200] Polymer-based delivery systems can be made from a variety of different natural and synthetic materials. DNA and other compounds can be entrapped into the polymeric matrix of polymeric nanoparticles or can be adsorbed or conjugated on the surface of the nanoparticles. Examples of commonly used polymers for nucleic acid delivery include poly(lactic-co-glycolic acid) (PLGA), poly lactic acid (PLA), poly(ethylene imine) (PEI) and PEI derivatives, chitosan, dendrimers, polyanhydride, polycaprolactone, polymethacrylates, poly-L-lysine, pullulan, dextran, and hyaluronic acid, poly- β -aminoesters. (Thomas *et al.*, (2019) *Molecules* 24, 3744.)

[0201] Polymeric-based nanoparticles can have different sizes, ranging from about 1 nm to about 1000 nm, from about 10 nm to about 500 nm, from about 50 nm to about 200 nm, from about 100 nm to about 150 nm, and from about 150 nm or less.

VI.C. Lipid Polymer Nanoparticles

[0202] Lipid polymer nanoparticles are hybrid nanoparticles providing both a lipid component and a polymer component, and as such can be considered to be an LNP or LPNP. The LPNP configuration can provide an outer polymer and inner lipid or an outer lipid and inner polymer. The presence of two different types of material facilitates designing nanoparticles to provide for delayed release of a component. Different lipid and polymer components can be selected taking into account the material to be delivered. (For example, see Teo *et al.*, *Advanced Drug Delivery Reviews* (2016) 98, 41; Boichicchio *et al.*, *Pharmaceutics* (2021) 13, 198; Mahzabin and Das, *IJPSR* (2021) 12(1), 65; and Teixeira *et al.*, (2017) *Prog. Lipid Res.* Oct; 68:1-11.)

IV.D. Protein and Peptide-Based Nanoparticles

[0203] Protein and peptide-based systems can employ a variety of different proteins and peptides. Examples of proteins that can be employed include gelatin and elastin. Peptide-based systems can employ, for example, cell-penetrating peptides (CPPs).

[0204] CPPs are short peptides (6–30 amino acid residues) potentially capable of intracellular penetration to deliver therapeutic molecules. The majority of CPPs consists mainly of arginine and lysine residues, making them cationic and hydrophilic, but CPPs can also be amphiphilic, anionic, or hydrophobic. CPPs can be derived from natural biomolecules (*e.g.*, HIV-1 Tat protein), or obtained by synthetic methods (*e.g.*, poly-L-lysine, polyarginine) (Singh *et al.*, *Drug Deliv.* 2018;25(1):1996-2006). Examples of CPPs include cationic CPPs (highly positively charged) such as the Tat peptide, penetratin, protamine, poly-L-lysine, and polyarginine; amphipathic CPPs (chimeric or fused peptides, constructed from different sources, containing both positively and negatively charged amino acid sequences), such as transportan, VT5, battenecin-7 (Bac7), proline-rich peptide (PPR), SAP (VRLPPP)₃, TP10, pep-1, and MPG), membranotropic CPPs (exhibit both hydrophobic and amphipathic nature simultaneously, and

comprise both large aromatic residues and small residues) such as H625, SPIONs-PEG-CPP and NPs; and hydrophobic CPPs (contain only non-polar motifs or residues) such as SG3, PFVYLI, pep-7, and fibroblast growth factors.

[0205] Protein and peptide nanoparticles can be provided in different sizes, for example, ranging from about 1 nm to about 1000 nm, from about 10 nm to about 500 nm, from about 50 nm to about 200 nm, from about 100 nm to about 150 nm, or from about 150 nm or less.

[0206] VI.E. Peptide Cage Nanoparticles

[0207] Peptide cage-based delivery systems can be produced from proteinaceous material able to assemble into a cage-like structure forming a constrained internal environment. Peptide cages can comprise a proteinaceous shell that self-assembles to form a protein cage (*e.g.*, a structure with an interior cavity that is either naturally accessible to the solvent or can be made so by altering solvent concentration, pH, or equilibria ratios). The monomers of the protein cages can be naturally occurring or variant forms, including amino acid substitutions, insertions, and deletions (*e.g.*, fragments).

[0208] Different types of protein “shells” can be assembled and loaded with different types of materials. Protein cages can be produced using viral coat protein(s) (*e.g.*, from the Cowpea Chlorotic Mottle Virus protein coat), as well non-viral proteins (*e.g.*, U.S. Patent Nos. 6,180,389 and 6,984,386, U.S. Patent Publication No. 20040028694, and U.S. Patent Publication No. 20090035389, each of which is incorporated by reference herein in their entirety).

[0209] Examples of protein cages derived from non-viral proteins include: eukaryotic or prokaryotic derived ferritins and apoferritins such as 12 and 24 subunit ferritins; and heat shock proteins (HSPs), such as the class of 24 subunit heat shock proteins that form an internal core space, the small HSP of *Methanococcus jannaschii*, the dodecameric Dsp HSP of *E. coli*; and the MrgA protein.

[0210] Protein cages can have different core sizes, such as ranging from about 1 nm to about 1000 nm, from about 10 nm to about 500 nm, from about 50 nm to about 200 nm, from about 100 nm to about 150 nm, or from about 150 nm or less.

VI.F. Exosomes

[0211] Exosomes are small biological membrane vesicles and have been utilized to deliver various cargoes including small molecules, peptides, proteins and nucleic acids. Exosomes generally range in size from about 30 nm to 100 nm and can be taken up by a cell and deliver its cargo. Cargoes can be associated with exosome surface structure or may be encapsulated within the exosome bilayer.

[0212] Various modifications can be made to exosomes facilitating cargo delivery and cell targeting. Modifications for facilitating cargo delivery include structures for associating with

cargoes such as protein scaffolds and polymers. Modifications for cell targeting include targeting ligands and modifying surface charge. Publications describing production, modification, and use of exosomes for delivery of different cargoes include Munagala *et al.*, Cancer Letters (2021), 505, 58; Fu *et al.*, (2020) NanoImpact 20, 100261; and Dooley *et al.*, (2021) Molecular Therapy 29(5), 1729 (each of which is hereby incorporated by reference herein).

VII. Pharmaceutical Compositions

[0213] Pharmaceutical compositions comprise a pharmaceutical acceptable carrier facilitating administration and/or storage of the PPT1 polypeptides, encoding polynucleotides, viral vectors or non-viral vectors. Reference to “pharmaceutically acceptable” indicates the components do not cause substantial undesirable biological effects at the amount utilized. Pharmaceutically acceptable carriers can contain different components such as one or more pharmaceutically acceptable excipients. Examples of pharmaceutically acceptable excipients include salt, sugar, buffer, solvent, preservative, protein and surfactant. A particular excipient can have more than one function. Examples of pharmaceutically acceptable excipients and carriers that can be used with viral vectors are provided in, for example, International Patent Publication No. WO2021/071835.

[0214] Pharmaceutical compositions can be formulated to be compatible with a particular route of administration or delivery. Compositions suitable for parenteral administration include aqueous and non-aqueous solutions, suspensions or emulsions, which preparations are typically sterile and can be isotonic with the blood of the intended recipient. Illustrative examples include water, buffered saline, Hanks’ solution, Ringer’s solution, dextrose, fructose, ethanol, animal vegetable and synthetic oils. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran.

[0215] In an embodiment, the pharmaceutical composition contains a formulation capable of injection into a subject. Examples of injectable formulation components include isotonic, sterile, saline solutions, salts (*e.g.*, monosodium or disodium phosphate, sodium, potassium, calcium or magnesium chloride and mixtures of such salts), buffered saline, sugars (*e.g.*, dextrose), and water for injection. Pharmaceutical compositions include dry, for example, freeze-dried compositions which upon addition of sterilized water or physiological saline, permit the constitution of solutions suitable for administration.

[0216] Additionally, suspensions can be prepared as appropriate oil injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension can

also contain suitable stabilizers or agents which increase compound solubility facilitating the preparation of concentrated solutions.

[0217] An “effective amount” or “sufficient amount” refers to an amount providing an indicated or desired effect. The effective amount can be administered, in single or multiple doses, alone or in combination, with one or more other compositions (*e.g.*, additional therapeutic or immunosuppressive agents), treatments, protocols, or therapeutic regimens; and provide for a long or short term response.

[0218] Pharmaceutical compositions comprising transgenes encoding PPT1 polypeptides can be delivered to a subject, so as to allow production of the encoded protein. Delivery can be *in vivo* or *ex vivo*. In certain embodiments, pharmaceutical compositions comprise sufficient genetic material to enable a recipient to produce a therapeutically effective amount of a protein in the subject.

[0219] A “therapeutically effective amount” refers to an amount of an active ingredient or component that elicits the desired or indicated biological or medicinal response in a subject. A therapeutically effective amount can be determined based on observed symptoms and/or through the use of biomarkers associated with a particular disease or disorder. Selection of a particular effective dose can be optimized taking into account different factors, including the disease or disorder to be treated or prevented, the symptoms involved, safety and effectiveness in animal models, the patient’s body mass, and the patient’s immune status. The optimal dose to be employed in the formulation will also depend on the route of administration, and the severity of the disease or disorder, and can be evaluated depending upon patient’s circumstances. Effective doses can be extrapolated from dose-response curves derived from *in vitro* or animal model test systems.

[0220] In certain embodiments, a pharmaceutical composition comprising a rAAV vector comprises empty AAV capsids. In certain embodiments, in a pharmaceutical composition comprising rAAV vectors and empty AAV capsids, the ratio of empty AAV capsids to rAAV vector is within or between about 100:1-50:1, from about 50:1-25:1, from about 25:1-10:1, from about 10:1-1:1, from about 1:1-1:10, from about 1:10-1:25, from about 1:25-1:50, or from about 1:50-1:100. In certain embodiments, the ratio of the empty AAV capsids to the rAAV vector is about 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, or 10:1.

[0221] Additional guidance and examples of pharmaceutical compositions and delivery systems are provided in, for example, Remington: The Science and Practice of Pharmacy (2020) 23th ed., University of the Sciences in Philadelphia, published by Elsevier; The Merck Index (2013) 15th ed., Whitehouse, NJ; Pharmaceutical Principles of Solid Dosage Forms (1993), Technomic

Publishing Co., Inc., Lancaster, Pa.; and Ansel and Stoklosa, Pharmaceutical Calculations (2001) 11th ed., Lippincott Williams & Wilkins, Baltimore, MD.

VIII. Administration and Treatment

[0222] PPT1 polypeptides and encoding polynucleotide constructs, viral vectors and non-viral vectors can be administered to a subject, preferably a human subject, to provide for prophylactic treatment reducing the likelihood or severity of a disease or order and/or treating a diagnosed disease or disorder. In certain embodiments, the particular therapeutic agent, route of administration, and/or pharmaceutical composition is selected taking into account the particular disease or disorder being treated.

[0223] Subjects having a particular disease or disorder, or at an increased risk of a particular disease or disorder, can be identified, for example, based on symptoms, PPT1 activity, biomarkers and genetic markers. Treatment can be carried, for example, on subjects at increased risk for developing NCL1 symptoms and subjects diagnosed with NCL1.

[0224] NCL1 is an autosomal recessive progressive neurodegenerative disease, whose symptoms vary depending upon onset. Major symptoms include: onset at birth (congenital), microcephaly, dysmorphic features, seizures, and hyperkinetic activity; onset 6-18 months (infantile), decreased head growth, neuro-developmental regression, and seizures; onset 2-4 years (variant), seizures, neuro-development regression, and behavioral disturbances; and onset 5-7 years (juvenile), visual loss and cognitive decline. (Simonati and Williams (2022) Front. Neurol. 11;13:811686.)

[0225] Different mutations are associated with NCL1. References providing examples of mutations associated with NCL1 include Sheth *et al.*, (2018) BMC Neurol. 12;18(1):203; Kumar *et al.*, Advances in Protein Chemistry and Structural Biology (2022) 132:89-109; Kousi *et al.*, (2012) Hum. Mutat. 33(1):42-63; and [hyperlink://www.uniprot.org/uniprotkb/P50897/entry#disease_variants](https://www.uniprot.org/uniprotkb/P50897/entry#disease_variants) (January 17, 2023); each of which are incorporated by reference herein in their entirety.

[0226] Because of the progressive nature of the NCL1, early treatment is very important. In certain embodiments, treatment is carried out prior to identification of major symptoms. Such patients can be identified, for example, based on the PPT1 enzymatic activity level and/or presence of PPT1 mutations.

[0227] In certain embodiments, treatment is carried out on a patient diagnosed with NCL1. Diagnose can be based on symptoms, genetic testing, and/or measuring enzymatic activity.

[0228] Potential routes of administration include subcutaneously, epidurally, intradermally, intrathecally, intraorbitally, intramucosally, intranasally, intraperitoneally, intravenously, intrapleurally, intraarterially, intracavitary, orally, intrahepatically, via the portal vein,

intramuscularly, intraparenchymal, intracisternal, intracranial, intra cisterna magna, intracerebroventricular or intraventricular administration. In certain embodiments viral or non-viral vectors are administered to a patient via infusion in a pharmaceutically carrier.

[0229] Suitable routes of administration should provide for therapeutic delivery to the CNS. Delivery to the CNS can be achieved using different routes of initial administration including outside of the CNS, for example, intravenous administration outside of the brain and spinal cord, and to the eye (*e.g.*, intravitreal and subretinal); and direct administration to the brain (*e.g.*, intraparenchymal, intracerebroventricular, and intracisternal) and/or spine (*e.g.*, intrathecal). (Zhu *et al.*, (2021) Trends Mol. Med. 27(6):524-537 hereby incorporated by reference herein in its entirety.)

[0230] When the initial site of administration is outside of the brain, administration to the brain can be facilitated using techniques providing transport across the blood brain barrier. Examples of such techniques include disruption of the blood brain barrier and making use blood brain barrier carriers (Chen *et al.*, (2021) J. Control. Release 333:129-138; Bellettato and Scrapa, Italian Journal of Pediatrics (2018) 44(Suppl 2):131; Haumann *et al.*, (2020) CNS Drugs 34, 1121-1131; and Cammalleri *et al.*, (2020) J. Clin. Neurophysiol. 37(2):104-117; each of which are hereby incorporated by reference herein in their entirety). Techniques facilitating crossing the blood brain barrier can be utilized on the delivery vehicle and/or the PPT1 protein.

[0231] In certain embodiments, treatment is carried out using expression systems providing for polypeptide expression outside of the CNS (*e.g.*, high liver expression) in combination with techniques facilitating PPT1 transport across the blood brain barrier.

[0232] In certain embodiments, treatment is carried out using techniques facilitating the delivery vehicle crossing the blood brain barrier. In a further embodiment, crossing of the blood brain barrier is facilitated using focused ultrasound in combination with microbubbles (See, for example, Cammalleri *et al.*, (2020) J Clin Neurophysiol. 37(2):104-117, hereby incorporated by reference herein in its entirety).

[0233] In certain embodiments, an AAV capsid providing for CNS or crossing the blood brain barrier is used. Examples of such capsids are provided in Chen *et al.*, (2021) J. Control. Release 333:129-138 (*e.g.*, AAV9, AAV-PHP-B, AAV-PHP.eB, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, and AAV1/rh.10), U.S. Patent No. 9,585,971, and U.S. Patent Publication No. US202/1214749, each of which are hereby incorporated by reference herein in their entirety.

[0234] CNS administration can also be carried out, for example, by direct administration to the brain using needles or catheters. (For example, International Publication No. WO 2021/108809; Cohen-Pferrer *et al.*, Pediatric Neurology 67 (2017) 23-35; and U.S. Patent No. 10,369,329; each of which are hereby incorporated by reference herein in their entirety.)

[0235] Another example of a technique for CNS administration is convection enhanced delivery. Convection enhanced delivery comprises surgical exposure of the brain followed by placement of a catheter directly into the target area, followed by infusion of a therapeutic agent (for example, U.S. Patent Publication No. 2022/010001; and Debinski *et al.*, (2009) Expert Rev Neurother. 9(10):1519-27; both of which are hereby incorporated by reference herein in their entirety).

[0236] CNS delivery devices, systems and techniques also include those described in, for example, U.S. Patent No. 8128600, U.S. Patent Publication No. 2020/0324089, U.S. Patent No. 11129643, U.S. Patent No. 11154377, U.S. Patent Publication No. 2021/0343397, U.S. Patent Publication No. 2021/0282866, U.S. Patent No. 9572928, U.S. Patent No. 8337458, U.S. Patent No. 10722265, and US Patent publication No. 2021/214749, each of which are incorporated by reference herein in their entirety.

[0237] Delivery of PPT1 polypeptide and encoding nucleic acid, may also provide a benefit in treating PPT1 deficiency or defective lysosomal storage outside of the CNS. In certain embodiments, administration provides for systemic delivery. In further embodiments, treatment comprises the use of an expression cassette or viral vectors comprising a ubiquitous or promiscuous promotor. In further embodiments, CNS entry is facilitated using techniques and/or vectors facilitating transport across the blood brain barrier.

[0238] In certain embodiments, the expression cassette comprises a PGK promoter, CBh promoter, or E1F alpha promoter.

[0239] Optimal doses can vary depending upon different factors such as a particular therapeutic, desired endpoint and administration route. The dose amount, number, frequency or duration can be proportionally increased or reduced, taking into account adverse side effects, complications or other risk factors of the treatment or therapy and the status of the subject.

[0240] A “unit dosage form” refers to a physically discrete unit containing a predetermined effective amount of active ingredient in combination with a pharmaceutically acceptable carrier. Unit dosage forms can be provided within, for example, ampules and vials, which can include a pharmaceutically acceptable carrier, or a composition in a freeze-dried or lyophilized state. In the case of a freeze-dried or lyophilized state, a sterile liquid carrier can be added prior to administration. Individual unit dosage forms can be included in multi-dose kits or containers.

[0241] An “effective amount” achieves the desired or indicated effect. For example, an effective amount for treatment decreases one or more adverse symptoms, reduces the likelihood of one or more symptoms associated with a disease or disorder, or reduces disease or disorder progression. Preferred effective amounts for treatment are effective to decrease multiple or all adverse symptoms.

[0242] In certain embodiments, a pharmaceutical composition is administered to a subject at a dose suitable to increase PPT1 activity. In certain embodiments, the dose is sufficient to increase PPT1 activity to a level of at least 5%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 100% normal activity; about 10% to 200% average activity, 20% to 150% average activity, 30% to 100% average activity; or about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 100%, about 110%, about 120%, about 130%, about 140% or about 150% average activity. In certain embodiments, PPT1 activity increases to 10X, 100X or 1000X average activity. Average activity refers to the average of activity occurring in the general population.

[0243] In different embodiments a suitable dosage is from about 0.01 mg/kg to about 10 mg/kg of vector per kg body weight of a subject, about 0.01 mg/kg to about 0.1 mg/kg of vector per kg body weight of a subject, about 0.1 mg/kg to about 1.0 mg/kg of vector per kg body weight of a subject, or about 1.0 mg/kg to about 10 mg/kg of vector per body weight of a subject.

[0244] Generally, rAAV doses range from at least 1×10^8 vector genomes per kilogram (vg/kg) of the weight of the subject, or more, for example, 1×10^9 , 1×10^{10} , 1×10^{11} , 1×10^{12} , 1×10^{13} or 1×10^{14} , or more, vector genomes per kilogram (vg/kg) of the weight of the subject, to achieve a therapeutic effect. In different embodiments the rAAV dose is about 5×10^{11} rAAV vg/kg or greater than about 5×10^{11} rAAV vg/kg; about 1×10^{12} rAAV vg/kg or greater than about 1×10^{12} rAAV vg/kg; about 2×10^{12} rAAV vg/kg or greater than about 2×10^{12} rAAV vg/kg; about 3×10^{12} rAAV vg/kg or greater than about 3×10^{12} rAAV vg/kg; about 4×10^{12} rAAV vg/kg or greater than about 4×10^{12} rAAV vg/kg; about 5×10^{12} rAAV vg/kg or greater than about 5×10^{12} rAAV vg/kg; about 1×10^{13} rAAV vg/kg or greater than about 1×10^{13} rAAV vg/kg; about 2×10^{13} rAAV vg/kg or greater than about 2×10^{13} rAAV vg/kg; about 3×10^{13} rAAV vg/kg or greater than about 3×10^{13} rAAV vg/kg; about 4×10^{13} rAAV vg/kg or greater than about 4×10^{13} rAAV vg/kg; about 5×10^{13} rAAV vg/kg or greater than about 5×10^{13} rAAV vg/kg; about 6×10^{13} rAAV vg/kg or greater than about 6×10^{13} rAAV vg/kg.

[0245] Examples of dose ranges of rAAV vg/kg include a dose range from about 5×10^{11} to about 6×10^{13} rAAV vg/kg; a dose range from about 5×10^{11} to about 5.5×10^{11} rAAV vg/kg; a dose range from about 5.5×10^{11} to about 6×10^{11} rAAV vg/kg; a dose range from about 6×10^{11} to about 6.5×10^{11} rAAV vg/kg; a dose range from about 6.5×10^{11} to about 7×10^{11} rAAV vg/kg; a dose range from about 7×10^{11} to about 7.5×10^{11} rAAV vg/kg; a dose range from about 7.5×10^{11} to about 8×10^{11} rAAV vg/kg; a dose range from about 8×10^{11} to about 8.5×10^{11} rAAV vg/kg; a dose range from about 8.5×10^{11} to about 9×10^{11} rAAV vg/kg; a dose range from about 9×10^{11} to about 9.5×10^{11} rAAV vg/kg; a dose range from about 9.5×10^{11} to about 1×10^{12} rAAV vg/kg; a

dose range from about 1×10^{12} to about 1.5×10^{12} rAAV vg/kg; a dose range from about 1.5×10^{12} to about 2×10^{12} rAAV vg/kg; a dose range from about 2×10^{12} to about 2.5×10^{12} rAAV vg/kg; a dose range from about 2.5×10^{12} to about 3×10^{12} rAAV vg/kg; a dose range from about 3×10^{12} to about 3.5×10^{12} rAAV vg/kg; a dose range from about 3.5×10^{12} to about 4×10^{12} rAAV vg/kg; a dose range from about 4×10^{12} to about 4.5×10^{12} rAAV vg/kg; a dose range from about 4.5×10^{12} to about 5×10^{12} rAAV vg/kg; a dose range from about 5×10^{12} to about 5.5×10^{12} rAAV vg/kg; a dose range from about 5.5×10^{12} to about 6×10^{12} rAAV vg/kg; a dose range from about 6×10^{12} to about 6.5×10^{12} rAAV vg/kg; a dose range from about 6.5×10^{12} to about 7×10^{12} rAAV vg/kg; a dose range from about 7×10^{12} to about 7.5×10^{12} rAAV vg/kg; a dose range from about 7.5×10^{12} to about 8×10^{12} rAAV vg/kg; a dose range from about 8×10^{12} to about 8.5×10^{12} rAAV vg/kg; a dose range from about 8.5×10^{12} to about 9×10^{12} rAAV vg/kg; a dose range from about 9×10^{12} to about 9.5×10^{12} rAAV vg/kg; a dose range from about 9.5×10^{12} to about 1×10^{13} rAAV vg/kg; a dose range from about 1×10^{13} to about 1.5×10^{13} rAAV vg/kg; a dose range from about 1.5×10^{13} to about 2×10^{13} rAAV vg/kg; a dose range from about 2×10^{13} to about 2.5×10^{13} rAAV vg/kg; a dose range from about 2.5×10^{13} to about 3×10^{13} rAAV vg/kg; a dose range from about 3×10^{13} to about 3.5×10^{13} rAAV vg/kg; a dose range from about 3.5×10^{13} to about 4×10^{13} rAAV vg/kg; a dose range from about 4×10^{13} to about 4.5×10^{13} rAAV vg/kg; a dose range from about 4.5×10^{13} to about 5×10^{13} rAAV vg/kg; a dose range from about 5×10^{13} to about 5.5×10^{13} rAAV vg/kg; a dose range from about 5.5×10^{13} to about 6×10^{13} rAAV vg/kg; a dose range from about 6×10^{13} to about 1×10^{14} rAAV vg/kg.

[0246] In certain embodiments, rAAV vg/kg are administered at a dose of about 5×10^{11} vg/kg, about 6×10^{11} vg/kg, about 7×10^{11} vg/kg, about 8×10^{11} vg/kg, about 9×10^{11} vg/kg, about 1×10^{12} vg/kg, about 2×10^{12} vg/kg, about 3×10^{12} vg/kg, about 4×10^{12} vg/kg, about 5×10^{12} vg/kg, about 6×10^{12} vg/kg, about 7×10^{12} vg/kg, about 8×10^{12} vg/kg, about 9×10^{12} vg/kg, about 1×10^{13} vg/kg, about 2×10^{13} vg/kg, about 3×10^{13} vg/kg, about 4×10^{13} vg/kg, about 5×10^{13} vg/kg, or about 6×10^{13} vg/kg.

[0247] In certain embodiments doses and dose ranges for other viral vectors is as provided herein with respect to rAAV. For example, in certain embodiments the dose and dose range for recombinant adenovirus vectors, recombinant retrovirus vectors (*e.g.*, lentivirus), and recombinant herpes simplex virus vectors is the same as illustrated above with respect to rAAV.

[0248] In different embodiments a suitable dosage for PPT1 administration is from about 0.01 mg/kg to about 25 mg/kg of protein per kg body weight of a subject or about 0.1 mg/kg to about 1.0 mg/kg of protein per kg body weight of a subject.

[0249] In certain embodiments, the polypeptide constructs, polynucleotide constructs, viral vectors and non-viral vectors described herein are administered in combination with additional

compounds or treatments for a particular disease or disorder; and/or in combination with a compound decreasing an immune response generated against the provided or produced polypeptide, polynucleotide, and/or delivery vehicle. Additional compounds or treatments can be provided in different modalities such as administered separately; and administered or performed prior to, substantially contemporaneously with or following administration of the polypeptide constructs, encoding polynucleotide constructs, viral vectors and non-viral vectors described herein.

[0250] In certain embodiments, administration of polypeptide constructs, encoding polynucleotide constructs, viral vectors and non-viral vectors described herein is in combination with an immunosuppressive agent or regimen. Such agents and regimens can be utilized, as needed, to achieve immune tolerance or mitigate the immune response to the produced PPT1 protein, the provided polynucleotides, or the provided delivery vehicles. Examples of immunosuppressive agents and regimens include methotrexate, rituximab, intravenous gamma globulin (IVIG), omalizumab, ImmTOR[®] (synthetic vaccine particle (SVP)-rapamycin (rapamycin encapsulated in a biodegradable nanoparticle)), ImmTOR-IL[™] (ImmTOR with Treg-selective IL-2 agonist), B-cell depletion, immunoadsorption, and plasmapheresis.

[0251] In certain embodiments, the viral vector or non-viral vector is administered in conjunction with one or more immunosuppressive agents, where one or more immunosuppressive agent is administered prior to, substantially at the same time as, or after, administering the vector or non-viral vector. In certain embodiments, the one or more immunosuppressive agent is administered concomitantly with the vector or non-viral vector. In certain embodiments, the one or more immunosuppressive agents is administered 1-12, 12-24 or 24-48 hours; or 2-4, 4-6, 6-8, 8-10, 10-14, 14-20, 20-25, 25-30, 30-50 days, or more than 50 days prior to viral or non-viral vector administration. In certain embodiments, the one or more immunosuppressive agent is administered 1-12, 12-24 or 24-48 hours; or 2-4, 4-6, 6-8, 8-10, 10-14, 14-20, 20-25, 25-30, 30-50 days, or more than 50 days, following viral or non-viral vector administration. Administration of immunosuppressive agents after a period of time following administering vector or non-viral vector can be done, for example, if there is a decrease in the encoded protein after the initial expression levels for a period of time, e.g., 20-25, 25-30, 30-50, 50-75, 75-100, 100-150, 150-200 or more than 200 days following vector or non-viral vector administration.

[0252] In certain embodiments, the immunosuppressive agent is an anti-inflammatory agent. In certain embodiments, the immunosuppressive agent is a steroid, e.g., a corticosteroid. In certain embodiments, the immunosuppressive agent is prednisone, prednisolone, calcineurin inhibitor (e.g., cyclosporine, tacrolimus), MMF (mycophenolic acid, e.g. CellCept[®], Myfortic[®]), CD52

inhibitor (*e.g.*, alemtuzumab), CTLA4-Ig (*e.g.*, abatacept, belatacept), anti-CD3 mAb, anti-LFA-1 mAb (*e.g.*, efalizumab), anti-CD40 mAb (*e.g.*, ASKP1240), anti-CD22 mAb (*e.g.*, epratuzumab), anti-CD20 mAb (*e.g.*, rituximab, orelizumab, ofatumumab, veltuzumab), proteasome inhibitor (*e.g.*, bortezomib), TACI-Ig (*e.g.*, atacicept), anti-C5 mAb (*e.g.*, eculizumab), mycophenolate, azathioprine, sirolimus everolimus, TNFR-Ig, anti-TNF mAb, tofacitinib, anti-IL-2R (*e.g.*, basiliximab), anti-IL-17 mAb (*e.g.*, secukinumab), anti-IL-6 mAb (*e.g.*, anti-IL-6 antibody sirukumab, anti-IL-6 receptor antibody tocilizumab (Actemra®), IL-10 inhibitor, TGF-beta inhibitor, a B cell targeting antibody (*e.g.*, rituximab), a mammalian target of rapamycin (mTOR) inhibitor (*e.g.*, rapamycin), synthetic vaccine particle (SVP™)-rapamycin (rapamycin encapsulated in a biodegradable nanoparticle), intravenous gamma globulin (IVIG), omalizumab, methotrexate, a tyrosine kinase inhibitor (*e.g.*, ibrutinib), cyclophosphamide, fingolimod, an inhibitor of B-cell activating factor (BAFF) (*e.g.*, anti-BAFF mAb, *e.g.*, belimumab), an inhibitor of a proliferation-inducing ligand (APRIL), anti-IL-1b mAb (*e.g.*, canakinumab (Haris®)), a C3a inhibitor, a Tregitope (see, *e.g.*, U.S. Patent No. 10,213,496), or a combination and/or derivative thereof.

[0253] Immune-suppression protocols, including the use of rapamycin, alone or in combination with IL-10, can be used to decrease, reduce, inhibit, prevent or block humoral and cellular immune responses to PPT1 protein. Hepatic gene transfer with viral vector (*e.g.*, rAAV) and non-viral vector can be used to induce immune tolerance to PPT1 protein through induction of regulatory T cells (Tregs).

[0254] Strategies to reduce (overcome) or avoid humoral immunity to viral vectors, such as rAAV in systemic gene transfer include, administering high vector doses; use of AAV empty capsids as decoys to adsorb anti-AAV antibodies; administration of immunosuppressive drugs to decrease, reduce, inhibit, prevent or eradicate the humoral immune response to rAAV; changing the rAAV capsid serotype or engineering the rAAV capsid to be less susceptible to neutralizing antibodies; use of plasma exchange cycles to adsorb anti-AAV immunoglobulins, thereby reducing anti-AAV antibody titer; and use of delivery techniques such as balloon catheters followed by saline flushing. Such strategies are described in Mingozi *et al.*, (2013) Blood, 122:23-36. Additional strategies include using AAV-specific plasmapheresis columns to selectively deplete anti-AAV antibodies without depleting the total immunoglobulin pool from plasma, as described in Bertin *et al.*, 2020, *Sci. Rep.* 10:864. Similar techniques and strategies can be used for other types of viral vectors.

[0255] Empty capsids used as decoy probes can be provided in different ratios to viral vectors. Amounts of empty capsids administered can be calibrated based upon the amount (titer) of antibodies produced in a particular subject. In certain embodiments, the ratio of the empty AAV

capsids to the rAAV vector is within or between about 100:1-50:1, from about 50:1 to 25 to 1, from about 25:1 to 10:1, from about 10:1 to 1:1, from about 1:1 to 1:10, from about 1:10 to 1:25, from about 1:25 to 1:50, or from about 1:50 to 1:100. In particular aspects, the ratio of the administered empty AAV capsids to rAAV vector is about 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, or 10:1. Preferably, the serotype of the empty capsids is the same as the rAAV serotype.

[0256] Strategies to reduce humoral immunity to rAAV (which can be applied to other viral vectors) include methods to remove, deplete, capture, and/or inactivate AAV antibodies, commonly referred to as apheresis and more particularly, plasmapheresis where blood products are involved. Apheresis or plasmapheresis, is a process in which a human subject's plasma is circulated *ex vivo* (extracorporeal) through a device that modifies the plasma through addition, removal and/or replacement of components before its return to the patient. Plasmapheresis can be used to remove human immunoglobulins (*e.g.*, IgG, IgE, IgA, IgD) from a blood product (*e.g.*, plasma). This procedure can be employed to deplete, capture, inactivate, reduce or remove immunoglobulins (antibodies) that bind AAV thereby reducing the titer of AAV antibodies in the treated subject that can contribute to rAAV neutralization. An example is using a device composed of an AAV capsid affinity matrix column, and passing blood product (*e.g.*, plasma) through an AAV capsid affinity matrix resulting in binding of AAV antibodies of different isotypes. (See, *e.g.*, Bertin *et al.*, (2020) Sci. Rep. 10, 864, hereby incorporated by reference herein in its entirety.)

[0257] In certain embodiments the polypeptide constructs, encoding polynucleotide constructs, viral vectors and non-viral vectors can be used in combination with an agent that blocks, inhibits, or reduces the interaction of IgG with the neonatal Fc receptor (FcRn), such as an anti-FcRn antibody, to reduce IgG recycling and enhance IgG clearance *in vivo*; and/or an agent that decreases the circulating antibodies that bind to a PPT1 polypeptide, encoding nucleic acid, or delivery vehicle. In certain embodiments, antibody binding is reduced or inhibited by an agent that reduces interaction of IgG with FcRn, a protease or a glycosidase.

[0258] In certain embodiments, the polypeptide constructs, polynucleotide constructs, viral vectors and non-viral vectors described herein are used in combination with an endopeptidase (*e.g.*, IdeS from *Streptococcus pyogenes*) or a modified variant thereof, or an endoglycosidase (*e.g.*, *S. pyogenes* EndoS) or a modified variant thereof. Such treatment can, for example, be carried out to reduce or clear neutralizing antibodies and enable treatment of patients previously viewed as not eligible for treatment. Such strategies are described in, for example, Leborgne *et al.*, (2020) Nat. Med., 26:1096–1101.

[0259] In certain embodiments, method of treatment in a subject is carried out in combination with a compound reducing native mutant PPT1 expression, which provides for a mutant PPT1

having decreased activity. Mutant PPT1 expression can be inhibited, for example, using inhibitory nucleic acid selectively targeting mutant PPT1 encoding sequences. Reference to “selectively targeting” mutant PPT1 activity indicates that expression of the polynucleotide encoding PPT1 protein providing increased activity is not significantly affected. The inhibitory nucleic acid can be provided on the same polynucleotide and/or vector that encodes for the PPT1 protein or using a separate viral or non-viral vector. Examples of inhibitory nucleic acid include a short hairpin RNA (shRNA), a small interfering RNA (siRNA), a microRNA (miRNA), a ribozyme, and an antisense RNA.

IX. Kits

[0260] The present invention includes kits with packaging material and one or more components therein. A kit typically includes a label or packaging insert including a description of the components or instructions for use *in vitro*, *in vivo*, or *ex vivo*, of the components therein. A kit can contain a collection of such components, *e.g.*, PPT1 polypeptide, a viral or a non-viral vector, and optionally a second active, such as another compound, agent, drug or composition.

[0261] A kit refers to a physical structure housing one or more components. Packaging material can maintain the components sterile, and can be made of material commonly used for such purposes such as paper, corrugated fiber, glass, plastic, foil, ampules, vials, and tubes.

[0262] Labels or inserts can include identifying information of one or more components therein, dose amounts, clinical pharmacology of the active ingredient(s) including mechanism of action, pharmacokinetics and pharmacodynamics. Labels or inserts can include information identifying manufacturer information, lot numbers, manufacture location and date, and expiration dates.

Labels or inserts can include information on a disease for which a kit component can be used. Labels or inserts can include instructions for the clinician or subject for using one or more of the kit components in a method, use, or treatment protocol or therapeutic regimen. Instructions can include dosage amounts, frequency or duration, and instructions for practicing any of the methods, uses, treatment protocols or prophylactic or therapeutic regimes described herein.

[0263] Labels or inserts can include information on one or more benefits a component can provide, such as a prophylactic or therapeutic benefit. Labels or inserts can include information on potential adverse side effects, complications or reactions, such as warnings to the subject or clinician regarding situations where it would not be appropriate to use a particular composition. Adverse side effects or complications could also occur when the subject has, will be or is currently taking one or more other medications that can be incompatible with the composition, or the subject has, will be or is currently undergoing another treatment protocol or therapeutic regimen which would be incompatible with the composition and, therefore, instructions could include information regarding such incompatibilities.

[0264] Labels or inserts include “printed matter,” *e.g.*, paper or cardboard, or separate or affixed to a component, a kit or packing material (*e.g.*, a box), or attached to an ampule, tube or vial containing a kit component. Labels or inserts can additionally include a computer readable medium, such as a bar-coded printed label, a disk, optical disk such as CD- or DVD-ROM/RAM, DVD, MP3, magnetic tape, or an electrical storage media such as RAM and ROM or hybrids of these such as magnetic/optical storage media, FLASH media or memory type cards.

X. mRNA Therapeutics

[0265] In certain embodiments, RNA versions of nucleic acid encoding for PPT1 polypeptides described herein is provided as an mRNA construct able to express the encoded polypeptide inside a cell. The mRNA construct comprises a 5'-cap, a 5'UTR, the encoding RNA, 3'UTR, and a poly(A) tail. The UTRs and poly(a) tail can provide for different functions such as participating in mRNA subcellular localization, regulating translation efficiency and mRNA stability. The design and production of mRNA constructs, including different modifications, are illustrated in different publications such as Ouranidis *et al.*, (2022) Biomedicines, 10, 50; Qin *et al.*, Signal Transduct Target Ther. (2022) 21;7(1):166, and U.S. Patent Publication No. U.S. 2013/0259924, each of which are hereby incorporated by reference herein in their entirety.

[0266] In certain embodiments, the mRNA construct is delivered to a cell or subject using nanoparticles. Examples of nanoparticles include those provided in Sections VI (including VI.A. to VI.E. *supra.*, Ouranidis *et al.*, (2022) Biomedicines, 10, 50, and U.S. Patent Publication No. U.S. 2013/0259924.

XI. Additional Aspects and Embodiments

[0267] Additional aspects, embodiments, and examples of combinations thereof include the following:

1) A polynucleotide comprising a nucleic acid sequence encoding a palmitoyl-protein thioesterase-1(PPT1) polypeptide, wherein said PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 95% identity, at least 97% identity, or at least 99% identity to the sequence of SEQ ID NO: 1, wherein:

(a) said PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or

(b) said PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or

(c) said PPT1 sequence comprises the amino acid sequence leucine-glutamine histidine-leucine at its N-terminus; and/or

(d) said nucleic acid sequence comprises a PPT1 encoding sequence having at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity, or 100% identity to any of SEQ ID NOs: 61-94.

2) The polynucleotide of 1, wherein said PPT1 polypeptide further comprises said signal sequence comprising the sequence of any of SEQ ID NOs: 16-27.

3) The polynucleotide of 2, wherein said nucleic acid comprises a signal encoding sequence of any of SEQ ID NOs: 43-58.

4) The polynucleotide of 2, wherein said polypeptide comprises said signal sequence of any of SEQ ID NOs: 16-21 and 24-27.

5) The polynucleotide of 4, wherein said signal sequence comprises the sequence of SEQ ID NOs: 16 or 19.

6) The polynucleotide of 5, wherein said signal sequence comprises the sequence of SEQ ID NO: 16 and said nucleic acid sequence comprises the signal encoding sequence of SEQ ID NO: 43 or said signal peptide comprises the sequence of SEQ ID NO: 19 and said nucleic acid sequence comprises the signal encoding sequence of SEQ ID NO: 50.

7) The polynucleotide of 2, wherein said polypeptide comprises the signal encoding sequence of SEQ ID NO: 23.

8) The polynucleotide of 7, wherein said nucleic acid sequence comprises the signal encoding sequence of SEQ ID NO: 54.

9) The polynucleotide of any one of 1-6, wherein said PPT1 amino acid sequence comprises a G, V, or L substitution for aspartic acid D at its amino terminus and said PPT1 amino acid sequence has at least 95% identity, at least 97% identity, at least 99% identity to SEQ ID NO: 1; or comprise SEQ ID NO: 1.

10) The polynucleotide of 9, wherein said PPT1 amino acid sequence comprises the sequence of SEQ ID NO: 2, wherein X is G, X is V, or X is L.

11) The polynucleotide of any one of 1-3, 7 or 8, wherein said PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus and said PPT1 amino acid sequence has at least 95% identity, at least 97% identity, or at least 99% identity to the sequence of SEQ ID NO: 1.

12) The polynucleotide of 11, wherein said PPT1 sequence comprises SEQ ID NO: 4.

13) The polynucleotide of any one of 1-12, wherein said nucleic acid comprising a PPT1 encoding sequence comprises a sequence having at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity to any of SEQ ID NOs: 61-94, or comprising any of SEQ ID NOs: 61-94; in further embodiments the PPT1 encoding sequence comprises a sequence having at least 85% identity, at least 90% identity, at least 95% identity, at least 99%

identity the sequence of any of SEQ ID NOs: 62-64, 71, 74, 78, 79, or 83 or comprising any of SEQ ID NOs: 62-64, 71, 74, 78, 79, or 83; in further embodiments the PPT1 encoding sequence comprises a sequence having at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity to, or comprises, the sequence of any of SEQ ID NOs: 62-64, 71, 74, 78, 79, or 83; in a further embodiment the PPT1 encoding sequence comprises a sequence having at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity to, or comprises the sequence of SEQ ID NO: 64; in a further embodiment the PPT1 encoding sequence comprises a sequence having at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity to, or comprises, the sequence of SEQ ID NO: 79.

14) The polynucleotide of 1, wherein said PPT1 polypeptide comprises a sequence with at least 99% identity to any of SEQ ID NOs: 31-42 or comprises any of SEQ ID NOs: 31-42.

15) The polynucleotide of 14, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 or SEQ ID NO: 34.

16) The polynucleotide of 15, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 and said nucleic acid comprises a sequence having a sequence identity of at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity to any of SEQ ID NOs: 107-125, and 168; or said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 wherein X is G and said nucleic acid comprises a sequence having a sequence identity of at least 85% identity, at least 90% identity, at least 95%, at least 99% identity to any of SEQ ID NOs: 107-125 or comprises any of SEQ ID NOs: 107-125 and 168. In a further embodiment, said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 wherein X is G and said nucleic acid comprises a sequence having a sequence identity of at least 85% identity, at least 90% identity, at least 95%, at least 99% identity to SEQ ID NO: 64 or comprises the sequence of SEQ ID NO: 64; In a further embodiment, said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 wherein X is G and said nucleic acid comprises a sequence having a sequence identity of at least 85% identity, at least 90% identity, at least 95%, at least 99% identity to SEQ ID NO: 79 or comprises the sequence of SEQ ID NO: 79.

17) The polynucleotide of 15, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 34 and said nucleic acid comprises a sequence having a sequence identity of at least 85%, at least 90%, at least 95%, at least 99% identity to any of SEQ ID NOs: 126-140, and 161-167; or said PPT1 polypeptide comprises the sequence of SEQ ID NO: 34 wherein X is G and said nucleic acid comprises a sequence having a sequence identity of at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity to any of SEQ ID NOs: 126-140, and 161-167 or comprises any of SEQ ID NOs: 126-140, and 161-167.

18) The polynucleotide of 1, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 38.

19) A palmitoyl-protein thioesterase-1(PPT1) polypeptide comprising a PPT1 amino acid sequence having at least 95% identity, at least 97% identity, at least 99% identity to the sequence of SEQ ID NO: 1, wherein:

(a) said PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or

(b) said PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or

(c) said PPT1 sequence comprises the amino acid sequence leucine-glutamine- histidine-leucine at its N-terminus.

20) The polypeptide of 19, wherein said PPT1 polypeptide further comprises a signal sequence comprising the sequence of any of SEQ ID NOs: 16-27.

21) The polypeptide of 20, wherein said polypeptide comprises a signal sequence of any of SEQ ID NOs: 16-21 and 24-27.

22) The polypeptide of 21, wherein said signal sequence comprises the sequence of any of SEQ ID NOs: 16 or 19.

23) The polypeptide of 20, wherein said polypeptide comprises a signal sequence of SEQ ID NO: 23.

24) The polypeptide of any one of 19-23, wherein said PPT1 amino acid sequence comprises a G, V, or L substitution for aspartic acid D at its amino terminus and said PPT1 amino acid sequence has at least 97% identity, or at least 99% identity to the sequence of SEQ ID NO: 1.

25) The polypeptide of 24, wherein said PPT1 amino acid sequence comprises the sequence of SEQ ID NO: 2, wherein X is G.

26) The polypeptide of 19, wherein said PPT1 sequence comprises the amino acid sequence leucine-glutamine histidine-leucine at its N-terminus and said PPT1 amino acid sequence has at least 97% or at least 99% identity to the sequence of SEQ ID NO: 1.

27) The polypeptide of 26, wherein said PPT1 sequence comprises SEQ ID NO: 4.

28) The polypeptide of 19, wherein said PPT1 polypeptide comprises a sequence with at least 99% identity or 100% identity to any of SEQ ID NOs: 31-42.

29) The polypeptide of 28, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 wherein X is G, SEQ ID NO: 34 wherein X is G, or SEQ ID NO: 38.

30) A polynucleotide comprising a PPT1 polypeptide encoding nucleic acid sequence, wherein said PPT1 encoding nucleic acid sequence encodes the PPT1 polypeptide of any one of 19-29.

31) A polynucleotide comprising two or more exons together encoding for the PPT1 polypeptide of any one of 19-29, and one or more introns.

32) The polynucleotide of any one of 1-18, 30 or 31, wherein said polynucleotide is an expression cassette comprising one or more expression control elements operably linked to said nucleic acid encoding said PPT1 polypeptide.

33) The polynucleotide of 32, wherein said nucleic acid encoding said PPT1 polypeptide is operably linked to an upstream promoter and a downstream polyadenylation signal.

34) The polynucleotide of 32, wherein said expression cassette comprises 5' to 3', operably linked to said nucleic acid encoding said PPT1 polypeptide, a promoter, a kozak sequence, said PPT1 polypeptide encoding nucleic acid sequence, and a polyadenylation signal.

35) The polynucleotide of 33 or 34, wherein said promoter comprises a sequence having at least 95% identity, at least 97% identity, or at least 99% identity to the sequence of SEQ ID NO: 5 or comprises SEQ ID NO: 5 or said promoter comprises a sequence having at least 95% identity, at least 97% identity, or at least 99% identity to the sequence of SEQ ID NO: 173 or comprises SEQ ID NO: 173.

36) The polynucleotide of any one of 33-35, wherein said polyadenylation signal operably linked to the PPT1 encoding nucleotide sequence comprises a sequence having at least 95% identity, at least 97% identity, or at least 99% identity to the sequence of SEQ ID NO: 6, or comprises SEQ ID NO: 6.

37) The polynucleotide of any one of 32-36, wherein said expression cassette comprises a nucleotide sequence having at least 95% identity, at least 97% identity to the sequence of any of SEQ ID NOs: 141-143, or comprises any of SEQ ID NOs: 141-143; or said expression cassette comprises a nucleotide sequence having at least 95% identity, at least 97% identity to SEQ ID NO: 169 or comprises SEQ ID NO: 169; or said expression cassette comprises a nucleotide sequence having at least 95% identity, at least 97% identity to SEQ ID NO: 170 or comprises SEQ ID NO: 170.

38) The polynucleotide of any one of 1-18, and 30-37, wherein said polynucleotide is DNA.

39) A recombinant viral vector nucleic acid comprising the polynucleotide of any one of 1-18 and 30-38, and 5' and/or 3' viral elements providing for viral packaging and replication.

40) The recombinant viral vector nucleic acid of 39, wherein said recombinant viral

vector nucleic acid is DNA and comprises an adeno-associated virus (AAV) inverted repeat (ITR) flanking the 5' terminus of said polynucleotide and an AAV ITR flanking the 3' terminus of said polynucleotide.

41) The recombinant viral vector nucleic acid of 40, wherein said recombinant viral vector nucleic acid comprises a 5' ITR and a 3' ITR of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.10, AAVrh.74 or AAV3B.

42) The recombinant viral vector nucleic acid of 40, wherein said 5' ITR comprises a sequence having at least 95% identity, at least 97% identity, at least 99% identity to the sequence of SEQ ID NO: 8 or comprises SEQ ID NO: 8; and said 3' ITR comprises a sequence having at least 95% identity, at least 97% identity, at least 99% identity to the sequence of SEQ ID NO: 9 or comprises SEQ ID NO: 9.

43) The recombinant viral vector nucleic acid of any one of 39-42, further comprising a polyadenylation sequence operably linked to said 3' ITR.

44) The recombinant viral vector nucleic acid of any one of 39-43, further comprising one or more stuffer sequences.

45) The recombinant viral vector nucleic acid of any one of 39-44, wherein said recombinant viral vector nucleic acid comprises a sequence having at least 95% identity, at least 97% identity, at least 99% identity, or 100% identity to the sequence of any of SEQ ID NOs: 144-154; or said recombinant viral vector nucleic acid comprises a sequence having at least 95% identity, at least 97% identity, at least 99% identity, or 100% identity to the sequence SEQ ID NO: 171; or said recombinant viral vector nucleic acid comprises a sequence having at least 95% identity, at least 97% identity, at least 99% identity, or 100% identity to the sequence SEQ ID NO: 172.

46) A gene delivery vehicle comprising a viral or a non-viral vector and the polynucleotide of any one of 1-18, 30-38, or the recombinant viral vector nucleic acid of any one of 39-45.

47) The gene delivery vehicle of 46, wherein said gene delivery vehicle is a viral vector.

48) The gene delivery vehicle of 47, wherein said viral vector is a recombinant AAV vector, a recombinant lentivirus vector, or a recombinant adenovirus vector.

49) The gene delivery vehicle of 48, wherein said viral vector is a recombinant AAV vector, and said recombinant AAV vector comprises a capsid comprising a VP1, VP2 or VP3 having at least 90% identity, at least 95% identity, or 100% identity to a VP1, VP2 or VP3 sequence of any of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.74, AAV3B, AAV-2i8, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, AAV1/rh.10, SEQ ID NO: 12 or SEQ ID NO: 15.

50) The gene delivery vehicle of 49, wherein said capsid is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.74, AAV3B, AAV-2i8, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, or AAV1/rh.10 capsid; or said capsid comprises VP1 of SEQ ID NO: 12 or SEQ ID NO: 15; in a further embodiment the capsid comprises VP1 comprising a sequence of SEQ ID NO: 12, VP2 comprising a sequence of SEQ ID NO: 13, and VP3 comprising a sequence of SEQ ID NO: 14.

51) The gene delivery vehicle of 46, wherein said gene delivery vehicle is the non-viral vector.

52) The gene delivery vehicle of 51, wherein said non-viral vector is a nanoparticle selected from the group consisting of a lipid nanoparticle (LNP), a polymeric nanoparticle, a lipid polymer nanoparticle (LPNP), a protein or peptide-based nanoparticle, a DNA dendrimer or DNA-based nanocarrier, a carbon nanotube, a microparticle, a microcapsule, an inorganic nanoparticle, a peptide cage nanoparticle, and an exosome.

53) The gene delivery of vehicle of 52, wherein said non-viral vector is an LNP or LPNP.

54) A pharmaceutical composition comprising the polynucleotide of any one of 1-18 or 30-38, the polypeptide of any one of 19-29, the recombinant viral vector nucleic acid of any one of 39-45, or the gene delivery vehicle of any one of 46-53, and a pharmaceutically acceptable carrier.

55) A method of increasing PPT1 in a subject comprising administering to the subject the polynucleotide of any one of 1-18 or 30-38, the polypeptide of any one of 19-29, the recombinant viral vector nucleic acid of any one of 39-45, or the gene delivery vehicle of any one of 46-53, or the pharmaceutical composition of 54.

56) A method of treating Neuronal Ceroid Lipofuscinosis 1 in a subject, comprising administering to the subject the polynucleotide of any one of 1-18 or 30-38, the polypeptide of any one of 19-29, the recombinant viral vector nucleic acid of any one of 39-45, or the gene delivery vehicle of any one of 46-53, or the pharmaceutical composition of 54.

57) The method of 55 or 56, wherein said administering comprises intraparenchymal, intracisternal or intraventricular administration. In further embodiments administration is intraventricular; and administration is intraventricular and results in significant rAAV delivery to at least the frontal cortex, parietal cortex, temporal cortex, occipital cortex, thalamus, cerebellar cortex, hippocampus, corpus callosum, spinal cord, caudate, choroid plexus, optic chiasm, fornix, periaqueductal gray, olfactory bulb, and optic nerve.

58) The method of 55 or 56, wherein said administering comprises initial administration outside the central nervous system (CNS).

- 59) The method of any one of 55-58, wherein said administration is systemic.
- 60) The method of any one of 55-59, wherein said subject is a human.
- 61) An AAV vector genome plasmid comprising the recombinant viral vector nucleic acid of any one of 39-45.
- 62) The AAV genome plasmid of 61, wherein said plasmid lacks *rep* and *cap* genes.
- 63) A method of producing a rAAV vector comprising the step of culturing an rAAV production cell line comprising rAAV helper virus activity, wherein the genome of said production cell comprises the recombinant viral vector nucleic acid of any one of 39-45, a *rep* gene and a *cap* gene, wherein said rAAV vector is produced.
- 64) A method of producing a rAAV vector comprising the step of culturing an rAAV permissive cell comprising the AAV genome plasmid of 61 or 62, wherein said rAAV permissive cell further comprises (a) *rep* and *cap* genes provided either as part of the cell genome and/or by one or more separate plasmids, and (b) helper virus activity provided by the cell genome and/or provided by one or more separate plasmids.
- 65) The method of 64, wherein said rAAV permissive cell is a packaging cell, wherein the genome of said packaging comprises a *cap* gene and a *rep* gene.
- 66) The method of 64, wherein either (a) said *rep* gene, said *cap* gene and said helper activity is provided in a single plasmid or (b) said *rep* gene and said *cap* gene is provided by a *rep/cap* plasmid and said helper activity is provided by a helper plasmid.
- 67) A method of obtaining an rAAV vector comprising the steps of (a) producing the rAAV using the method of anyone of 63-67 and (b) purifying the rAAV.

XII. Sequences

[0268] Table 2 provides different nucleic acid and amino acid sequences. In some cases variable sequences are noted in the description. Underlying notes signal sequence when present in a full-length PPT1 sequence. Some nucleic acid sequences indicated in bold provide a codon. Reference to “derived” with respect to a signal amino acid sequence indicates a modification was made to the native sequence.

[0269] In different embodiments a polynucleotide comprises a nucleic acid sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% to any of the nucleic acid sequences provided in Table 2; a polynucleotide comprises a nucleic acid sequence having a sequence identity of at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% to any of the nucleic acid sequences provided in Table 2, wherein the stop codon shown in bold is not present and/or is replaced with a different stop codon; or a polypeptide comprises an amino acid sequence having a sequence identity of at least 95%, at

least 96%, at least 97%, at least 98%, at least 99%, or 100% to any of the amino acid sequences provided in Table 2.

[0270] Table 2

SEQ ID NO	
Mature PPT1	
1	Native Mature DPPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFFLNVN SQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINLISVGGQHQQGVFGLPR CPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVYRNHSIFLADINQERGINE SYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPLQETSLYTQDRLGLKEMD NAGQLVFLATEGDHLQLSEEWFYAHIIIFFLG
2	Non-Native Mature XPPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFFLNVN SQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINLISVGGQHQQGVFGLPR CPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVYRNHSIFLADINQERGINE SYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPLQETSLYTQDRLGLKEMD NAGQLVFLATEGDHLQLSEEWFYAHIIIFFLG Where X is G, V, or L
3	Non-Native Mature LPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFFLNVNSQVTT VCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINLISVGGQHQQGVFGLPRCPGES SHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVYRNHSIFLADINQERGINESYKKN LMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPLQETSLYTQDRLGLKEMDNAGQL VFLATEGDHLQLSEEWFYAHIIIFFLG
4	Non-Native Mature LQHLDPPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFF LNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINLISVGGQHQQGVF GLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVYRNHSIFLADINQER GINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPLQETSLYTQDRLGL KEMDNAGQLVFLATEGDHLQLSEEWFYAHIIIFFLG
Components	
5	EF1A promoter GGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGG AGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATG TCGTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTC GCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAG
6	BGHpA CTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCTTGCTTCCTTGACC CTGGAAGGTGCCACTCCCACTGTCCTTTTCTAATAAAATGAGGAAATGCATCACATTGTCT GAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGG AAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGG
7	Synthetic pA CACACAAAAAACCACACACAGATGTAATGAAAATAAAGATCCTTTATT
8	variant 5' ITR v1 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCGCGGCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCTCT
9	3' ITR AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCC GGGCGACCAAAGGTGCGCCGACGCGGGGCTTTGCCGGGCGGCTCAGTGAGCGAGCGAGC GCGCAGCTGCCTGCAGG

10	5' ITR CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGGCCCTCAGTGAGCGAG CGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTTCCT
11	Intron GTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTATGGCCCTTGCGTGCC TTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTG GGTGGGAGAGTTCGAGGCCCTTGCGCTTAAGGAGCCCCCTTCGCCCTCGTGCTTGAGTTGAGGCC TGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGC TTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTCTGGC AAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCGCG GGCGGCGACGGGGCCCGTGCGTCCCGAGCGCACATGTTTCGGCGAGGCGGGGCTGCGAGCGCG GCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCG CGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTGCGCACCAAGTTGCGTGA GCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTC GGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTTCGTCTCAGCCGTCG CTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTGGAGCTTT TGGAGTACGTCTGCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGA GTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCC TTTTTGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTT CCATTTTCAG
Capsid Protein	
12	VP1 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDNGRGLVLPGYKYLGPFGNGLDKG EPVNAADAAALEHDKAYDQQLQAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFAQKR VLEPLGLVESPVKTAPGKKRPVEPSPQRSPTSSTGIGKKGQPPAKKRLNFGQTGDSESVDPDP QPIGEPPAAPSGVGPNTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRIITSTRTW ALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDFNRFHCHFSRQWRLINNNWGF PKRLNFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADV FMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFEFSYNFEDVPHSSYAHQSOLD RLMNPLIDQYLYLSRTQSTGGTAGTQQLLFSQAGPNNMSAQAKNWLPGPCYRQQRVSTTSL QNNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDEERFFPSSGVLPMFGKQGAGKDNVDYS SVMLTSEEEIKTTNPVATEQYGVVADNLQQQNAAPIVGAVNSQGALPGMVWQNRDVYLQGP WAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPADPPTTFNQAKLASFITQYSTGQVS VEIEWELQKENS KRWNPEIQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL
13	VP2 MAPGKKRPVEPSPQRSPTSSTGIGKKGQPPAKKRLNFGQTGDSESVDPDPPIGEPPAAPSGV GPNTMAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRIITSTRTWALPTYNNHLYKQI SNGTSGGSTNDNTYFGYSTPWGYFDFNRFHCHFSRQWRLINNNWGFPRKRLNFKLFNIQV KEVTQNEGTKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNN GSQAVGRSSFYCLEYFPSQMLRTGNNFEFSYNFEDVPHSSYAHQSOLDRLMNPLIDQYLY LSRTQSTGGTAGTQQLLFSQAGPNNMSAQAKNWLPGPCYRQQRVSTTSLQNNNSNFAWTGAT KYHLNGRDSLVPNGVAMATHKDDEERFFPSSGVLPMFGKQGAGKDNVDYSSVMLTSEEEIKTT NPVATEQYGVVADNLQQQNAAPIVGAVNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHP SPLMGGFGLKHPPPQILIKNTPVPADPPTTFNQAKLASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL
14	VP3 MAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRIITSTRTWALPTYNNHLYKQISNGT SGGSTNDNTYFGYSTPWGYFDFNRFHCHFSRQWRLINNNWGFPRKRLNFKLFNIQVKEVT QNEGTKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQA VGRSSFYCLEYFPSQMLRTGNNFEFSYNFEDVPHSSYAHQSOLDRLMNPLIDQYLYLSRT QSTGGTAGTQQLLFSQAGPNNMSAQAKNWLPGPCYRQQRVSTTSLQNNNSNFAWTGATKYHL NGRDSLVPNGVAMATHKDDEERFFPSSGVLPMFGKQGAGKDNVDYSSVMLTSEEEIKTTNPVA TEQYGVVADNLQQQNAAPIVGAVNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPLM GGFGLKHPPPQILIKNTPVPADPPTTFNQAKLASFITQYSTGQVSVEIEWELQKENS KRWN PEIQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL
15	VP1 MAADGYLPDWLEDNLSEGIREWALQPGAPKPKANQQHQDNARGLVLPGYKYLGPFGNGLDKG

	EPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKR LLEPLGLVEEAAKTAPGKKRPVDQSPQEPDSSSGVGKSGKQPARKRLNFGQTGDSESVDPDQ PLGEPPAAPTSLSGNTMASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDREVITTTSTRTWA LPTYNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCHFSRPDWQRLINNNWGFPRPKK LSFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMV PQYGYLTNNQSQA VGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHSSQSLDRIM NPLIDQYLYLNRTOGTTSGTTNQSRLLFSQAGPQMSLQARNWLPGPCYRQQRLSKTANDN NNSNFPWTAASKYHLNGRDSLNVNPGPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNV MITDEEEIRTTNPVATEQYGTVANNLQSSNTAPTTRTVNDQGALPGMVWQDRDVYLQGP IWA KIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVE IEWELQKENS KRWNPEIQYTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRPL
Signal Peptides	
16	Sp7 (human chymotrypsinogen B2 signal peptide) MAFLWLLSCWALLGTTFG
17	NGF (nerve growth factor signal peptide) MSMLFYTLITAFLLIGIQA
18	Insulin signal peptide MALWMRLRLPLLALLALWGPDPAAA
19	SPARC (secreted protein acidic and cysteine rich signal peptide) MRAWIFFLLCLAGRALA
20	CD33 signal peptide MPLLLLLLPLLWAGARA
21	IL34 (Interleukin 34 signal peptide) MPRGFTWLRYYYYLLGIFLGVARG
22	OSM (Oncostatin M signal peptide) MGVLLTQRTLLSLVLALLFPDARA
23	tPA (Tissue plasminogen activator derived signal peptide) MDAMKRGLCCVLLLCGAVFVSPSARA
24	IL2 (Interleukin 2 derived signal peptide) MYRMQLLSICIALSLALVLLSARA
25	Trypsin derived signal peptide MNLILLITFVAAAVVIRA
26	BDNF-1 (brain derived neurotrophic factor signal peptide) MTILFLTMVISYFGCARA
27	BDNF-2 (brain derived neurotrophic factor derived signal peptide) MTILFLTMVISYFGCAKA
28	Native PPT1 signal peptide MASPGCLWLLAVALLPWTCASRALQHL
Full-Length PPT1 (signal sequence is underlined)	
29	Native full length PPT1 protein MASPGCLWLLAVALLPWTCASRALQHLDPPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKI PGIYVLSLEIGKTLMEDVENSFFLNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAV AQRCPSPPMINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYW HDPIKEDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYR SGQAKETIPLQETSLYTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWFYAHIIPFLG
30	MASPGCLWLLAVALLPWTCASRALQHLXPPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKI PGIYVLSLEIGKTLMEDVENSFFLNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAV AQRCPSPPMINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYW HDPIKEDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYR SGQAKETIPLQETSLYTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWFYAHIIPFLG wherein X is G, V or L
31	MAFLWLLSCWALLGTTFGXPPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLE IGKTLMEDVENSFFLNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPM

	INLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IKEDVY RNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIP LQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG wherein X is D, G, V or L
32	MSMLFYTLITAFILIGIQA X PPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLE IGKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPM INLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IKEDVY RNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIP LQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG wherein X is D, G, V or L
33	MALWMRLPLLLALALWGPDPAA X PPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGI YVLSLEIGKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQR CPSPPMINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IKEDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQ AKETIPLQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG wherein X is D, G, V or L
34	MRAWIFFLLCLAGRALA X PPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEI GKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMI NLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IKEDVYR NHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPL QETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG wherein X is D, G, V or L
35	MPLLLLLPLWAGARA X PPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEIG KTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMIN LISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IKEDVYRN HSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPLQ ETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG where X is D, G, V or L
36	MPRGFTWLRyLLLGIFLGVARG X PPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYV LSLEIGKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPS PPMINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IK EDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAK ETIPLQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG where X is D, G, V or L
37	MGVLLTQRTLLSLVLALLFPDARALPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIYVLSL EIGKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPP MINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP IKEDV YRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETI PLQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG
38	MDAMKRGLCVLLLCCGAVFVSPSARALQHLDPAPLPLVIWHGMGDSCCNPLSMGAIKKMVE KKIPGIYVLSLEIGKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFL RAVAQRCPSPPMINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQA EYWHDP IKEDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFG FYRSGQAKETIPLQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG
39	MYRMQLLSIALSLALVLLSARA X PPAPLPLVIWHGMGDSCCNPLSMGAIKKMVEKKIPGIY VLSLEIGKTLMEDVENSFFLNVSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRC PSPPMINLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDP I KEDVYRNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQA KETIPLQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWfYAHIIIPFLG where X is D, G, V or L

40	MNLLLLILTFVAAAIVIRAVXLPLVIWHMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLEIGK TLMEDVENSFFLNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINL ISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVYRNH SIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIPLQE TSLYTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWFYAHIIIPFLG where X is D, G, V or L
41	MTILFLTMVISYFGCARAXPPAPLPLVIWHMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLE IGKTLMEDVENSFFLNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPM INLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVY RNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIP LQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWFYAHIIIPFLG where X is D, G, V or L
42	MTILFLTMVISYFGCAKAXPPAPLPLVIWHMGDSCCNPLSMGAIKKMVEKKIPGIYVLSLE IGKTLMEDVENSFFLNVNSQVTTVCQALAKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPM INLISVGGQHQQGVFGLPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQAEYWHDPKEDVY RNHSIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYRSGQAKETIP LQETSlyTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEWFYAHIIIPFLG where X is D, G, V or L
Signal Sequence Encoding Nucleic Acid	
43	Sp7-F ATGGCCTTTCTGTGGCTGCTGTCCTGCTGGGCCCTGCTGGGGACCACCTTTGGC
44	Sp7-P ATGGCCTTCCTGTGGCTGCTGTCCTGCTGGGCACTGCTGGGCACCACATTTGGC
45	Sp7-3 ATGGCCTTCCTGTGGCTGCTGAGCTGCTGGGCCCTGCTGGGGACCACCTTTGGG
46	Sp7-4 ATGGCCTTCCTGTGGCTCCTGAGCTGCTGGGCCCTCCTGGGGACCACCTTTGGG
47	Sp7-5 ATGGCCTTCCTGTGGCTGCTGAGCTGCTGGGCCCTGCTGGGGACCACATTTGGC
48	NGF-1 ATGTCCATGTTGTTCTACACTCTGATCACAGCTTTTCTGATCGGCATACAGGCG
49	Insulin-1 ATGGCCCTGTGGATGCGCCTCCTGCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC AGCCGCAGCC
50	SPARC-1 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGCA
51	CD33-1 ATGCCGCTGCTGCTACTGCTGCCCCTGCTGTGGGCAGGGGCCAGGGCT
52	IL-34-1 ATGCCCCGGGGCTTCACCTGGCTGCGCTATCTTCTGCTGGGGATCTTCCTTGGCGTGGCCAG GGGG
53	OSM-1 ATGGGGGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAGA CGCTAGAGCC
54	tPA-1 ATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTC GCCCAGCGCCAGAGCC
55	IL2-1 ATGTACAGGATGCAACTCCTGTCTTGCAATTGCACTAAGTCTTGCACTTGTCTCTGCTGTCTGC CAGGGCC
56	Trypsin-1 ATGAATCTACTTCTGATCCTTACCTTTGTTGCAGCTGCTGTTGTGATTAGGGCA
57	BDNF-1 ATGACAATCCTCTTCCTGACCATGGTGATTAGCTACTTTGGCTGCGCTAGGGCC

58	BDNF-2 ATGACAATCCTCTTCCTGACCATGGTGATTAGCTACTTTGGCTGCGCTAAGGCC
59	Native PPT1 ATGGCGTCGCCCCGGCTGCCTGTGGCTCTTGGCTGTGGCTCTCCTGCCATGGACCTGCGCTTC TCGGGCGCTGCAGCATCTG
Mature PPT1 Encoding Sequences	
60	Native Mature PPT1 GACCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCC CTTAAGCATGGGTGCTATTAATAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTAT CTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAAT TCCCAAGTAACAACAGTGTGTGTCAGGCACCTTGCTAAGGATCCTAAATGCAGCAAGGCTACAA TGCTATGGGATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCAC CTCCCATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGA TGCCCAAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTA CTCCAAAGTTGTTTCAAGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGG ATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAG TCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAATTCCTCAATGATTC CATTGTGGACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAAGCCAAGGAAA CCATTCCCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAAATGGAC AATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATG GTTTTATGCCACATCATACCATTCTTGGAT TAA In variants thereof nucleotides 1-3 encodes G, D, V or L
61	7-PPT1co1 GGCCCCCAGCACCTCTTCCCCTGGTGATCTGGCATGGCATGGGGGATAGCTGCTGCAATCC TCTCTCAATGGGAGCCATTAAGAAGATGGTGGAGAAGAAAATCCCAGGGATATATGTTCTGT CCTTGGAGATTGGAAAAACCTTGATGGAAGATGTGGAAAATAGCTTTTTCTTAATGTGAAC TCTCAAGTCACCACAGTGTGCCAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAAGGCTACAA TGCAATGGGCTTCTCCCAAGGAGGTCAATTCCCTGAGGGCTGTTGCACAGAGGTGTCCCAGTC CCCCATGATCAACCTGATCTCAGTGGGTGGCCAGCACCAAGGTGTCTTTGGGCTCCCCAGG TGTCAGGGGAGAGCAGCCACATCTGTGACTTCATTAGAAAGACACTAAATGCAGGAGCCTA CAGCAAGGTGGTGCAGGAGAGGCTTGTTTCAGGCTGAGTACTGGCATGACCCCATTAAGGAAG ATGTCTATAGAAACCACAGCATCTTCTTGGCAGATATAAACCAGAAAGAGGCATTAATGAA AGTTACAAGAAGAATCTTATGGCCCTCAAGAAGTTTGTGATGGTGAAGTTCTCTGAATGACTC CATTGTTGACCTGTGGACTCTGAATGGTTTGGTTTCTACAGATCAGGACAAGCCAAGGAAA CAATCCCCCTGCAAGAGACCTCCCTGTACACCCAGGACAGGCTGGGCCTGAAAGAAAATGGAT AATGCTGGCCAACCTGGTCTTCTGGCAACTGAAGGGGACCACCTGCAACTCTCTGAGGAATG GTTCTATGCTCACATCATCCCATTTTTTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
62	PPT1co2 GGCCCACCTGCACCACTTCCCTTGGTCATCTGGCATGGCATGGGTGACTCTTGCTGCAATCC CTTATCCATGGGGGCTATTAAGAAGATGGTGGAAAAGAAGATTCCCTGGAATTTATGTGCTGA GTCTGGAAATTGGAAAGACCCTCATGGAAGATGTTGAAAACAGCTTCTTCTTGAATGTCAAT TCCAGGTACACCACAGTGTGTGTCAGGCTCTGGCAAAGGACCCCAAGCTCCAGCAAGGATATAA TGCCATGGGATTTTCCCAAGGTGGCCAGTTCCCTGAGGGCAGTGGCTCAGAGATGCCCTTCAC CCCTATGATCAACCTCATCTCTGTGGGTGGGCAGCATCAGGGTGTCTTTGGGCTCCCCAAGA TGCCCAAGGAGAAAAGCTCTCACATCTGTGACTTTATTAGAAAGACCCTGAATGCTGGGGCTTA CAGCAAGGTGTCCAGGAGAGGCTGGTGCAGGCTGAATACTGGCATGACCCCATTAAGGAAG ATGTGTACAGAAACCACAGTATCTTCTTGGCTGATATCAACCAAGAAAGAGGCATCAATGAG AGCTACAAGAAGAATCTCATGGCCCTGAAGAAGTTTGTGATGGTGAATTCCTCAATGATAG CATTGTGGACCCTGTAGACTCAGAGTGGTTTGGATTCTACAGGTCTGGCCAGGCTAAGGAGA CCATCCCCTGCAAGAGACAAGCCTTTACACACAGGACAGGCTTGGTCTGAAAGAGATGGAT AATGCTGGACAGCTGGTGTCTTCTGGCCACTGAAGGGGACCACCTGCAGCTGAGTGAAGAATG GTTCTATGCCACATCATCCCCTTCTTAGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L

63	PPT1co3 GGGCCCCCAGCACCCCTGCCTCTGGTCATTTGGCATGGCATGGGGGACTCCTGCTGCAATCC CTTATCCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAATTTATGTTCTGT CTTTAGAAATAGGAAAAACACTGATGGAAGATGTGGAAAATTCCTTCTCCTCAATGTCAAC TCCCAGGTAACAACAGTGTGTGAGGCCCTGGCCAAAGACCCCAAATTACAGCAGGGCTACAA TGCTATGGGATTCTCACAGGGAGGCCAATTTCTCAGAGCTGTTGCTCAGAGATGCCCTTCAC CCCCATGATCAACCTGATTTTCAGTGGGTGGACAACATCAAGGAGTTTTTGGATTACCAAGG TGCCCTGGGGAGAGCTCTCACATTTGTGATTTTATAAGAAAAACCTTGAATGCTGGGGCCTA CTCAAAGGTGGTCCAGGAGAGGTTAGTGCAGGCTGAGTATTGGCATGACCCCATTAAGGAAG ATGTGTACAGAAAACCACTCCATTTTTTTGGCTGACATCAACCAGGAGAGAGGAATTAATGAA TCCTACAAGAAGAACCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCTCAATGACTC CATTGTGGATCCAGTGGATTACAGAGTGGTTTGGATTCTACAGAAGTGGCCAGCCAAGGAAA CAATACCACTTCAAGAGACCAGCCTGTACACCCAGGACAGACTGGGCCTAAAGGAAATGGAC AATGCAGGTCAGCTTGTCTTGGCCACTGAGGGGGACCACCTTCAACTCTCTGAAGAATG GTTTTATGCCACATCATACCCCTTTCTGGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
64	PPT1co4 GGCCCTCCAGCACCTCTTCCTTTAGTGATCTGGCATGGCATGGGTGACAGCTGCTGCAACCC ACTGTCCATGGGAGCCATCAAGAAAATGGTGGAGAAGAAAATTCAGGAATCTATGTCCTAT CCCTGGAGATTGGGAAGACTCTGATGGAAGATGTGGAGAACTCCTTCTCCTGAATGTCAAT TCCCAAGTCACCACAGTCTGCCAGGCCCTGGCCAAAGACCCCAAGCTGCAGCAAGGGTACAA TGCTATGGGCTTCTCCCAAGGGGGGCAGTTCTTGAGAGCAGTTGCCCAGAGATGCCCCAGCC CTCCAATGATTAATCTGATCTCTGTGGGAGGTGAGCATCAAGGAGTCTTTGGGTTGCCCAGA TGTCAGGAGAGAGCAGTCACATCTGTGACTTCATCAGGAAAACACTGAATGCTGGGGCATA CAGCAAGGTGGTGCAGGAAAGGCTTGTGCAGGCAGAGTACTGGCATGACCCCATTAAGAGG ATGTGTACAGAAAACCAAGCATCTTTCTTGACAGACATAAATCAGGAGAGAGGTATCAATGAG TCCTACAAGAAGAAATCTGATGGCCCTAAAGAAATTTGTGATGGTAAATTCCTCAATGACTC CATTGTGGACCCAGTTGACTCTGAGTGGTTTGGATTTTACAGAAGTGGCCAGCAAGGAGAGA CCATTCCCCTTCAAGAGACATCCCTGTACACCAAGATAGGCTGGGCCTGAAAGAGATGGAT AATGCAGGGCAATTGGTGTTCCTTGCTACAGAAGGAGACCACCTCCAGCTATCTGAAGAGTG GTTCTATGCCACATCATCCCTTTTCTAGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
65	7-PPT1co5 GGCCCCCTGCCCCACTTCCCTTGGTGATATGGCATGGCATGGGAGATTCTTGTTGTAACCC ACTGTCAATGGGGGCCATTAAGAAGATGGTGGAAAAGAAGATCCCTGGAATCTATGTCCTCT CCCTGGAAATTTGGGAAGACCCCTCATGGAGGATGTGGAAAACCTATTCCTCCTCAATGTGAAC AGCCAGGTACCACTGTTTGTGAGGCATTAGCAAAGGACCCCAAGCTTCAGCAGGGATACAA TGCTATGGGCTTTAGCCAAGGTGGGCAATTTCTTGAGAGCAGTGGCACAGAGATGCCCCTCAC TCCCATGATCAACCTCATCAGTGTAGGAGGCCAGCACCAGGGGGTTTTTGGGCTCCCCAGA TGCCCTGGTGAATCATCTCACATCTGTGACTTCATCAGAAAAACCTTGAATGCTGGAGCCTA CAGTAAAGTAGTGCAAGAGAGGCTGGTGCAAGCAGAATACTGGCATGATCCCATCAAGGAAG ATGTCTACAGGAACCATAGCATCTTCCTGGCTGACATCAACCAGGAGAGGGGAATCAATGAG TCCTACAAGAAAAACCTGATGGCACTGAAGAAGTTTGTGATGGTCAAATTCCTCAATGACTC CATTGTGGACCCTGTGATTCTGAATGGTTTGGATTTTACAGGTCTGGCCAAGCAAAGGAAA CCATCCCTCTTCAGGAAACAGCCTGTACACACAAGACAGGCTGGGCCTAAAGGAAATGGAC AATGCTGGACAATTGGTGTCTTTTGGCTACTGAGGGTGACCACCTCCAGCTCTCAGAAGAATG GTTTTATGCTCACATCATCCCTTTCTGGGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
66	7-PPT1co6 GGCCACCTGCTCCATTGCCCCCTAGTCATTTGGCATGGGATGGGGGACAGCTGCTGCAATCC TTTATCTATGGGAGCCATTAAGAAGATGGTGGAGAAAAAGATCCCTGGCATATATGTGTTAA GCCTGGAGATTGGAAAGACCCCTCATGGAAGATGTTGAGAATTCCTTCTCCTGAATGTAAAC AGCCAAGTCACCACAGTGTGCAAGCTCTGGCCAAGGACCCCAAGCTCCAGCAGGGTTATAA TGCTATGGGATTCTCTCAGGGTGGGCAGTTTCTGAGGGCTGTGGCACAGAGATGCCCAAGCC CACCATGATCAACCTCATCAGTGTGGTGGACAGCATCAGGGAGTGTGGCTGGCCAGCA

	TGTCCAGGTGAGTCTCTCATATCTGTGACTTTATCAGGAAAACCTTAAATGCAGGGGCCTA TTCCAAGGTGGTGCAGGAGAGACTTGTCCAGGCTGAGTATTGGCATGATCCCATTAAAGAAG ATGTCTATAGAAAACCATTCATCTTCCTGGCTGATATCAACCAGGAAAAGGGGATCAATGAG AGCTACAAGAAGAATCTAATGGCACTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATAG TATTGTGGATCCTGTAGACAGTGAAGTGGTTTGGCTTCTACAGAACTGGACAAGCCAAAGAGA CCATTCCCTTACAAGAGACCTCACTCTACACACAGGACAGACTGGGCCTGAAGGAAATGGAC AATGCAGGTCAGCTGGTGTTCCTTGCAACAGAGGGGGATCACCTGCAGCTGTCAGAGGAATG GTTTTATGCCCACATCATCCCTTTTCCTTGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
67	7-PPT1co7 GGTCCCCCTGCTCCATTGCCTCTGGTCATCTGGCATGGCATGGGTGACTCCTGTTGCAACCC CTTAAGCATGGGGCCATCAAGAAGATGGTGGAGAAGAAGATTCTTGAATCTATGTCTTGT CTTTAGAGATTGGTAAAACCTCATGGAAGATGTGGAGAATTCTTCTTCTTGAATGTGAAC TCACAAGTCACAACAGTCTGCCAGGCCCTGGCAAAGGACCCCAAGCTCCAGCAGGGCTACAA TGCCATGGGCTTCTCCCAGGCTGGGCAATTTCTGAGGGCTGTGGCCAGAGATGTCCCTCAC CCCCATGATCAACTTGATATCAGTTGGAGGCCAGCACCAGGGAGTATTTGGGCTTCCCAGG TGCCCAGGAGAATCCAGCCATATCTGTGATTTTATTAGAAAAGACCTCAATGCTGGGGCTTA CTCCAAGTGGTGCAGGAGAGGTTAGTGCAAGCTGAATACTGGCATGATCCCATAAAGGAGG ATGTTTATAGGAACCACAGCATCTTCTAGCAGATATTAATCAAGAAAGAGGTATCAATGAG TCCTACAAGAAAAACCTCATGGCCCTGAAGAAGTTTGTATGGTCAAGTTTCTCAATGACTC CATTGTGGATCCTGTAGACTCTGAATGGTTTGGCTTTTACAGGAGTGGCCAAGCAAAAGAAA CCATTCTCTGCAGGAAACAAGTCTGTACACTCAGGACAGACTGGGCCTGAAAGAAATGGAT AATGCTGGCCAGTTAGTGTTCCTGGCCACTGAGGGGGACCACCTGCAACTCAGTGAGGAATG GTTCTATGCTCACATAATTCCCTTTCCTTGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
68	7-PPT1co8 GGACCTCCTGCTCCTCTGCCCCCTGGTCATCTGGCATGGAATGGGAGACAGCTGTTGTAACCC ACTCAGTATGGGAGCCATAAAGAAGATGGTTGAGAAGAAGATTCTTGGCATATATGTACTCA GCTTGGAGATTGGGAAGACCTTGATGGAAGATGTGGAGAACTCATTTCTTCTTGAATGTCAAC AGCCAGGTCAACCACTGTTTGCCAGGCCCTTGCTAAGGACCCCAACTGCAACAGGGCTACAA TGCCATGGGGTTCTCTCAGGGAGGACAGTTCTCAGAGCAGTGGCAGAGAGGTGCCCCAGCC CACCCATGATTAATCTGATCAGTGTGGAGGGCAGCACCAAGGGGTCTTTGGACTTCCAAGG TGCCCTGGGGAATCCAGCCATATTTGTGATTTTATCAGAAAGACTCTGAATGCAGGGGCCTA TAGTAAAGTTGTTCAAGAAAGGCTGGTGCAGGCAGAATACTGGCATGATCCCATCAAGAAG ATGTATACAGAAACCACAGCATTTTCTGGCAGATATAAACCAGAAAGAGGCATTAATGAG TCCTACAAGAAACCTCATGGCCCTGAAGAATTTGTGATGGTGAAATTTTTGAATGATTC TATAGTGGACCCTGTGGATAGTGAGTGGTTTGGTTTCTACAGGAGTGGGCAGGCCAAGGAGA CCATTCCCTTGCAGGAGACCTCCCTGTACACACAGGACAGACTGGGCCTAAAGGAAATGGAC AATGCAGGCCAGCTAGTGTTCCTGGCCACTGAGGGGGACCACCTCCAGCTCTCTGAAGAATG GTTTTATGCCCATATTATCCCTTTTCCTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
69	7-PPT1co9 GGCCCTCCAGCTCCCTTGCCTTTAGTCATTTGGCATGGAATGGGAGACTCCTGCTGCAATCC TCTCTCTATGGGAGCCATAAAGAAGATGGTAGAGAAAAAATTCCTGGAATATATGTCTTAA GCCTGGAATTTGGGAAGACACTCATGGAAGATGTGGAGAACTCCTTCTTCTCAATGTGAAC AGCCAGGTGACCACAGTCTGCCAGGCCTTGGCCAAAGACCCCTAAGCTTCAGCAAGGCTACAA TGCCATGGGCTTCAGCCAGGGTGGTCAATTCCTGAGAGCAGTGGCCAGAGATGCCCCCTCAC CACCTATGATTAACCTCATCTCTGTGGGTGGTCAGCACCAAGGTGTGTTTGGTCTCCCTAGG TGCCCAGGAGAGAGTAGCCATATTTGTGACTTTCATCAGGAAAACATTGAATGCTGGAGCCTA CTCAAAAGTTGTACAAGAAAGGCTAGTGCAGGCTGAGTACTGGCATGATCCCATCAAGGAAG ATGTGTACAGAAACCACAGCATCTTTCTGGCTGATATCAATCAGGAGAGGGGGATCAATGAA AGCTACAAGAAACCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCTCAATGACAG CATTGTTGATCCTGTTGATTCTGAATGGTTTGGATTCTACAGGTCTGGACAGGCCAAAGAGA CCATCCCTCTTCAGGAAACCTCCCTTTACACACAGGACAGACTGGGGCTCAAGGAGATGGAC AATGCTGCTCAGCTGGTTTTTCTGGCTACAGAGGGAGACCACCTACAGTTGAGTGAAGAGTG

	GTTCTATGCCCACATCATCCCTTTCCTGGGTTG
	In variants thereof nucleotides 1-3 encodes G, D, V or L
70	7-PPT1co10 GGCCCCACCAGCTCCCCCTCCCCCTGGTCATTTGGCATGGTATGGGAGACTCCTGTTGTAATCC TCTCTCCATGGGAGCCATTAAGAAAATGGTGGAGAAGAAAATTCCTGGGATCTATGTGCTCT CTCTAGAGATTGGGAAGACCCTGATGGAAGATGTGGAGAACAGCTTTTTCTTAAATGTGAAC AGCCAGGTGACAACCTGTCTGTCAAGCATTGGCCAAGGACCCCAAGCTGCAGCAAGGCTACAA TGCTATGGGATTCTCTCAGGGGGGCCAGTTCTGAGGGCTGTGGCCAGAGATGCCCATCCC CTCCCATGATCAATCTGATATCTGTTGGGGGCCAGCACCAGGGAGTTTTTGGGCTGCCCAGA TGCCCTGGGAGAGCTCTCACATCTGTGATTTTCATCAGGAAAACCTCAATGCTGGAGCATA TAGCAAGGTGGTGCAGGAGAGGTTGGTGCAGGCAGAGTACTGGCATGACCCCATCAAGGAGG ATGTGTACAGAAACCACTCAATTTTCTGGCAGATATCAACCAAGAGAGGGGCATCAATGAG TCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTTCTGAATGATTC TATTGTGGACCCTGTGGATAGTGAGTGGTTTGGCTTTTACAGAAGTGGCCAAGCTAAGGAAA CCATTCCCCTACAAGAACTTCCCTGTACACCCAGGACAGGTTGGGACTCAAAGAGATGGAC AATGCTGGACAGCTGGTGTTCCTGGCCACAGAAGGGGACCACCTGCAGCTGAGTGAAGAGTG GTTCTATGCTCATATCATCCCTTTCCTTGGATG In variants thereof nucleotides 1-3 encodes G, D, V or L
71	PPT1co11 GGCCCCCTGCCCCCTCTTCCCCTGGTCATTTGGCATGGGATGGGTGACTCCTGCTGCAACCC CCTAAGCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATTCCTGGAATTTATGTCCTTA GCCTGGAAATTGGGAAGACCTTAATGGAAGATGTGGAAAACAGCTTTTTCTTAAATGTCAAT TCTCAGGTAACAACAGTGTGCCAAGCCCTGGCCAAGGACCCCTAACTGCAGCAGGGCTACAA TGCTATGGGCTTCAGCCAGGGTGGACAGTTTCTGAGGGCAGTGGCTCAGAGGTGCCCCAGTC CTCCCATGATCAATCTCATCTCTGTGGGAGGCCAGCACCAGGGTGTTTTTGGCCTGCCCAGG TGTCCTGGAGAGTCCTCACATATCTGTGACTTCATCAGGAAGACACTGAATGCAGGGGCCTA TAGCAAGGTGGTGCAGAGAGACTGGTTCAGGCTGAATACTGGCATGACCCCATCAAGGAAG ATGTTTACAGGAACCATAGCATCTTCCTTGCTGATATTAATCAAGAGAGGGGCATAAATGAG AGCTACAAGAAAGAAATTTAATGGCTCTAAAGAAATTTGTGATGGTGAAGTTCTGAATGACTC CATTGTGGATCCTGTGGACTCTGAGTGGTTTGGGTTCTACAGAAGTGGCCAGGCCAAGGAGA CCATCCCCCTGCAGGAGACAAGCCTGTACACACAGGACAGGCTGGGACTGAAGGAGATGGAC AATGCAGGACAGCTGGTGTTCCTGGCTACAGAGGGAGACCACCTGCAGCTCTCAGAGGAATG GTTCTATGCACACATCATTCCTTTTTTAGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
72	7PPT1co12 GGACCTCCTGCCCCCTCTTCCCTTTGGTTATCTGGCATGGGATGGGTGACTCCTGCTGCAATCC TTTAAGCATGGGGGCTATTAAGAAGATGGTGGAGAAAAAAATTCCTGGAATCTATGTGCTTA GCTTAGAAATAGGGAAGACATTAATGGAAGATGTGGAAAACAGCTTTTTCTTGAATGTCAAC TCACAGGTGACTACTGTCTGTGAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAAGGTTACAA TGCTATGGGATTGAGCCAGGGTGGTCAGTTTCTGAGGGCTGTGGCACAGAGATGCCCCAC CTCCCATGATCAACCTCATCAGTGTGGGAGGACAGCACCAGGTGTTTTTGGTCTGCCAGG TGTCCTGGAGAGTCCTCACATATCTGTGACTTCATCAGGAAGACTCTGAATGCTGGGGCCTA TAGCAAAGTAGTCCAGGAGAGGTTGGTGCAGGCTGAATACTGGCATGACCCCATTAAGAAG ATGCTATAGGAACCATAGCATATTCCTTGAGATATCAACCAGGAGAGGGGGATCAATGAG AGCTACAAGAAAAACTTGATGGCTCTCAAGAAGTTTGTGATGGTAAAGTTCTTAAATGACTC CATTGTGGATCCTGTAGACTCAGAGTGGTTTGGGTTCTACAGGTCTGGCCAAGCCAAAGAGA CCATTCCCTCTGCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGAC AATGCTGGCCAGCTGGTGTTCCTGGCTACAGAGGGAGACCACCTGCAGCTCTCAGAAGAATG GTTCTATGCTCACATCATTCCTTTTTTAGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
73	7-co13 GGTCCCCCTGCTCCTCTGCCTTTGGTGATTTGGCATGGCATGGGTGACTCCTGCTGCAACCC TCTGTCTATGGGTGCCATTAAGAAAATGGTGGAAAAGAAGATCCCTGGAATTTATGTCCTGT CTTTGGAGATAGGCAAGACTTTGATGGAAGATGTTGAGAATTCCCTTTTTCTTAAATGTAAAT

	AGCCAGGTAACCACAGTTTGTTCAGGCACTTGCCAAGGACCCCAAGCTGCAGCAAGGCTACAA TGCCATGGGCTTCTCCCAGGGAGGGCAGTTTCTGAGAGCAGTGGCTCAGAGGTGCCCCAGCC CTCCAATGATCAATTTGATCTCAGTGGGGGGCCAGCATCAGGGGGTCTTTGGCCTACCTAGG TGCCCTGGAGAGAGCTCTCATATCTGTGACTTCATTAGGAAGACATTAAATGCTGGGGCATA CAGTAAGGTGGTGCAGGAGAGATTAGTGCAGGCTGAATACTGGCATGATCCCATCAAGGAAG ATGTCTACAGGAACCACAGCATTTTCTGGCAGACATTAACCAAGAGAGGGGGATCAATGAG AGCTACAAGAAAAACCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAGTTTCTCAATGACAG CATTGTGGATCCAGTGGATTCTGAGTGGTTTGGGTTCTACAGAAGTGGCCAGGCCAAAGAAA CCATCCCATTGCAGGAAACCAGTTTGTACACTCAGGACAGGTTGGGCCCTGAAGGAGATGGAC AATGCTGGCCAACTGGTGTCTTGGCCACTGAGGGGGACCACCTACAGCTGAGTGAGGAATG GTTCTATGCCACATCATCCCTTTTCTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
74	PPT1co14 GGACCCCGAGCTCCACTCCCCCTGGTGATCTGGCATGGCATGGGAGATAGCTGTTGCAACCC ATTATCCATGGGGGCCATTAAGAAGATGGTGGAGAAGAAAATCCCTGGCATTATGTCTCTTA GCCTGGAAATTGGAAAGACCTTAATGGAGGATGTGGAAACTCCTTCTCTCCTCAATGTCAAC TCCCAAGTCACCACAGTGTGTCAAGCCCTTGCCAAGGACCCCAAGTTGCAACAGGGCTATAA TGCTATGGGCTTCTCTCAGGGAGGCCAGTTTCTGAGGGCTGTGGCCAGAGGTGCCCTCCC CACCTATGATCAATCTTATCTCTGTGGGTGGACAACACCAAGGTGTCTTTGGGCTGCCAGG TGCCCTGGGGAGTCCAGCCACATTTGTGACTTCATCAGGAAGACACTAAATGCTGGTGCCTA CTCCAAAGTTGTCCAGGAGAGACTGGTGCAAGCAGAATACTGGCATGACCCTATTAAAGAGG ATGTCTACAGGAACCACAGCATCTTTCTGGCAGACATCAACCAGGAGAGAGGGATCAATGAA TCCTACAAGAAGAACCTCATGGCCCTGAAGAAATTTGTGATGGTTAAGTTTCTCAATGACTC CATTGTGGACCCCTGTGGATTCTGAATGGTTTGGATTTTACAGAAGTGGCCAGGCCAAAGAAA CCATCCCCTCTGCAGGAGACATCCCTGTACACACAGGACAGACTGGGCCCTGAAGGAGATGGAC AATGCTGGGCAGCTAGTGTTCTTGGCTACTGAGGGGGATCATCTTCAGTTGTCTGAAGAATG GTTCTATGCCACATCATTCCTTTCTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
75	7-PPT1co15 GGGCCCCCAGCTCCTCTTCCCCTGGTGATCTGGCATGGCATGGGAGACTCCTGTTGCAACCC ATTAAGCATGGGGCTATTAAGAAAATGGTGGAGAAGAAAATCCCTGGAATTTATGTCTCTCT CTCTGGAGATAGGCAAAACCTTAATGGAAGATGTGGAGAACTCCTTCTCTCCTCAATGTCAAC TCCCAAGTCACCAGTGTATGTCAAGCTCTGGCAAAGGACCCCAAATGCAACAGGGATATAA TGCCATGGGATTCTCCCAGGGTGGCCAGTTCCTGAGGGCTGTGGCACAGAGATGCCCTCCC CACCTATGATCAATCTCATCTCTGTGGGTGGACAACATCAAGGTGTTTTTGGGCTGCCAGA TGCCCTGGGGAGTCCCTCTCACATTTGTGACTTCATCAGGAAGACCCCTCAATGCTGGAGCCTA CTCTAAAGTTGTTTCAAGAAAGGCTGGTCCAAGCAGAATACTGGCATGACCCTATAAAAGAGG ATGTTTACAGGAACCACAGTATCTTCTTGGCAGACATCAACCAGGAGAGGGGCATCAATGAG TCCTACAAGAAAAACCTCATGGCACTCAAGAAATTTGTGATGGTCAAATTCCTGAATGACTC CATTGTGGACCCCTGTGGATTCTGAGTGGTTTGGATTCTACAGAAGTGGCCAGGCCAAAGAAA CCATCCCCTCTGCAGGAGACAAGTCTGTACACCCAGGATAGGTTGGGTCTGAAGGAAATGGAT AATGCTGGGCAGCTAGTCTTCTTGGCCACAGAAGGGGACCACCTGCAGCTTCTGAAGAGTG GTTCTATGCTCACATCATTCCTTTCTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
76	7-PPT1co16 GGCCCCCTGCTCCTCTGCCCCTGGTCATCTGGCATGGGATGGGAGACAGCTGCTGTAACCC TTTATCAATGGGAGCCATCAAGAAAATGGTGGAGAAGAAGATCCCTGGGATATATGTTCTTA GCTTGGAAATTGGCAAGACTCTCATGGAAGATGTTGAAAACAGCTTCTTCTGAATGTCAAT AGCCAAGTGACCACAGTGTGCCAGGCATTGGCTAAAGACCCCAAGTTGCAGCAAGGCTACAA TGCCATGGGCTTTTCCCAGGGTGGTCAGTTCCTGAGGGCAGTGGCTCAGAGGTGCCCTTCAC CACCTATGATAAACCTCATCTCTGTGGGAGGGCAACACCAAGGAGTATTTGGACTACCCAGG TGCCCTGGGGAGTCCCTCCCATATTTGTGACTTTATCAGGAAGACACTGAATGCAGGAGCCTA CAGCAAGGTGGTTTCAAGAGAGACTTGTGCAGGCAGAGTACTGGCATGATCCCATTAAGGAAG ATGTGTACAGAAACCATAGTATTTTCTGGCAGACATCAATCAGGAGAGGGGCATCAATGAG AGCTACAAGAAAAATCTGATGGCCCTAAAGAAGTTTGTGATGGTCAAGTTCTGAATGATTC

	TATAGTGGACCCTGTGGATTACAGAGTGGTTTGGTTTTTACAGATCAGGGCAAGCCAAGGAAA CCATCCCCTGTCAGGAAACAAGCCTGTACACTCAAGACAGGTTAGGCCTAAAAGAGATGGAC AATGCTGGCCAGCTTGTGTTCTTGGCAACAGAGGGGGATCATTTACAGTTATCTGAGGAATG GTTCTATGCCCACATCATCCCCTTCCTTGGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
77	7-PPT1co17 GGGCCCCCTGCCCCCTCTCCCTTTGGTTATTTGGCATGGCATGGGGGACTCATGCTGTAACCC TTTGAGCATGGGGGCCATTAAGAAAATGGTGGAGAAGAAGATCCCAGGTATTTATGTCTTGT CTCTAGAAATAGGGAAGACTCTGATGGAAGATGTGGAGAATAGCTTCTCTTAAATGTGAAT TCACAGGTTACCACTGTGTGCCAGGCCCTGGCCAAAGACCCCAAGTTACAGCAGGGGATACAA TGCTATGGGATTTTCTCAGGGTGGTCAGTTCTGAGGGCAGTGGCCAGAGATGCCCCAAGCC CTCCAATGATTAACCTCATTAGTGTGGGAGGGCAGCATCAAGGGGTATTTGGCCTGCCCAGG TGCCCTGGAGAGTCTCTCATATCTGTGACTTTATCAGGAAGACTTTAAATGCTGGGGCCTA CAGCAAGGTGGTCCAGGAGAGGCTGGTGCAGCAGAGTATTGGCATGATCCCATAAAGGAGG ATGTGTATAGGAACCACTCCATCTTCTGGCAGACATCAACCAGGAGAGGGGGATCAATGAG TCTTACAAGAAAAATTTAATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTCTCAATGATTC TATTGTGGACCCTGTGGATTCTGAGTGGTTTGGCTTTTACAGGAGTGGCCAGGCCAAGGAAA CCATCCCCTGTCAGGAACTTCTCTATACACACAGGATAGGCTGGGCCTGAAAGAGATGGAC AATGCAGGGCAGCTGGTGTTCCTTGGCACTGAGGGTGACCACCTTCAGCTATCTGAGGAATG GTTTTATGCTCACATCATCCCCTTCCTGGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
78	PPT1co18 GGTCCACCAGCTCCTCTGCCCTTGGTCATCTGGCATGGGATGGGGGACTCTTGTTGTAACCC TCTTAGCATGGGAGCCATTAAGAAGATGGTGGAGAAGAAAATACCTGGCATATATGTCCTGA GCCTGGAGATTGGCAAGACCCTGATGGAGGATGTGGAGAACAGCTTCTTCTGAATGTCAAT TCCCAGGTGACCACAGTGTGTGAGGCCCTGGCCAAGGACCCCAAGTTGCAGCAAGGCTACAA TGCTATGGGCTTTTCAAGGTGGGCAGTTCTGAGAGCAGTGGCACAGAGATGTCCTTCAC CTCCAATGATCAACCTTATCTCAGTGGGAGGCCAACACCAGGGTGTTTTTGGGTTGCCAGG TGTCCTGGTGAGAGCAGCCATATCTGTGACTTTATCAGGAAGACACTGAATGCAGGGGCCTA CAGCAAGGTGTTCAAGAGAGGCTGGTGCAGGCTGAGTATTGGCATGATCCCATTAAGGAGG ATGTGTATAGGAACCACTCCATCTTTTTGGCAGACATCAATCAGGAGAGGGGCATCAATGAG TCTTACAAGAAGAACCTGATGGCCCTCAAGAAGTTTGTGATGGTCAAGTTCTGAATGATTC TATTGTGGATCCAGTGGACTCTGAATGGTTTGGGTTCTACAGATCTGGCCAGGCCAAGGAGA CCATCCCCTTCAAGAGACCAGCCTGTATACCCAGGACAGATTAGGCCTGAAGGAGATGGAC AATGCTGGCCAGCTTGTCTTCTGGCTACAGAGGGGGACCACCTCCAGCTGAGTGAAGAGTG GTTTTATGCCCACATCATACCCTTCCTGGGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
79	PPT1co19 GGCCCCCAGCCCCCTCTGCCCTTGGTCATCTGGCATGGCATGGGAGACAGCTGCTGCAACCC TCTCTCCATGGGAGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAATCTATGTCCTCA GCCTGGAGATTGGGAAGACCCTGATGGAGGATGTGGAGAACTCCTTCTTCTGAATGTGAAC AGCCAGGTGACCACAGTGTGCCAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAA TGCCATGGGCTTCTCCAGGGAGGCCAGTTCTTGAGGGCTGTGGCCAGAGATGCCCTCAC CTCCCATGATCAACCTCATCTCTGTGGGAGGCCAGCACCAAGGGGTCTTTGGCCTGCCCAGG TGTCCTGGGGAGAGCAGCCACATCTGTGACTTCATCAGGAAGACCCTGAATGCTGGTGCCTA CAGCAAGGTGGTGCAGGAGAGGCTGGTGCAGGCTGAATACTGGCATGACCCCATCAAGGAGG ATGTGTACAGGAACCACAGCATCTTCTGGCAGACATCAACCAGGAGAGAGGCATCAATGAG AGCTACAAGAAGAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTCTCAATGACTC CATTGTGGACCCTGTGGACTCTGAGTGGTTTGGCTTCTACAGGAGTGGCCAGGCCAAGGAGA CCATCCCCCTGTCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGAC AATGCTGGGCAGCTGGTGTTCCTGGCCACAGAGGGAGACCACCTGCAGCTCTCAGAGGAGTG GTTTTATGCCCACATCATCCCCTTCCTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L

80	C-PPT1co1 GGACCTCCTGCCCCCTTGCCTCTGGTGATCTGGCATGGCATGGGAGACAGCTGTTGCAACCC CCTCAGCATGGGTGCCATTAAAGAAGATGGTAGAGAAGAAGATACCTGGTATCTATGTGCTCT CCCTAGAGATTGGGAAAACCTTTGATGGAGGATGTGGAAAATTCCTTCTTCTGAATGTCAAT TCCCAGGTCACCACAGTGTGCCAGGCTCTGGCAAAGGACCCAAAGCTACAGCAGGGCTATAA TGCCATGGGCTTCAGCCAGGGTGGCCAATTCTTGAGGGCAGTGGCTCAGAGGTGTCCATCAC CCCCATGATTAATCTAATATCTGTGGGTGGCCAACACCAAGGTGTCTTTGGTCTCCCCAGG TGTCCTGGGGAGTCCTCACACATATGTGACTTCATCAGGAAGACTCTAAATGCAGGGGCTTA TAGCAAGGTGGTTCAAGAAAGGCTGGTTCAAGCTGAGTATTGGCATGACCCCATCAAGGAAG ATGTGTACAGAAACCACTCAATTTTCTGGCTGACATCAACCAGGAGAGGGGTATCAATGAG AGTTACAAGAAGAACCTCATGGCACTGAAGAAATTTGTGATGGTGAAATTCCTTAATGACTC CATTGTGGATCCTGTGGACTCAGAATGGTTTGGCTTCTACAGGAGTGGCCAAGCCAAAGAGA CCATCCCCCTGCAGGAGACTAGCCTGTACACCAGGACAGGCTTGGTCTGAAGGAAATGGAC AATGCTGGGCAGCTGGTCTTCTGGCCACAGAAGGGGACCACCTTCAATTATCAGAGGAGTG GTTCTATGCCCACATCATACCTTTCCTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
81	C-PPT1co2 GGGCCCCCTGCCCCCTTCCCTTGGTCATATGGCATGGGATGGGAGATTCTGCTGCAACCC TCTGTCCATGGGAGCCATTAAAGAAGATGGTGGAAAAGAAGATCCCTGGAATTTATGTCCTAA GCCTGGAAATTGGGAAAACACTCATGGAAGATGTTGAAAATTCCTTCTTCTCAATGTCAAT AGTCAGGTAACCTACAGTTTGTGAGGCTCTGGCAAAGGACCCCAAGCTGCAGCAGGGCTACAA TGCCATGGGCTTCTCTCAGGGAGGACAATTCTTAAGGGCAGTAGCTCAGAGATGCCCATCTC CTCCAATGATCAACCTCATCTCTGTTGGAGGGCAGCACCAAGGGGTTTTTGGCCTCCCCAGA TGTCCTGGAGAGAGCAGCCATATCTGTGACTTCATCAGGAAGACTCTAAATGCTGGAGCCTA CAGTAAGGTTGTACAGGAGAGGCTTGTTCAGGCTGAGTACTGGCATGACCCCATTAAGGAGG ATGTGTATAGGAATCATAGCATATTCTTGGCTGATATAAACAAGAAAGGGGAATCAATGAA AGCTACAAGAAGAACCTCATGGCCCTAAAGAAGTTTGTGATGGTGAAGTTTCTCAATGACTC CATTGTTGACCCTGTGGACTCTGAATGGTTTGGCTTCTACAGGTCTGGCCAAGCCAAGGAGA CAATTCTCTCCAGGAGACTTCTCTGTACACACAAGATAGGCTGGGACTAAAGGAAATGGAC AATGCAGGTGAGCTGGTGTCTTGGCCACAGAGGGGACCACCTGCAGCTTCTGAGGAGTG GTTCTATGCCCATATTATACCCTTCTTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
82	C-PPT1co3 GGTCCCCCAGCTCCTCTCCCTCTTGTGTCATCTGGCATGGAATGGGTGACAGCTGCTGCAATCC ATTATCCATGGGAGCCATTAAAGAAGATGGTGGAGAAGAAGATCCCAGGCATCTATGTATTGT CTTTGGAAATTGGGAAGACACTGATGGAGGATGTGGAGAACTCCTTCTTCTTAAATGTCAAC TCACAGGTTACCACAGTGTGCCAAGCCCTGGCCAAGGACCCCAAGCTGCAACAGGGTTACAA TGCTATGGGGTTCTCTCAAGGTGGCCAGTTCCCTCAGGGCTGTGGCAGAGGTTGCCCTCAC CACCTATGATTAATCTCATCAGTGTGGAGGACAGCATCAAGGTGTCTTTGGCCTCCCAAGA TGTCCAGGGGAAAGTTCTCACATCTGTGATTTTATAAGGAAGACCCTCAATGCTGGAGCTTA CAGTAAAGTGGTCCAGGAGAGGCTGGTCCAGGCTGAATACTGGCATGACCCAATCAAGGAGG ATGTTTATAGAAATCATTCCATCTTCTGGCAGACATCAACCAGGAGAGGGGCATCAATGAG AGCTACAAGAAAAATCTCATGGCCCTGAAGAAGTTTGTGATGGTGAAATTTCTAAATGATAG TATTGTGGACCCTGTAGACTCTGAGTGGTTTGGCTTCTACAGGAGTGGGCAGGCCAAGGAAA CCATCCCTCTGCAGGAACTAGCCTTTACACACAGGACAGGTTGGGCTTGAAGGAAATGGAT AATGCTGGGCAGCTTGTCTTCTTGGCCACAGAGGGGGACCACCTTGCAGCTCAGTGAGGAGTG GTTTTATGCCCACATCATCCCCTTCTTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
83	PPT1co20 GGGCCCCCAGCACCCCTGCCTCTGGTCATCTGGCATGGGATGGGAGATTCTGCTGCAACCC TCTCAGCATGGGGGCCATTAAAGAAAATGGTGGAGAAGAAGATCCCTGGCATTTATGTCTTGT CCTTGGAGATAGGCAAAACACTAATGGAAGATGTGGAGAACAGCTTCTTCTTAAATGTGAAC AGCCAGGTCACCACAGTATGTGAGGCCCTTGGCCAAGGACCCCAAGCTCCAGCAGGGCTACAA TGCCATGGGTTTTTCTCAGGGAGGTGAGTTCCCTCAGGGCTGTGGCTCAGAGATGCCCTCAC CACCTATGATCAATCTGATCTCTGTGGGAGGGCAGCATCAAGGTGTCTTTGGTCTTCTAGG

	TGCCCTGGGGAGTCATCCCACATCTGTGACTTCATCAGGAAAACCTTTAAATGCTGGAGCCTA CAGCAAGGTTGTCCAAGAGAGGCTTGTCCAAGCAGAGTACTGGCATGATCCCATCAAAGAGG ATGTTTACAGAAAACACAGCATCTTCTGGCAGACATCAATCAAGAGAGGGGCATTAATGAG TCCTACAAGAAAAACCTCATGGCCCTAAAGAAATTTGTGATGGTGAAATTCCTGAAATGACTC CATTGTTGACCCTGTAGACTCTGAGTGTTTGGCTTCTATAGATCAGGCCAGGCCAAGGAAA CCATTCTCTCCAAGAAACCTCCCTGTACACACAAGATAGGCTGGGGTTGAAAGAGATGGAT AATGCTGGGCAACTGGTGTTCTTGCCACTGAGGGTGACCATTACAGCTTTCAGAGGAATG GTTCTATGCCACATCATACCCTTCTTGATG In variants thereof nucleotides 1-3 encodes G, D, V or L
84	C-PPT1co5 GGGCCTCCAGCCCCCTTACCTTTGGTCATATGGCATGGAATGGGAGACAGTTGTTGCAACCC CCTCTCTATGGGTGCTATTAAGAAAATGGTGGAGAAGAAGATACCTGGCATTATGTCTCT CCCTGGAAATAGGCCAAAACACTGATGGAAGATGTAGAAAACAGCTTTTCTCCTCAATGTCAAT TCCCAGGTAACAACCTGTGTGCCAGGCCCTGGCCAAGGATCCCAATTGCAGCAAGGTTATATA TGCAATGGGATTCTCCCAAGGAGGCCAATTTCTGAGGGCTGTGGCCAGAGGTGTCCTATCTC CTCCCATGATCAACCTCATCAGTGTGGGGGGGCAGCACCAAGGAGTTTTTGGACTTCCAAGG TGCCCTGGGGAATCCAGCCACATTTGTGACTTCATTAGAAAGACACTCAATGCTGGTGCTTA TTCCAAAGTTGTTTCAAGAGAGGCTTGTGCAGGCTGAGTATTGGCATGACCCCATCAAGGAGG ATGTGTACAGGAATCATTCATCTTCTTGGCTGACATCAACCAGGAGAGAGGCATCAATGAG AGCTACAAGAAAAACCTCATGGCCTTAAAGAAGTTTGTGATGGTGAAGTTCCTAAATGATAG CATAGTTGACCCAGTGGATTCTGAGTGGTTTGGTTTCTATAGGAGTGGCCAAGCCAAGGAGA CCATTCCTTGCAGGAGACCTCCCTGTACACCCAGGACAGGCTTGGTCTCAAGGAGATGGAC AATGCAGGACAGCTAGTTTCTTGCCACTGAGGGGGATCACCTCCAACCTCAGTGAAGAATG GTTCTATGCACACATCATCCCCCTTCTTGGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
85	C-PPT1co6 GGGCCCCCAGCACCCCTTCTTTGGTGATATGGCATGGCATGGGTGATAGCTGTTGCAACCC CCTGAGCATGGGGGCCATCAAGAAAATGGTGGAGAAAAGATTCCAGGCATATATGTGCTCT CCCTGGAGATTGGAAAACCCCTCATGGAAGATGTGGAGAACAGCTTCTTCTTAAATGTGAAT AGTCAAGTTACCACAGTGTGCCAGGCACTGGCCAAGGACCCCAAGCTCCAACAGGGGTACAA TGCCATGGGCTTCTCTCAGGGAGGGCAGTTTCTCAGAGCAGTGGCCAGAGATGCCCTTCAC CCCCCATGATCAATCTCATTTCTGTTGGAGGGCAACACCAGGGGGTATTTGGTCTACCCAGG TGCCCAGGTGAATCCAGCCACATCTGTGACTTCATCAGGAAGACACTGAATGCAGGAGCCTA CAGCAAAGTGGTGCAAGAGAGGCTGGTGAGGCAGAGTACTGGCATGACCCCATCAAGGAGG ATGTCTACAGAAACCATTCCATCTTCTTGGCAGACATCAACCAGGAAAGAGGAATCAATGAG TCCTACAAGAAGAACCTTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCTGAATGATTC TATTGTTGACCCTGTGGATTCTGAGTGGTTTGGATTTTACAGAAGTGGCCAGGCCAAGGAAA CAATTCCTACTACAAGAGACCTCCCTCTACACCCAGGACAGGCTGGGCCTCAAAGAAATGGAC AATGCTGGCCAGCTGGTGTTTCTGGCTACAGAGGGAGACCATCTGCAGCTGTCAGAGGAGTG GTTCTATGCTCACATCATACCTTCTTAGGATG In variants thereof nucleotides 1-3 encodes G, D, V or L
86	C-PPT1co7 GGCCCTCCTGCTCCCCTGCCTTTGGTTATCTGGCATGGAATGGGTGACAGCTGTTGCAACCC CTTAAGCATGGGAGCTATTAAGAAAATGGTTGAGAAGAAGATCCCTGGGATCTATGTCTCTGA GCTTGGAATTTGGGAAGACACTCATGGAAGATGTGGAGAACTCCTTCTTCTGAATGTGAAT AGCCAAGTGACTACAGTGTGTCAAGCCTTGCCCAAGGACCCCAAGCTGCAGCAGGGCTACAA TGCCATGGGCTTCTCCAGGGAGGACAGTTTCTGAGAGCAGTAGCTCAGAGTGCCCTTCAC CTCCCATGATCAACCTCATCAGTGTGGGGGGCCAGCACAGGGAGTGTGGGCTGCCAGG TGCCCTGGGGAGTCATCTCACATCTGTGACTTCATCAGGAAAACCTCAATGCTGGGGCCTA CAGCAAGGTGGTCCAGGAGAGACTGGTGAGGCAGAGTATTGGCATGACCCCATCAAGGAGG ATGTCTATAGGAACCACAGCATTTTCTTAGCTGACATCAATCAAGAGAGAGGCATCAATGAG TCCTACAAGAAGAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTCCTCAATGATTC TATTGTGGATCCTGTGGACTCAGAGTGGTTTGGCTTTTACAGATCTGGGCAAGCAAAGGAGA CCATTCCTTGCAGGAAACATCCCTGTACACCCAGGATAGGCTGGGTCTGAAAGAGATGGAC AATGCTGGCCAGCTTGTTTCTTCTGGCTACAGAAGGGGACCATTACAGCTATCTGAAGAATG

	GTTTTATGCTCACATTATCCCTTTCCTTGGATG In variants thereof nucleotides 1-3 encodes G, D, V or L
87	C-PPT1co8 GGGCCCCCTGCACCTCTGCCCCCTGGTCATTTGGCATGGCATGGGTGACTCCTGTTGCAATCC CCTGAGCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATCCCAGGAATCTATGTTCTGT CCTTAGAAATTGGGAAGACTCTGATGGAAGATGTGGAAAACAGCTTCTTCCTGAATGTGAAT TCCCAGGTGACCACAGTCTGCCAGGCTCTGGCTAAGGACCCCAAACCTCCAGCAGGGCTACAA TGCTATGGGATTCTCCCAGGGTGGGCAGTTCCTCAGAGCAGTGGCTCAGAGATGCCCATCAC CCCCAATGATCAACCTCATCTCTGTTGGTGGTCAGCATCAAGGTGTCTTTGGACTCCCCAGG TGCCCTGGGGAGTCTCTCACATCTGTGACTTTATCAGAAAAACCTTAAATGCTGGGGCCTA CAGCAAGGTGGTCCAGGAAAGGCTGGTCCAGGCTGAGTACTGGCATGATCCCATCAAGGAAG ATGTGTACAGGAACCACTCCATTTTTCTGGCTGACATAAACCAAGAGAGGGGCATCAATGAG TCATATAAGAAAAACCTGATGGCTCTGAAAAAGTTTGTCTATGGTTAAATTCCTCAATGACTC TATTGTGGACCCTGTTGACTCTGAGTGGTTTGGCTTCTACAGGTCAGGCCAAGCTAAAGAGA CCATCCCCCTGCAGGAGACATCCCTGTACACACAGGACAGGTTAGGACTGAAGGAGATGGAC AATGCTGGACAGCTGGTGTCTCTGGCCACAGAGGGTGACCACCTACAGTTGTCAGAAGAATG GTTCTATGCTCACATCATACCCTTTCCTTGGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
88	C-PPT1co9 GGCCCCACCTGCTCCCCCTCCCCTTGGTGATATGGCATGGCATGGGAGATAGCTGTTGTAATCC CCTCTCCATGGGTGCCATTAAGAAGATGGTGGAAAAGAAGATTCTTGGAAATCTATGTGCTGT CCCTGGAGATAGGGAAGACTCTGATGGAAGATGTGGAAAATTCCTTCTTCTGAATGTCAAC TCCCAGGTGACAACTGTATGCCAGGCCTTGGCAAAGGACCCCAAGCTGCAGCAGGGATACAA TGCCATGGGATTCTAGCCAAGGTGGACAGTTCCTGAGAGCAGTAGCCCAGAGATGCCCCAAGTC CCCCCATGATCAACCTCATCTCTGTGGGTGGACAGCACCAGGGAGTGTCTTGGCCTGCCCAGA TGTCAGGGGAGAGCAGCCACATTTGTGACTTCATCAGGAAAACACTGAATGCTGGGGCCTA CAGCAAGGTGGTGCAGGAGAGGCTAGTGCAGGCAGAGTACTGGCATGACCCAATCAAAGAGG ATGTTTATAGGAACCAAGTATCTTTTTGGCAGATATTAACCAGGAAAGAGGCATCAATGAG TCTTACAAGAAGAACCTAATGGCCTTGAAGAAATTTGTCTATGGTGAAGTTCTCTCAATGATTC CATTGTGGATCCTGTTGACTCAGAATGGTTTGGATTTTATAGATCAGGCCAGGCCAAAGAAA CCATACCACTGCAGGAAACATCCCTCTACACTCAAGATAGGCTGGGGCTGAAGGAAATGGAT AATGCTGGGCAGCTGGTATTCTCTGGCTACTGAAGGGGATCACCTCCAGCTAAGTGAGGAGTG GTTCTATGCTCACATCATCCCATTTCTTGGCTG In variants thereof nucleotides 1-3 encodes G, D, V or L
89	C-PPT1co10 GGACCCCCCTGCCCTTTTGCCCCCTGGTCATTTGGCATGGGATGGGTGACAGCTGCTGCAATCC CCTAAGCATGGGAGCCATCAAGAAGATGGTGGAAAAGAAGATTCTTGGCATTATGTCCTCTCT CCCTGGAGATTGGGAAGACCCTGATGGAGGATGTGGAGAACAGCTTCTTCCTGAATGTGAAC TCCCAGGTTACAACAGTTTGCCAGGCCTTAGCCAAGGACCCCTAAACTCCAGCAAGGCTACAA TGCAATGGGCTTCTCTCAGGGTGGGCAGTTCCTGAGGGCTGTGGCACAGAGGTGCCCCAGTC CTCCCATGATCAACCTCATCTCAGTGGGAGGCCAGCACCAGGGTGTCTTTGGGTTGCCAAGA TGTCCTGGAGAGAGCTCTCATATCTGTGACTTCATCAGGAAGACACTGAATGCTGGGGCTTA TAGCAAGGTGGTACAGGAGAGGCTGGTGCAGGCTGAGTACTGGCATGACCCCATCAAGGAAG ATGTGTACAGGAACCATTAATCTTCTTGGCAGATATTAACCAGGAGAGGGGCATCAATGAA AGCTACAAGAAGAACCTGATGGCCTTGAAGAAATTTGTCTATGGTAAAGTTTCTGAATGACTC CATTGTGGATCCTGTAGACTCTGAGTGGTTTGGCTTTTACAGAAGTGGCCAAGCCAAGGAAA CCATCCCCCTGCAGGAGACATCTCTGTACACTCAGGACAGGTTGGGACTGAAGGAGATGGAC AATGCAGGACAGCTGGTCTTTCTGGCCACAGAGGGAGACCACCTGCAGCTCTCAGAAGAGTG GTTTTATGCCACATCATTCCTTCTTAGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
90	C-PPT1co11 GGACCCCCCTGCACCTCTGCCCCCTTGTTATATGGCATGGGATGGGAGACAGCTGCTGCAATCC CCTCAGCATGGGTGCCATCAAGAAGATGGTGGAGAAAAAATTCCTGGAATTTATGTCCTCA

	GCTTGGAGATAGGGAAGACACTTATGGAAGATGTGGAAAACAGTTTCTTCTCCTGAATGTGAAC TCCCAGGTGACCACTGTGTGCCAGGCACTGGCCAAAGACCCTAAGCTGCAGCAAGGCTACAA TGCTATGGGCTTCAGCCAGGGGGGTTCAGTTTCTGAGAGCAGTGGCTCAGAGATGCCCCCTCAC CTCCCATGATCAACCTGATCTCTGTGGGAGGCCAGCACCAAGGGGTTTTTGGGCTACCAAGG TGTCCTGGAGAGTCTTCACATATCTGTGACTTCATCAGAAAAGACCTTAAATGCAGGGGCCTA TAGCAAAGTGGTTCAGGAGAGGCTTGTTCAGGCTGAATACTGGCATGACCCCATCAAAGAGG ATGTCTACAGGAACCATTCATCTTCTTAGCTGACATCAATCAGGAGAGGGGGATAAATGAA AGCTACAAGAAAAATCTCATGGCCCTGAAGAAGTTTGTCTATGGTGAAGTTTCTCAATGATTC CATTGTGGATCCAGTGGATTGAGAGTGGTTTGGCTTTTACAGGTCAGGCCAAGCTAAAGAGA CCATCCCCCTCCAGGAGACCTCCCTCTACACACAGGACAGACTGGGATTGAAAGAGATGGAC AATGCTGGGCAGCTAGTTTTTTTGGCCACAGAGGGAGACCACCTGCAGCTCAGTGAGGAATG GTTTTATGCTCACATCATTCCCTTCCTTGGTTG In variants thereof nucleotides 1-3 encodes G, D, V or L
91	C-PPT1co12 GGGCCACCTGCTCCCCCTGCCCCCTTGTCTATCTGGCATGGGATGGGAGACAGCTGCTGCAACCC CCTGTCCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAAAATACCTGGAATATATGTCTCTCT CTCTGGAGATTGGCAAGACACTGATGGAAGATGTGGAGAATTCATTTTCTGAATGTCAAC TCCCAAGTGACCACAGTGTGCCAGGCCTTAGCCAAGGACCCCAAGCTTCAGCAGGGCTACAA TGCAATGGGATTGAGCCAGGGAGGGCAATTTCTGAGAGCAGTTGCTCAGAGATGCCCATCTC CACCTATGATCAATTTGATCTCTGTGGGTGGACAGCATCAAGGTGTGTTTGGGCTACCCAGA TGCCCTGGGGAGTCTAGCCACATCTGTGACTTCATCAGGAAGACTCTGAATGCTGGAGCCTA CTCAAAGTGGTGCAGGAGAGGCTGGTGCAAGCTGAGTACTGGCATGATCCATAAAGGAGG ATGTGTACAGGAACACAGCATCTTCTTAGCAGACATCAACCAAGAGAGAGGCATCAATGAG TCCTACAAGAAAGAACCTCATGGCTCTGAAGAAGTTTGTGATGGTCAAGTTCTTAAATGATTC TATTGTGGATCCTGTGGACTCTGAATGGTTTGGATTTTACAGATCAGGCCAAGCCAAGGAGA CCATCCCTCTGCAGGAGACCTCTCTGTACACACAAGACAGGCTGGGACTCAAGGAGATGGAC AATGCTGGACAGCTTGTCTTTCTGGCCACAGAAGGGGACCACCTGCAGCTGTCAGAAGAATG GTTCTATGCTCACATCATAACCATTCTTGGTTG Variants thereof wherein nucleotides 1-3 encodes G, D, V or L
92	C-PPT1co13 GGGCCCCCTGCACCTCTCCCCCTTGTCTATCTGGCATGGGATGGGAGACTCCTGCTGTAACCC TTTATCAATGGGGGCCATCAAGAAGATGGTAGAGAAGAAAAATCCAGGAATTTATGTCTCTCA GCTTGGAATAGGCAAGACTCTCATGGAAGATGTGGAAAACAGCTTCTTCTGAATGTGAAT AGCCAGGTCACAACAGTGTGCCAGGCCTTAGCTAAAGATCCCAAATTCAGCAGGGCTACAA TGCCATGGGATTTTCCAGGGAGGGCAATTCCTCAGGGCAGTGGCACAGAGATGCCCTAGCC CTCCAATGATCAACCTTATCAGTGTGGGTGGGCAGCATCAAGGGGTTTTTGGCTTGCCCAGG TGCCCTGGGGAATCTAGCCACATCTGTGACTTTATAAGGAAAACCTGAATGCAGGAGCCTA CTCAAAAGTGGTCCAGGAGAGGCTGGTGCAAGCTGAATATTGGCATGATCCCATCAAAGAGG ATGTGTATAGGAACACAGCATCTTTTTAGCTGACATCAATCAAGAAAGGGGTATCAATGAA AGCTACAAGAAAAATTTAATGGCTCTGAAGAAGTTTGTCTATGGTAAAGTTTCTCAATGACTC CATTGTGGACCTGTGGATTCTGAGTGGTTTGGATTCTATAGAAGTGGACAAGCCAAAGAAA CCATCCCCTGCAGGAGACCAGCCTCTACACACAAGACAGATTGGGACTCAAAGAGATGGAC AATGCTGGACAGCTTGTCTTTCTGGCTACAGAGGGTGACCATCTGCAGCTGAGTGAAGAATG GTTTTATGCTCACATCATAACCATTCTTGGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
93	C-PPT1co14 GGGCCACCTGCCCCCTCTGCCATTGGTGATTGGCATGGCATGGGAGATAGCTGCTGCAATCC TCTCTCCATGGGGGCCATTAAGAAAATGGTGGAAAAAAGATTCCAGGGATTTATGTGTTGT CTTTAGAGATTGGGAAGACTCTCATGGAAGATGTGGAGAACAGCTTCTTCTGAATGTCAAC TCACAGGTCACTACTGTCTGTGTCAGGCTTTAGCTAAGGACCCCAACTCCAGCAGGGGATACAA TGCTATGGGCTTCTCTCAGGGAGGGCAATTTCTGAGAGCAGTGGCCCAGAGGTGCCCCAGCC CTCCCATGATCAATCTTATCTCTGTGGGTGGACAGCACCAAGGTGTTTTTGGCTTGCCCAGG TGTCCTGGGGAGAGCTCTCACATCTGTGATTTTCATCAGGAAGACACTCAATGCAGGAGCTTA CTCCAAAGTGGTCCAAGAGAGGCTGGTCCAAGCAGAATACTGGCATGATCCAATTAAAGAAG ATGTTTATAGGAACACAGCATCTTCTGGCAGACATCAATCAGGAAAGAGGGATAAATGAG

	TCCTACAAGAAAACTTAATGGCCCTGAAGAAGTTTGTGATGGTCAAATTCCTTAATGACTC CATTGTGGACCTGTAGATTACAGAGTGGTTTGGTTTTATAGAAGTGGCCAGGCTAAAGAGA CAATCCCCCTGCAAGAGACCAGCCTCTACACCCAAGACAGGCTAGGCCTGAAAGAAATGGAC AATGCTGGACAACCTGGTCTTTCTTGCCACTGAAGGTGACCATCTACAGCTGTCAGAAGAATG GTTCTATGCTCACATCATCCCCCTTCCTGGGGTG In variants thereof nucleotides 1-3 encodes G, D, V or L
94	C-PPT1co16 GGCCCTCCAGCCCCCTGCCCCCTGGTCATCTGGCATGGCATGGGAGACAGCTGCTGTAACCC TCTGTCCATGGGAGCCATTAAGAAGATGGTGGAGAAGAAGATCCCTGGAATCTATGTCCTGT CCCTGGAGATTGGGAAGACACTGATGGAGGATGTGGAGAATTCCTTCTCTTAAATGTGAAC AGCCAGGTCACCACAGTTTGCCAGGCATTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAA TGCCATGGGCTTCTCCCAGGGAGGCCAGTTCCTGAGGGCTGTGGCCAGAGGTGCCCCCTCAC CTCCCATGATCAACCTGATCTCTGTGGGAGGCCAGCACCAAGGTGTGTTTGGCCTGCCCAGG TGCCCTGGGGAGTCCAGCCACATCTGTGACTTCATCAGGAAGACTTTAAATGCTGGAGCCTA CAGCAAGGTGGTCCAGGAGAGGCTGGTGCAGGCAGAATACTGGCATGATCCCATCAAGGAGG ATGTGTACAGGAACACAGCATCTTCCTGGCAGACATCAACCAAGAGAGAGGGGATCAATGAG AGCTACAAGAAGAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAATTCCTCAATGACTC CATTGTGGACCTGTGGATTCTGAGTGGTTTGGCTTTTACAGGTCTGGCCAGGCCAAGGAGA CCATCCCTCTGCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGAC AATGCTGGGCAGCTGGTGTTCCTGGCCACAGAGGGAGACCATCTGCAGCTGAGTGAGGAGTG GTTTTATGCTCACATCATCCCCCTTCCTGGGGTG In variants thereof nucleotides 1-3 encode G, D, V or L
Full-Length PPT1 Encoding sequences	
95	Sp7-F signal_PPT1 ATGGCCTTTCTGTGGCTGCTGTCTGCTGGGCCCTGCTGGGGACCACCTTTGGCGGGCCGCC GGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCA TGGGTGCTATTAAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAG ATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTCAATTCCTCAAGT AACACAGTGTGTGAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGG GATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATG ATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGG AGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAG TTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTAT CGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAA GAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAATTCCTCAATGATTCCATTGTGG ACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTTCCC TTACAGGAGACCTCCCTGTACACACAGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGG ACAGCTAGTGTCTTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATG CCCACATCATACCATTCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encodes G, D, V or L
96	Sp7-P signal_PPT1 ATGGCCTTCTGTGGCTGCTGTCTGCTGGGCACTGCTGGGCACCACATTTGGCGGGCCGCC GGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCA TGGGTGCTATTAAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAG ATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTCAATTCCTCAAGT AACACAGTGTGTGAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGG GATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATG ATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGG AGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAG TTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTAT CGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAA GAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAATTCCTCAATGATTCCATTGTGG ACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTTCCC TTACAGGAGACCTCCCTGTACACACAGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGG

	ACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATG CCCACATCATACCATTCCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encode G, D, V or L
97	NGF-1 derived Signal_PPT1 ATGTCCATGTTGTTCTACACTCTGATCACAGCTTTTCTGATCGGCATACAGGCGGGGCCGCC GGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCA TGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAG ATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTCAATTCCTAAGT AACAACAGTGTGTCAAGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGG GATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATG ATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGG AGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAG TTGTTCAAGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTAT CGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCTTACAA GAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGG ACCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTTCCC TTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGG ACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATG CCCACATCATACCATTCCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encode G, D, V or L
98	Insulin-1 signal_PPT1 ATGGCCCTGTGGATGCGCCTCCTGCCCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC AGCCGACAGCCGTGCCGCCCGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCT GTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATT TACGTCTTATCTTTAGAGATTGGGAAGACCCCTGATGGAGGACGTGGAGAACAGCTTCTTCTT GAATGTCAATTCCTAAGTAACAACAGTGTGTCAAGGCACTTGCTAAGGATCCTAAATTGCAGC AAGGCTACAATGCTATGGGATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGA TGCCCTTCACCTCCCATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGG ACTCCCTCGATGCCAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATG CTGGGGCGTACTCCAAAGTTGTTCAAGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCC ATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGG TATCAATGAGTCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCC TCAATGATTCCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAA GCCAAGGAAACCATTCCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAA GGAAATGGACAATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGT CTGAAGAATGGTTTTATGCCCACATCATACCATTCCCTTGGATTAA In variants thereof nucleotides GTG shown in bold encode G, D, V or L
99	SPARC-1 signal_PPT1 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGCAGGGGCCGCCGGC GCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGG GTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATT GGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCTAAGTAAC AACAGTGTGTCAAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGAT TCTCCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATGATC AATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGGAGA GAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTG TTACAGGAACCGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGC AACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAA AAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACC CTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCTTA CAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACA GCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCC

	ACATCATACCATTTCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encode G, D, V or L
100	CD33-1 derived signal_PPT1 ATGCCGCTGCTGCTACTGCTGCCCCCTGCTGTGGGCAGGGGCCAGGGCTGGGCCGCCGGCGCC GCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTG CTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGG AAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAAC AGTGTGTCAAGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCT CCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATGATCAAT CTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGGAGAGAG CTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTT AGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAAC CACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCTTACAAGAAAA CCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTG TAGATTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAG GAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCT AGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACA TCATACCATTTCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encode G, D, V or L
101	IL34-1 derived signal_PPT1 ATGCCCCGGGGCTTCACCTGGCTGCGCTATCTTCTGCTGGGGATCTTCCTTGGCGTGGCCAG GGGGGGGCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCA ATCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTC TTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGT CAATTCCCAAGTAACAACAGTGTGTCAAGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCT ACAATGCTATGGGATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCT TCACCTCCCATGATCAATCTGATCTCGGTTGGGGACAACATCAAGGTGTTTTTGGACTCCC TCGATGCCCAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGG CGTACTCCAAAGTTGTTTCAAGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAG GAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAA TGAGTCTTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATG ATTCCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAG GAAACCATTCCCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAAT GGACAATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAG AATGGTTTTATGCCACATCATACCATTTCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encode G, D, V or L
102	OSM-1 signal_PPT1 ATGGGGGTACTGCTCACACAGAGGACGCTGCTCAGTCTGGTCCCTTGCACTCCTGTTTCCAGA CGCTAGAGCCCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAA GCATGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTA GAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCCA AGTAACAACAGTGTGTCAAGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTA TGGGATTCTCCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCC ATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCC AGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCA AAGTTGTTCAAGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTG TATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTA CAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTG TGGACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATT CCCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGC AGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTT ATGCCACATCATACCATTTCCTTGGATTAA

	In variants thereof nucleotides CTG shown in bold encode G, D, V or L
103	tPA-1 derived signal PPT1 ATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTC GCCCAGCGCCAGAGCCCTGCAGCACCTGGACCGCCGGCGCCGCTGCCGTTGGTGATCTGGC ATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAG AAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACGT GGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTGTGTCAGGCACTTGCTA AGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCCAGGGAGGCCAATTTCT AGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGATCTCGGTTGGGGGACA ACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGCTCTCACATCTGTGACTTCA TCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGCC GAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGA TATAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCTGATGGCCCTGAAGAAGT TTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTGCGAGTGGTTTGGGA TTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAGGAGACCTCCCTGTACACACA GGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAG GGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCCCTTGGATTAA In variants thereof nucleotides CTG shown in bold encode G, D, V or L
104	IL-2-1 (derived) _PPT1 ATGTACAGGATGCAACTCCTGTCTTGCAATTGCACTAAGTCTTGCACTTGTCTCTGCTGTCTGC CAGGGCCGGGCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTT GCAATCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTAC GTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAA TGTCAATTCCCAAGTAACAACAGTGTGTCAGGCACTTGCTAAGGATCCTAAATTGCAGCAAG GCTACAATGCTATGGGATTCTCCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGC CCTTCACCTCCCATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACT CCCTCGATGCCCAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTG GGGCGTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCAT AAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTAT CAATGAGTCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAATTCCTCA ATGATTCCATTGTGAGCCCTGTAGATTGCGAGTGTTGGATTTTACAGAAGTGGCCAAGCC AAGGAAACCATTCCCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGA AATGGACAATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTG AAGAATGGTTTTTATGCCACATCATACCATTCCCTTGGATTAA In variants thereof nucleotides GGG shown in bold encode G, D, V or L
105	Trypsin-1_PPT1 ATGAATCTACTTCTGATCCTTACCTTTGTTGCAGCTGCTGTTGTGATTAGGGCAGTGGACCT GCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTA TTAATAAATGGTGGAGAAGAAAAATACCTGGAATTTACGCTTATCTTTAGAGATTGGGAAG ACCCTGATGGAGGACGTGGAGAAGCAGCTTCTTCTTGAATGTCAATTTCCCAAGTAACAACAGT GTGTGAGGCACTTGCTAAGGATCCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCC AGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTG ATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTC TCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGG AACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCAC AGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCT GATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAG ATTCGGAGTGGTTTGGATTTTACAGAAGTGCCAAGCCAAGGAAACCATTCCCTTACAGGAG ACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGT GTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTTATGCCACATCA TACCATTCCCTTGGATTAA

	In variants thereof nucleotides GTG shown in bold encode G, D, V or L
106	BDNF-1-1/2_PPT1 ATGACAATCCTCTTCTGACCATGGTGATTAGCTACTTTGGCTGCGCTAXGGCCGGGCCGCC GGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCA TGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAG ATTGGGAAGACCTGATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTCAATTCCCAAGT AACAAACAGTGTGTCAAGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGG GATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATG ATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGG AGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAG TTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTAT CGCAACCACAGCATCTTCTTGCCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAA GAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGG ACCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTTCCC TTACAGGAGACCTCCCTGTACACACAGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGG ACAGCTAGTGTCTTGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTATG CCCACATCATACCATTCCTTGGATAA Wherein X is A or G; In variants thereof nucleotides GGG shown in bold encode G, D, V or L
107	sp7_7-PPT1co1 ATGGCCTTCTCTGGCTGCTCAGCTGTTGGGCCTTGCTGGGGACCACCTTTGGAGGCCCCC AGCACCTCTTCCCTTGGTGATCTGGCATGGCATGGGGGATAGCTGCTGCAATCCTCTCTCAA TGGGAGCCATTAAGAAGATGGTGGAGAAGAAAATCCCAGGGATATATGTTCTGTCTTGGAG ATTGGAAAAACCTTGATGGAAGATGTGGAATAAGCTTTTTCCTTAATGTGAACTCTCAAGT CACCACAGTGTGCCAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAAGGCTACAATGCAATGG GCTTCTCCCAAGGAGGTCAATTCTGAGGGCTGTTGCACAGAGGTGTCCAGTCCCCCATG ATCAACCTGATCTCAGTGGGTGGCCAGCACCAAGGTGTCTTTGGGCTCCCCAGGTGTCCAGG GGAGAGCAGCCACATCTGTGACTTCATTAGAAAGACACTAAATGCAGGAGCTACAGCAAGG TGGTGCAGGAGAGGCTTGTTCAGGCTGAGTACTGGCATGACCCCATTAAGGAAGATGTCTAT AGAAACCACAGCATCTTCTGGCAGATATAAACCAGAAAGAGGCATTAATGAAAGTTACAA GAAGAATCTTATGGCCCTCAAGAAGTTTGTGATGGTGAAGTTCTGAATGACTCCATTGTTG ACCTGTGGACTCTGAATGGTTTGGTTTCTACAGATCAGGACAAGCCAAGGAAACAATCCCC CTGCAAGAGACCTCCCTGTACACCCAGGACAGGCTGGGCCTGAAAGAAATGGATAATGCTGG CCAACTGGTCTTCTGGCAACTGAAGGGGACCACCTGCAACTCTCTGAGGAATGGTTCTATG CTCACATCATCCCATTTTTTGGGCTGA
108	sp7_PPT1co2 ATGGCCTTTCTCTGGTTGCTCTCTTGTGGGCCCTGCTGGGCACAACTTTTGGGGGCCACC TGCACCACTTCCCTTGGTCATCTGGCATGGCATGGGTGACTCTTGCTGCAATCCCTTATCCA TGGGGGCTATTAAAGAAGATGGTGGAAAAGATTCCTGGAATTTATGTGCTGAGTCTGGAA ATTGGAAAGACCTCATGGAAGATGTTGAAAACAGCTTCTTCTGAATGTCAATTCCCAAGT CACCACAGTGTGTCAAGGCTCTGGCAAAGGACCCCAAGCTCCAGCAAGGATATAATGCCATGG GATTTTCCCAAGGTGGCCAGTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCCCTATG ATCAACCTCATCTCTGTGGGTGGGCAGCATCAGGGTGTCTTTGGGCTCCCAAGATGCCCAGG AGAAAGCTCTCACATCTGTGACTTTATTAGAAAGACCCTGAATGCTGGGGCTTACAGCAAGG TTGTCCAGGAGAGGCTGGTGCAGGCTGAATACTGGCATGACCCCATTAAGGAAGATGTGTAC AGAAACCACAGTATCTTCTTGCTGATATCAACCAAGAAAGAGGCATCAATGAGAGCTACAA GAAGAATCTCATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATAGCATTGTGG ACCTGTAGACTCAGAGTGGTTTGGATTCTACAGGTCTGGCCAGGCTAAGGAGACCATCCCA CTGCAAGAGACAAGCCTTTACACACAGGACAGGCTTGGTCTGAAAGAGATGGATAATGCTGG ACAGCTGGTGTCTTGCCACTGAAGGGGACCACCTGCAGCTGAGTGAAGAATGGTTCTATG CCCACATCATCCCTTCTTAGGCTGA
109	sp7_PPT1co3 ATGGCCTTTCTGTGGCTCCTCAGCTGCTGGGCCCTGCTGGGCACCACTTTTGGAGGGCCCCC AGCACCCCTGCCTCTGGTCATTTGGCATGGCATGGGGGACTCCTGCTGCAATCCCTTATCCA TGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAATTTATGTTCTGTCTTTAGAA ATAGGAAAAACACTGATGGAAGATGTGGAATAATTCCTTCTTCTCAATGTCAACTCCAGGT

	AACAACAGTGTGTCTAGGCCCTGGCCAAAGACCCCAAATTACAGCAGGGCTACAATGCTATGG GATTCTCACAGGGAGGCCAATTTCTCAGAGCTGTTGCTCAGAGATGCCCTTACACCCCATG ATCAACCTGATTTTCAGTGGGTGGACAACATCAAGGAGTTTTTGGATTACCAAGGTGCCCTGG GGAGAGCTCTCACATTTGTGATTTTATAAGAAAAACCTTGAATGCTGGGGCCTACTCAAAGG TGGTCCAGGAGAGGTTAGTGCAGGCTGAGTATTGGCATGACCCCATTAAGGAAGATGTGTAC AGAAACCACTCCATTTTTTTGGCTGACATCAACCAGGAGAGAGGAATTAATGAATCCTACAA GAAGAACCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCCTCAATGACTCCATTGTGG ATCCAGTGGATTACAGAGTGGTTTGGATTCTACAGAAGTGGCCAAGCCAAGGAAACAATACCA CTTCAAGAGACCAGCCTGTACACCCAGGACAGACTGGGCCCTAAAGGAAATGGACAATGCAGG TCAGCTGTCTTCTGGCCACTGAGGGGGACCACCTTCAACTCTCTGAAGAATGGTTTTATG CCCACATCATACCCCTTCTGGGGTGA
110	sp7_PPT1co4 ATGGCCTTTTTATGGCTGCTGAGCTGCTGGGCTTTGCTGGGCACCACATTTGGGGGCCCTCC AGCACCTCTTCCCTTTAGTGATCTGGCATGGCATGGGTGACAGCTGCTGCAACCCACTGTCCA TGGGAGCCATCAAGAAAATGGTGGAGAAGAAAATTCAGGAATCTATGTCTTATCCCTGGAG ATTGGGAAGACTCTGATGGAAGATGTGGAGAATCCTTCTTCTCTGAATGTCAATTCCTCAAGT CACCACAGTCTGCCAGGCCCTGGCCAAAGACCCCAAGCTGCAGCAAGGGTACAATGCTATGG GCTTCTCCCAAGGGGGGCAGTTCTTGAGAGCAGTTGCCAGAGATGCCCCAGCCCTCCAATG ATTAATCTGATCTCTGTGGGAGGTGAGCATCAAGGAGTCTTTGGGTTGCCCCAGATGTCCAGG AGAGAGCAGTCACATCTGTGACTTCATCAGGAAAACACTGAATGCTGGGGCATAACAGCAAGG TGGTGCAGGAAAGGCTTGTGCAGGCAGAGTACTGGCATGACCCCATTAAGAGGATGTGTAC AGAAACCACAGCATCTTCTTGCAGACATAAATCAGGAGAGAGGTATCAATGAGTCTTACAA GAAGAATCTGATGGCCCTAAAGAAATTTGTGATGGTAAAATTCCTCAATGACTCCATAGTGG ACCCAGTTGACTCTGAGTGGTTTTGGATTTTACAGAAGTGGCCAAGCAAAGAGACCATTCCC CTTCAAGAGACATCCCTGTACACCCAAGATAGGCTGGGCCCTGAAAGAGATGGATAATGCAGG GCAATTGGTGTCTCTTGTACAGAAGGAGACCACCTCCAGCTATCTGAAGAGTGGTTCTATG CCCACATCATCCCTTTTCTAGGGTGA
111	sp7_7-PPT1co5 ATGGCCTTTCTGTGGCTGCTGAGCTGCTGGGCCCTGCTGGGGACAACATTTGGAGGCCCCC TGCCCCACTTCCCTTGGTGATATGGCATGGCATGGGAGATTCTTGTGTAAACCCACTGTCAA TGGGGGCCATTAAAGAAGATGGTGGAAAAGAAGATCCCTGGAATCTATGTCTCTCCCTGGAA ATTGGGAAGACCCTCATGGAGGATGTGGAAAACCTCATTTCTTCTCAATGTGAACAGCCAGGT CACCCTGTTTGTGAGGCATTAGCAAAGGACCCCAAGCTTCAGCAGGGATACAATGCTATGG GCTTTAGCCAAGGTGGGCAATTCTTGAGAGCAGTGGCACAGAGATGCCCTCACCTCCCATG ATCAACCTCATCAGTGTAGGAGGCCAGCACAGGGGGTTTTTGGGCTCCCCAGATGCCCTGG TGAATCATCTCACATCTGTGACTTCATCAGAAAAACCTTGAATGCTGGAGCCTACAGTAAAG TAGTGCAAGAGAGGCTGGTGCAAGCAGAATACTGGCATGATCCCATCAAGGAAGATGTCTAC AGGAACCATAGCATCTTCTGGCTGACATCAACCAGGAGAGGGGAATCAATGAGTCTTACAA GAAAAACCTGATGGCACTGAAGAAGTTTGTGATGGTCAAATTCCTCAATGACTCCATTGTGG ACCCTGTGGATTCTGAATGGTTTGGATTTTACAGGTCTGGCCAAGCAAAGGAAACCATCCCT CTTCAGGAAACAGCCTGTACACACAAGAGGCTGGGCCCTAAAGGAAATGGACAATGCTGG ACAATTGGTGTTTTTGGCTACTGAGGGTGACCACCTCCAGCTCTCAGAAGAATGGTTTTATG CTCACATCATCCCTTCTGGGTTGA
112	sp7_7-PPT1co6 ATGGCCTTCTTTGGCTACTGAGCTGCTGGGCCCTTCTGGGCACCACTTTTGGTGGCCCACC TGCTCCATTGCCCCTAGTCATTTGGCATGGGATGGGGGACAGCTGCTGCAATCCTTTATCTA TGGGAGCCATTAAAGAAGATGGTGGAGAAAAAGATCCCTGGCATATATGTGTTAAGCCTGGAG ATTGGAAAGACCCTCATGGAAGATGTTGAGAATTTCTTCTTCTGATGTAAACAGCCAAGT CACCACAGTGTGCCAAGCTCTGGCCAAGGACCCCAAGCTCCAGCAGGGTTATAATGCTATGG GATTCTCTCAGGGTGGGCAGTTCCTGAGGGCTGTGGCACAGAGATGCCCAAGCCCACCAATG ATCAACCTCATCAGTGTGGTGGACAGCATCAGGGAGTGTGGCCTGCCAGATGTCCAGG TGAGTCTCTCATATCTGTGACTTTATCAGGAAAACCTTAAATGCAGGGGCCTATTCCAAGG TGGTGCAGGAGAGACTGTCCAGGCTGAGTATTGGCATGATCCCATTAAGAAGATGTCTAT AGAAACCATTCCATCTTCTGGCTGATATCAACCAGGAAAGAGGGATCAATGAGAGCTACAA GAAGAATCTAATGGCACTGAAGAAGTTTGTGATGGTCAAATTCCTCAATGATAGTATTGTGG ATCCTGTAGACAGTGAAGTGGTTTGGCTTCTACAGAAGTGGACAAGCCAAAGAGACCATTCCC TTACAAGAGACCTCACTCTACACACAGGACAGACTGGGCCCTGAAGGAAATGGACAATGCAGG TCAGCTGGTGTCTTGTGAACAGAGGGGGATCACCTGCAGCTGTGAGAGGAATGGTTTTATG

	CCCACATCATCCCTTTCTCTGGGT TGA
113	sp7_7-PPT1co7 ATGGCTTTCTCTGGCTCTTGAGTTGCTGGGCACTGCTGGGCACCACCTTTGGAGGTCCCC TGCTCCATTGCTCTGGTCATCTGGCATGGCATGGGTGACTCCTGTTGCAACCCCTTAAGCA TGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATTCCTGGAATCTATGTCTTGTCTTTAGAG ATTGGTAAAACCCCTCATGGAAGATGTGGAGAATTCCTTCTTCTGAATGTGAACCTCACAAGT CACAAAGTCTGCCAGGCCCTGGCAAAGGACCCCAAGCTCCAGCAGGGCTACAATGCCATGG GCTTCTCCAGGGTGGGCAATTTCTGAGGGCTGTGGCCAGAGATGTCCCTCACCCCCCATG ATCAACTTGATATCAGTTGGAGGCCAGCACCAGGGAGTATTTGGGCTTCCAGGTGCCAGG AGAATCCAGCCATATCTGTGATTTTCATTAGAAAGACCCCTCAATGCTGGGGCTTACTCCAAGG TGGTGCAGGAGAGGTTAGTGCAAGCTGAATACTGGCATGATCCCATAAAGGAGGATGTTTAT AGGAACCACAGCATCTTCTTAGCAGATATTAATCAAGAAAGAGGTATCAATGAGTCCACAA GAAAACTCATGGCCCTGAAGAAGTTTGTTATGGTCAAGTTTCTCAATGACTCCATTGTGG ATCCTGTAGACTCTGAATGGTTTGGCTTTTACAGGAGTGGCCAAGCAAAAGAAACCATTCCT CTGCAGGAAACAAGTCTGTACACTCAGGACAGACTGGGCCCTGAAAGAAATGGATAATGCTGG CCAGTTAGTGTCTCTGGCCACTGAGGGGGACCACCTGCAACTCAGTGAGGAATGGTTCTATG CTCACATAATTCCCTTCTCTGGGTTGA
114	sp7_7-PPT1co8 ATGGCTTTCTCTGGCTGCTCTCTCTGTTGGGCTCTTCTTGGGACCACATTTGGGGGACCTCC TGCTCCTCTGCCCTGGTCATCTGGCATGGAATGGGAGACAGCTGTTGTAACCCACTCAGTA TGGGAGCCATAAAGAAGATGGTTGAGAAGAAGATTCCTGGCATATATGTACTCAGCTTGGAG ATTGGGAAGACCTTGATGGAAGATGTGGAGAATCATTCTTCTTGAATGTCAACAGCCAGGT CACCCTGTTTCCAGGCCCTTGCTAAGGACCCCAACTGCAACAGGGCTACAATGCCATGG GGTCTCTCAGGGAGGACAGTTCCTCAGAGCAGTGGCACAGAGGTGCCCCAGCCCACCCATG ATTAATCTGATCAGTGTGGAGGGCAGCACCAAGGGGTCTTTGGACTTCCAAGGTGCCCTGG GGAATCCAGCCATATTTGTGATTTTCATCAGAAAGACTCTGAATGCAGGGGCTATAGTAAAG TTGTTCAAGAAAGGCTGGTGCAGGCAGAATACTGGCATGATCCCATCAAAGAAGATGTATAC AGAAACCACAGCATTTTCTGGCAGATATAAACCAAGAAAGAGGCATTAATGAGTCCACAA GAAGAACCTCATGGCCCTGAAGAAATTTGTGATGGTGAAATTTTGAATGATTCTATAGTGG ACCTGTGGATAGTGTGAGTGGTTTGGTTTCTACAGGAGTGGGCAGGCCAAGGAGACCATTCCT TTGCAGGAGACCTCCCTGTACACACAGGACAGACTGGGCCCTAAAGGAAATGGACAATGCAGG CCAGCTAGTGTCTCTGGCCACTGAGGGGGACCACCTCCAGCTCTCTGAAGAATGGTTTATG CCCATATTATCCCTTTCTCTGGGTTGA
115	sp7_7-PPT1co9 ATGGCTTCTCTGGCTGCTGAGCTGTTGGGCACTGCTGGGCACCACCTTTGGGGGCCCTCC AGCTCCCTTGCCCTTTAGTCATTTGGCATGGAATGGGAGACTCCTGCTGCAATCCTCTCTCTA TGGGAGCCATAAAGAAGATGGTAGAGAAAAAATTCCTGGAATATATGTCTTAAGCCTGGAA ATTGGGAAGACACTCATGGAAGATGTGGAGAATCCTTCTTCTCAATGTGAACAGCCAGGT GACCACAGTCTGCCAGGCCTTGCCCAAAGACCCTAAGCTTCAGCAAGGCTACAATGCCATGG GCTTCAGCCAGGGTGGTCAATTCTGAGAGCAGTGGCCAGAGATGCCCCCACCACCTATG ATTAACCTCATCTCTGTGGTGGTCAGCACCAAGGTGTGTTGGTCTCCCTAGGTGCCCAGG AGAGAGTAGCCATATTTGTGACTTCATCAGGAAAACATTGAATGCTGGAGCCTACTCAAAAG TTGTACAAGAAAGGCTAGTGCAGGCTGAGTACTGGCATGATCCCATCAAGGAAGATGTGTAC AGAAACCACAGCATCTTCTGGCTGATATCAATCAGGAGAGGGGGATCAATGAAAGCTACAA GAAGAACCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCTCAATGACAGCATTGTTG ATCCTGTGATTCTGAATGGTTTGGATTCTACAGGTCTGGACAGGCCAAAGAGACCATCCCT CTTCAGGAAACCTCCCTTTACACACAGGACAGACTGGGGCTCAAGGAGATGGACAATGCTGG TCAGCTGGTTTTTCTGGCTACAGAGGGAGACCACCTACAGTTGAGTGAAGAGTGGTTCTATG CCCACATCATCCCTTTCTCTGGGTTGA
116	sp7_7-PPT1co10 ATGGCATTCCTCTGGCTACTCTCTTGCTGGGCTCTGCTGGGCACAACCTTTGGGGGCCACC AGTCCCTTCCCTGGTCATTTGGCATGGTATGGGAGACTCCTGTTGTAATCCTCTCTCCA TGGGAGCCATTAAGAAAATGGTGGAGAAGAAAATTCCTGGGATCTATGTCTCTCTAGAG ATTGGGAAGACCCTGATGGAAGATGTGGAGAACAGCTTTTTCTTAAATGTGAACAGCCAGGT GACAACTGTCTGTCAAGCATTGGCCAAGGACCCCAAGCTGCAGCAAGGCTACAATGCTATGG GATTCTCTCAGGGGGGCCAGTTCCTGAGGGCTGTGGCCAGAGATGCCCATCCCTCCCATG ATCAATCTGATATCTGTTGGGGGCCAGCACCAGGGAGTTTTTGGGCTGCCCAGATGCCCTGG GGAGAGCTCTCACATCTGTGATTTTCATCAGGAAAACCTCTCAATGCTGGAGCATATAGCAAGG

	TGGTGCAGGAGAGGTTGGTGCAGGCAGAGTACTGGCATGACCCCATCAAGGAGGATGTGTACAGAAACCCTCAATTTTCTGGCAGATATCAACCAAGAGAGGGGCATCAATGAGTCCTACAAGAAAACTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTTCTGAATGATTCTATTGTGGACCTGTGGATAGTGAGTGGTTTGGCTTTTACAGAAGTGGCCAAGCTAAGGAAACCATTCCCCTACAAGAAACTTCCCTGTACACCCAGGACAGGTTGGGACTCAAAGAGATGGACAATGCTGGACAGCTGGTGTTCCTGGCCACAGAAGGGGACCACCTGCAGCTGAGTGAAGAGTGGTTCTATGCTCATATCATCCCCCTTCCTTGGATGA TGA
117	sp7_PPT1co11 ATGGCCTTCTCTGGCTCCTGTCTGTTGGGCTCTCCTAGGCACCACTTTTGGGGGCCCCCTGCCCCCTTCCCTGGTCATTTGGCATGGGATGGGTGACTCCTGCTGCAACCCCTAAGCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATTCTGGAATTTATGTCTTAGCCTGGAAATTGGGAAGACCTTAATGGAAGATGTGGAAAACAGCTTTTTCTTAAATGTCAATTCTCAGGTAAACAACAGTGTGCCAAGCCCTGGCCAAGGACCCTAAACTGCAGCAGGGCTACAATGCTATGGCTTCAGCCAGGGTGGACAGTTTCTGAGGGCAGTGGCTCAGAGGTGCCCCAGTCTCCCATGATCAATCTCATCTCTGTGGGAGGCCAGCACCAGGGTGTTTTGGCCTGCCCAGGTGTCTGGAGAGTCCTCACATATCTGTGACTTCATCAGGAAGACACTGAATGCAGGGGCCTATAGCAAGGTGGTGCAGAGAGACTGGTTCAGGCTGAATACTGGCATGACCCCATCAAGGAAGATGTTTACAGGAACCATAGCATCTTCCCTTGCTGATATTAATCAAGAGAGGGGCATAAATGAGAGCTACAAAGAATTTAATGGCTCTAAAGAAATTTGTCATGGTGAAGTTCTGAATGACTCCATTGTGGATCCTGTGGACTCTGAGTGGTTTGGGTTCTACAGAAGTGGCCAGGCCAAGGAGACCATCCCCCTGCAGGAGACAAGCCTGTACACACAGGACAGGCTGGGACTGAAGGAGATGGACAATGCAGGACAGCTGGTGTTCCTGGCTACAGAGGGAGACCACCTGCAGCTCTCAGAGGAATGGTTCTATGCACACATCATTCCTTTTTAGGG TGA
118	sp7_7-PPT1co12 ATGGCCTTCTCTGTGGCTCTTGTCTGTTGGGCATTGCTAGGCACCACTTTTGGGGGACCTCTGCCCCCTTCTCTTTGGTTATCTGGCATGGGATGGGTGACTCCTGCTGCAATCCTTTAAGCATGGGGGCTATTAAGAAGATGGTGGAGAAAAAATTCCTGGAATCTATGTGCTTAGCTTAGAAATAGGGAAGACATTAATGGAAGATGTGGAAAACAGCTTTTTCTTGAATGTCAACTCACAGGTGACTACTGTCTGTGAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAAGGTTACAATGCTATGGATTTCAGCCAGGGTGGTCAGTTTCTGAGGGCTGTGGCACAGAGATGCCCCCTCACCTCCCATGATCAACCTCATCAGTGTGGGAGGACAGCACCAAGGTGTTTTGGTCTGCCCAGGTGTCTGGAGAGTCCTCACATATCTGTGACTTCATCAGGAAGACTCTGAATGCTGGGGCCTATAGCAAAGTAGTCCAGGAGAGGTTGGTGCAGGCTGAATACTGGCATGACCCCATTAAGAAGATGTCTATAGGAACCATAGCATATTCCCTTGAGATATCAACCAGGAGAGGGGGATCAATGAGAGCTACAAGAAAACTTGATGGCTCTCAAGAAGTTTGTGATGGTAAAGTTCTTAAATGACTCCATTGTGGATCCTGTAGACTCAGAGTGGTTTGGGTTCTACAGGTCTGGCCAAGCCAAAGAGACCATTCCCTGTCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGACAATGCTGGCCAGCTGGTGTTCCTGGCTACAGAGGGAGACCACCTGCAGCTCTCAGAAGAATGGTTCTATGCTCACATCATTCCTTTTTAGGG TGA
119	sp7_7-PPT1co13 ATGGCTTTCCTGTGGTTACTGAGCTGCTGGGCACCTGCTGGGCACCACTTTTGGAGGTCCCCCTGCTCCTCTGCCTTTGGTGATTTGGCATGGCATGGGTGACTCCTGCTGCAACCCCTCTGTCTATGGGTGCCATTAAGAAAAATGGTGGAAAAGAAGATCCCTGGAATTTATGTCTGTCTTTGGAGATAGGCAAGACTTTGATGGAAGATGTTGAGAATTCTTTTCTTAAATGTAAATAGCCAGGTAAACCACAGTTTGTGAGGCACTTGCCAAGGACCCCAAGCTGCAGCAAGGCTACAATGCCATGGGCTTCTCCAGGGAGGGCAGTTTCTGAGAGCAGTGGCTCAGAGGTGCCCCAGCCCTCCAATGATCAATTTGATCTCAGTGGGGGGCCAGCATCAGGGGGTCTTTGGCCTACCTAGGTGCCCTGGAGAGAGCTCTCATATCTGTGACTTCATTAGGAAGACATTAAATGCTGGGGCATACAGTAAGGTGGTGCAGGAGAGATTAGTGCAGGCTGAATACTGGCATGATCCCATCAAGGAAGATGTCTACAGGAACCACAGCATTTTCCCTGGCAGACATTAACCAAGAGAGGGGGATCAATGAGAGCTACAAGAAAACTTGATGGCCCTCAAGAAGTTTGTGATGGTGAAGTTTCTCAATGACAGCATTGTGGATCCAGTGGATTCTGAGTGGTTTGGGTTCTACAGAAGTGGCCAGGCCAAGAAACCATCCCAATTGCAGGAAACCAGTTTGTACACTCAGGACAGGTGGGGCCTGAAGGAGATGGACAATGCTGGCCAAGTGGTGTTCCTGGCCACTGAGGGGGACCACCTACAGCTGAGTGAGGAATGGTTCTATGCCACATCATCCCTTTTCTGGGC TGA
120	PPT1co14 ATGGCCTTCTTGTGGTTGCTCTCCTGCTGGGCCCTCTTGGGCACAACCTTTGGGGGACCCCCAGCTCCACTCCCCCTGGTGATCTGGCATGGCATGGGAGATAGCTGTTGCAACCCATTATCCA

	TGGGGGCCATTAAAGAAGATGGTGGAGAAGAAAATCCCTGGCATTATGTCTTAGCCTGGAA ATTGGAAGACCTTAATGGAGGATGTGGAAACTCCTTCTTCTCAATGTCAACTCCCAAGT CACCACAGTGTGTCAAGCCCTTGCCAAGGACCCCAAGTTGCAACAGGGCTATAATGCTATGG GCTTCTCTCAGGGAGGCCAGTTTCTGAGGGCTGTGGCCAGAGGTGCCCCCTCCCCACCTATG ATCAATCTTATCTCTGTGGGTGGACAACACCAAGGTGTCTTTGGGCTGCCAGGTGCCCTGG GGAGTCCAGCCACATTTGTGACTTCATCAGGAAGACACTAAATGCTGGTGCCTACTCCAAAG TTGTCCAGGAGAGACTGGTGCAAGCAGAATACTGGCATGACCCTATTAAAGAGGATGTCTAC AGGAACCACAGCATCTTTCTGGCAGACATCAACCAGGAGAGAGGGATCAATGAATCCTACAA GAAGAACCCTCATGGCCCTGAAGAAATTTGTGATGGTTAAGTTTCTCAATGACTCCATTGTGG ACCTGTGGATTCTGAATGGTTTGGATTTTACAGAAGTGGCCAGGCCAAGGAAACCATCCCT CTGCAGGAGACATCCCTGTACACACAGGACAGACTGGGCCTGAAGGAGATGGACAATGCTGG GCAGCTAGTGTCTCTGGCTACTGAGGGGGATCATCTTCAGTTGTCTGAAGAATGGTTCTATG CCCACATCATTCCCTTCTCTGGGCTGA
121	sp7_7-PPT1co15 ATGGCCTTTCTCTGGTTGCTTTCTCTGTTGGGCCCTCTTGGGGACAACCTTTGGGGGGCCCCC AGCTCCTCTTCCCCTGGTGATCTGGCATGGCATGGGAGACTCCTGTTGCAACCCATTAAAGCA TGGGGGCTATTAAAGAAAATGGTGGAGAAGAAAATCCCTGGAATTTATGTCTCTCTCTGGAG ATAGGCAAAACCTTAATGGAAGATGTGGAGAACTCCTTCTTCTCAATGTCAACTCCCAAGT CACCACCTGTATGTCAAGCTCTGGCAAAGGACCCCAAATTGCAACAGGGATATAATGCCATGG GATTCTCCCAGGGTGGCCAGTTCTGAGGGCTGTGGCACAGAGATGCCCTCCCCACCTATG ATCAATCTCATCTCTGTGGGTGGACAACATCAAGGTGTTTTTGGGCTGCCAGATGCCCTGG GGAGTCTCTCACATTTGTGACTTCATCAGGAAGACCCTCAATGCTGGAGCCTACTCTAAAG TTGTTTCAGGAAAGGCTGGTCCAAGCAGAATACTGGCATGACCCTATAAAAGAGGATGTTTAC AGGAACCACAGTATCTTCTTGGCAGACATCAACCAGGAGAGGGGCATCAATGAGTCTTACAA GAAAAACCTCATGGCACTCAAGAAATTTGTGATGGTCAAATTCCTGAATGACTCCATTGTGG ACCTGTGGATTCTGAGTGGTTTGGATTCTACAGAAGTGGCCAGGCCAAGGAAACCATCCCT CTGCAGGAGACAAGTCTGTACACCCAGGATAGGTTGGGTCTGAAGGAAATGGATAATGCTGG GCAGCTAGTCTTCTGGCCACAGAAGGGGACCACCTGCAGCTTTCTGAAGAGTGGTTCTATG CTCACATCATTCCCTTTCTGGGCTGA
122	sp7_7-PPT1co16 ATGGCCTTTCTCTGGCTCTTGTCTTGCTGGGCTCTGCTGGGCACCACCTTTGGTGGCCCCC TGCTCCTCTGCCCTGGTCATCTGGCATGGGATGGGAGACAGCTGCTGTAACCCCTTTATCAA TGGGAGCCATCAAGAAAATGGTGGAGAAGAAGATCCCTGGGATATATGTTCTTAGCTTGGA ATTGGCAAGACTCTCATGGAAGATGTTGAAAACAGCTTCTTCTGAATGTCAATAGCCAAGT GACCACAGTGTGCCAGGCATTGGCTAAAGACCCCAAGTTGCAGCAAGGCTACAATGCCATGG GCTTTTCCCAGGGTGGTCAGTTCTGAGGGCAGTGGCTCAGAGGTGCCCTTCACCACCTATG ATAAACCTCATCTCTGTGGGAGGGCAACACCAAGGAGTATTTGGACTACCCAGGTGCCCTGG GGAGTCCCTCCCATATTTGTGACTTTATCAGGAAGACACTGAATGCAGGAGCCTACAGCAAGG TGTTTCAGGAGAGACTTGTGCAGGCAGAGTACTGGCATGATCCCATTAAGGAAGATGTGTAC AGAAACCATAGTATTTTCTTGGCAGACATCAATCAGGAGAGGGGCATCAATGAGAGCTACAA GAAAAATCTGATGGCCCTAAAGAAGTTTGTCTGATGGTCAAGTTCTTGAATGATTCTATAGTGG ACCTGTGGATTCTAGAGTGGTTTGGTTTTTACAGATCAGGGCAAGCCAAGGAAACCATCCCA CTGCAGGAAAACAAGCCTGTACACTCAAGACAGGTTAGGCCATAAAGAGATGGACAATGCTGG CCAGCTTGTGTTCTTGGCAACAGAGGGGGATCATTTACAGTTATCTGAGGAATGGTTCTATG CCCACATCATCCCTTCTCTTGGTTGA
123	sp7_7-PPT1co17 ATGGCATTTCTGTGGTTGCTCAGCTGTTGGGCCCTCCTTGGAACCACCTTTGGAGGGCCCCC TGCCCCCTCCTCTTTGGTTATTTGGCATGGCATGGGGGACTCATGCTGTAACCCCTTTGAGCA TGGGGGCCATTAAAGAAAATGGTGGAGAAGAAGATCCCAGGTATTTATGTCTTGTCTCTAGAA ATAGGGAAGACTCTGATGGAAGATGTGGAGAATAGCTTCTTCTTAAATGTGAATTCACAGGT TACCACTGTGTGCCAGGCCCTGGCCAAAGACCCCAAGTTACAGCAGGGATACAATGCTATGG GATTTTCTCAGGGTGGTCAGTTCTGAGGGCAGTGGCCAGAGATGCCCAAGCCCTCCAATG ATTAACCTCATTAGTGTGGGAGGGCAGCATCAAGGGGTATTTGGCCTGCCAGGTGCCCTGG AGAGTCTCTCATATCTGTGACTTTATCAGGAAGACTTTAAATGCTGGGGCCTACAGCAAGG TGGTCCAGGAGAGGCTGGTGCAAGCAGAGTATTGGCATGATCCCATAAAGGAGGATGTGTAT AGGAACCACTCCATCTTCTTGGCAGACATCAACCAGGAGAGGGGGATCAATGAGTCTTACAA GAAAAATTTAATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTCTCAATGATTCTATTGTGG ACCTGTGGATTCTGAGTGGTTTGGCTTTTACAGGAGTGGCCAGGCCAAGGAAACCATCCCT

	CTGCAGGAACTTCTCTATACACACAGGATAGGCTGGGCCTGAAAGAGATGGACAATGCAGG GCAGCTGGTGTCTTCTTGCCACTGAGGGTGACCACCTTCAGCTATCTGAGGAATGGTTTTATG CTCACATCATCCCCCTTCTGGGG TGA
124	sp7_PPT1co18 ATGGCTTTCCTGTGGCTGCTTAGCTGCTGGGCCCTTCTTGGGACCACCTTTGGGGGTCCACC AGCTCCTCTGCCCTTGCTCATCTGGCATGGGATGGGGGACTCTTGTGTAAACCCTCTTAGCA TGGGAGCCATTAAAGAAGATGGTGGAGAAGAAAATACCTGGCATATATGTCTTGAGCCTGGAG ATTGGCAAGACCCTGATGGAGGATGTGGAGAACAGCTTCTTCTGAATGTCAATTCCCAGGT GACCACAGTGTGTGAGGCCCTGGCCAAGGACCCCAAGTTGCAGCAAGGCTACAATGCTATGG GCTTTTTCACAAGGTGGGCAGTTCCTGAGAGCAGTGGCACAGAGATGTCTTTCACCTCCAATG ATCAACCTTATCTCAGTGGGAGGCCAACACCAGGGTGTTTTTGGGTGCCCAGGTGTCTTG TGAGAGCAGCCATATCTGTGACTTTATCAGGAAGACACTGAATGCAGGGGCTTACAGCAAGG TTGTTCAAGAGAGGCTGGTGCAGGCTGAGTATTGGCATGATCCCATTAAGGAGGATGTGTAT AGGAACCACTCCATCTTTTTTGGCAGACATCAATCAGGAGAGGGGCATCAATGAGTCTTACAA GAAGAACCTGATGGCCCTCAAGAAGTTTGTGATGGTCAAGTTTCTGAATGATTCTATTGTGG ATCCAGTGGACTCTGAATGGTTTGGGTTCTACAGATCTGGCCAGGCCAAGGAGACCATCCCC CTTCAAGAGACCAGCCTGTATACCCAGGACAGATTAGGCCTGAAGGAGATGGACAATGCTGG CCAGCTTGTCTTCTGGCTACAGAGGGGACCACCTCCAGCTGAGTGAAGAGTGGTTTTATG CCCACATCATACCCCTTCTGGGT TGA
125	sp7_PPT1co19 ATGGCCTTCTGTGGCTGCTCTCCTGCTGGGCCCTGCTGGGCACCACCTTTGGAGGCCCCC AGCCCCCTCTGCCCTGGTCATCTGGCATGGCATGGGAGACAGCTGCTGCAACCCTCTCTCCA TGGGAGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAATCTATGTCTCAGCCTGGAG ATTGGGAAGACCCTGATGGAGGATGTGGAGAACTCCTTCTTCTGAATGTGAACAGCCAGGT GACCACAGTGTGCCAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAATGCCATGG GCTTCTCCAGGGAGGCCAGTTCCTGAGGGCTGTGGCCAGAGATGCCCCCTCACCTCCCATG ATCAACCTCATCTCTGTGGGAGGCCAGCACCAAGGGTCTTTGGCCTGCCCAGGTGTCTTG GGAGAGCAGCCACATCTGTGACTTCATCAGGAAGACCCTGAATGCTGGTGCCTACAGCAAGG TGGTGCAGGAGAGGCTGGTGCAGGCTGAATACTGGCATGACCCCATCAAGGAGGATGTGTAC AGGAACCACAGCATCTTCTGGCAGACATCAACCAGGAGAGAGGCATCAATGAGAGATACAA GAAGAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTCTCAATGACTCCATTGTGG ACCTGTGGACTCTGAGTGGTTTGGCTTCTACAGGAGTGGCCAGGCCAAGGAGACCATCCCC CTGCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGACAATGCTGG GCAGCTGGTGTCTTCTGGCCACAGAGGGAGACCACCTGCAGCTCTCAGAGGAGTGGTTTTATG CCCACATCATCCCCCTTCTGGGC TGA
126	spSPARC_S-PPT1co1 ATGAGGGCCTGGATCTTTTTTCTGCTTTGCCTGGCAGGAAGGGCCCTAGCAGGACCTCCTGC CCCCCTTGCTCTGGTGATCTGGCATGGCATGGGAGACAGCTGTTGCAACCCCTCAGCATGG GTGCCATTAAAGAAGATGGTAGAGAAGAAGATACCTGGTATCTATGTGCTCTCCCTAGAGATT GGGAAAACCTTTGATGGAGGATGTGGAAAATTCCTTCTTCTGAATGTCAATTCCCAGGTAC CACAGTGTGCCAGGCTCTGGCAAAGGACCCAAAGCTACAGCAGGGCTATAATGCCATGGGCT TCAGCCAGGGTGGCCAATTCTTGAGGGCAGTGGCTCAGAGGTGTCCATCACCCCATGATT AATCTAATATCTGTGGGTGGCCAACACCAAGGTGTCTTTGGTCTCCCCAGGTGTCTTGGGGA GTCTTCACACATATGTGACTTCATCAGGAAGACTCTAAATGCAGGGGCTTATAGCAAGGTGG TTCAAGAAAGGCTGGTTCAAGCTGAGTATTGGCATGACCCCATCAAGGAAGATGTGTACAGA AACCCTCAATTTTCTGGCTGACATCAACCAGGAGAGGGGTATCAATGAGAGTTACAAGAA GAACCTCATGGCACTGAAGAAATTTGTGATGGTGAATTCCTTAATGACTCCATTGTGGATC CTGTGGACTCAGAATGGTTTGGCTTCTACAGGAGTGGCCAAGCCAAAGAGACCATCCCCCTG CAGGAGACTAGCCTGTACACCCAGGACAGGCTTGGTCTGAAGGAAATGGACAATGCTGGGCA GCTGGTCTTTCTGGCCACAGAAGGGGACCACCTTCAATTATCAGAGGAGTGGTTCTATGCCC ACATCATACCTTTCTGGGC TGA
127	spSPARC_S-PPT1co2 ATGAGGGCCTGGATCTTCTTCTTCTTGTGCTGGCTGGCAGAGCTTTGGCAGGGCCCCCTGC CCCTCTTCCCTTGGTCATATGGCATGGGATGGGAGATTCTGTGCAACCCTCTGTCCATGG GAGCCATTAAAGAAGATGGTGGAAAAGAAGATCCCTGGAATTTATGTCTTAAGCCTGGAAATT GGGAAAACACTCATGGAAGATGTTGAAAATTCCTTCTTCTCAATGTCAATAGTCAGGTAAC TACAGTTTGTGAGGCTCTGGCAAAGGACCCCAAGCTGCAGCAGGGCTACAATGCCATGGGCT TCTCTCAGGGAGGACAATTCTTAAGGGCAGTAGCTCAGAGATGCCATCTCTCCAATGATC

	AACCTCATCTCTGTTGGAGGGCAGCACCAAGGGGTTTTTGGCCTCCCCAGATGTCTCGGAGA GAGCAGCCATATCTGTGACTTCATCAGGAAGACTCTAAATGCTGGAGCCTACAGTAAGGTTG TACAGGAGAGGCTTGTTTCAGGCTGAGTACTGGCATGACCCCATTAAGGAGGATGTGTATAGG AATCATAGCATATTCTTGGCTGATATAAACCAAGAAAGGGGAATCAATGAAAGCTACAAGAA GAACCTCATGGCCCTAAAGAAGTTTTGTGATGGTGAAGTTTTCTCAATGACTCCATTGTTGACC CTGTGGACTCTGAATGGTTTTGGCTTCTACAGGTCTGGCCAAGCCAAGGAGACAATTCTCTC CAGGAGACTTCTCTGTACACACAAGATAGGCTGGGACTAAAGGAAATGGACAATGCAGGTCA GCTGGTGTCTTGGCCACAGAGGGGGACCACCTGCAGCTTCTGAGGAGTGGTTCTATGCCC ATATTATACCCTTTCTTGGT TGA
128	spSPARC_S-PPT1co3 ATGAGAGCCTGGATCTTCTTCCTGCTTTGCCTTGCAGGCAGAGCCTTAGCAGGTCCCCCAGC TCCTCTCCCTCTTGTCTATCTGGCATGGAATGGGTGACAGCTGCTGCAATCCATTATCCATGG GAGCCATTAAGAAGATGGTGGAGAAGAAGATCCCAGGCATCTATGTATTGTCTTTGGAAATT GGGAAGACACTGATGGAGGATGTGGAGAACTCCTTCTTCTTAAATGTCAACTCACAGGTTAC CACAGTGTGCCAAGCCCTGGCCAAGGACCCCAAGCTGCAACAGGGTTACAATGCTATGGGGT TCTCTCAAGGTGGCCAGTTCTCAGGGCTGTGGCACAGAGGTGTCCCTCACCACCTATGATT AATCTCATCAGTGTGGAGGACAGCATCAAGGTGTCTTTGGCCTCCCAAGATGTCCAGGGGA AAGTTCTCACATCTGTGATTTCTATAAGGAAGACCCTCAATGCTGGAGCTTACAGTAAAGTGG TCCAGGAGAGGCTGGTCCAGGCTGAATACTGGCATGACCCATCAAGGAGGATGTTTATAGA AATCATTCCATCTTCTGGCAGACATCAACCAGGAGAGGGGCATCAATGAGAGCTACAAGAA AAATCTCATGGCCCTGAAGAAGTTTTGTGATGGTGAATTTCTAAATGATAGTATTGTGGACC CTGTAGACTCTGAGTGGTTTTGGCTTCTACAGGAGTGGGCAGGCCAAGGAAACCATTCCCTG CAGGAAACTAGCCTTTACACACAGGACAGGTTGGGCTTGAAGGAAATGGATAATGCTGGGCA GCTTGTCTTCTTGGCCACAGAGGGGGACCCTTGCAGCTCAGTGAGGAGTGGTTTTATGCCC ACATCATCCCTTCTTGGC TGA
129	spSPARC_PPT1co20 ATGAGGGCTGGATATTCTTCCTGCTTTGTCTAGCTGGAAGAGCTCTGGCAGGGCCCCCAGC ACCCCTGCCCTCTGGTCATCTGGCATGGGATGGGAGATTCTGTGCAACCCCTCTCAGCATGG GGGCCATTAAGAAAATGGTGGAGAAGAAGATCCCTGGCATTATGTCTTGTCTTGGAGATA GGCAAAACACTAATGGAAGATGTGGAGAACAGCTTCTTCTTAAATGTGAACAGCCAGGTCAC CACAGTATGTCAGGCCTTGGCCAAGGACCCCAAGCTCCAGCAGGGCTACAATGCCATGGGTT TTTCTCAGGGAGGTGAGTTCTCAGGGCTGTGGCTCAGAGATGCCCCCTCACCACCTATGATC AATCTGATCTCTGTGGGAGGGCAGCATCAAGGTGTCTTTGGTCTTCTAGGTGCCCTGGGGA GTCTCCACATCTGTGACTTTCATCAGGAAAACCTTTAAATGCTGGAGCCTACAGCAAGGTTG TCCAAGAGAGGCTTGTCCAAGCAGAGTACTGGCATGATCCCATCAAAGAGGATGTTTACAGA AACCACAGCATCTTCTGGCAGACATCAATCAAGAGAGGGGCATTAATGAGTCTTACAAGAA AAACCTCATGGCCCTAAAGAAATTTGTGATGGTGAATTTCTGAATGACTCCATTGTTGACC CTGTAGACTCTGAGTGGTTTTGGCTTCTATAGATCAGGCCAGGCCAAGGAAACCATTCTCTC CAAGAAACCTCCCTGTACACACAAGATAGGCTGGGGTTGAAAGAGATGGATAATGCTGGGCA ACTGGTGTCTTGGCCACTGAGGGTGACCATTACAGCTTTCAGAGGAATGGTTCTATGCCC ACATCATACCCTTTCTTGGAT TGA
130	spSPARC_S-PPT1co5 ATGAGAGCCTGGATCTTCTTTCTCCTTTGCCTTGCCTGGAAGAGCTTTGGCAGGGCCTCCAGC CCCCTTACCTTTGGTCATATGGCATGGAATGGGAGACAGTTGTTGCAACCCCTCTCTATGG GTGCTATTAAGAAAATGGTGGAGAAGAAGATACCTGGCATTATGTCTCTCCCTGGAATA GGCAAAACACTGATGGAAGATGTAGAAAACAGCTTTTTCTCAATGTCAATTCCCAGGTAAC AACTGTGTGCCAGGCCCTGGCCAAGGATCCCAAATTGCAGCAAGGTTATAATGCAATGGGAT TCTCCCAAGGAGGCCAATTTCTGAGGGCTGTGGCCAGAGGTGTCCATCTCCTCCCATGATC AACCTCATCAGTGTGGGGGGGCAGCACCAAGGAGTTTTTGGACTTCCAAGGTGCCCTGGGGA ATCCAGCCACATTTGTGACTTCATTAGAAAGACACTCAATGCTGGTGCTTATCCAAAGTTG TTCAGGAGAGGCTTGTGCAGGCTGAGTATTGGCATGACCCCATCAAGGAGGATGTGTACAGG AATCATTCAATCTTCTTGGCTGACATCAACCAGGAGAGAGGCATCAATGAGAGCTACAAGAA AAACCTCATGGCCCTAAAGAAGTTTTGTGATGGTGAAGTTCTTAAATGATAGCATAGTTGACC CAGTGGATTCTGAGTGGTTTTGGTTTCTATAGGAGTGGCCAAGCCAAGGAGACCATTCCCCTG CAGGAGACCTCCCTGTACACCCAGGACAGGCTTGGTCTCAAGGAGATGGACAATGCAGGACA GCTAGTTTTCTTGGCCACTGAGGGGATCACCTCCAACCTCAGTGAAGAATGGTTCTATGCAC ACATCATCCCTTCTTGGGC TGA

131	spSPARC_S-PPT1co6 ATGAGGGCCTGGATCTTCTTCTCCTGTGCCTGGCAGGGAGAGCCTTGGCTGGGCCCCCAGC ACCCCTTCTTTGGTGATATGGCATGGCATGGGTGATAGCTGTTGCAACCCCTGAGCATGG GGGCCATCAAGAAAATGGTGGAGAAAAAGATTCCAGGCATATATGTGCTCTCCCTGGAGATT GGAAAAACCTCATGGAAGATGTGGAGAACAGCTTCTTCTTAAATGTGAATAGTCAAGTTAC CACAGTGTGCCAGGCACCTGGCCAAGGACCCCAAGCTCCAACAGGGGTACAATGCCATGGGCT TCTCTCAGGGAGGGCAGTTTCTCAGAGCAGTGGCCAGAGATGCCCTTCAACCCCCATGATC AATCTCATTTCTGTTGGAGGGCAACACCAGGGGGTATTTGGTCTACCCAGGTGCCAGGTGA ATCCAGCCACATCTGTGACTTCATCAGGAAGACACTGAATGCAGGAGCCTACAGCAAAGTGG TGCAAGAGAGGCTGGTGCAGGCAGAGTACTGGCATGACCCCATCAAGGAGGATGTCTACAGA AACCATTCCATCTTCTGGCAGACATCAACCAGGAAAGAGGAATCAATGAGTCTACAAGAA GAACCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCTGAATGATTCCATTGTTGACC CTGTGGATTCTGAGTGGTTTGGATTTTACAGAAGTGGCCAGGCCAAGGAAACAATTCCACTA CAAGAGACCTCCCTCTACACCCAGGACAGGCTGGGCCTCAAAGAAATGGACAATGCTGGCCA GCTGGTGTCTTGGCTACAGAGGGAGACCATCTGCAGCTGTCAGAGGAGTGGTTCTATGCTC ACATCATACCTTTCTTAGGTGA
132	spSPARC_S-PPT1co7 ATGAGAGCCTGGATCTTCTTCTCCTTGTGTCTGGCTGGGAGGGCCTTGGCTGGCCCTCCTGC TCCCTGCCCTTTGGTTATCTGGCATGGAATGGGTGACAGCTGTTGCAACCCCTTAAGCATGG GAGCTATTAAGAAAATGGTTGAGAAGAAGATCCCTGGGATCTATGTCTGAGCTTGGAAATT GGGAAGACACTCATGGAAGATGTGGAGAACTCCTTCTTCTGAATGTGAATAGCCAAGTGAC TACAGTGTGTCAAGCCTTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAATGCCATGGGCT TCTCCCAGGGAGGACAGTTTCTGAGAGCAGTAGCTCAGAGGTGCCCTTCACTCCCATGATC AACCTCATCAGTGTGGGGGGCCAGCACCAGGGAGTGTGGGGTGGCCAGGTGCCCTGGGGA GTCATCTCACATCTGTGACTTCATCAGGAAAACCTCAATGCTGGGGCTACAGCAAGGTGG TCCAGGAGAGACTGGTGCAGGCAGAGTATTGGCATGACCCCATCAAGGAGGATGTCTATAGG AACCACAGCATTTTCTTAGCTGACATCAATCAAGAGAGAGGCATCAATGAGTCTACAAGAA GAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAAGTTCCTCAATGATTCTATTGTGGATC CTGTGGACTCAGAGTGGTTTGGCTTTTACAGATCTGGGCAAGCAAAGGAGACCATTCCCCTG CAGGAAACATCCCTGTACACCCAGGATAGGCTGGGTCTGAAAGAGATGGACAATGCTGGCCA GCTTGTTTTCTGGCTACAGAAGGGGACCATTACAGCTATCTGAAGAATGGTTTATGCTC ACATTATCCCTTTCTTGGATTGA
133	spSPARC_S-PPT1co8 ATGAGGGCCTGGATCTTCTTTCTGTTGTGTTTGGCAGGAAGGGCCTTGGCTGGGCCCCCTGC ACCTCTGCCCCCTGGTCATTTGGCATGGCATGGGTGACTCCTGTTGCAATCCCCTGAGCATGG GGGCCATCAAGAAGATGGTGGAGAAGAAGATCCCAGGAATCTATGTTCTGTCTTAGAAATT GGGAAGACTCTGATGGAAGATGTGGAAAACAGCTTCTTCTGAATGTGAATTTCCAGGTGAC CACAGTCTGCCAGGCTCTGGCTAAGGACCCCAAACTCCAGCAGGGCTACAATGCTATGGGAT TCTCCCAGGGTGGGCAGTTCTCAGAGCAGTGGCTCAGAGATGCCCATCAACCCCAATGATC AACCTCATCTCTGTTGGTGGTCAGCATCAAGGTGTCTTTGGACTCCCAGGTGCCCTGGGGA GTCCCTCTCACATCTGTGACTTTATCAGAAAAACCTTAAATGCTGGGGCTACAGCAAGGTGG TCCAGGAAAGGCTGGTCCAGGCTGAGTACTGGCATGATCCCATCAAGGAAGATGTGTACAGG AACCCTCCATTTTCTGGCTGACATAAACCAAGAGAGGGGCATCAATGAGTCATATAAGAA AAACCTGATGGCTCTGAAAAAGTTTGTGATGTTAAATTCCTCAATGACTCTATTGTGGACC CTGTTGACTCTGAGTGGTTTGGCTTCTACAGGTGAGGCCAAGCTAAAGAGACCATCCCCCTG CAGGAGACATCCCTGTACACACAGGACAGGTTAGGACTGAAGGAGATGGACAATGCTGGACA GCTGGTGTCTTGGCCACAGAGGGTGACCACCTACAGTTGTGAGAAGAATGGTTCTATGCTC ACATCATACCTTCTCTTGGTTGA
134	spSPARC_S-PPT1co9 ATGAGGGCCTGGATCTTCTTTCTCCTGTGCCTGGCTGGAAGGGCCTTGGCAGGCCCCACCTGC TCCCTTCCCTTGGTGATATGGCATGGCATGGGAGATAGCTGTTGTAATCCCCTCTCCATGG GTGCCATTAAGAAGATGGTGGAAAAGAAGATTCTTGAATCTATGTGCTGTCCCTGGAGATA GGGAAGACTCTGATGGAAGATGTGGAAAATTCCTTCTTCTGAATGTCAACTCCAGGTGAC AACTGTATGCCAGGCCTTGGCAAAGGACCCCAAGCTGCAGCAGGGATACAATGCCATGGGAT TCAGCCAAGGTGGACAGTTTCTGAGAGCAGTAGCCAGAGATGCCCCAGTCCCCCATGATC AACCTCATCTCTGTGGGTGGACAGCACCAGGGAGTGTGGCCCTGCCAGATGTCCAGGGGA GAGCAGCCACATTTGTGACTTCATCAGGAAAACACTGAATGCTGGGGCTACAGCAAGGTGG TGCAGGAGAGGCTAGTGCAGGCAGAGTACTGGCATGACCCAATCAAGAGGATGTTTATAGG

	AACCACAGTATCTTTTTGGCAGATATTAACCAGGAAAGAGGCATCAATGAGTCCTACAAGAA GAACCTAATGGCCTTGAAGAAATTTGTCATGGTGAAGTTCCTCAATGATTCCATTGTGGATC CTGTTGACTCAGAATGGTTTGGATTTTATAGATCAGGCCAGGCCAAAGAAACCATACCACTG CAGGAAACATCCCTCTACACTCAAGATAGGCTGGGGCTGAAGGAAATGGATAATGCTGGGCA GCTGGTATTCCTGGCTACTGAAGGGGATCACCTCCAGCTAAGTGAGGAGTGGTTCTATGCTC ACATCATCCCATTCTTGGCTGA
135	spSPARC_S-PPT1co10 ATGAGGGCCTGGATATTCTTTCTCCTCTGTCTGGCTGGGAGAGCCTTGGCAGGACCCCCCTGC CCCTTTGCCCCCTGGTCATTTGGCATGGGATGGGTGACAGCTGCTGCAATCCCCTAAGCATGG GAGCCATCAAGAAGATGGTGGAAAAGAAGATTCTTGGCATTTATGTCTCTCCCTGGAGATT GGGAAGACCCTGATGGAGGATGTGGAGAACAGCTTCTTCCTGAATGTGAATCCCAGGTTAC AACAGTTTGCCAGGCCTTAGCCAAGGACCCTAAACTCCAGCAAGGCTACAATGCAATGGGCT TCTCTCAGGGTGGGCAGTTTCTGAGGGCTGTGGCACAGAGGTGCCCCAGTCTCTCCCATGATC AACCTCATCTCAGTGGGAGGCCAGCACCAGGGTGT'TTTTGGGTTGCCAAGATGTCTGGAGA GAGCTCTCATATCTGTGACTTCATCAGGAAGACACTGAATGCTGGGGCTTATAGCAAGGTGG TACAGGAGAGGCTGGTGCAGGCTGAGTACTGGCATGACCCCATCAAGGAAGATGTGTACAGG AACCATTCAATCTTCTTGGCAGATATTAACCAGGAGAGGGGCATCAATGAAAGCTACAAGAA GAACCTGATGGCCTTGAAGAAATTTGTCATGGTAAAGTTTCTGAATGACTCCATTGTGGATC CTGTAGACTCTGAGTGGTTTGGCTTTTACAGAAGTGGCCAAGCCAAGGAAACCATCCCCCTG CAGGAGACATCTCTGTACACTCAGGACAGGTTGGGACTGAAGGAGATGGACAATGCAGGACA GCTGGTCTTCTGGCCACAGAGGGAGACCACCTGCAGCTCTCAGAAGAGTGGTTTTATGCC ACATCATTCCTTCTTAGGGTGA
136	spSPARC_S-PPT1co11 ATGAGAGCCTGGATATTCTTTCTCCTCTGTCTTGTCTGGAAGAGCCCTGGCTGGACCCCCCTGC ACCTCTGCCCCCTTGTTATATGGCATGGGATGGGAGACAGCTGCTGCAATCCCCTCAGCATGG GTGCCATCAAGAAGATGGTGGAGAAAAAATTCCTGGAATTTATGTCTCAGCTTGGAGATA GGGAAGACACTTATGGAAGATGTGGAAAACAGTTTCTTCTGAATGTGAATCCCAGGTGAC CACTGTGTGCCAGGCACTGGCCAAAGACCCTAAGCTGCAGCAAGGCTACAATGCTATGGGCT TCAGCCAGGGGGTCAAGTTTCTGAGAGCAGTGGCTCAGAGATGCCCTCACCTCCCATGATC AACCTGATCTCTGTGGGAGGCCAGCACCAAGGGGTTTTTGGGCTACCAAGGTGTCTGGAGA GTCTTCACATATCTGTGACTTCATCAGAAAGACCTTAAATGCAGGGGCCTATAGCAAAGTGG TTCAGGAGAGGCTTGTTTCAAGCTGAATACTGGCATGACCCCATCAAAGAGGATGTCTACAGG AACCATTCAATCTTCTTAGCTGACATCAATCAGGAGAGGGGGATAAATGAAAGCTACAAGAA AAATCTCATGGCCCTGAAGAAGTTTGTCTATGGTGAAGTTTCTCAATGATTCCATTGTGGATC CAGTGGATTGAGAGTGGTTTGGCTTTTACAGGTGAGGCCAAGCTAAAGAGACCATCCCCCTC CAGGAGACCTCCCTCTACACACAGGACAGACTGGGATTGAAAGAGATGGACAATGCTGGGCA GCTAGTTTTTTTTGGCCACAGAGGGAGACCACCTGCAGCTCAGTGAGGAATGGTTTTATGCTC ACATCATTCCTTCTTGGTGA
137	spSPARC_S-PPT1co12 ATGAGAGCCTGGATCTTCTTTCTTCTGTGCCTTGTCTGGGAGGGCCCTTGTCTGGGCCACCTGC TCCCCTGCCCCCTTGTCTATCTGGCATGGGATGGGAGACAGCTGCTGCAACCCCCCTGTCCATGG GGGCCATCAAGAAGATGGTGGAGAAGAAAATACCTGGAATATATGTCTCTCTCTGGAGATT GGCAAGACACTGATGGAAGATGTGGAGAATTCATTTTTCTTGAATGTCAACTCCCAAGTGAC CACAGTGTGCCAGGCCTTAGCCAAGGACCCCAAGCTTCAGCAGGGCTACAATGCAATGGGAT TCAGCCAGGGAGGGCAATTTCTGAGAGCAGTTGCTCAGAGATGCCATCTCCACCTATGATC AATTTGATCTCTGTGGGTGGACAGCATCAAGGTGTGTTTGGGCTACCCAGATGCCCTGGGGA GTCTAGCCACATCTGTGACTTCATCAGGAAGACTCTGAATGCTGGAGCCTACTCAAAGGTGG TGCAGGAGAGGCTGGTGAAGCTGAGTACTGGCATGATCCCATAAAGGAGGATGTGTACAGG AACCACAGCATCTTCTTAGCAGACATCAACCAAGAGAGAGGCATCAATGAGTCCTACAAGAA GAACCTCATGGCTCTGAAGAAGTTTGTGATGGTCAAGTTCTTAAATGATTCTATTGTGGATC CTGTGGACTCTGAATGGTTTGGATTTTACAGATCAGGCCAAGCCAAGGAGACCATCCCTCTG CAGGAGACCTCTCTGTACACACAAGACAGGCTGGGACTCAAGGAGATGGACAATGCTGGACA GCTTGTCTTTCTGGCCACAGAAGGGGACCACCTGCAGCTGTCTAGAAGAATGGTTCTATGCTC ACATCATACCATTCTTGGTGA
138	spSPARC_S-PPT1co13 ATGAGAGCCTGGATCTTCTTCTCTGTGTCTTGTCTGGCAGAGCCCTGGCAGGGCCCCCTGC ACCTCTCCCCCTTGTCTATCTGGCATGGGATGGGAGACTCCTGCTGTAACCCCTTATCAATGG GGGCCATCAAGAAGATGGTAGAGAAGAAAATTCAGGAATTTATGTCTCAGCTTGAAATA

	GGCAAGACTCTCATGGAAGATGTGGAAAACAGCTTCTTCCTGAATGTGAATAGCCAGGTCAC AACAGTGTGCCAGGCCTTAGCTAAAGATCCCAAATTGCAGCAGGGCTACAATGCCATGGGAT TTTCCCAGGGAGGGCAATTCTCAGGGCAGTGGCACAGAGATGCCCTAGCCCTCCAATGATC AACCTTATCAGTGTGGGTGGGCAGCATCAAGGGGTTTTTGGCTTGCCAGGTGCCCTGGGGA ATCTAGCCACATCTGTGACTTTATAAGGAAAACCCTGAATGCAGGAGCCTACTCAAAAGTGG TCCAGGAGAGGCTGGTGCAAGCTGAATATTGGCATGATCCCATCAAAGAGGATGTGTATAGG AACCACAGCATCTTTTTAGCTGACATCAATCAAGAAAGGGGTATCAATGAAAGCTACAAGAA AAATTTAATGGCTCTGAAGAAGTTTGTCTATGGTAAAGTTTCTCAATGACTCCATTGTGGACC CTGTGGATTCTGAGTGGTTTGGATTCTATAGAAGTGGACAAGCCAAAGAAACCATCCCCTG CAGGAGACCAGCCTCTACACACAAGACAGATTGGGACTCAAAGAGATGGACAATGCTGGACA GCTTGTTTTTCTGGCTACAGAGGGGTGACCATCTGCAGCTGAGTGAAGAATGGTTTTATGCTC ACATCATACCATTCTGGGG TGA
139	spSPARC_S-PPT1c014 ATGAGGGCTTGATCTTCTTCCTCCTGTGTCTTGCAGGGAGAGCCTTGGCTGGGGCCACCTGC CCCTCTGCCATTGGTGATTTGGCATGGCATGGGAGATAGCTGCTGCAATCCTCTCTCCATGG GGGCCATTAAGAAAATGGTGGAAAAAAGATTCCAGGGATTTATGTGTGTCTTTAGAGATT GGGAAGACTCTCATGGAAGATGTGGAGAACAGCTTCTTCCTGAATGTCAACTCACAGGTAC TACTGTCTGTGAGGCTTTAGCTAAGGACCCCAAACTCCAGCAGGGATACAATGCTATGGGCT TCTCTCAGGGAGGGCAATTTCTGAGAGCAGTGGCCAGAGGTGCCCCAGCCCTCCCATGATC AATCTTATCTCTGTGGGTGGACAGCACCAAGGTGTTTTTGGCTTGCCAGGTGTCTGGGGA GAGCTCTCACATCTGTGATTTTCATCAGGAAGACACTCAATGCAGGAGCTTACTCCAAAGTGG TCCAAGAGAGGCTGGTCCAAGCAGAATACTGGCATGATCCAATTAAAGAAGATGTTTATAGG AACCACAGCATCTTCCTGGCAGACATCAATCAGGAAAGAGGGATAAATGAGTCTTACAAGAA AACTTAATGGCCCTGAAGAAGTTTGTGATGGTCAAATTCCTTAATGACTCCATTGTGGACC CTGTAGATTTCAGAGTGGTTTGGTTTTTATAGAAGTGGCCAGGCTAAAGAGACAATTCCTCTG CAAGAGACCAGCCTCTACACCAAGACAGGCTAGGCCTGAAAGAAATGGACAATGCTGGACA ACTGGTCTTTCTTGCCACTGAAGGTGACCATCTACAGCTGTCAGAAGAATGGTCTATGCTC ACATCATCCCCTTCTGGGG TGA
140	spSPARC_S-PPT1c016 ATGAGGGCTTGATATTCTTCCTCCTCTGCCTGGCAGGGAGAGCCTTGGCAGGCCCTCCAGC CCCTCTGCCCCCTGGTCATCTGGCATGGCATGGGAGACAGCTGCTGTAACCCTCTGTCCATGG GAGCCATTAAGAAGATGGTGGAGAAGAAGATCCCTGGAATCTATGTCTGTCCCTGGAGATT GGGAAGACACTGATGGAGGATGTGGAGAATTCCTTCTTCTTAAATGTGAACAGCCAGGTAC CACAGTTTGCCAGGCATTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAATGCCATGGGCT TCTCCCAGGGAGGCCAGTTCTGAGGGCTGTGGCCAGAGGTGCCCCCTCACCTCCCATGATC AACCTGATCTCTGTGGGAGGCCAGCACCAAGGTGTGTTTGGCTGCCCAGGTGCCCTGGGGA GTCCAGCCACATCTGTGACTTCATCAGGAAGACTTTAAATGTGGAGCCTACAGCAAGGTGG TCCAGGAGAGGCTGGTGCAGGCAGAATACTGGCATGATCCCATCAAGGAGGATGTGTACAGG AACCACAGCATCTTCCTGGCAGACATCAACCAAGAGAGAGGGATCAATGAGAGCTACAAGAA GAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGACTCCATTGTGGACC CTGTGGATTCTGAGTGGTTTGGCTTTTACAGGTCTGGCCAGGCCAAGGAGACCATCCCTCTG CAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGACAATGCTGGGCA GCTGGTGTCTCTGGCCACAGAGGGAGACCATCTGCAGCTGAGTGAGGAGTGGTTTTATGCTC ACATCATCCCCTTCTGGGG TGA
Expression Cassettes	
141	Expression cassette (Sp7F) GGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGG AGGGGTTCGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATG TCGTGTAAGTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTC GCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGG TTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCT GGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGG CCTTGCGCTTAAGGAGCCCCCTTCGCTCTGCTTGAAGTGGAGGCTGGCCCTGGGCGCTGGGG CCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGTGCCTTTTCGATAAGTCTCTAG CCATTTAAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAAT GCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCGCGGGCGGCGACGGGGCCG TGCGTCCAGCGCACATGTTTCGGCGAGGCGGGGCTGCGAGCGCGGCCACCGAGAATCGGAC GGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCGGTGTATCGCC

	CCGCCCTGGGCGGCAAGGCTGGCCCCGTCGGCACCAAGTTGCGTGAGCGGAAAAGATGGCCGCT TCCCCGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTG AGTACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCAGGCACCTCGATTAGTTCTGGAGCTTTTGGAGTACGTGCTCTTT AGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAAG TTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCCCTTTTGGATTTGGATCT TGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTCCATTTCAGGTGTCGTG ACGCGTGCCACCATGGCCTTTCTGTGGCTGCTGTCCTGCTGGGCCCTGCTGGGGACCACCTT TGGCGGGCCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCA ATCCCTTAAGCATGGGTGCTATTAATAAATAATGGTGGAGAAGAAAATACCTGGAATTTACGTC TTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGT CAATTCCCAAGTAACAACAGTGTGTGTCAGGCACCTGCTAAGGATCCTAAATTGCAGCAAGGCT ACAATGCTATGGGATTCTCCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCT TCACCTCCCATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCC TCGATGCCCAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGG CGTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAG GAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAA TGAGTCCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATG ATTCCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAG GAAACCATTCCCTTACAGGAGACCTCCCTGTACACACAGGACCGCTGGGGCTAAAGGAAAT GGACAATGCAGGACAGCTAGTGTCTTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAG AATGGTTTTATGCCACATCATAACCATTCTTGGATAAAAGATCTAGAGCTGAATTCCTGCA GCCAGGGGATCAGCCTCTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCC CCTTGCCCTTCTTGACCCTGGAAGGTGCCACTCCACTGTCTTTCCTAATAAATGAGGAA ATTGCATCACATTGTCTGAGTAGGTGTCTATTCTATTTCTGGGGGGTGGGGTGGGGCAGGACAG CAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGG
142	Expression cassette (SPARC) GGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGG AGGGGTCCGCAATTGAACCGGTGCCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATG TCGTGTAATGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTC GCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGG TTCCCCGCGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCT GGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGG CCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGG CCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAG CCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAAT GCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCCGGGCGGCGACGGGGCCCG TGCGTCCCAGCGCACATGTTTCGGCGAGGCGGGGCCCTGCGAGCGCGGCCACCGAGAATCGGAC GGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCGCTGTATCGCC CCGCCCTGGGCGGCAAGGCTGGCCCGGTCCGACACAGTTGCGTGAGCGGAAAGATGGCCGCT TCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTG AGTACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCAGGCACCTCGATTAGTTCTGGAGCTTTTGGAGTACGTGCTCTTT AGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAAG TTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCCCTTTTGGATTTGGATCT TGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTCCATTTCAGGTGTCGTG ACGCGTGCCACCATGAGGGCCTGGATCTTCTTTCTCCTTTGCCGCGGGAGGGCCTTGGC AGGGCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATC CCTTAAGCATGGGTGCTATTAATAAATAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTA TCTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAA TTCCCAAGTAACAACAGTGTGTGTCAGGCACCTTCTTAAGGATCCTAAATGCAGCAAGGCTACA ATGCTATGGGATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCA CCTCCCATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCG ATGCCCAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGT ACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAG GATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGA GTCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATT CCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGCCAAGCCAAGGAA

	ACCATTCCCTTACAGGAGACCTCCCTGTACACACAGGACCGCTGGGGCTAAAGGAAATGGA CAATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAAT GGTTTTATGCCACATCATACCATTCCTTGGATAAAAGATCTAGAGCTGAATTCCTGCAGCC AGGGGATCAGCCTCTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCT TGCCTTCCTTGACCCTGGAAGGTGCCACTCCCAGTGTCTTTCCCTAATAAAATGAGGAAATT GCATCACATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGG
143	Expression cassette tPA GGTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGG AGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATG TCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTC GCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGG TTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCCTGAATTACTTCCACCT GGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTTCGAGG CCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGG CCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAG CCATTTAAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAAT GCGGGCCAAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCGCGGGCGCGGGGCCCG TGCGTCCCAGCGCACATGTTTCGGCGAGGCGGGGCTGCGAGCGCGGCCACCGAGAATCGGAC GGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCC CCGCCCTGGGCGGCAAGGCTGGCCCGGTTCGGCACCAAGTTGCGTGAGCGGAAAGATGGCCGCT TCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTG AGTCACCACACAAAGGAAAAGGGCCTTTCCGTCTCAGCCGTGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTGGAGCTTTTGAGGTACGTGCTCTTT AGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGGAGACTGAAG TTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCCCTTTTTGAGTTTGGATCT TGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTTCTTCCATTTAGGTGTGCTG ACGCGTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGC AGTCTTCGTTTTCGCCAGCGCCAGAGCCCTGAGCACCTGGACCCGCCGCGCCGCTGCCGT TGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAA AAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCT GATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTCAATCCCCAAGTAACAACAGTGTGTC AGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCAGGGA GGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATGATCAATCTGATCTC GGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTCTCACA TCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGGAACGC CTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCAT CTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCTGATGG CCCTGAAGAAGTTTGTGATGGTGAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTCTG GAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCTTACAGGAGACCTC CCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTTC TGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCA TTCTTGGATAAAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCTACTGTGC CTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCTTGCCTTCCTTGACCCTGGAAGGT GCCACTCCCAGTGTCTTTCCCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTG TCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATA GCAGGCATGCTGGGGATGCAGTGGGCTCTATGG
Viral Vector Nucleic Acid (signal sequence designation)	
144	Viral Vector Sp7-F CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCGCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGCCCAACTC CATCACTAGGGGTTCTGCGCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTTCGGCAATTGAACCGGTGCTTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCCCTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCCTTAAGGAGCCCTTCGCTCGTGCTTGAG

	TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGTGTCTTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGCCCGCGGGCGGCGACGGGGCCCGTGCCTCCAGCGCACATGTTTCGGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCTGCTCTGGTGCC TGGCCTCGCGCCCGCTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAAGGCCTTTCCGTCTC AGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGGAGTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGGTTCAATTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTCTTCCATTTTCAGGTGTCTGACGCGTGCCACCATGGCCTTTCTGTGGCTGCTGTCTC GCTGGGCCCTGCTGGGGACCACCTTTGGCGGGCCGCCGGCGCCGCTGCCGTTGGTGATCTGG CATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAAATGGTGGA GAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACG TGGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTGTGTGAGGCACTTGCT AAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCAGGGAGGCCAATTTCT GAGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGATCTCGGTTGGGGGAC AACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTCTCACATCTGTGACTTC ATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAAGAACGCTCGTGCAAGC CGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAG ATATAAATCAGGAGCGGGGTATCAATGAGTCTTACAAGAAAAACCTGATGGCCCTGAAGAAG TTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGG ATTTTACAGAAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAGGAGACCTCCCTGTACACAC AGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCTTGGCTACAGAA GGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCTTGGGATA AAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGATCAGCCTCTACTGTGCCTTCTAGTTGC CAGCCATCTGTTGTTTTGCCCTCCCCCTTGCCTTCCCTTGACCCTGGAAGGTGCCATCCCAC TGTCTTTTCTTAATAAAATGAGGAAATGTCATCACATTGTCTGAGTAGGTGTCATTCTATTTC TGGGGGTGGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCT GGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTACCTAGTCATG TAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTTAGG CAGTGGTGCCAGGTCACACCTCACTCATCAGGGGTGGCCTGACTCAGTGGTGATCTCTACT CACCTCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCTAAGGTGCTAATAGC AGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCATAAGGTAGGGA AATCCCTGAAGCTATCCTAATTAAAAGTAAAGTCTCTCCTAGAACTCTATCTAGTCCAGGTTA TTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAG GGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGTAGGGGGGTATGT AGGGATAACCTTACACTACCAGTAGTGACAGTGCACACTCACTTGGGCTGAGGCCCCAGGAAT AAGTAGATAGTGAGGGTCCCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCAC CATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCCCTCACTCCAGGG CCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTATAGGAGTGGGTACACAGATCCCTGTTC CACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACACAGATACTCAGCCAGCCC AGTGGTGTAAGTGGTGGCAGTTAGGTAGACAGCTTCAGGAACTGGTCAGCAAAGAATAGGG GAGCTTAAGTAGTGTTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCTCAG GGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGGT CTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAGGCCTCTTAATT CACACAAAAAACACACACAGATGTAATGAAAATAAAGATCCTTTATCTCGAGATCCACT AGGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCA CTGAGGCCGGCGACCAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGC GAGCGAGCGCGCAGCTGCCTGCAGG
145	Viral vector SPARC CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCTTGC GGCCGCGGGCTCCGGTGCCCGTCACTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCCGCAATTGAACCGGTGCCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCTTTTTTCCCGAGGGTGGGGG

	AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCCTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAAGTGGGTGGGAGAGTTTCGAGGCCCTTGCCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAAATTTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCCTCCAGCGCACATGTTTCGGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCCGCTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTTCGGCACCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAAGGGCCTTTCCGTCCTC AGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGGAGTACGTGCTCTTTAGGTTGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGTTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTTCAGGTGTCTGACGCGTGCCACCATGAGGGCCTGGATCTTCTTTCTCC TTTGCCTGGCCGGGAGGGCCTTGGCAGGGCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCAT GGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAGAA GAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGG AGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTGTGTCAGGCACTTCTAAG GATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCAGGGAGGCCAATTTCTGAG GGCAGTGGCTCAGAGATGCCCTTCACCTCCCATGATCAATCTGATCTCGGTTGGGGGACAAC ATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTCTCACATCTGTGACTTCATC CGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGCCGA ATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATA TAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTT GTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGGATT TTACAGAAGTGGCCAAGCCAAGGAAACCATTCCTTTACAGGAGACCTCCCTGTACACACAGG ACCGCTTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTTCCTGGCTACAGAAGGG GACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCTTGGATAAAA GATCTAGAGCTGAATTCCTGCAGCCAGGGGATCAGCCTCTACTGTGCCTCTAGTTGCCAG CCATCTGTGTTTTGCCCTCCCCCTTGCCCTTCTTGACCCTGGAAGTGCCACTCCCAGTGT CCTTTCCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTCATTCTATTCTGG GGGGTGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGG GATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTCACCTAGTCATGTAA GTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTTAGGCAG TGGTGCCAGGTCACACCTCACTCATCAGGGGGTGGCCTGACTCAGTGGTGATCTCTACTCAC CTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCTAAGGTGCTAATAGCAGG GTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCACTAAGGTAGGGAAAT CCCTGAAGCTATCCTAATTAAGTAAAGTAAAGTCTCTCCTAGAATCTATCTAGTCCAGGTTATTC CACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAGGGT AGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGTAGGGGGGTATGTAGG GATACCTTACACTACCAGATAGTGGTGACCACTCACTTGGGCTGAGGCCCCAGGAATAAG TAGATAGTGAGGGTGCCCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCACCAT ATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCCCTCACTCCAGGGCCA GGTAAGAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTGGGTACAGATCCCTGTTCCAC AGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAGATACTCAGCCAGCCAGT GGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAAGTGGTCAGCAAAGAATAGGGGAG CTTAAGTAGTGTTCAGCAGCACCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCTCAGGGC TCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGGTCTC CAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAAGCCTCTTAATTCAC AAAAAAACCAACACACAGATGTAATGAAAATAAAGATCCTTTATTCTCGAGATCCACTAGG GCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTG AGGCCGGCGACCAAAGGTCGCCCGACGCCGGGCTTTGCCGGGCGGCCCTCAGTGAGCGAG CGAGCGCGCAGCTGCCTGCAGG
146	Viral Vector tPA CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCCG

	GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCTCTGCGGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCCGCAATTGAACCGGTGCCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTATGGCCCT TGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTT TTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGGCGAGGCGGGGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTGGCACACG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCACACAAAGGAAAAGGCCCTTTCGCTCCTC AGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGGAGTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACTGAGTGGGTGGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTGTAGTTTGGATCTTGGTTTATTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTCTTCCATTTCAAGTGTCTGACGCGTGCCACCATGGATGCAATGAAGAGAGGGCTCT CTGTGTGCTGCTGTGTGGAGCAGTCTTCTGTTTCGCCCAGCGCCAGAGCCCTGCGACAC CTGGACCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAA TCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCT TATCTTTAGAGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTC AATTTCCCAAGTAACAACAGTGTGTCAAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTA CAATGCTATGGGATTCTCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTT CACCTCCCATGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCT CGATGCCCCAGGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGC GTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGG AGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAAT GAGTCTTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAATTTCTCAATGA TTCCATTGTGGACCTGTAGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGG AAACCATTTCCCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATG GACAAATGCAGGACAGCTAGTGTCTTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGA ATGGTTTTATGCCCACATCATACCATTCCTTGGATAAAAGATCTAGAGCTGAATTCCTGCAG CCAGGGGGATCAGCCTCTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCC CTTGCCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCTTAATAAAATGAGGAAA TTGCATCACATTGTCTGAGTAGGTGTATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGC AAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGGGGTG GTCAATCTAGATCCCATCAAAGGTCACCTAGTCAATGTAAGTGAGCCACTATCTCCTCCTACA AGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTTAGGCAGTGGTGCCAGGTACACCTCACTC ATCAGGGGGTGGCCTGACTCAGTGGTGATCTCTACTCACCTCCTCAGGCCCTCATCATTTAG CCTGTAAGACCCAGCAGGGCCTAAGGTGCTAATAGCAGGGTCCACTCACATCATGATGACTCC TTCCCCCTGGTGGAGCTCCTGCACTAAGGTAGGGAAATCCCTGAAGCTATCCTAATTAAAA GTAAAGTCTCTCCTAGAATCTATCTAGTCCAGGTTATTCCACTATATATCCAGTGAGAAAGTG GGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAGGGTAGTGCTGCTGTAGCCTCACCTGC ACAGGTGATTACCCACTCCCCAGTAGGGGGGTATGTAGGGATACCCCTACACTACCAGATAG TGGTGCACCACTCACTTGGGCTGAGGCCCCAGGAATAAGTAGATAGTGAAGGTGCCCCTAGT GCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCACCATATCTAATTAATTGCTGTAAAGCA GGAACACCCCTATGTGAAGAGCCCCCTCACTCCAGGGCCAGGTAGAAGGGAGGGCTCCTATAG AGGTTTTCTTAGGAGTGGGTACAGATCCCTGTTCACAGATTATATTAAGTCCATCAGGT GCTCAGTATCAGGGGTACAGATACTCAGCCAGCCCAGTGGTGTACTAGCTGGGCAGTTAGG TAGACAGCTTCAGGAAGTGGTCAGCAAGAATAAGGGAGCTTAAGTAGTGTTCAGCAGCACC CTCCCATCTAGGTGTCTCTCTCAGTGACCCCTCAGGGCTCCTTAAGAACTCCACAGAATA TAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGTCTCCAGAGGCACCTAATTACTCCAAT TGGGAGGGAATGAGTATGGTGTCAAGCCTCTTAATTACACAAAAAACCAACACACAGATGT AATGAAAATAAAGATCCTTTATTCTCGAGATCCACTAGGGCCGCGAGGAACCCCTAGTGATGG
--	--

	AGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCC CGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
147	Viral Vector NGF-1 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGGCGGCAAGCCCGGGCGTCCG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGAGTGGCCAACTC CATCACTAGGGGTTCTCTGCGGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCCTTTTCCCGAGGGTGCGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAAGTGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCGCGCGCTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTCCGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCCGCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCCGCACCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTTCCGTCTC AGCCGTGCTTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGAGTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGTTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTTCAGGTGTGCTGACGCGTGCCACCATGTCCATGTTGTTCTACACTCTGA TCACAGCTTTTCTGATCGGCATACAGGCGGGGCGCGCGCGCGCTGCCGTGGTGATCTGG CATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAATAAATGGTGGA GAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACG TGGAGAACAGCTTCTTCTTGAATGTCAATTCCTAAGTAACAACAGTGTGTGAGGCACTTGCT AAGGATCCTAAATTCAGCAAGGCTACAATGCTATGGGATTCTCCAGGGAGGCCAATTTCT GAGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGATCTCGGTTGGGGGAC AACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGGAGAGAGCTCTCACATCTGTGACTTC ATCCGAAAAACACTGAATGCTGGGGGCTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGC CGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGCCAG ATATAAATCAGGAGCGGGGTATCAATGAGTCTTACAAGAAAAACCTGATGGCCCTGAAGAAG TTTGTGATGGTGAAATTCTCAATGATTCCATTGTGGACCCTGTAGATTCCGAGTGGTTTTGG ATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCTTACAGGAGACCTCCCTGTACACAC AGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCTGGCTACAGAA GGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCCCTGGGATA AAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCTACTGTGCCCTTCTAGTTGC CAGCCATCTGTTGTTTGGCCCTCCCCCTTGCCCTTCCTTGACCCTGGAAGGTGCCACTCCCAC TGTCCTTTTCTAATAAATGAGGAAATGCATCACATTGTCTGAGTAGGTGTCATTCTATTTC TGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT GGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTCACCTAGTCATG TAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTTAGG CAGTGGTGCCAGGTACACCTCACTCATCAGGGGGTGGCCTGACTCAGTGGTGATCTCTACT CACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCTAAGGTGCTAATAGC AGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCATAAGGTAGGGA AATCCCTGAAGCTATCCTAATTAAGTAAGTCTCTCCTAGAAATCTATCTAGTCCAGGTTA TTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAG GGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGTAGGGGGGTATGT AGGGATACCCTTACACTACCAGATAGTGGTGCACCACTCACTTGGGCTGAGGCCCCAGGAAT AAGTAGATAGTGAGGGTGCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCAC CATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCCCTCACTCCAGGG CCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTGGGTACAGATACTCAGCCGCC CACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAGATACTCAGCCGCC AGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAACCTGGTCAGCAAGAATAGGG GAGCTTAAGTAGTGTTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCTCAG

	GGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGGT CTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAAGCCTCTTAATT CACACAAAAACCAACACACAGATGTAATGAAAATAAAGATCCTTTATCTCGAGATCCACT AGGGCCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCA CTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGC GAGCGAGCGCGCAGCTGCCTGCAGG
148	Viral Vector NGF-2 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCCG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCTCTGCGGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTTCGAGGCCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTCCGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGCGAGGCGGGGCCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCTGCTCTGGTGCC TGCCCTCGCGCCCGCTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCCGCGCACCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAAGGGCCTTTCCGTCTC AGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGAGTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGGTTCAATTCTCAAGCCTCAGACAGTGTTCAAAGTT TTTTCTTCCATTTACAGGTGTCTGACGCGTGCCACCATGTCCATGTTGTTCTACACTCTGA TCACAGCTTTTCTGATCGGCATACAGGCGGTGCCGCCGGCGCCGCTGCCGTTGGTGATCTGG CATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAAATGGTGGA GAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGGACG TGGAGAACAGCTTCTTCTGAATGTCAATTCCCAAGTAACAACAGTGTGTGAGGCACTTGCT AAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCAGGGAGGCCAATTTCT GAGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGATCTCGGTTGGGGGAC AACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTCTCACATCTGTGACTTC ATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTACGGAACGCCTCGTGCAAGC CGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAG ATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCTGATGGCCCTGAAGAAG TTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTCCGAGTGTTTGG ATTTTACAGAAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAGGAGACCTCCCTGTACACAC AGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCTGGCTACAGAA GGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCCTTGATA AAAGATCTAGAGCTGAATTCTTGCAGCCAGGGGATCAGCCTCTACTGTGCCTTCTAGTTGC CAGCCATCTGTGTTTGCCCCCTCCCCCTTGCCCTTCTTGACCTGGAAGGTGCCACTCCAC TGTCCTTTTCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTCAATTCTATTC TGGGGGTGGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCT GGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTACCTAGTCATG TAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTTAGG CAGTGGTGCCAGGTCACACCTCACTCATCAGGGGGTGGCCTGACTCAGTGGTGATCTCTACT CACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCTAAGGTGCTAATAGC AGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCCTAAGGTAGGGA AATCCCTGAAGCTATCCTAATTAAAAGTAAAGTCTCTCCTAGAATCTATCTAGTCCAGGTTA TTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAG GGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGTAGGGGGGTATGT AGGGATAACCCTTACACTACAGATAGTGGTGCACCACTCACTTGGGCTGAGGCCCCAGGAAT AAGTAGTAGTGAGGGTGCCCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCAC CATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCCCTCACTCCAGGG

	CCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTGGGTCACAGATCCCTGTTCCACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTCACAGATACTCAGCCAGCCCAGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAAGTGGTCAGCAAAGAATAGGGGAGCTTAAGTAGTGTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCTCAGGGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGGTCTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAAGCCTCTTAATTACACAAAAAACCAACACACAGATGTAATGAAAATAAAGATCCCTTTATTCTCGAGATCCACTAGGGCCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCATGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGCGCAGCTGCCTGCAGG
149	Viral Vector Insulin CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGGCAAATCATCACTAGGGGTTCTGCGGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTGCGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTGCGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCGCGCTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTGCGCACCA GTTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGCCTTTCCGTCCCTC AGCCGTGCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGAGTACGTGCTCTTAGGTTGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTGTAGTTTGGATCTTGGTTTATTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTCTTCCATTTTCAGTGTCGTGACGCGTGCCACCATGGCCCTGTGGATGCGCCTGTGTC CCGTGTGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCCGCGAGCCGTGCCCGCGCGCCG CTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGC TATTAATAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGA AGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACA GTGTGTCAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTC CCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATGATCAATC TGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCGAGGAGAGC TCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTCA GGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACC ACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCTTACAAGAAAAAC CTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGT AGATTTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCTTACAGG AGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTA GTGTTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACAT CATAACCATTCCTTGGATAAAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCT ACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCTTGCCCTTCTTGACCCT GGAAGGTGCCACTCCCCTGTCCTTTCTTAATAAAATGAGGAAATGCATCACATTGTCTGA GTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAA GACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATC AAAGGTCACCTAGTCATGTAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGT TGGTGAAGTCTAGTTAGGCAGTGGTGCCAGGTCACACCTCACTCATCAGGGGGTGGCCTGAC TCAGTGGTGATCTCTACTCACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCGAGG GCCTAAGGTGCTAATAGCAGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGCTGGAGCT CCTGCACTAAGGTAGGGAAATCCCTGAAGCTATCCTAATTAAGTAAGTCTCTCCTAGAA TCTATCTAGTCCAGGTTATTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTC

	CACCACCTGCCTGGTTAGGGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTC CCCAGTAGGGGGGTATGTAGGGATACCCTTACACTACCAGATAGTGGTGCACCACTCACTTG GGCTGAGGCCCCAGGAATAAGTAGATAGTGAGGGTGCCCCCTAGTGCTCCTTCAGCTCAGGGC TTAGCTAGGCTGGAGCACCATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAG AGCCCCCTCACTCCAGGGCCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTG GGTCACAGATCCCTGTTCCACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTCA CAGATACTCAGCCAGCCAGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAACT GGTCAGCAAAGAATAGGGGAGCTTAAGTAGTGTTTACAGCAGCACCCCTCCCATCTAGGTGTCTC CTCTCAGTGACCCCTCAGGGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCT GCTAAGCACTTAGGGGGTCTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATG GTGTCAAGCCTCTTAATTCACACAAAAAACCACACACAGATGTAATGAAAAATAAAGATCCT TTATTCTCGAGATCCACTAGGGCCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTC TGCGCGCTCGCTCGCTCACTGAGGCCGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCC CGGGCGGCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
150	Viral Vector CD33 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCAAAGCCCCGGGCGTCGG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTTCTGCGGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGAACCGGTGCTTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCCCTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCGCGCTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCCGTGGCACCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTTCCGTCTCTC AGCCGTGCTTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTTGGAGTACGTCTGCTTTTAGGTTGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGGTTTCAATCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTAGGTGTCTGACGCGTGCCACCATGCCGCTGCTGCTACTGCTGCCCC TGCTGTGGGCAGGGGCCAGGGCTGGGCCGCCGGCGCCGCTGCCGTTGGTGATCTGGCATGGG ATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAATGGTGGAGAAGAA AATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCGTATGGAGGACGTGGAGA ACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTGTGTCAAGGCACTTGCTAAGGAT CCTAAATTCAGCAAGGCTACAATGCTATGGGATTTCTCCAGGGAGGCCAATTTCTGAGGGC AGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGATCTCGGTGGGGGACAACATC AAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGCTCTCACATCTGTGACTTCACTCCGA AAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTACAGGAACGCCCTCGTGCAAGCCGAATA CTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAGATATAA ATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCTGATGGCCCTGAAGAAGTTTGTG ATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTCCGAGTGGTTTGGATTTTA CAGAAGTGGCCAAGCCAAGGAAACCATTCCTTACAGGAGACCTCCCTGTACACACAGGACC GCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCTTCTGGCTACAGAAGGGGAC CATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCCCTTGGATAAAAGAT CTAGAGCTGAATTCTTGCAGCCAGGGGGATCAGCCTCTACTGTGCCCTTCTAGTTGCCAGCCA TCTGTTGTTTGGCCCTCCCCCTTGCCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTCT TTCCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTCATTCTATTCTGGGGG GTGGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGAGT GCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTCACTAGCTAGGTAGTAAGTG AGCCACTATCTCTCTACAAGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTTAGGCAGTGG TGCCAGGTCACACCTCACTCATCAGGGGTGGCCTGACTCAGTGGTGATCTCTACTCACCTC

	CTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCCTAAGGTGCTAATAGCAGGGTC CACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCACTAAGGTAGGGAAATCCC TGAAGCTATCCTAATTAAGTAAAGTCTCTCCTAGAATCTATCTAGTCCAGGTTATTCCAC TATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAGGGTAGT GCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGTAGGGGGGTATGTAGGGAT ACCCCTACACTACCAGATAGTGGTGCACCACCTCACTTGGGCTGAGGCCCCAGGAATAAGTAG ATAGTGAGGGTGCCCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCACCATATC TAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCCCTCACTCCAGGGCCAGGT AGAAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTGGGTACAGATCCCTGTTCCACAGA TTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAGATACTCAGCCAGCCCAGTGGT GTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAACCTGGTCAGCAAAGAATAGGGGAGCTT AAGTAGTGTTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCTCAGGGCTCC TCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGGTCTCCAG AGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAAGCCTCTTAATTCACACA AAAAACCAACACACAGATGTAATGAAAATAAAGATCCTTTATTCTCGAGATCCACTAGGGCC GCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGG CCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGA GCGCGCAGCTGCCTGCAGG
151	Viral Vector IL34 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCCG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCTTGCGGCCGCGCTCCGGTGCCGTCAGTGGGACAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCCGCAATTGAACCGGTGCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAGTAGTCCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCCCTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTCCGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGCGAGGCGGGGCCCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGCCCTCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTTCGCGACCCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAAGGGCCTTTCCGTCTC AGCCGTGCGTTCATGTGACTCCACGGAGTACCGGGCGCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGAGTACGTCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGGAGACTGAAGTTAGGCCAGCTTGGCACCTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGGTTCAATTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTTCAGGTGTCTGTGACGCGTGCCACCATGCCCCGGGGCTTCACCTGGCTGC GCTATCTTCTGCTGGGGATCTTCCTTGCGTGGCCAGGGGGGGGCCCGCGCGCCGCTGCCG TTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAA AAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCC TGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTGTGT CAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCCAGGG AGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCATGATCAATCTGATCT CGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTCTCAC ATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGGAACG CCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCA TCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCTTACAAGAAAAACCTGATG GCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGATTCT GGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCTTACAGGAGACCT CCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCT CTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACC ATTCCTTGATAAAAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCTACTGTG CCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCTTGCCCTTCTTGACCTTGAAGG TGCCACTCCCCTGTCTTTTCTAATAAATGAGGAAATGCATCACATTGTCTGAGTAGGT

	GTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAAT AGCAGGCATGCTGGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGT CACCTAGTCATGTAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGGTGA AGTCTAGTTAGGCAGTGGTGCCAGGTACACCTCACTCATCAGGGGGTGGCCTGACTCAGTG GTGATCTCTACTCACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCCTAA GGTGCTAATAGCAGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCA CTAAGGTAGGGAAATCCCTGAAGCTATCCTAATTAAAAGTAAAGTCTCTCCTAGAATCTATC TAGTCCAGGTTATTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCAC CTGCCTGGTTAGGGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGT AGGGGGGTATGTAGGGATACCCCTTACACTACCAGATAGTGGTGCACCCTCACTTGGGCTGA GGCCCCAGGAATAAGTAGATAGTGAGGGTGCCCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCT AGGCTGGAGCACCATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCC CTCACTCCAGGGCCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTATAGGAGTGGGTAC AGATCCCTGTTCCACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAGATA CTCAGCCAGCCAGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAACTGGTCAG CAAAGAATAGGGGAGCTTAAGTAGTGTTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCA GTGACCCCTCAGGGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAG CACTTAGGGGGTCTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCA AGCCTCTTAATTCACACAAAAACCAACACACAGATGTAATGAAAAATAAGATCCTTTATTC TCGAGATCCACTAGGGCCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCG CTCGCTCGCTCACTGAGGCCGGGCGACCAAGGTGCCCCGACGCCCCGGGCTTTGCCCGGGCG GCCTCAGTGAGCGAGCGAGCGCAGCTGCCTGCAGG
152	Viral Vector OSM CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCGG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCTTGC GGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTGCGCAATTGAACCGGTGCCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCTCCGAGGGTGGGGG AGAACCGTATATAAGTGCAAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCGGTGCGAATCTGGTGGCACCTTCGCGCCTGT TCTCGCTGCTTTTCGATAAGTCTCTAGCCATTTTAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCGCGCTGTATCGCCCCGCCCTGGGCGGCAAGGTGGCCCGGTGGCCACCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCC GGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAAGGGCCTTTCCGTCTCTC AGCCGTGCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGAGTACGTGCTCTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGATCTTGTTTCACTCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTTCAGGTGTCGTGACGCGTGCCACCATGGGGGTACTGCTCACACAGAGGA CGTGCTCAGTCTGGTCCTTGCACTCCTGTTTCCAGACGCTAGAGCCCTGCCGTTGGTGATC TGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAAAAAATGGT GGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGACCCTGATGGAGG ACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTGTGTACAGGCACTT GCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCAGGGAGGCCAATT TCTGAGGCACTGGCTCAGAGATGCCCTTCACCTCCCATGATCAATCTGATCTCGGTTGGGG GACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCAGGAGAGAGCTCTCACATCTGTGAC TTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCA AGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTG CAGATATAAATCAGGAGCGGGGTATCAATGATTCCTACAAGAAAAACCTGATGGCCCTGAAG AAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCTGTAGATTGCGGAGTGGTT TGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAGGAGACCTCCCTGTACA CACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTCTTGCTACA

	GAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCTTGG AAAAAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCTACTGTGCCTTCTAGT TGCCAGCCATCTGTTGTTTGGCCCTCCCCCTTGCCCTTCCTTGACCTGGAAGGTGCCACTCC CACTGTCTCTTCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTCATTCTA TTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT GCTGGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTCACCTAGTC ATGTAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGGTGAAGTCTAGTT AGGCAGTGGTGCCAGGTACACCTCACTCATCAGGGGGTGGCCTGACTCAGTGGTGATCTCT ACTCACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCTAAGGTGCTAAT AGCAGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCCTAAGGTAG GGAAATCCCTGAAGCTATCCTAATTAAAGTAAAGTCTCTCCTAGAATCTATCTAGTCCAGG TTATTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGT TAGGGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCCAGTAGGGGGGTA TGTAGGGATACCCTTACACTACCAGATAGTGGTGCACCACTCACTTGGGCTGAGGCCCCAGG AATAAGTAGATAGTGAGGGTGGCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAG CACCATATCTAATTAATTGCTGTAAAGCAGGAACACCCCTATGTGAAGAGCCCCCTCACTCCA GGCCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTGGGTACAGATCCCTG TTCCACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAGATACTCAGCCAG CCCAGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAAGTGGTCAGCAAAGAATA GGGGAGCTTAAGTAGTGTTTCAAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCT CAGGGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGG GGTCTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAAGCCTCTTA ATTACACAAAAAACCAACACACAGATGTAATGAAAATAAAGATCCTTTATTCTCGAGATCC ACTAGGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGC TCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTG AGCGAGCGAGCGCGCAGCTGCCTGCAGG
153	Viral Vector IL-2 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCGG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTC CATCACTAGGGGTTCCTGCGGCCGCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGTTCGGCAATTGAACCGGTGCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCGAGGGTGGGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGGCTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAGTGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGCGAGGCGGGGCGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCC TGGCCTCGCGCCGCGCTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTGCGCACCA TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTTCCGTCCTC AGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGAGTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTTGAGTTTGGATCTTGGTTTCAATCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTAGGTGTCGTGACGCGTGCCACCATGTACAGGATGCAACTCCTGTCTT GCATTGCACTAAGTCTTGCACTTGTCTGTCTGTCTGCCAGGGCCGGGCCGCCGCGCGCTG CCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTAT TAAAAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGA CCCTGATGGAGGACGTGGAGAACAGCTTCTTCTTGAATGTCAATTCCCAAGTAACAACAGTG TGTCAGGCACTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTATGGGATTCTCCCA GGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGA TCTCGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGGAGAGAGCTCT CACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAAAAGTTGTTTCAGGA ACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACA

	GCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTACAAGAAAAACCTG ATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGTGGACCCTGTAGA TTCGGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAGGAGA CCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAAATGCAGGACAGCTAGTG TTTCTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTATGCCCACATCAT ACCATTCCTTGGATAAAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCTACT GTGCCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCTTGCCCTTCCTTGACCCTGGA AGGTGCCACTCCCCTGTCTTTCTTAATAAAATGAGGAAATGCATCACATTGTCTGAGTA GGTGTCAATCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC AATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAA GGTCACCTAGTCATGTAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTTGG TGAAGTCTAGTTAGGCAGTGGTGCCAGGTACACCTCACTCATCAGGGGGTGGCCTGACTCA GTGGTGATCTCTACTCACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCC TAAGGTGCTAATAGCAGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGAGCTCCT GCACTAAGGTAGGGAAATCCCTGAAGCTATCCTAATTAAAAGTAAAGTCTCTCCTAGAATCT ATCTAGTCCAGGTTATTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCAC CACCTGCCTGGTTAGGGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCC AGTAGGGGGTATGTAGGGATACCCTTACACTACCAGATAGTGGTGCACCACTCACTTGGGC TGAGGCCCCAGGAATAAGTAGATAGTGAGGGTGGCCCTAGTGCTCCTTCAGCTCAGGGCTTA GCTAGGCTGGAGCACCATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGC CCCCTCACTCCAGGGCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCCTTAGGAGTGGGT CACAGATCCCTGTTCCACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAG ATACTCAGCCAGCCAGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAACTGGT CAGCAAAGAATAGGGGAGCTTAAGTAGTGTTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTC TCAGTGACCCCTCAGGGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCT AAGCACTTAGGGGGTCTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTG TCAAGCCTCTTAATTACACAAAAAACCAACACACAGATGTAATGAAAATAAAGATCCTTTA TTCCTCAGATCCACTAGGGCCGCAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGC GCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGG GCGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
154	Viral Vector Sp7-P CCTGCAGGCAGCTGCGCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCCG GCGACCTTTGGTCGCCCGGCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAATC CATCACTAGGGGTTTCTGCGGCCCGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGAGGGGTGCGCAATTGAACCGGTGCCCTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCT TGCGTGCCCTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAAGTGGGTGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTGTATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTCCGTTTGTG GGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCAGCGCACATGTTTCGGCGAGGCGGGCCCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCTGCTCTGGTGCC TGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCCGTGGCACCAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCACACAAAGGAAAAGGGCCTTTCCGTCTC AGCCGTGCTTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGGAGTACGTCTCTTTAGGTTGGGGGAGGGGTTTATGCGATGGAGTTTCCC CACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTGTAGTTTGGATCTTGTTTCAATCTCAAGCCTCAGACAGTGGTTCAAAGTT TTTTTCTTCCATTTAGGTGTCTGTACGCGTGCCACCATGGCCTTCCTGTGGCTGCTGTCT GCTGGGCACTGCTGGGCACCACATTTGGCGGGCCGCCGCGCGCTGCCGTTGGTGATCTGG CATGGGATGGGAGACAGCTGTTGCAATCCCTTAAGCATGGGTGCTATTAATAAATAGGTGGA GAAGAAAATACCTGGAATTTACGTCTTATCTTTAGAGATTGGGAAGAGCTGTGAGGACG TGGAGAACAGCTTCTTCTTGAATGTCAATTCCTCAAGTAACAACAGTGTGTAGGCACCTGCT AAGGATCCTAAATGCAGCAAGGCTACAATGCTATGGGATTCTCCCAGGGAGGCCAATTTCT

	GAGGGCAGTGGCTCAGAGATGCCCTTACCTCCCATGATCAATCTGATCTCGGTTGGGGGAC AACATCAAGGTGTTTTTGGACTCCCTCGATGCCCAGGAGAGAGCTCTCACATCTGTGACTTC ATCCGAAAAAACTGAATGCTGGGGCGTACTCCAAAGTTGTTTCAGGAACGCCTCGTGCAAGC CGAATACTGGCATGACCCCATAAAGGAGGATGTGTATCGCAACCACAGCATCTTCTTGGCAG ATATAAATCAGGAGCGGGGTATCAATGAGTCCACAAAGAAAAACCTGATGGCCCTGAAGAAG TTTGTGATGGTGAAATTCTCAATGATTCCATTGTGGACCCTGTAGATTTCGGAGTGGTTTGG ATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTCCCTTACAGGAGACCTCCCTGTACACAC AGGACCGCTGGGGCTAAAGGAAATGGACAATGCAGGACAGCTAGTGTTCCTGGCTACAGAA GGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTATGCCACATCATACCATTCCCTGGATA AAAGATCTAGAGCTGAATTCTGCAGCCAGGGGATCAGCCTCTACTGTGCCTTCTAGTTGC CAGCCATCTGTTGTTTGCCCCCTCCCCCTTGCCCTTCCTTGACCCTGGAAGGTGCCACTCCAC TGTCCTTTTCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTCATTCTATTC TGGGGGTGGGGTGGGGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCT GGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTAGATCCCATCAAAGGTACCTAGTCATG TAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGGCTATGGGGTGGTGAAGTCTAGTTAGG CAGTGGTGCCAGGTCACACCTCACTCATCAGGGGTGGCCTGACTCAGTGGTGATCTCTACT CACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGACCCAGCAGGGCCTAAGGTGCTAATAGC AGGGTCCACTCACATCATGTAGCTCCTTCCCCCTGGTGGAGCTCCTGCACTAAGGTAGGGA AATCCCTGAAGCTATCCTAATTAAAAGTAAAGTCTCTCCTAGAACTATCTAGTCCAGGTTA TTCCACTATATATCCAGTGAGAAGTGGGCCCTGGTTATGTAGTCCACCACCTGCCTGGTTAG GGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGATTACCCACTCCCAGTAGGGGGGTATGT AGGGATACCCCTTACACTACCAGATAGTGGTGCACCACCTCACTTGGGTGAGGCCCCAGGAAT AAGTAGATAGTGAGGGTGCCCCCTAGTGCTCCTTCAGCTCAGGGCTTAGCTAGGCTGGAGCAC CATATCTAATTAATTGCTGTAAAGCAGGAACACCCTATGTGAAGAGCCCCCTCACTCCAGGG CCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCTTAGGAGTGGGTACAGATCCCTGTTC CACAGATTATATTAAGTCCATCAGGTGCTCAGTATCAGGGGTACAGATACTCAGCCAGCCC AGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCTTCAGGAAGTGGTCAAGAAAGTATAGGG GAGCTTAAGTAGTGTTTCAGCAGCACCCCTCCCATCTAGGTGTCTCCTCTCAGTGACCCCTCAG GGCTCCTCTAAGAACTCCACAGAATATAGGGGGCTGGGAGCCCTGCTAAGCACTTAGGGGGT CTCCAGAGGCACCTAATTACTCCAATTGGGAGGGAATGAGTATGGTGTCAAGCCTCTTAATT CACACAAAAAACACACACAGATGTAATGAAAATAAAGATCCTTTATCTCGAGATCCACT AGGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCA CTGAGGCCGGCGACCAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGC GAGCGAGCGCGCAGCTGCCTGCAGG
155	Viral Vector Native CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGCCAACTC CATCACTAGGGGTTCTGCGGCCGCGGCTCCGGTGCCCGTCACTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTGCGCAATTGAACCGGTGCTTAGAGAAGGTGG CGCGGGGTAAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTCCCGAGGGTGGGGG AGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTCGCAACGGGTTTGCCGCCA GAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTATGGCCCT TGCGTGCCTTGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGT TGGAAGTGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCTCGTGCTTGAG TTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGT CTCGTGCTTTTCGATAAGTCTCTAGCCATTTAAATTTTTTGATGACCTGCTGCGACGCTTTT TTTCTGGCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTG GGGCCGCGGGCGGCGACGGGGCCCGTGCCTCCAGCGCACATGTTTCGGCGAGGCGGGGCTG CGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGCGCTGCTCTGGTGCC TGGCCTCGCGCCCGCGTGTATCGCCCCGCCCTGGGCGGCAAGGTGGCCCGGTGCGCACAG TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACG CGGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAGGGCCTTTCCGTCTC AGCCGTGCTTTCATGTGACTCCACGGAGTACCGGGCGCGTCCAGGCACCTCGATTAGTTCT GGAGCTTTTGGAGTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCC CACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGA ATTTGCCCTTTTGGAGTTGGATCTTGTTTCTTCAAGCTCAGACAGTGGTTTCAAAGTT TTTTTCTTCCATTTTCAGGTGTGTCGACGCGTGCCACCATGGCGTCGCCCCGCTGCTGTGGC TCTTGGCTGTGGCTCTCCTGCCATGGACCTGCGCTTCTCGGGCGCTGCAGCATCTGGACCCG

	CCGGCGCCGCTGCCGTTGGTGATCTGGCATGGGATGGGAGACAGCTGTTGCAATCCCTTAAG CATGGGTGCTATTAATAAATGGTGGAGAAGAAAATACCTGGAATTTACGTCTTATCTTTAG AGATTGGGAAGACCCTGATGGAGGACGTGGAGAACAGCTTCTTCTGAATGTCAATTCCCAA GTAACAACAGTGTGTGAGGCACCTTGCTAAGGATCCTAAATTGCAGCAAGGCTACAATGCTAT GGGATTCTCCCAGGGAGGCCAATTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCACCTCCCA TGATCAATCTGATCTCGGTTGGGGGACAACATCAAGGTGTTTTTGGACTCCCTCGATGCCCA GGAGAGAGCTCTCACATCTGTGACTTCATCCGAAAAACACTGAATGCTGGGGCGTACTCCAA AGTTGTTTCAGGAACGCCTCGTGCAAGCCGAATACTGGCATGACCCCATAAAGGAGGATGTGT ATCGCAACCACAGCATCTTCTTGGCAGATATAAATCAGGAGCGGGGTATCAATGAGTCCTAC AAGAAAAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATTCCATTGT GGACCTGTAGATTTCGGAGTGTTTTGGATTTTACAGAAGTGGCCAAGCCAAGGAAACCATTC CCTTACAGGAGACCTCCCTGTACACACAGGACCGCCTGGGGCTAAAGGAAATGGACAATGCA GGACAGCTAGTGTCTTGGCTACAGAAGGGGACCATCTTCAGTTGTCTGAAGAATGGTTTTTA TGCCACATCATACCATTCTTGGATAAAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGA TCAGCCTCTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCTTGCCTTC CTTGACCTGGAAGGTGCCACTCCCCTGTCTTTCCTAATAAAATGAGGAAATTCATCAC ATTGTCTGAGTAGGTGTCTATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAG GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCTATGGGGTGGTCAATCTA GATCCCATCAAAGGTACCTAGTCATGTAAGTGAGCCACTATCTCCTCCTACAAGGCCAGGG CTATGGGGTGGTGAAGTCTAGTTAGGCAGTGGTGCCAGGTACACCTCACTCATCAGGGGG TGGCTGACTCAGTGCTGATCTCTACTCACCTCCTCAGGCCCTCATCATTTAGCCTGTAAGA CCCAGCAGGGCCTAAGGTGCTAATAGCAGGGTCCACTCACATCATGTAGCTCCTTCCCCCT GGTGGAGCTCCTGCACTAAGGTAGGGAAATCCCTGAAGCTATCCTAATTAAGTAAAGTCT CTCCTAGAATCTATCTAGTCCAGGTTATTCCTACTATATATCCAGTGAGAAGTGGGCCCTGGT TATGTAGTCCACCACCTGCCTGGTTAGGGTAGTGCTGCTGTAGCCTCACCTGCACAGGTGAT TACCCACTCCCAGTAGGGGGGTATGTAGGGATACCCTTACACTACCAGATAGTGGTGCACC ACTCACTTGGGCTGAGGCCCCAGGAATAAGTAGATAGTGAGGGTGGCCCTAGTGCTCCTTCA GCTCAGGGCTTAGCTAGGCTGGAGCACCATATCTAATTAATTGCTGTAAAGCAGGAACACCC TATGTGAAGAGCCCCCTCACTCCAGGGCCAGGTAGAAGGGAGGGCTCCTATAGAGGTTTTCC TTAGGAGTGGGTACAGATCCCTGTTCCACAGATTATATTAAGTCCATCAGGTGCTCAGTAT CAGGGGTACAGATACTCAGCCAGCCAGTGGTGTACTAGCTGGGCAGTTAGGTAGACAGCT TCAGGAACTGGTCAGCAAAGAATAGGGGAGCTTAAGTAGTGTTTCAGCAGCACCCCTCCCCT AGGTGTCTCCTCTCAGTGACCCCTCAGGGCTCCTCTAAGAACTCCACAGAATATAGGGGGCT GGGAGCCCTGCTAAGCACTTAGGGGGTCTCCAGAGGCACCTAATTACTCCAATTGGGAGGGA ATGAGTATGGTGTCAAGCCTCTTAATTCACACAAAAAACCAACACACAGATGTAATGAAAAT AAAGATCCTTTATTCTCGAGATCCACTAGGGCCGCGAGGAACCCCTAGTGATGGAGTTGGCCA CTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCG GGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
Components	
156	Variant 5' ITR v2 CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTTCGCCCCGGCCT CAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAAC'TCCATCACTAGGGGTTCTCT
157	Variant 5' ITR v3 (self-complementary) CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCGGGCGACCTTTGG TCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG
158	Variant 3' ITR v1 AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCC GCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG
159	Variant 3' ITR v2 AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCC GGGCGACCAAAGGTCGCCCCGACGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG
160	Variant 3' ITR v3 (self-complementary) CACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCC CGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG
Full-Length signal + PPT	
161	SPARC PPT1co2 ATGAGGGCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA

	GGCCCACCTGCACCACCTTCCCTTGGTCATCTGGCATGGCATGGGTGACTCTTGCTGCAATCC CTTATCCATGGGGGCTATTAAGAAGATGGTGGAAAAGAAAGATTCTGGAATTTATGTGCTGA GTCTGGAAATTGGAAAGACCCCTCATGGAAGATGTTGAAAACAGCTTCTTCTTGAATGTCAAT TCCCAGGTCACCACAGTGTGTCTAGGCTCTGGCAAAGGACCCCAAGCTCCAGCAAGGATATAA TGCCATGGGATTTTCCCAAGGTGGCCAGTTCTGAGGGCAGTGGCTCAGAGATGCCCTTCAC CCCCATGATCAACCTCATCTCTGTGGGTGGGCAGCATCAGGGTGTCTTTGGGCTCCCAAGA TGCCCAGGAGAAAGCTCTCACATCTGTGACTTTATTAGAAAGACCCTGAATGCTGGGGCTTA CAGCAAGGTTGTCCAGGAGAGGCTGGTGCAGGCTGAATACTGGCATGACCCCATTAAGGAAG ATGTGTACAGAAACCACAGTATCTTCTTGGCTGATATCAACCAAGAAAGAGGCATCAATGAG AGCTACAAGAAGAATCTCATGGCCCTGAAGAAGTTTGTGATGGTGAAATTCCTCAATGATAG CATTGTGGACCCTGTAGACTCAGAGTGGTTTGGATTCTACAGGTCTGGCCAGGCTAAGGAGA CCATCCCCTGCAAGAGACAAGCCTTTACACACAGGACAGGCTTGGTCTGAAAGAGATGGAT AATGCTGGACAGCTGGTGTCTTCTGGCCACTGAAGGGGACCACCTGCAGCTGAGTGAAGAATG GTTCTATGCCCACATCATCCCCCTTCTTAGGCTGA
162	SPARC_PPT1co3 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA GGGCCCCCAGCACCCCTGCCTCTGGTCATTTGGCATGGCATGGGGGACTCCTGCTGCAATCC CTTATCCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAATTTATGTTCTGT CTTTAGAAATAGGAAAAACACTGATGGAAGATGTGGAAAATTCCTTCTTCTCAATGTCAAC TCCCAGGTAACAACAGTGTGTCTAGGCCCTGGCCAAAGACCCCAAATTACAGCAGGGCTACAA TGCTATGGGATTCTCACAGGGAGGCCAATTTCTCAGAGCTGTTGCTCAGAGATGCCCTTCAC CCCCCATGATCAACCTGATTTTCAGTGGGTGGACAACATCAAGGAGTTTTTGGATTACCAAGG TGCCCTGGGGAGAGCTCTCACATTTGTGATTTTATAAGAAAAACCTTGAATGCTGGGGCCTA CTCAAAGGTGGTCCAGGAGAGGTTAGTGCAGGCTGAGTATTGGCATGACCCCATTAAGGAAG ATGTGTACAGAAACCCTCCATTTTTTTGGCTGACATCAACCAGGAGAGAGGAATTAATGAA TCCTACAAGAAGAACCCTGATGGCCCTCAAGAAGTTTGTGATGGTGAAATTCCTCAATGACTC CAATTGTGGATCCAGTGGATTTCAGAGTGGTTTGGATTCTACAGAAGTGGCCAAGCCAAGGAAA CAATACCACTTCAAGAGACCAGCCTGTACACCCAGGACAGACTGGGCCTAAAGGAAATGGAC AATGCAGGTACGCTTGTCTTCTGGCCACTGAGGGGGACCACCTTCAACTCTCTGAAGAATG GTTTTATGCCCACATCATACCCTTCTGGGGTGA
163	SPARC_PPT1co4 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA GGCCCTCCAGCACCTCTTCTTTAGTGATCTGGCATGGCATGGGTGACAGCTGCTGCAACCC ACTGTCCATGGGAGCCATCAAGAAAATGGTGGAGAAGAAAATTCAGGAATCTATGTCCTAT CCCTGGAGATTGGGAAGACTCTGATGGAAGATGTGGAGAATCCTTCTTCTGAATGTCAAT TCCCAAGTCACCACAGTCTGCCAGGCCCTGGCCAAAGACCCCAAGCTGCAGCAAGGGTACAA TGCTATGGGCTTCTCCCAAGGGGGGCAGTTCTTGAGAGCAGTTGCCAGAGATGCCCCAGCC CTCCAATGATTAATCTGATCTCTGTGGGAGGTGAGCATCAAGGAGTCTTTGGGTGCCCAGA TGTCCAGGAGAGAGCAGTCACATCTGTGACTTCATCAGGAAAACACTGAATGCTGGGGCATA CAGCAAGGTGGTGCAGGAAAGGCTTGTGCAGGCAGAGTACTGGCATGACCCATTAAAGAGG ATGTGTACAGAAACCACAGCATCTTTCTTGACAGACATAAATCAGGAGAGAGGTATCAATGAG TCCTACAAGAAGAATCTGATGGCCCTAAAGAAAATTTGTGATGGTAAAATTCCTCAATGACTC CATAGTGGACCCAGTTGACTCTGAGTGGTTTGGATTTTACAGAAGTGGCCAAGCAAAAGAGA CCATTCCCCTTCAAGAGACATCCCTGTACACCCAAGATAGGCTGGGCCTGAAAGAGATGGAT AATGCAGGGCAATTGGTGTCTCTGTACAGAAGGAGACCACCTCCAGCTATCTGAAGAGTG GTTCTATGCCCACATCATCCCTTTTCTAGGGTGA
164	SPARC_PPT1co11 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA GGCCCCCCTGCCCCCTCTTCCCCGGTTCATTTGGCATGGGATGGGTGACTCCTGCTGCAACCC CCTAAGCATGGGGGCCATCAAGAAGATGGTGGAGAAGAAGATTCTGGAATTTATGTCCTTA GCCTGGAAATTGGGAAGACCTTAATGGAAGATGTGGAAAACAGCTTTTTCTTAAATGTCAAT TCTCAGGTAACAACAGTGTGCCAAGCCCTGGCCAAGGACCTAAACTGCAGCAGGGCTACAA TGCTATGGGCTTCAGCCAGGGTGGACAGTTTCTGAGGGCAGTGGCTCAGAGGTGCCCCAGTC CTCCCATGATCAATCTCATCTCTGTGGGAGGCCAGCACCAGGGTGTCTTTGGCCTGCCCAGG TGTCTGGAGAGTCTCACATATCTGTGACTTCATCAGGAAGACACTGAATGCAGGGGCCCTA TAGCAAGGTGGTGAAGAGAGACTGGTTCAGGCTGAATACTGGCATGACCCCATCAAGGAAG ATGTTTACAGGAACCATAGCATCTTCTTGTGATATTAATCAAGAGAGGGGCATAAATGAG AGCTACAAGAAGAATTTAATGGCTCTAAAGAAAATTTGTGATGGTGAAGTTCCTGAATGACTC

	CATTGTGGATCCTGTGGACTCTGAGTGGTTTGGGTTCTACAGAAGTGGCCAGGCCAAGGAGACCATCCCCCTGCAGGAGACAAGCCTGTACACACAGGACAGGCTGGGACTGAAGGAGATGGACAATGCAGGACAGCTGGTGTTCCTGGCTACAGAGGGAGACCACCTGCAGCTCTCAGAGGAATGGTTCTATGCACACATCATTCCTTTTATAGGG TGA
165	SPARC_PPT1c014 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA GGACCCCCAGCTCCACTCCCCCTGGTGATCTGGCATGGCATGGGAGATAGCTGTTGCAACCCATTATCCATGGGGGCCATTAAGAAGATGGTGGAGAAGAAAATCCCTGGCATTATGTCCTTAGCCTGGAATTGGAAAGACCTTAATGGAGGATGTGGAAAACCTTCTTCCTCAATGTCAACTCCCAAGTCACCACAGTGTGTCAAGCCCTTGCCAAGGACCCCAAGTTGCAACAGGGCTATAATGCTATGGGCTTCTCTCAGGGAGGCCAGTTTCTGAGGGCTGTGGCCAGAGGTGCCCTCCCACCTATGATCAATCTTATCTCTGTGGGTGGACAACACCAAGGTGTCTTTGGGCTGCCCAGGTGCCCTGGGGAGTCCAGCCACATTTGTGACTTCATCAGGAAGACACTAAATGCTGGTGCCTACTCCAAAGTTGTCCAGGAGAGACTGGTGCAAGCAGAATACTGGCATGACCCTATTAAAGAGGATGTCTACAGGAACCACAGCATCTTCTGGCAGACATCAACCAGGAGAGAGGGATCAATGAATCCTACAAGAAGAACCTCATGGCCCTGAAGAAATTTGTGATGGTTAAGTTTCTCAATGACTCATTGTGGACCCTGTGGATTCTGAATGGTTTGGATTTTACAGAAGTGGCCAGGCCAAGGAAACCATCCCTCTGCAGGAGACATCCCTGTACACACAGGACAGACTGGGCCTGAAGGAGATGGACAATGCTGGGCAGCTAGTGTTCCTGGCTACTGAGGGGGATCATCTTCAGTTGTCTGAAGAATGATTCTATGCCACATCATTCCTTCTCGGG TGA
166	SPARC_PPT1c018 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA GGTCCACCAGCTCCTCTGCCCTTGGTCATCTGGCATGGGATGGGGGACTCTTGTTGTAACCCTCTTAGCATGGGAGCCATTAAGAAGATGGTGGAGAAGAAAATACCTGGCATATATGTCCTGAGCCTGGAGATTGGCAAGACCCTGATGGAGGATGTGGAGAACAGCTTCTTCCTGAATGTCAATTCCCAGGTGACCACAGTGTGTGAGGCCCTGGCCAAGGACCCCAAGTTGCAGCAAGGCTACAAATGCTATGGGCTTTTACAAAGGTGGGCAGTTCCTGAGAGCAGTGGCACAGAGATGTCCTTCACCTCCAATGATCAACCTTATCTCAGTGGGAGGCCAACACCAGGGTGTTTTTGGGTTGCCCAGGTGTCCTGGTGAGAGCAGCCATATCTGTGACTTTATCAGGAAGACACTGAATGCAGGGGCCACAGCAAGGTTGTTCAAGAGAGGCTGGTGAGGCTGAGTATTGGCATGATCCATTAAAGGAGGATGTGTATAGGAACCACTCCATCTTTTGGCAGACATCAATCAGGAGAGGGGCATCAATGAGTCTCTACAAGAAGAACCTGATGGCCCTCAAGAAGTTTGTGATGGTCAAGTTTCTGAATGATTCATTGTGGATCCAGTGGACTCTGAATGGTTTGGGTTCTACAGATCTGGCCAGGCCAAGGAGAACCATCCCCTTCAAGAGACCAGCCTGTATACCCAGGACAGATTAGGCCTGAAGGAGATGGACAATGCTGGCCAGCTTGTCTTCCTGGCTACAGAGGGGGACCACCTCCAGCTGAGTGAAGAGTGGTTTTATGCCACATCATACCTTCTCGGG TGA
167	SPARC_PPT1c019 ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCCTTGGA GGCCCCCAGCCCCCTCTGCCCTGGTCATCTGGCATGGCATGGGAGACAGCTGCTGCAACCCTCTCTCCATGGGAGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAATCTATGTCCTCAGCCTGGAGATTGGGAAGACCCTGATGGAGGATGTGGAGAACTCCTTCTTCCTGAATGTGAACAGCCAGGTGACCACAGTGTGCCAGGCCCTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAAATGCCATGGGCTTCTCCCAGGGAGGCCAGTTCCTGAGGGCTGTGGCCAGAGATGCCCTCACCTCCCATGATCAACCTCATCTCTGTGGGAGGCCAGCACCAAGGGGTCTTTGGCCTGCCCAGGTGTCCTGGGGAGAGCAGCCACATCTGTGACTTCATCAGGAAGACCCTGAATGCTGGTGCCTACAGCAAGGTGGTGAGGAGAGGCTGGTGAGGCTGAATACTGGCATGACCCCATCAAGGAGGATGTGTACAGGAACCACAGCATCTTCCTGGCAGACATCAACCAGGAGAGAGGCATCAATGAGAGCTACAAGAAGAACCCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTTCTCAATGACTCATTGTGGACCCTGTGGACTCTGAGTGGTTTGGCTTCTACAGGAGTGGCCAGGCCAAGGAGAACCATCCCCCTGCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTGAAGGAGATGGACAATGCTGGGCAGCTGGTGTTCCTGGCCACAGAGGGGAGACCACCTGCAGCTCTCAGAGGAGTGGTTTTATGCCACATCATCCCCCTTCTCGGG TGA
168	sp7_PPT1c020 ATGGCCTTTCTGTGGCTGCTGTCTGCTGGGCCCTGCTGGGGACCACCTTTGGC GGGCCCCCAGACCCCTGCCTCTGGTCATCTGGCATGGGATGGGAGATTCTCTGTCGAACCTGTCAGCATGGGGGCCATTAAGAAAATGGTGGAGAAGAAGATCCCTGGCATTATGTCCTGTCTTGGAGATAGGCAAAACACTAATGGAAGATGTGGAGAACAGCTTCTTCTTAAATGTGAACAGCCAGGTCAACCACAGTATGTCAGGCCTTGCCAAGGACCCCAAGCTCCAGCAGGGCTACAA

	TGCCATGGGTTTTTCTCAGGGAGGTCAGTTCCTCAGGGCTGTGGCTCAGAGATGCCCTCAC CACCTATGATCAATCTGATCTCTGTGGGAGGGCAGCATCAAGGTGTCTTTGGTCTTCTAGG TGCCCTGGGGAGTCATCCCACATCTGTGACTTCATCAGGAAAACCTTAAATGCTGGAGCCTA CAGCAAGGTTGTCCAAGAGAGGCTTGTCCAAGCAGAGTACTGGCATGATCCCATCAAAGAGG ATGTTTACAGAAAACCACAGCATCTTCTGGCAGACATCAATCAAGAGAGGGGCATTAATGAG TCCTACAAGAAAAACCTCATGGCCCTAAAGAAATTTGTGATGGTGAAATTCCTGAATGACTC CATTGTTGACCCTGTAGACTCTGAGTGGTTTGGCTTCTATAGATCAGGCCAGGCCAAGGAAA CCATTCTCTCCAAGAAACCTCCCTGTACACACAAGATAGGCTGGGGTTGAAAGAGATGGAT AATGCTGGGCAACTGGTGTCTTGGCCACTGAGGGTGACCATTACAGCTTTCAGAGGAATG GTTCTATGCCCACATCATACCCTTCTTGGATGA
	Expression Cassette
169	sp7_PPTlco4 TACAGAAAACCAAGAAAGCACCATAGCGTTTGCCTTCAACTGCTTCCCAACTTGTTTTCTCT TCCAATTTGATTGTGGTTTCTCTCCAGAAGGAACCTCCACAGTACTTAGCGTTGGCCACATA GTAGGTTCTCAAATACTTGTTAATAATAAGTTTGTTCGAGAAGCTGGGCAATGATATTCTA CAGCTGGAAGAAGAAACATAATGATCTAGTAATTAGCTCATTAAAAATAAACGTTCTTCTTT CCTCAGAGGAGCATTTCCCAAGGCCTGCCTTGATAGCCATCCAAAAGGCCAAGCTCATCCA ATCTTGCCCTAGATTTATGCTAAAATGCAGTTACAATCGATAGGATGACAGAAAACGACAGC ACTTATTTAAATATAATAGGCACTTATTTAAATAGGAGAAGCTGTGACTTCATAGCAAGTGT TGGGGTTAGGAAACTGGGTGGATAAACTTGCTGATGCTGTAGATCTTAGCCTCTACATGAGA TCATGTGGAAAATCTGAAAGCATTTTAGGTTCTTATGTTTGCAATCAAATAACTGTACACC TTTTAATTTAAAAAGTACCATGAGGCACACACACACTCGCAGGAACCTTTTGGCGTAACA AAACTAGAAATTAGATCTAAAGCTAACTGTAGGACTGAGTCTATTCTAAACTGAAAGCCTGGA CATCTGGAGTACCAGGGGGAGATGACGTGTTACGGGCTTCCATAAAAGCAGCTGGCTTTGAA TGGAAGGAGCCAAGAGGCCAGCACAGGAGCGGATTCGTGCTTTCACGGCCATCGAGCCGAA CCTCTCGCAAGTCCGTGAGCCGTTAAGGAGGCCCCAGTCCCGACCCTTCGCCCCAAGCCCC TCGGGGTCCCCGGGCCTGGTACTCCTTGCCACACGGGAGGGGCGCGGAAGCCGGGGCGGAGG AGGAGCCAACCCCGGGCTGGGCTGAGACCCGCAGAGGAAGACGCTCTAGGGATTTGTCCCGG ACTAGCGAGATGGCAAGGCTGAGGACGGGAGGCTGATTGAGAGGCCAAGGTACACCCTAATC TCAATACAACCTTTGGAGCTAAGCCAGCAATGGTAGAGGGAAGATTCTGCACGTCCCTTCCA GGCGGCTCCCCGTCACCAACCCCCCAACCCGCCCCGACCGGAGCTGAGAGTAATTCATAC AAAAGGACTCGCCCTGCCTTGGGGAATCCCAGGGACCGTCGTTAACTCCCCTAACGTAAG AACCAGAGATCGCTGCGTTCCCGCCCCCTCACCCGCCGCTCTCGTCATCAGCTGAGGTGGA GAAGAGCATGCGTGAGGCTCCGGTGCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCC CGAGAAGTTGGGGGGAGGGGTGCGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAA ACTGGGAAAGTGATGTGCTGTAAGTGGCTCCGCTTTTTCGAGGGTGGGGGAGAACCCTAT ATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT AAGTGCCGTGTGTGGTTCCCGCGGGCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTT GAATTACTTCCACGCCCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAG TGGGTGGGAGAGTTTCGAGGCCTTGCGCTTAAGGAGCCCTTCGCCTCGTGCTTGAGTTGAGG CCTGGCTTGGGCGCTGGGGCCGCCGCTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCT GCTTTCGATAAGTCTCTAGCCATTTAAATTTTGTATGACCTGCTGCGACGCTTTTTTCTG GCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCG CGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTTCGGCGAGGCGGGGCTGCGAGCG CGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCTGCTCTGGTGCCCTGGCCT CGCGCCGCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTGCGCACAGTTGCGT GAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGC TCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGGCCTTTCGCTCTCAGCCGT CGCTTCATGTGACTCCACGGAGTACCGGGCGCGTCCAGGCACCTCGATTAGTCTTTTGGGA GTACGTGCTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTCCCCACACTGAGTGG GTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTT TGAGTTTGGATCTTGGTTTCAATCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTTCTTCCAT TTCAGGTGTGCTGAAAACCTACCCCTAAAAGCCAAAACGCGTGCCACCATGGCCTTTCTGTGG CTGCTGTCTGCTGGGCCCTGCTGGGGACCACCTTTGGCGGCCCTCCAGCACCTCTTCTTTT AGTGATCTGGCATGGCATGGGTGACAGCTGCTGCAACCCACTGTCCATGGGAGCCATCAAGA AAATGGTGGAGAAGAAATTCAGGAATCTATGTCCCTATCCCTGGAGATTGGGAAGACTCTG ATGGAAGATGTGGAGAATCCTTCTTCTGCAATGTCAATTCCTCAAGTACCACAGTCTGCCA GGCCCTGGCCAAAGACCCCAAGCTGCAGCAAGGGTACAATGCTATGGGCTTCTCCCAAGGGG

	GGCAGTTCTTGAGAGCAGTTGCCAGAGATGCCCCAGCCCTCCAATGATTAATCTGATCTCTGTGGGAGGTCAGCATCAAGGAGTCTTTGGGTTGCCAGATGTCCAGGAGAGAGCAGTCACATCTGTGACTTCATCAGGAAAACACTGAATGCTGGGGCATAACAGCAAGGTGGTGAGGAAAGGC TTGTGCAGGCAGAGTACTGGCATGACCCCATTAAGAGGATGTGTACAGAAACCACAGCATC TTTCTTGACACATAAATCAGGAGAGAGGTATCAATGAGTCCACAGAAGAATCTGATGGC CCTAAAGAAATTTGTGATGGTAAAATTCCTCAATGACTCCATAGTGGACCCAGTTGACTCTG AGTGGTTTGGAATTTACAGAAGTGGCCAAGCAAAAGAGACCATTCCCTTCAAGAGACATCC CTGTACACCCAAGATAGGCTGGGCCTGAAAGAGATGGATAATGCAGGGCAATTGGTGTTCCCT TGCTACAGAAGGAGACCACCTCCAGCTATCTGAAGAGTGGTTCATGCCCACATCATCCCTT TTCTAGGGTGAAGATCTAGAGCTGAATTCCTGCAGCCAGGGGGATCAGCCTCTACTGTGCCT CTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCTTGCCTTCCTTGACCTGGAAGGTGC CACTCCCCTGTCTTTCTAATAAAATGAGGAAATTGCATCACATTGTCTGAGTAGGTGTC ATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGC AGGCATGCTGGGGATGCAGTGGGCTCTATGG
170	sp7_PPT1c019 TACAGAAACCAAGAAAGCACCATAGCGTTTGCCTTCAACTGCTTCCCAACTTGTTTTCTCTTCCAAATTTGATTGTGGTTTCTCTCCAGAAGGAACCTCCACAGTACTTAGCGTTGGCCACATAGTAGGTTCTCAAACTTGTTAATAAATAAGTTTGTTCGAGAAGCTGGGCAATGATATTCTACAGCTGGAAGAAGAAACATAATGATCTAGTAATTAGCTCATTAATAAATAAACGTTCTTCTTT CCTCAGAGGAGCATTTCCCAAGGCCTGCCTTGATAGCCATCCAAAAAGGCCAAGCTCATCCA ATCTTGCCCTAGATTTATGCTAAAATGCAGTTACAATCGATAGGATGACAGAAAACGACAGC ACTTATTATAATATAATAGGCACCTATTTAAATAGGAGAAGCTGTGACTTCATAGCAAGTGT TGGGGTTAGGAAACTGGGTGGATAAACTTGCTGATGCTGTAGATCTTAGCCTCTACATGAGA TCATGTGGAAAAATCTGAAAGCATTTTAGGTTCCCTATGTTTGCAATCAAATAACTGTACACC TTTTAATTTAAAAAGTACCATGAGGCACACACACACTCGCAGGAACTTTTTGGCGTAACA AACTAGAAATTAGATCTAAAGCTAACTGTAGGACTGAGTCTATTCTAAACTGAAAGCCTGGA CATCTGGAGTACCAGGGGGAGATGACGTGTTACGGGCTTCCATAAAAGCAGCTGGCTTTGAA TGGAAAGGAGCCAAGAGGCCAGCACAGGAGCGGATTTCGTGCTTTCACGGCCATCGAGCCGAA CCTCTCGCAAGTCCGTGAGCCGTTAAGGAGGCCCCAGTCCCGACCCCTTCGCCCCAAGCCCC TCGGGTCCCCGGGCCTGGTACTCCTTGCCACACGGGAGGGGCGCGGAAGCCGGGGCGGAGG AGGAGCCAACCCCGGGCTGGGCTGAGACCCGCAGAGGAAGACGCTCTAGGGATTGTCCCGG ACTAGCGAGATGGCAAGGCTGAGGACGGGAGGCTGATTGAGAGGCGAAGGTACACCCTAATC TCAATACAACCTTTGGAGCTAAGCCAGCAATGGTAGAGGGAAGATTCTGCACGTCCTTCCCA GCGGCTTCCCGCTCACCACCCCCCAACCCGCCCCGACCGGAGCTGAGAGTAATTCATAC AAAAGGACTCGCCCCCTGCCTTGGGGAATCCCAGGGACCGTCGTTAAACTCCCCTAACGTAG AACCCAGAGATCGCTGCGTTCCCGCCCCCTCACCCGCCCCGCTCTCGTCATCACTGAGGTGGA GAAGAGCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCC CGAGAAGTTGGGGGGAGGGTTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAA ACTGGGAAAGTGATGTCGTGTACTGGCTCCGCTTTTTTCCCGAGGGTGGGGGAGAACCCTAT ATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTCGCCGCCAGAACACAGGT AAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTATGGCCCTTGCGTGCCCTT GAATTACTTCCACGCCCCCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTGGAAG TGGGTGGGAGAGTTCGAGGCCCTTGCGCTTAAGGAGCCCCCTTCGCTCGTGCTTGAGTTGAGG CCTGGCTTGGGCGCTGGGGCCCGCGCTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCT GCTTTCGATAAGTCTCTAGCCATTTAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTG GCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCG CGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGGCGAGCGGGGGCTGCGAGCG CGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCTGCTCTGGTGCTTGGCCT CGCGCCGCGGTGTATCGCCCCGCCCTGGGCGCAAGGCTGGCCCGTGGCACCAAGTTGCGT GAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAATGGAGGACGCGGCGC TCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGCCGT CGCTTCATGTGACTCCACGGAGTACCGGGCGCGTCCAGGCACCTCGATTAGTTCTTTTGGA GTACGTCGTCTTTAGGTTGGGGGGAGGGTTTTATGCGATGGAGTTCCCCACACTGAGTGG GTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGAATTTGCCCTTTT TGAGTTTGGATCTTGTTTCTTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCAT TTCAGGTGTCGTGAAAACCTACCCCTAAAAGCCTGCGTGGCACCCTTCTGCTGG CTGCTGTCTGCTGGGCCCTGCTGGGGACCACCTTTGGCGGCCCCCAGCCCCCTTGCCTT GGTTCATCTGGCATGGCATGGGAGACAGCTGCTGCAACCCTCTCTCCATGGGAGCCATCAAGA

	AGATGGTGGAGAGAAGATCCCTGGAACTCTATGTCTCAGCCTGGAGATTGGGAAGACCCTG ATGGAGGATGTGGAGAACCTCCTTCTTCCCTGAATGTGAACAGCCAGGTGACCACAGTGTGCCA GGCCCTGGCCAAGGACCCCAAGCTGCAGCAGGGCTACAATGCCATGGGCTTCTCCCAGGGAG GCCAGTTCCTGAGGGCTGTGGCCAGAGATGCCCTCACCTCCCATGATCAACCTCATCTCT GTGGGAGGCCAGCACCAAGGGTCTTTGGCCTGCCAGGTGTCTGGGGAGAGCAGCCACAT CTGTGACTTCATCAGGAAGACCCTGAATGCTGGTGCCTACAGCAAGGTGGTGCAGGAGAGGC TGGTGCAGGCTGAATACTGGCATGACCCCATCAAGGAGGATGTGTACAGGAACCACAGCATC TTCCTGGCAGACATCAACCAGGAGAGAGGCATCAATGAGAGCTACAAGAAGAACCTGATGGC CCTGAAGAAGTTTGTGATGGTGAAGTTCCCTCAATGACTCCATTGTGGACCCTGTGGACTCTG AGTGGTTTGGCTTCTACAGGAGTGGCCAGGCCAAGGAGACCATCCCCCTGCAGGAGACCAG CTGTACACAGGAGACAGACTGGGACTGAAGGAGATGGACAATGCTGGGACATGGTGTCTCCCT GGCCACAGAGGGAGACACCTCGAGCTCTCAGAGGAGTGTTTTTATGCCACATCATCCCCCT TCCTGGGCTGAAGATCTAGAGCTGAATTCTCGAGCCAGGGGGATCAGCCTCTACTGTGCCT TCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCTTGCCCTTCCTTGACCCTGGAAGGTGC CACTCCCCTGTCTTTCTTAATAAAATGAGGAAATGCATCACATTGTCTGAGTAGGTGTC ATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGC AGGCATGCTGGGGATGCAGTGGGCTCTATGG
Viral	nucleic acid
171	sp7_PPT1co4 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCG TCGGGCGACCTTTGGTCGCCCGGCCCTCAGTGAGCGGAGCGAGCGCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTCCCTGATATCGCGGCCGCTACAGAAACCAAGAAAGCACCATAGCG TTTGCCCTCAACTGCTTCCCCAAGTTGTTTTCTCTTCCAATTTGATTGTGGTTTCTCTCCA GAAGGAACCTCAACAGTACTTAGCTTGCCACATAGTAGGTTCTCAAACTACTTGTTAATAAA TAAGTTTGTTTCGAGAAGCTGGGCAATGATATTCTACAGCTGGAAGAAGAAACATAATGATCT AGTAATTAGCTCATTAATAAATAAACGTTCTTCTTTCTCAGAGGAGCATTTCCCAAGGCCGTG CCTTGATAGCCATCCAAAAGGCCAAGCTCATCCAATCTTGCCCTAGATTTATGCTAAAATG CAGTTACAATCGATAGGATGACAGAAAACGACAGCACTTATTTAAATATAATAGGCACTTAT TTAAATAGGAGAAGCTGTGACTTCATAGCAAGTGTGGGGTTAGGAAACTGGGTGGATAAAC TTGCTGATGCTGTAGATCTTAGCCTCTACATGAGATCATGTGGAAAATCTGAAAGCATTTTA GGTTTCCTTATGTTTGCAATCAAATAACTGTACACCTTTTAAATTTAAAAAGTACCATGAGGCA CACACACACACTCGCAGGAACTTTTTGCGGTAAACAAAACCTAGAATTAGATCTAAAGCTAACT GTAGGACTGAGTCTATTCTAAACTGAAAGCCTGGACATCTGGAGTACCAGGGGGAGATGACG TGTTACGGGCTTCCATAAAAGCAGCTGGCTTTGAATGGAAGGACCAAGAGGCCAGCAGAG AGCGGATTCGTGCCTTTACGGCCTCGAGCCGAACCTCTCGCAAGTCCGTGAGCCGTTAAG GAGGCCCCAGTCCCGACCTTTCGCCCCAAGCCCCCTCGGGGTCCCCGGGCTGGTACTCCTT GCCACACGGGAGGGGCGCGGAAGCCGGGCGGAGGAGGCCAACCCCGGGCTGGGCTGAGA CCCGCAGAGGAAGACGCTCTAGGGATTTGTCCCGGACTAGCGAGATGGCAAGGCTGAGGACG GGAGGCTGATTGAGAGGCGAAGGTACACCTAATCTCAATACAACCTTTGGAGCTAAGCCAG CAATGGTAGAGGGAAGATTCTGCACGTCCCTTCCAGCGGCCCTCCCCGTCAACACCCCCC AACCCGCCCCGACCGGAGCTGAGAGTAATTATACAAAAGGACTCGCCCTGCCTTGGGGAA TCCCAGGGACCGTCGTTAAACTCCCCTAACGTAGAACCCAGAGATCGCTGCGTTCCCGCCC CCTCACCCGCCCCGCTCTCGTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCC CGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTTCGGCAA TTGAACCCGCTTCCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTATGTCTGTACTGGC TCCGCTTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTT CTTTTTTCGAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCC TGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACGCCCTGGCTGCAG TACGTGATTCTTGATCCCGAGCTTCGGGTGGAAGTGGGTGGGAGAGTTCGAGGCCCTTGCGC TTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCTTGGGCGCTGGGGCCGCCGCG TGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTGATAAGTCTCTAGCCATTTAA AATTTTTGATGACCTGCTGCGACGCTTTTTTCTGGCAAGATAGTCTTGTAATGCGGGCCA AGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCCGGGGCGGCGACGGGGCCCGTGCGTCCC AGCGCACATGTTTCGGCAGGGCGGGGCCCTGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAG TCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTG GGCGGCAAGGCTGGCCCGGTTCGGCACCAGTTGCGTGAGGACGGAGAGTGGCCGCTTCCGGCC CTGCTGCAGGAGGATCAAAATGGAGGACGCGCGCTCGGGAGAGCGGGCGGCTGAGTCAACC

	GGCGCCGTCCAGGCACCTCGATTAGTTCTTTTGGAGTACGTCGTCCTTTAGGTTGGGGGGAGG GGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGG CACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCACTTCTCAA GCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTTTCAGGTGTCGTGAAAACCTACCCCTAA AAGCCAAAACGCGTGCCACCATGGCCTTTCTGTGGCTGCTGTCTGCTGGGCCCTGCTGGGG ACCACCTTTGGCGGCCCTCCAGCACCTCTTCCCTTTAGTGATCTGGCATGGCATGGGTGACAG CTGCTGCAACCCACTGTCCATGGGAGCCATCAAGAAAATGGTGGAGAAGAAAATTCCAGGAA TCTATGTCTATCCCTGGAGATTGGGAAGACTCTGATGGAAGATGTGGAGAACTCCTTCTTC CTGAATGTCAATTCCCAAGTCACCACAGTCTGCCAGGCCCTGGCCAAAGACCCCAAGCTGCA GCAAGGGTACAATGCTATGGGCTTCTCCCAAGGGGGGAGTTCTTGAGAGCAGTTGCCCAGA GATGCCCCAGCCCTCCAATGATTAATCTGATCTCTGTGGGAGGTCAAGCATCAAGGAGTCTTT GGGTTGCCCAGATGTCCAGGAGAGAGCAGTCACATCTGTGACTTCATCAGGAAAAACTGAA TGCTGGGGCATAACAGCAAGGTGGTGCAGGAAAGGCTTGTGCAGGCAGAGTACTGGCATGACC CCATTAAAGAGGATGTGTACAGAAACCACAGCATCTTTCTTGCAGACATAAATCAGGAGAGA GGTATCAATGAGTCTTACAAGAAGAATCTGATGGCCCTAAAGAAATTTGTGATGGTAAATTT CCTCAATGACTCCATAGTGGACCCAGTTGACTCTGAGTGGTTTGGATTTTACAGAAGTGGCC AAGCAAAAGAGACCATTCCCTTCAAGAGACATCCCTGTACACCCAAGATAGGCTGGGCCTG AAAGAGATGGATAATGCAGGGCAATTGGTGTTCCTTGCTACAGAAGGAGACCACCTCCAGCT ATCTGAAGAGTGGTTCTATGCCACATCATCCCTTTTCTAGGGTGAAGATCTAGAGCTGAAT TCCTGCAGCCAGGGGGATCAGCCTCTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGC CCCTCCCCCTTGCTTCCCTTGACCCTGGAAGGTGCCACTCCCCTGCTCTTCTAATAAAAA TGAGGAAATTGCATCACATGTCTGAGTAGGTGTCTATTCTATTCTGGGGGTGGGGTGGGGC AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCT ATGGCTCGAGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCT CACTGAGGCCGGGCGACCAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGA GCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA
172	sp7_PPT1co19 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCCGGGCG TCGGGCGACCTTTGGTGCGCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTTCTGATATCGCGGCCGCTACAGAAACCAAGAAAGCACCATAGCG TTTGCCTTCAACTGCTTCCCACTTGTTTTCTCTTCCAATTTGATTGTGGTTTCTCTCCA GAAGGAACTCCACAGTACTTAGCGTTGGCCACATAGTAGGTTCTCAAATACTTGTTAATAAAA TAAGTTTGTTCGAGAAGCTGGGCAATGATTTCTACAGCTGGAAGAAGAAACATAATGATCT AGTAATTAGCTCATTAAAAATAAACGTTCTTCTTCTCTCAGAGGAGCATTTCCTCAAGGCGT CCTTGATAGCCATCCAAAAAGGCCAAGCTCATCCAATCTTGCCCTAGATTTATGCTAAAATG CAGTTACAATCGATAGGATGACAGAAAACGACAGCACTTATTTAAATATAATAGGCACTTAT TTAAATAGGAGAAGCTGTGACTTCATAGCAAGTGTGGGGTTAGGAACTGGGTGGATAAAC TTGCTGATGCTGTAGATCTTAGCCTCTACATGAGATCATGTGAAAAATCTGAAAGCATTTTA GGTTCTTATGTTTGCAATCAAATAACTGTACACCTTTTAAATTTAAAAAGTACCATGAGGCA CACACACACACTCGCAGGAACTTTTTGGCGTAACAAAACCTAGAAATAGATCTAAAGCTAACT GTAGGACTGAGTCTATTCTAAACTGAAAGCCTGGACATCTGGAGTACCAGGGGGAGATGACG GTTCACGGGCTTCCATAAAAGCAGCTGGCTTTGAATGGAAGGAGCCAAGAGGCCAGCACAGG AGCGGATTTCGTGCTTTTACGGCCATCGAGCCGAACCTCTCGCAAGTCCGTGAGCCGTTAAG GAGGCCCCCAGTCCCCGACCCTTCGCCCCAAGCCCCCTGGGGTCCCCGGGCCTGGTACTCCTT GCCACACGGGAGGGGCGCGGAAGCCGGGGCGGAGGAGGAGCCAACCCCGGGCTGGGCTGAGA CCCGCAGAGGAAGACGCTCTAGGGATTTGTCCCGGACTAGCGAGATGGCAAGGCTGAGGACG GGAGGCTGATTGAGAGGCGAAGGTACACCCTAATCTCAATACAACCTTTGGAGCTAAGCCAG CAATGGTAGAGGGAAGATTCTGCACGTCCCTTCCAGGCGGCCCTCCCCGTACACACCCCCC AACCCGCCCCGACCGGAGCTGAGAGTAATTCATACAAAAGGACTCGCCCCCTGCCTTGGGGAA TCCCAGGGACCGTCGTTAACTCCCCTAACGTAGAACCCAGAGATCGCTGCGTTCCCGCCC CCTCACCCGCCCCGCTCTCGTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCC CGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTGCGCAA TTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAACTGGGAAAGTGATGTCGTGTACTGGC TCCGCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCGGTGAACGTT CTTTTTCGCAACGGGTTTGCCTCCAGAACACAGGTAAAGTCCGCTGTGTGGTTCCTCCGCGG TGCCCTTTTACGGGTTATGGCCCTTGCCTGCTTGAATTACTTCCACGCCCCCTGGCTGCAG TACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTTCAGGCGCTTGC TTAAGGAGCCCCCTTCGCTCGTGCTTGTGAGTTGAGGCTGGCTTGGGCGCTGGGGCGCCGCG

	TGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTTCGATAAGTCTCTAGCCATTTAA AATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGGCCA AGATCTGCACACTGGTATTTTCGGTTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCCTCCC AGCGCACATGTTTCGGCGAGGCGGGGCTGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAG TCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTG GGCGGCAAGGCTGGCCCCGGTCGGCACCAAGTTGCGTGAGCGGAAAAGATGGCCGCTTCCCGGCC CTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCC ACACAAAGGAAAAGGGCCTTTCCGTCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCG GGCGCGTCCAGGCACCTCGATTAGTTCTTTTGGAGTACGTCTCTTAGGTTGGGGGGAGG GGTTTTATGCGATGGAGTTTCCCACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGG CACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCATTCTCAA GCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTTTCAGGTGTCGTGAAAACTACCCCTAA AAGCCAAAACGCGTGCCACCATGGCCTTTCTGTGGCTGCTGTCTGCTGGGCCCTGCTGGGG ACCACCTTTGGCGCCCCCAGCCCCCTCTGCCCCGGTCATCTGGCATGGCATGGGAGACAG CTGCTGCAACCCCTCTCTCCATGGGAGCCATCAAGAAGATGGTGGAGAAGAAGATCCCTGGAA TCTATGTCTCAGCCTGGAGATTGGGAAGACCCTGATGGAGGATGTGGAGAACTCCTTCTTC CTGAATGTGAACAGCCAGGTGACCACAGTGTGCCAGGCCCTGGCCAAGGACCCCAAGCTGCA GCAGGGCTACAATGCCATGGGCTTCTCCCAGGGAGGCCAGTTCTGAGGGCTGTGGCCAGA GATGCCCTCACCTCCCATGATCAACCTCATCTCTGTGGGAGGCCAGCACCAAGGGGTCTTT GGCCTGCCAGGTGTCTGGGGAGAGCAGCCACATCTGTGACTTCATCAGGAAGACCCTGAA TGCTGGTGCCTACGCAAGGTGGTGCAGGAGAGGCTGGTGCAGGCTGAATACTGGCATGACC CCATCAAGGAGGATGTGTACAGGAACACAGCATCTTCTGGCAGACATCAACCAGGAGAGA GGCATCAATGAGAGCTACAAGAAGAACCTGATGGCCCTGAAGAAGTTTGTGATGGTGAAGTT CCTCAATGACTCCATTGTGGACCCTGTGGACTCTGAGTGGTTTGGCTTCTACAGGAGTGGCC AGGCCAAGGAGACCATCCCCCTGCAGGAGACCAGCCTGTACACACAGGACAGACTGGGACTG AAGGAGATGGACAATGCTGGGCAGCTGGTGTCTGGCCACAGAGGGAGACCACCTGCAGCT CTCAGAGGAGTGGTTTTATGCCCACATCATCCCCCTTCTGGGCTGAAGATCTAGAGCTGAAT TCTGAGCCAGGGGGATCAGCCTCTACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGC CCCTCCCCCTTGCTTCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCTTAATAAAAA TGAGGAAATGTCATCACATTGTCTGAGTAGGTGTATTCTATTCTGGGGGGTGGGGTGGGGC AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCAGTGGGCTCT ATGGCTCGAGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCT CACTGAGGCCGGCGACCAAGGTCGCCCCGAGCCCGGGCTTTGCCGGGCGGCCCTCAGTGA GCGAGCGAGCGCGCAGAGAGGAGTGGCCAA
Promoter and signal	
173	EF1α (longer length) TACAGAAAACCAAGAAAGCACCATAGCGTTTGCCTTCAACTGCTTCCCAACTTGTTTTCTCT TCCAATTTGATTGTGGTTTCTCTCCAGAAGGAAGTCCACAGTACTTAGCGTTGGCCACATA GTAGGTTCTCAAATACTTGTTAATAAATAAGTTTGTTCGAGAAGCTGGGCAATGATATTCTA CAGCTGGAAGAAGAAACATAATGATCTAGTAATTAGCTCATTAATAAATAAACGTTCTTCTTT CCTCAGAGGAGCATTTCCCAAGGCCTGCCTTGATAGCCATCCAAAAAGGCCAAGCTCATCCA ATCTTGCCCTAGATTTATGCTAAAATGCAGTTACAATCGATAGGATGACAGAAAACGACAGC ACTTATTTAAATATAATAGGCACTTATTTAAATAGGAGAAGCTGTGACTTCATAGCAAGTGT TGGGGTAGGAACTGGGTGGATAAACTTGCTGATGCTGTAGATCTTAGCCTCTACATGAGA TCATGTGGAAAATCTGAAAGCATTTTAGGTTCTTATGTTTGCAATCAAATAACTGTACACC TTTTAATTTAAAAAGTACCATGAGGCACACACACACTCGCAGGAACTTTTTGGCGTAACA AACTAGAAATAGATCTAAAGCTAACTGTAGGACTGAGTCTATTCTAAACTGAAAGCCTGGA CATCTGGAGTACCAGGGGGAGATGACGTGTTACGGGCTTCCATAAAAGCAGCTGGCTTTGAA TGGAAGGAGCCAAGAGGCCAGCACAGGAGCGGATTTCGTGCTTTTACGGCCATCGAGCCGAA CCTCTCGCAAGTCCGTGAGCCGTTAAGGAGGCCCCAGTCCCAGCCCTCGCCCCAAGCCCC TCGGGGTCCCCGGGCCTGGTACTCCTTGCCACACGGGAGGGGCGCGGAAGCCGGGGCGGAGG AGGAGCCAACCCGGGCTGGGCTGAGACCCGAGAGGAAGACGCTCTAGGGATTGTCCCGG ACTAGCGAGATGGCAAGGCTGAGGACGGGAGGCTGATTGAGAGGCGAAGGTACACCCTAATC TCAATACAACCTTTGGAGCTAAGCCAGCAATGGTAGAGGGAAGATTCTGCACGTCCCTTCCA GGCGGCTCCCCGTACACACCCCCCAACCCGCCCCGACCGGAGCTGAGAGTAATTATAC AAAAGGACTCGCCCCGTGCTTGGGGAATCCCAGGACCGTCGTTAACTCCCCTAAGCTAG AACCAGAGATCGCTGCGTTCCCGCCCCCTCACCCGCCCCGCTCTCGTCATCACTGAGTGGA GAAGAGCATGCGTGAGGCTCCGGTGCCGTGAGTGGGCAGAGCGCACATCGCCACAGTCCC

	CGAGAAGTTGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAA ACTGGGAAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTAT ATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGT AAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGGCCCTTGCCTGCCTT GAATTACTTCCACGCCCCCTGGCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAG TGGGTGGGAGAGTTTCGAGGCCTTGCCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGG CCTGGCTTGGGCGCTGGGGCCGCCGCTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCT GCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTCTG GCAAGATAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCG CGGGCGGCGACGGGGCCCGTGCCTCCAGCGCACATGTTTCGGCGAGGCGGGGCTGCGAGCG CGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCCTGGCCT CGCGCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTGCGCACCAAGTTGCGT GAGCGGAAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGACGCGGCGC TCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGCCTTTCGTCCTCAGCCGT CGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTTTTGA GTACGTCGTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGG GTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCCTTGGAATTTGCCCTTTT TGAGTTTGGATCTTGGTTTCAATCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCAT TTCAGGTGTCGTGAAAACCTACCCCTAAAAGCCAAA
174	sp7 codon optimized ATGGCCTTCTGTGGCTCCTCAGCTGCTGGGCCCTGCTGGGCACCACTTTTGA
175	SPARC codon optimized ATGAGGGCCTGGATATTCTTCCTGCTTTGTCTAGCTGGAAGAGCTCTGGCA

EXAMPLES

[0271] Examples are provided below further illustrating different features of the present invention and methodology for practicing the invention. The provided examples do not limit the claimed invention.

[0272] Example 1: PPT1 Expression Constructs Having Exogenous Signal Peptides

[0273] Intracellular and secreted PPT1 activity were measured in cells transfected with plasmids containing different PPT1 rAAV expression constructs. To assess the depalmitoylation activity of palmitoyl-protein thioesterase (PPT1), 4-methylumbelliferyl-6-thiopalmityl- β -D-glucopyranoside (MUTG) was used as a substrate. Release of MU from MU-6SPalm- β Gal cannot be accomplished by the sole action of PPT1, since this enzyme only hydrolyses the palmitoyl thioester linkage, yielding the non-fluorescent reaction intermediate. Therefore, exogenous β -galactosidase or β -glucosidase was used to hydrolyze the intermediate resulting in release of fluorescent 4 MU that can be detected by fluorescence plate reader. Known amounts of free 4MU were run in parallel and used to plot a standard curve. The level of PPT1 in a sample was either quantified using the standard curve or just expressed as relative fluorescence using (RFU).

[0274] FIG. 1 is a basic schematic of a PPT1 rAAV expression construct with the following nucleic acid regions noted: 5'-ITR, EF-1 α promoter (EF1 α), Kozak sequence (Kozak), signal sequence (SS), mature PPT1 sequence (human PPT1), bovine polyadenylation sequence (bGH-pA), stuffer (stuffer sequence), synthetic polyadenylation sequence (synthetic pA) and 3'-ITR.

[0275] Table 3 provides construct designations for the AAV nucleic acids and references the SEQ ID NOs for the encoded signal peptide sequence, the encoded mature PPT1 amino acid sequence, the encoding polynucleotide signal sequence, the encoding mature PPT1 polynucleotide sequence, and the complete rAAV sequence.

[0276] Table 3

Designation	Signal Peptide SEQ ID NO:	PPT1 SEQ ID NO:	Encoding Signal Peptide SEQ ID NO:	Viral Vector Nucleic Acid SEQ ID NO:	Group
Sp7-F	16	2 X is G	43	144	1
Sp7-P	16	2 X is G	44	154	1
NGF-1	17	2 X is G	48	147	1
NGF-2	17	2 X is V	48	148	1
Insulin	18	2 X is V	49	149	1
SPARC	19	2 X is G	50	145	1
CD33	20	2 X is G	51	150	2
IL34	21	2 X is G	52	151	2
OSM	22	3	53	152	3
tPA	23	4	54	146	3
IL-2	24	2 X is G	55	153	2
Native	28	1	59	155	

[0277] Table 3 characterizes the different constructs as either Group 1, Group 2, or Group 3. Group 1 provides polypeptides where (a) the mature PPT1 N-terminus is substituted and (b) a heterologous signal peptide. Group 2 provides polypeptides where (a) the mature PPT1 N-terminus is substituted and (b) the heterologous signal peptide was modified (derived). Group 3 provides polypeptides comprising (a) a mature PPT1 with an N-terminus addition (tPA) or deletion (OSM) and (b) the heterologous signal peptide was modified.

[0278] FIG. 2A provides a bar graph showing PPT1 enzymatic activity level detected in total cell lysate (intracellular) from PPT1 knock out HeLa cells transfected with different AAV plasmid constructs. The x axis notes the signal sequence present in the full-length PPT1. The data illustrate the ability of the different constructs to express functional PPT1 protein. The values are expressed relative to native (unmodified PPT1) which was set at 100%. The data is

normalized based on the transfection efficiency. Number of experiments n=3-4, except for tPA, n=2. The values are mean $\pm$ S.E.M.

[0279] FIG. 2B provides a bar diagram showing PPT1 enzymatic activity level detected in the secreted media (secreted) collected from PPT1 knock out HeLa cells transfected with AAV plasmids carrying different engineered human PPT1 candidates. The x axis notes the secretion signal present in the full-length PPT1. The data illustrate the ability of the different constructs to express functional PPT1 protein. The values are expressed relative to native (unmodified PPT1) which was set at 100%. The data is normalized based on the transfection efficiency. Number of experiments n=3-4. The values are mean $\pm$ S.E.M. *p<0.05, native vs. others, Mann-Whitney test.

[0280] Example 2: PPT1 Uptake by Un-Transfected Cells

[0281] Uptake of secreted PPT1 as expressed from different constructs was measured using a plasmid containing different rAAV nucleic acid and a TdTomato (a red fluorescent protein) reporter. FIG. 3A presents a schematic of the plasmid, where the following rAAV nucleic acid regions are noted: 5'-ITR, EF-1 α promoter (EF1 α), Kozak sequence (Kozak), signal sequence (SS), mature PPT1 sequence (human PPT1), bovine growth hormone polyadenylation sequence (bGH-pA), stuffer (stuffer sequence), synthetic polyadenylation sequence (synthetic pA) and 3'-ITR; and the following tdTomato regions are noted: UbC (ubiquitin C promoter), 5' UTR, tdTomato, 3' UTR and RbGpA (rabbit beta globin polyadenylation sequence)

[0282] Table 4 provides construct designations for the rAAV nucleic acids and references the SEQ ID NOs for the encoded signal peptide sequence, the encoded mature PPT1 amino acid sequence, the encoding polynucleotide signal sequence, the encoding mature PPT1 polynucleotide sequence, and the complete rAAV sequence.

[0283] Table 4

Designation	Signal Peptide SEQ ID NO:	PPT1 SEQ ID NO:	Encoding Signal Peptide SEQ ID NO:	Viral Vector Nucleic Acid SEQ ID NO:
Sp7-F	16	2 X is G	43	144
Sp7-P	16	2 X is G	44	154
SPARC	19	2 X is G	50	145
tPA	23	4	54	146
Native	28	1	59	155

[0284] PPT1 KO HeLa cells were transfected with different plasmids, fixed 48 hours post-transfection with 4% paraformaldehyde and immunostained with mouse anti-PPT1 (Novus

Biologicals (OTI1F10) and Alexa 488 conjugated anti-mouse secondary antibodies (green fluorescence). Nuclei were labelled with Hoechst and were seen as blue.

[0285] Images were captured in Opera Phenix Plus HCS system (Perkin Elmer) (image data not shown). Cells lighting up green indicated presence of PPT1 and cells lighting up yellow indicated the presence of both PPT1 and TdTomato (plasmid transfected cells). Cells lighting up only green suggested uptake of PPT1 present in media that was secreted by transfected (red) cells. FIG. 3B is a bar graph reflecting the image analyses of the immunostaining and shows the ratio of cells with PPT1 (green) to cells that were transfected (red). Increased ratio for Sp7-F, SPARC and tPA compared to PPT1 with native signal peptide suggests an increased number of PPT1 positive cells resulting from increased uptake of engineered PPT1 carrying exogenous signal sequences. Each circle (average of 30 fields) is independent transfection (from 2-3 exp). The values are mean $\pm$ SD. P values by one way ANOVA, **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$.

[0286] FIG. 3C provides a bar diagram showing results from FIG. 3B, normalized for the native construct. The white circles show independent transfections. Each point (average of 30 fields) is independent transfection (from 2-3 exp). The values are mean $\pm$ SD. P values by one way ANOVA, **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$.

[0287] Example 3: In vitro expression of PPT1 by cortical neurons transduced by recombinant AAVs

[0288] Dose dependent PPT1 expression and secretion from AAV-transduced rat primary cortical neurons was evaluated. The overall design of the recombinant AAV vectors is shown FIG. 1. Recombinant rAAV vectors were designated “Sp7-F.PPT1” or “Dead PPT1”. Sp7-F.PPT1 corresponds to Sp7-F as described in Table 3. Dead PPT1 encodes a mutated full-length PPT1, where the mutation made the protein catalytically inactive.

[0289] Recombinant rAAV vectors were produced by triple transfection using capsids comprising VP1 of SEQ ID NO: 12, VP2 of SEQ ID NO: 13 and VP3 of SEQ ID NO: 14 in human embryonic kidney cells and purified by CsCl purification.

[0290] Cultured neurons were transduced with three MOI (low, $1E+5$; mid, $5E+5$ and high, $1E+6$) of rAAV and PPT1 activity of the culture media was analyzed at 3-, 4- or 6-days post transduction. The results are shown in FIG. 4. Cells transduced with rAAV encoding catalytically dead PPT1 did not show any activity. Untreated or cells treated with diluent served as negative control. Purified recombinant PPT1 was used as a positive control for the assay. Each circle represents the value from independently transduced cells in a single well. Data are mean $\pm$ SD.

[0291] Secreted PPT1 bands from rat primary cortical neurons transduced at mid and high doses were tested for glycosylation on day 6. No PPT1 protein was observed in the vehicle or

untransduced cells. Glycosylated PPT1 was detected in cells transduced with rAAV vector (data not shown).

[0292] Example 4: In vivo Expression of PPT1 by Recombinant AAVs

[0293] Expression and secretion of PPT1 was measured in C57BL6/J mice dosed with rAAV comprising Sp7-1.PPT1 viral vector nucleic acid (SEQ ID NO: 143, see Table 4) or viral vector nucleic acid encoding Dead PPT1 (enzymatically inactive PPT1). Recombinant rAAV vectors were produced using capsids comprising VP1 of SEQ ID NO: 12, VP2 of SEQ ID NO: 13, and VP3 of SEQ ID NO: 14 and purified by CsCl purification. Administration was done directly into the hippocampus through stereotactic injections or through IV route. Table 5 summarizes the experiment protocols.

[0294] Table 5

Group #	Group ID	N	Age (weeks)	Route, Volume	Dose vg/site	Total Dose vg/animal
1	Diluent_IPa	4	6-8	Bilateral, dentate gyrus, 2 μ L	NA	NA
2	SpB_IPa (Dead)	4	6-8	Bilateral, dentate gyrus, 2 μ L	1E10	2E10
3	SpJ_IPa (Sp7-F)	4	6-8	Bilateral, dentate gyrus, 2 μ L	1E10	2E10
4	Diluent_IV	4	6-8	IV, 200 μ L	NA	NA
5	SpB_IV (Dead)	4	6-8	IV, 200 μ L	2E12	2E12
6	SpJ_IV (Sp7-F)	4	6-8	IV, 200 μ L	2E12	2E12

N = number of animals

[0295] Brains were analyzed 6 weeks post injection. Animals dosed only with diluent served as negative control.

[0296] FIG. 5A, FIG. 5B and FIG. 5C illustrate serum PPT1 expression and activity at different time points from mice dosed with rAAV comprising Sp7-F.PPT1 viral vector nucleic acid. FIG. 5A illustrates production of PPT1 in the serum of mice dosed IV with AAV comprising Sp7-F.PPT1 viral vector nucleic acid or Dead PPT1 viral vector nucleic acid. Serum samples were collected at 2-, 4- or 6-weeks post administration and analyzed by automated capillary-based immunoassay (WES; Protein Simple) to detect PPT1. Mice dosed by diluent served as a negative control. The presence of PPT1 (indicated by arrowhead in Fig. 5A) was detected in mice dosed with rAAV comprising Sp7-F.PPT1 viral vector nucleic acid at 2 weeks with further increased level at 4 weeks. The signal saturated at 6 weeks. Purified recombinant PPT1 used as a positive control. FIG. 5B shows serum activity of mice dosed IV. FIG. 5C shows serum activity of mice dosed IPa (in the hippocampus through stereotactic injection). ND, not detected. Data are mean $\pm$ SD. P values by ANOVA *P<0.05, ***P<0.001, ****P<0.0001.

[0297] FIG. 6 illustrates liver PPT1 activity from mice (n=4, shown by circle.) dosed IV with rAAV comprising Sp7-F.PPT1 viral vector nucleic acid. Mice injected with diluent served as negative controls.

[0298] FIG. 7, FIG. 8, FIG. 9 and Fig. 10, illustrate localization, expression, glycosylation, and activity of PPT1 expressed in brain following hippocampal administration of rAAV comprising Sp7-F.PPT1 viral vector nucleic acid in mice. FIG. 7 illustrates immunohistochemical analysis showing increased PPT1 staining (indicated by asterisk) in the hippocampus. FIG 8 shows detection of PPT1 protein in hippocampal protein lysate analyzed by JESS. The doublet seen in low exposure (indicated as low) indicated two glycosylated forms of PPT1. FIG. 9 confirms the PPT1 expressed in hippocampus was glycosylated. Hippocampal protein extract was treated with deglycosylase and analyzed by JESS assay. Reduced molecular weight upon deglycosylase treatment (lanes marked with +) suggested PPT1 was glycosylated. FIG. 10 provides a PPT1 activity assay of hippocampal lysate showing PPT1 expressed in the brain was biologically active as described in Example 1.

[0299] Example 5: CNS Distribution of rAAV Encoding PPT1

[0300] Wild-type C57BL/6J male mouse pups at postnatal day 1 were dosed by bilateral intracerebroventricular (ICV) injection of rAAV comprising Sp7-F.PPT1 viral vector nucleic acid. Recombinant AAV comprising Sp7-F.PPT1 viral vector nucleic acid is described in Example 3. The injections involved administering $1\text{E}+10$ vector genomes (vg) in the low-dose group and $1\text{E}+11$ vg in the high-dose group, with 3 μL per side, totaling 6 μL . At 6 weeks post-injection, brains were extracted, and the left hemisphere was dissected into specific regions: cortex, hippocampus, thalamus, brain stem, and cerebellum. Each of these regions was further divided into two equal sections. One section underwent processing for vector genome analysis, while the other was used for testing transgene (PPT1 protein) expression. The right hemisphere of the brain was fixed and subjected to histological assessment. The spinal cord was dissected coronally into cervical, thoracic, and lumbar segments, with half of each segment used for vector genome analysis. Cerebrospinal fluid (CSF) was collected through cisternae magna puncture and subsequently analyzed for PPT1 levels.

[0301] ICV delivery of viral vector showed a dose dependent distribution in brain and spinal cord of mice. FIG. 11 shows vector genome copy number (VGCN) an indicator of viral transduction in different brain and spinal cord regions. FIG. 12 shows fold change (FC) in PPT1 activity in different brain regions of rAAV-injected animals relative to diluent-injected mice. PPT1 activity in tissue lysate was quantified using MUTG as described in the Example 1. FIG. 13 illustrates FC in PPT1 activity in the cerebrospinal fluid (CSF) of mice subjected to rAAV injection compared to those injected with the diluent. The FC is relative to the average activity

level of the diluent group. FIG. 14 is a scatter plot illustrating the correlation between VGCN and PPT1 enzymatic activity in the brain. The Spearman correlation coefficient was $r=0.66$, $p<0.0001$.

[0302] FIG. 15 shows detection of glycosylated and deglycosylated PPT1 protein in brain lysate analyzed by JESS. Protein extract was treated with deglycosylase and analyzed by JESS assay. The doublet seen in lanes 1-2 indicated two glycosylated forms of PPT1. The reduced molecular weight upon deglycosylase treatment (lanes 3-4) suggested PPT1 was glycosylated.

[0303] Histological assessment was carried out using antibodies to PPT1, followed by detection using fluorophore-conjugated secondary antibodies. Images were captured using fluorescence microscopy. The animal receiving diluent showed baseline PPT1 signal. Animals given the low dose rAAV showed a slightly increased PPT1 expression compared to diluent, while those receiving the high dose exhibited significantly higher expression. PPT1 expression was noticeable throughout the cortex, spanning from rostral to caudal areas, and was also prominent in hippocampus, striatum, and olfactory bulb. The PPT1 signal surrounded the NeuN signal, implying PPT1 was primarily expressed within neurons. (Data not shown.)

[0304] Example 6: Biodistribution of rAAV and Encoded PPT1 in Sheep

[0305] CNS distribution of rAAV comprising Sp7-1.PPT1 (Example 3) and the encoded PPT1 was evaluated in sheep. Nine to ten months old PPT1 $\pm$ male sheep were dosed by unilateral intracerebroventricular (ICV) injection of rAAV comprising Sp7-F.PPT1 (n=4 sheep) or GFP (n=2 sheep) encoding viral vector nucleic acid. A dose of $1E+14$ vg in 2mL was infused. After 8 weeks post-injection, the brains were extracted, coronally slabbed and brain punches (3 mm wide) from the area of interest were analyzed for vector genome and PPT1 protein. The contralateral hemisphere brain slab were fixed and used for histological assessment. The spinal cord was dissected coronally into cervical, thoracic, and lumbar segments, and subjected to vector genome and PPT1 protein analysis. Cerebrospinal fluid (CSF) was analyzed for PPT1 levels. Liver was collected and analyzed for vector genome.

[0306] ICV administration resulted in vector delivery throughout different brain regions (rostral to caudal) and spinal cord in sheep. Vector genome copy number (VGCN) as a measure of viral transduction was measured, across various regions in sheep: cortical areas (FIG. 16A), (FIG. 16B) thalamus, (FIG. 16C) cerebellar cortex, (FIG. 16D) hippocampus, (FIG. 16E) corpus callosum, (FIG. 16F) additional indicated brain areas, (FIG. 16G) spinal cord and (FIG. 16H) liver. Quantitative PCR (qPCR) utilizing a standard curve was employed to quantify the vector genome from frozen tissue DNA. Vertical arrows indicate the sample was from the contralateral brain region. Open circles correspond to sheep dosed with Sp7-F.PPT1 vector, while filled circles depict results from sheep dosed with a vector expressing GFP. Data is provided for the

following regions: Frontal Cortex (FC), Motor Cortex (CM), Somato sensory cortex (SSC), Piriform Cortex (PC), Suprasylvian Gyrus (SSG), Ectolateral Gyrus (EcG), Entolateral Gyrus (EnG), Caudate (Cau), Choroid plexus (Ch Ple), Optic chiasm (Opt chi), Fornix (For), Periaqueductal gray (Periaq G), Olfactory bulb (Ol), bulb Optic nerve (Op) nerve, Hippocampus (HPC), Thalamus (Tha), Corpus callosum (Cca), Cerebellar cortex (Cer ctx) Spinal cord_Cervical (SC), Cer Spinal cord_Thoracic (SC_Tho), and Spinal cord_Lumbar (SC_Lum).

[0307] PPT1 activity assay demonstrates expression and secretion of functional PPT1 in brain and CSF. PPT1 activity was measured using MUTG as described in Example 1. FIGs. 17A-17D illustrate PPT1 activity of Sp7-F.PPT1 (n=4 sheep) or GFP (n=2 sheep) rAAV-injected sheep in the cortex (FIG. 17A), thalamus (FIG. 17B), cerebellar cortex (FIG. 17C), and caudate (FIG. 17D). Each circle is the result of one tissue punch from the brain region. N denotes number of regions from 2 GFP dosed animals and 4 from PPT1 dosed animals. The data is Mean $\pm$ SEM. Statistical analysis by Mann-Whitney U test, *P<0.05, *** P<0.001.

[0308] FIG. 18 provides 95% confidence intervals showing mean PPT1 activity by overall treatment group. After taking the log-transformed activity results of all data in FIGs. 17A-17D and accounting for numbers of punches, and brain region differences in average value, a hypothesis test for average Treatment Type difference was conducted at the 0.05 alpha level. The test resulted in finding a significant difference in average activity levels in PPT1 over the GFP. The average estimated fold change in PPT1 over GFP was 3.9, and with medians of 73 nmoles/mg/hr for the PPT1 vs. 19 nmoles/mg/hr for the GFP. **** P<0.0001, Weighted 2-way ANOVA.

[0309] FIG. 19 is a bar diagram showing percent change in PPT1 activity in the cerebrospinal fluid of (CSF) of sheep dosed with rAAV carrying Sp7-F.PPT1 (n=4 sheep) or GFP (n=2 sheep). The percent changes is relative to mean activity of control (GFP animals). Each circle is one animal. N denotes number of animals. The data is Mean $\pm$ SEM.

[0310] FIG. 20 provides results of a JESS assay detecting the expression of PPT1 in the tissue lysate of the spinal cord of sheep injected with a rAAV vector comprising either PPT1 (sheep 1-4) or GFP (sheep 1-2) encoding nucleic acid.

[0311] FIG. 21 illustrates results showing an elevated level of mean PPT1 activity in the thoracic and lumbar segment of the spinal cord of sheep dosed with rAAV comprising nucleic acid encoding PPT1.

[0312] Immunofluorescence micrographs were taken from sheep brain cerebellar lobes for histological assessment to evaluate transduction and transgene expression. The contralateral hemisphere brain slabs were fixed, and thirty micro thick sections were cryo-sectioned and processed for immunohistochemical analysis. DAPI (4',6-diamidino-2-phenylindole) was used

to produce is a blue-fluorescent DNA stain indicating dsDNA. PPT1 was immunostained using antibodies to purified PPT1 protein followed by subsequent staining by fluorophore conjugated secondary antibodies. Fluorescence microscopy was utilized to visualize and capture the images showing the expression and localization of GFP of PPT1. Widespread expression of transgenes (GFP reporter and PPT1) was observed across cerebellar lobes of sheep brain.

[0313] Example 7: Improvement in Motor Function

[0314] The impact of rAAV providing (1) Sp7-PPT1 or (2) SPARC.PPT1 on PPT1 knock out (KO) mice was evaluated using a rotarod assessment. Rotarod assessment is described, for example, in Deacon J. Vis. Exp. 2013 May 29;(75):e2609 (hereby incorporated by reference herein in its entirety). Sp7-PPT1, SPARC.PPT1 and the rAAV are described in Table 3. KO mice were dosed with rAAV carrying a PPT1 gene, through bilateral ICV injection at postnatal day 1 and assessed at 7 months of age. The rAAV were produced as described in Example 3, and were administered at a dose of 1×10^{11} vg/animal.

[0315] The results are shown in FIG. 22. Untreated (Un) or vehicle (Veh) treated KO mice served as negative control. Native refers to rAAV vector encoding native human PPT1 (SEQ ID NO: 29). Sp7-PPT1 is noted by “1” and SPARC.PPT1 is noted by “2”. Each circle represent one mouse. The bars are mean $\pm$ SEM. One-way ANOVA, Tukey’s post hoc test, * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$. The CNS-targeted rAAV therapy delivered functional human PPT1 in PPT1 KO mice, successfully rescuing coordination and balance.

[0316] Example 8: Improvement in Motor Function and Balance

[0317] The ability of CNS-targeted rAAV vector comprising nucleic acid encoding Sp7-PPT1 (rAAV-Sp7-PPT1) and SPARC.PPT1 (rAAV- SPARC.PPT1) to rescue motor coordination and balance in Ppt1^{-/-} mice as assessed on accelerated rotarod equipment. rAAV-Sp7-PPT1 and rAAV-SPARC.PPT1 were produced as described in Example 3. rAAV-Sp7-PPT1 and rAAV-SPARC.PPT1 vector nucleic acid is described in Table 4.

[0318] Ppt1^{-/-} mice were dosed with rAAV-Sp7-PPT1 or rAAV-SPARC.PPT1 through bilateral intra-cerebro ventricular injections at PND 1 (low dose, 1×10^{11} vg/animal) or PND 1 and 3 (high dose, 3.82×10^{11} vg/animal) and assessed at different timepoints. Mice were tested over 4 trials at each time point. Each session included a training trial of 5 minutes at 4 RPM on the rotarod apparatus (Rotamex, OH). One hour after the training trial, the animals were tested for 3 consecutive accelerating trials of 5 minutes each, during which speed changed over 300 seconds. The inter trial interval was at least 30 minutes. The latency to fall from the accelerating rod was recorded and quantitative analysis was performed. Untreated or vehicle treated Ppt1^{-/-} mice served as negative control. ‘Native’ refers to AAV vectors containing the unmodified

human PPT1 gene. In each group, the number of mice varied from 10 to 18, except for the WT, 11 mo group where n=7. Gender distribution was approximately balanced within each group.

[0319] FIG. 23 illustrates the ability of rAAV- Sp7-F.PPT1 and rAAV- SPARC.PPT1 to improved motor coordination and balance as assessed by latency to fall from an accelerating rod. (See, for example, Kovács and Pearce, Dis Model Mech. 2015 Apr;8(4):351-61): PMC4381334. rAAV- Sp7-F.PPT1 and rAAV-SPARC.PPT1 both reduced latency to fall.

[0320] Example 9: Muscle Strength

[0321] The ability of CNS-targeted rAAV vectors comprising nucleic acid encoding Sp7-PPT1 (rAAV-Sp7-PPT1) or SPARC.PPT1 (rAAV- SPARC.PPT1) to improve muscle strength in Ppt1^{-/-} mice was assessed based on grip strength. rAAV-Sp7-PPT1 and rAAV-SPARC.PPT1 were produced as described in Example 3. rAAV-Sp7-PPT1 and rAAV-SPARC.PPT1 vector nucleic acid is described in Table 4.

[0322] Ppt1^{-/-} mice were dosed with rAAV-Sp7-PPT1 and rAAV-SPARC.PPT1 through bilateral intra-cerebro ventricular injections at PND 1 (low dose, 1.00×10^{11} vg/animal) or PND 1 and 3 (high dose, 3.82×10^{11} vg/animal) and assessed at indicated timepoints. Grip strength was used to assess muscular strength in forelimb muscles (San Diego Instruments, San Diego, CA) with five consecutive trials. Average of all five trials was used for plotting. The animal was lowered toward the platform and gently pulled backwards with consistent force by the experimenter until it released its grip.

[0323] FIG. 24 illustrates the effect of rAAV- Sp7-F.PPT1 and rAAV-SPARC.PPT1 on grip strength at different timepoints. Ppt1^{-/-} mice that received no treatment or were treated with a vehicle served as the negative control group. The term 'Native' pertains to AAV vectors carrying the unmodified human PPT1 gene. In each group, the number of mice at each timepoint varied from 7-18, gender approximately balanced. rAAV- Sp7-F.PPT1 and rAAV- SPARC.PPT1 both improved grip strength.

[0324] Example 10: PPT1 Activity

[0325] Ppt1^{-/-} mice were dosed with rAAV-Sp7-PPT1 or rAAV-SPARC.PPT1 through bilateral intra-cerebro ventricular injections at PND 1 (low dose, 1.00×10^{11} vg/animal) or PND 1 and 3 (high dose, 3.82×10^{11} vg/animal). PPT1 activity at indicated timepoints in the serum was quantified using 4-methylumbelliferyl-6-thiopalmityl- β -D-glucopyranoside (MUTG) as a substrate as described in Example 1. The standard curve (prepared by using known concentrations of 4 MU) was used to determine PPT1 activity.

[0326] The results of the PPT1 assay are shown in FIGs. 25A and 25B. FIG. 25A illustrates activity at different time points. FIG. 25B illustrates activity at 8 months. An increased level of serum PPT1 in AAV-dosed mice, indicating sustained long-term expression of PPT1 was

observed. Untreated or vehicle treated Ppt1^{-/-} mice served as negative control. ‘Native’ refers to AAV vectors containing the unmodified human PPT1 gene.

[0327] Example 11: Cortex, Brainstem, and Cerebellum PPT1 Activity

[0328] Ppt1^{-/-} mice were dosed with rAAV-Sp7-PPT1 or rAAV-SPARC.PPT1 through bilateral intra-cerebro ventricular injections at PND 1 (low dose, 1.00×10^{11} vg/animal). PPT1 activity was measured in the cortex, brainstem and cerebellum using 4-methylumbelliferyl-6-thiopalmityl- β -D-glucopyranoside (MUTG) as a substrate as described in Example 1. The standard curve (prepared by using known concentrations of 4 MU) was used to determine PPT1 activity. Tissues were analyzed at 10 months of age except the untreated Ppt1^{-/-} mouse, which was tested at 8 months of age.

[0329] FIGs. 26A-C provide bar diagrams showing PPT1 activity in the cortex (FIG. 26A), brainstem (FIG. 26B) and cerebellum (FIG. 26C). The elevated PPT1 activity at 10 months of age in rAAV-dosed mice indicates durable expression of PPT1 in Ppt1^{-/-} mouse brain. ‘Native’ refers to AAV vectors containing the unmodified human PPT1 gene. The bars are mean $\pm$ SEM. Each circle represent a mouse.

[0330] Example 12: Codon Optimized PPT1 Encoding Sequences

[0331] Expression of different codon optimized constructs were evaluated in vitro. A plasmid comprising the rAAV nucleic acid is illustrated in FIG. 27. The constructs are summarized in Tables 6 and 7.

[0332] Table 6

	Sp7F sequence SEQ ID NO:	Sp7F optimized SEQ ID NO:	PPT Encoding Sequence SEQ ID NO:
SpF7-co2	43	174	62
SpF7-co3	43	174	63
SpF7-co4	43	174	64
SpF7-co11	43	174	71
SpF7-co14	43	174	74
SpF7-co18	43	174	78
SpF7-co19	43	174	79
SpF7-co20	43	174	83

[0333] Table 7

	SPARC sequence SEQ ID NO:	SPARC optimized SEQ ID NO:	PPT Encoding Sequence SEQ ID NO:
SPARC-co2	50	175	62

SPARC-co3	50	175	63
SPARC-co4	50	175	64
SPARC-co11	50	175	71
SPARC-co14	50	175	74
SPARC-co18	50	175	78
SPARC-co19	50	175	79
SPARC-co20	50	175	83

[0334] Examples of full-length rAAV nucleic acid constructs are provided by SEQ ID NO: 171 and SEQ ID NO: 172. SEQ ID NO: 171 provides the full length rAAV nucleic acid sequence of the construct designated “SpF7-co4” and SEQ ID NO: 172 provides the full length rAAV nucleic acid sequence of the construct designated SpF7-co19. The other rAAV sequence were produced by swapping out the signal and PPT encoding sequences.

[0335] PPT1 knockout HeLa cells were transfected with different AAV plasmids carrying codon-optimized encoding PPT1 sequences. PPT1 expression was assessed by measuring enzymatic activity in the culture media 48 hours post-transfection.

[0336] FIGs. 28A and 28B are bar graphs depicting the expression levels of PPT1 in the media of in vitro-cultured cells transfected with various codon-optimized variants of human PPT1. Different codon-optimized PPT1 cDNA constructs (excluding the signal sequence), labeled as CO followed by the number on the x-axis, were paired with signal sequences (B) Sp7F, SP7F (codon optimized), (C) SpSPARC, or SpSPARC (codon optimized) and cloned downstream of long EF1a promoter, as illustrated in FIG. 27. Each circle represents an independently transfected well of a 96-well plate. The values are presented as mean $\pm$ SEM.

[0337] Example 13: Survival Data

[0338] Ppt1^{-/-} mice were dosed with rAAV comprising nucleic acid encoding Sp7F.PPT1 or SPARC.PPT1, through bilateral intra-cerebro ventricular injections at PND 1 (low dose, 1×10^{11} vg/animal) or PND 1 and 3 (high dose, 3.82×10^{11} vg/animal). The rAAV Sp7F.PPT1 and SPARC.PPT1 constructs are summarized in Table 4. The number of mice enrolled in the study were, WT, n=20; KO, Un=16; KO, Veh and Native Hi, n=17. The remaining treatment groups comprised 18 mice each. Untreated or vehicle-treated Ppt1^{-/-} mice served as negative control. ‘Native’ refers to AAV vectors containing the unmodified human PPT1 gene.

[0339] The survival data is shown in FIG. 29. Mouse data was fitted with a Cox Proportional Hazards model to determine if survival probability differed between treatment groups. The KO, Veh group had significantly lower survival than WT and all AAV treated groups (all $p < 0.01$). The KO, Veh, and KO, Un groups were not different from each other ($p = 0.22$). AAV=adeno-

associated virus; CNS=central nervous system; lo=low dose (1×10^{11} vg/animal); hi= high dose (3.82×10^{11} vg/animal); KO= Ppt1^{-/-}; PND=post-natal day; Un=untreated; Veh=vehicle; WT=wild-type.

[0340] Example 14: Brain Weight Loss

[0341] Ppt1^{-/-} mice were dosed with rAAV comprising nucleic acid encoding Sp7F.PPT1 or SPARC.PPT1 through bilateral intra-cerebro ventricular injections at PND 1 (low dose, 1×10^{11} vg/animal) or PND 1 and 3 (high dose, 3.82×10^{11}). The rAAV Sp7F.PPT1 and SPARC.PPT1 constructs are summarized in Table 4. Untreated or vehicle-treated Ppt1^{-/-} mice served as negative control. 'Native' refers to AAV vectors containing the unmodified human PPT1 gene. The mice were euthanized at 8-10 months of age, brains were collected, and weights were recorded. Number of mice, n=6/group except for KO, Veh, n=4 and Lead 1, hi, n=9.

[0342] FIG. 30 is a bar diagram illustrating the effect of rAAV comprising nucleic acid encoding Sp7F.PPT1 or SPARC.PPT1 on brain weight in Ppt1^{-/-} mice. Sp7F.PPT1 and SPARC.PPT1 at the low and high doses inhibited loss of brain weight in the Ppt1^{-/-} mice.

[0343] While the invention has been described and illustrated with reference to certain particular embodiments thereof, those skilled in the art will appreciate that various adaptations, changes, modifications, substitutions, deletions, or additions of procedures and protocols may be made without departing from the spirit and scope of the invention.

CLAIMS

I/we claim:

1. A polynucleotide comprising a nucleic acid sequence encoding a palmitoyl-protein thioesterase-1(PPT1) polypeptide, wherein said PPT1 polypeptide comprises a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1, wherein:

(a) said PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or

(b) said PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or

(c) said PPT1 sequence comprises the amino acid sequence leucine-glutamine histidine-leucine at its N-terminus; and/or

(d) said nucleic acid sequence comprises a PPT1 encoding sequence having at least 85% identity to any of SEQ ID NOs: 61-94.

2. The polynucleotide of claim 1, wherein said PPT1 polypeptide further comprises said signal sequence comprising the sequence of any of SEQ ID NOs: 16-27.

3. The polynucleotide of claim 2, wherein said nucleic acid comprises a signal encoding sequence of any of SEQ ID NOs: 43-58.

4. The polynucleotide of claim 2, wherein said polypeptide comprises a signal sequence of any of SEQ ID NOs: 16-21 and 24-27.

5. The polynucleotide of claim 4, wherein said signal sequence comprises the sequence of any of SEQ ID NOs: 16 or 19.

6. The polynucleotide of claim 5, wherein said signal sequence comprises the sequence of SEQ ID NO: 16 and said nucleic acid sequence comprises the signal encoding sequence of SEQ ID NO: 43; or said signal peptide comprises the sequence of SEQ ID NO: 19 and said nucleic acid sequence comprises the signal encoding sequence of SEQ ID NO: 50.

7. The polynucleotide of claim 2, wherein said polypeptide comprises the signal sequence of SEQ ID NO: 23.

8. The polynucleotide of claim 7, wherein said nucleic acid sequence comprises the signal encoding sequence of SEQ ID NO: 54.

9. The polynucleotide of any one of claims 1-6, wherein said PPT1 amino acid sequence comprises a G, V, or L substitution for aspartic acid D at its amino terminus and said PPT1 amino acid sequence has at least 97% identity to the sequence of SEQ ID NO: 1.

10. The polynucleotide of claim 9, wherein said PPT1 amino acid sequence comprises the sequence of SEQ ID NO: 2, wherein X is G.

11. The polynucleotide of any one of claims 1-3, 7 or 8, wherein said PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus and said PPT1 amino acid sequence has at least 97% identity to the sequence of SEQ ID NO: 1.
12. The polynucleotide of claim 11, wherein said PPT1 sequence comprises the sequence of SEQ ID NO: 4.
13. The polynucleotide of any one of claims 1-12, wherein said nucleic acid comprises a sequence having at least 85% identity to any of SEQ ID NOs: 61-94.
14. The polynucleotide of claim 1, wherein said PPT1 polypeptide comprises a sequence having at least 99% identity to any of SEQ ID NOs: 31-42.
15. The polynucleotide of claim 14, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 or SEQ ID NO: 34.
16. The polynucleotide of claim 15, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 wherein X is G and said nucleic acid comprises a sequence having at least 85% identity with the sequence of any of SEQ ID NOs: 107-125 and 168.
17. The polynucleotide of claim 16, wherein said nucleic acid comprises a sequence having at least 95% identity with the sequence of any of SEQ ID NOs: 107-125 and 168.
18. The polynucleotide of claim 17, wherein said nucleic acid comprises the sequence of any of SEQ ID NOs: 107-125 and 168.
19. The polynucleotide of claim 15, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 34 wherein X is G and said nucleic acid comprises a sequence having at 85% identity with the sequence of any of SEQ ID NOs: 126-140 and 161-167.
20. The polynucleotide of claim 19, wherein said nucleic acid comprises a sequence having at 95% identity with the sequence of any of SEQ ID NOs: 126-140 and 161-167.
21. The polynucleotide of claim 20, wherein said PPT1 encoding sequence comprises any of SEQ ID NOs: 126-140 and 161-167.
22. The polynucleotide of claim 1, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 38.
23. A palmitoyl-protein thioesterase-1 (PPT1) polypeptide comprising a PPT1 amino acid sequence having at least 95% identity to the sequence of SEQ ID NO: 1, wherein:
- (a) said PPT1 polypeptide further comprises a signal sequence of any of SEQ ID NOs: 16-27 or a variant thereof having a one amino acid substitution, deletion or insertion; and/or
 - (b) said PPT1 amino acid sequence comprises a glycine (G), valine (V) or leucine (L) substitution for aspartic acid (D) at its amino terminus; and/or
 - (c) said PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus.

24. The polypeptide of claim 23, wherein said PPT1 polypeptide further comprises a signal sequence comprising the sequence of any of SEQ ID NOs: 16-27.
25. The polypeptide of claim 24, wherein said polypeptide comprises a signal sequence of any of SEQ ID NOs: 16-21 and 24-27.
26. The polypeptide of claim 25, wherein said signal sequence comprises the sequence of any of SEQ ID NOs: 16 or 19.
27. The polypeptide of claim 24, wherein said polypeptide comprises a signal sequence of SEQ ID NO: 23.
28. The polypeptide of any one of claims 23-27, wherein said PPT1 amino acid sequence comprises a G, V, or L substitution for aspartic acid D at its amino terminus and said PPT1 amino acid sequence has at least 97% identity to the sequence of SEQ ID NO: 1.
29. The polypeptide of claim 28, wherein said PPT1 amino acid sequence comprises the sequence of SEQ ID NO: 2, wherein X is G.
30. The polypeptide of 23, wherein said PPT1 sequence comprises the amino acid sequence leucine-glutamine-histidine-leucine at its N-terminus and said PPT1 amino acid sequence has at least 97% identity to the sequence of SEQ ID NO: 1.
31. The polypeptide of claim 30, wherein said PPT1 sequence comprises SEQ ID NO: 4.
32. The polypeptide of claim 23, wherein said PPT1 polypeptide comprises a sequence with at least 99% identity to any of SEQ ID NOs: 31-42.
33. The polypeptide of claim 32, wherein said PPT1 polypeptide comprises the sequence of SEQ ID NO: 31 wherein X is G, SEQ ID NO: 34 wherein X is G, or SEQ ID NO: 38.
34. A polynucleotide comprising a PPT1 polypeptide encoding nucleic acid sequence, wherein said PPT1 encoding nucleic acid sequence encodes the PPT1 polypeptide of any one of claims 23-33.
35. A polynucleotide comprising two or more exons together encoding for the PPT1 polypeptide of any one of claims 23-34, and one or more introns.
36. The polynucleotide of any one of claims 1-22, 34 or 35, wherein said polynucleotide is an expression cassette comprising one or more expression control elements operably linked to said nucleic acid encoding said PPT1 polypeptide.
37. The polynucleotide of claim 36, wherein said nucleic acid encoding said PPT1 polypeptide is operably linked to an upstream promoter and a downstream polyadenylation signal.
38. The polynucleotide of claim 36, wherein said expression cassette comprises 5' to 3', operably linked to said nucleic acid encoding said PPT1 polypeptide, a promoter, a kozak sequence, said PPT1 polypeptide encoding nucleic acid sequence, and a polyadenylation signal.
39. The polynucleotide of claims 37 or 38, wherein said promoter comprises a sequence

having at least 95% identity to the sequence of SEQ ID NOs: 5 or 173.

40. The polynucleotide of any one of claims 37-39, wherein said polyadenylation signal operably linked to the PPT1 encoding nucleotide sequence comprises a sequence having at least 95% identity to the sequence of SEQ ID NO: 6.

41. The polynucleotide of any one of claims 36-40, wherein said expression cassette comprises a nucleotide sequence having at least 95% identity to the sequence of any of SEQ ID NOs: 141-143, 169 and 170.

42. The polynucleotide of any one of claims 1-22 and 34-41, wherein said polynucleotide is DNA.

43. A recombinant viral vector nucleic acid comprising the polynucleotide of any one of claims 1-22 and 34-42, and 5' and/or 3' viral elements providing for viral packaging and replication.

44. The recombinant viral vector nucleic acid of claim 43, wherein said recombinant viral vector nucleic acid is DNA and comprises an adeno-associated virus (AAV) inverted repeat (ITR) flanking the 5' terminus of said polynucleotide and an AAV ITR flanking the 3' terminus of said polynucleotide.

45. The recombinant viral vector nucleic acid of claim 44, wherein said recombinant viral vector nucleic acid comprises the 5' ITR and the 3' ITR of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.10, AAVrh.74 or AAV3B.

46. The recombinant viral vector nucleic acid of claim 44, wherein said 5' ITR comprises a sequence having at least 95% identity to the sequence of SEQ ID NO: 8, and said 3' ITR comprises a sequence having at least 95% identity to the sequence of SEQ ID NO: 9.

47. The recombinant viral vector nucleic acid of any one of claims 43-46, further comprising a polyadenylation sequence operably linked to said 3' ITR.

48. The recombinant viral vector nucleic acid of any one of claims 43-47, further comprising one or more stuffer sequences.

49. The recombinant viral vector nucleic acid of any one of claims 43-48, wherein said recombinant viral vector nucleic acid comprises a sequence having at least 95% identity to the sequence of any of SEQ ID NOs: 144-154, 171 and 172.

50. A gene delivery vehicle comprising a viral or a non-viral vector and the polynucleotide of any one of claims 1-22 and 34-42, or the recombinant viral vector nucleic acid of any one of claims 43-49.

51. The gene delivery vehicle of claim 50, wherein said gene delivery vehicle is a viral vector.

52. The gene delivery vehicle of claim 51, wherein said viral vector is a recombinant AAV vector, a recombinant lentivirus vector, or a recombinant adenovirus vector.

53. The gene delivery vehicle of claim 52, wherein said viral vector is a recombinant AAV vector, and said recombinant AAV vector comprises a capsid comprising a VP1, VP2 or VP3 having at least 90% identity to a VP1, VP2 or VP3 sequence of any of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.74, AAV3B, AAV-2i8, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, AAV1/rh.10, SEQ ID NO: 12 or SEQ ID NO: 15.

54. The gene delivery vehicle of claim 53, wherein said capsid is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAVrh.74, AAV3B, AAV-2i8, AAVrh.10, AAVrh.8, AAVHSC, AAV-B1, AAV-AS, or AAV1/rh.10 capsid; or said capsid comprises VP1 of SEQ ID NO: 12 or SEQ ID NO: 15.

55. The gene delivery vehicle of claim 54, wherein the capsid comprises VP1 comprising a sequence of SEQ ID NO: 12, VP2 comprising a sequence of SEQ ID NO: 13, and VP3 comprising a sequence of SEQ ID NO: 14.

56. The gene delivery vehicle of claim 50, wherein said gene delivery vehicle is the non-viral vector.

57. The gene delivery vehicle of claim 56, wherein said non-viral vector is a nanoparticle selected from the group consisting of a lipid nanoparticle (LNP), a polymeric nanoparticle, a lipid polymer nanoparticle (LPNP), a protein or peptide-based nanoparticle, a DNA dendrimer or DNA-based nanocarrier, a carbon nanotube, a microparticle, a microcapsule, an inorganic nanoparticle, a peptide cage nanoparticle, and an exosome.

58. The gene delivery of vehicle of claim 56, wherein said non-viral vector is an LNP or LPNP.

59. A pharmaceutical composition comprising the polynucleotide any one of claims 1-22, and 34-42, the PPT1 polypeptide of any one of claims 23-33, the recombinant viral vector nucleic acid of any one of claims 43-49, or the gene delivery vehicle of any one of claims 50-58, and a pharmaceutically acceptable carrier.

60. A method of increasing PPT1 in a subject comprising administering to the subject the polynucleotide any one of claims 1-22 and 34-42, the PPT1 polypeptide of any one of claims 23-33, the recombinant viral vector nucleic acid of any one of claims 43-49, the gene delivery vehicle of any one of claims 50-58, or the pharmaceutical composition of claim 59.

61. A method of treating Neuronal Ceroid Lipofuscinosis 1 in a subject, comprising administering to the subject the polynucleotide any one of claims 1-22 and 34-42, the PPT1 polypeptide of any one of claims 23-33, the recombinant viral vector nucleic acid of any one of

claims 43-49, the gene delivery vehicle of any one of claims 50-58, or the pharmaceutical composition of claim 59.

62. The method of claims 60 or 61, wherein said administering comprises intraparenchymal, intracisternal or intraventricular administration.

63. The method of claim 62, wherein said administration is intraventricular and results in significant rAAV delivery to at least the frontal cortex, parietal cortex, temporal cortex, occipital cortex, thalamus, cerebellar cortex, hippocampus, corpus callosum, spinal cord, caudate, choroid plexus, optic chiasm, fornix, periaqueductal gray, olfactory bulb, and optic nerve.

64. The method of claims 60 or 61, wherein said administering comprises initial administration outside the central nervous system (CNS).

65. The method of claim 63, wherein said subject is a sheep.

66. The method of any one of claims 60-62, wherein said administration is systemic.

67. The method of any one of claims 60-63, wherein said subject is a human.

68. An AAV vector genome plasmid comprising the recombinant viral vector nucleic acid of any one of claims 43-49.

69. The AAV genome plasmid of claim 68, wherein said plasmid lacks *rep* and *cap* genes.

70. A method of producing a rAAV vector comprising the step of culturing an rAAV production cell line comprising rAAV helper virus activity, wherein the genome of said production cell comprises the recombinant viral vector nucleic acid of any one of claims 43-49, a *rep* gene and a *cap* gene, wherein said rAAV vector is produced.

71. A method of producing rAAV vector comprising the step of culturing an rAAV permissive cell comprising the AAV genome plasmid of claims 68 or 69, wherein said rAAV permissive cell further comprises (a) *rep* and *cap* genes provided either as part of the cell genome and/or by one or more separate plasmids, and (b) helper virus activity provided by the cell genome and/or provided by one or more separate plasmids.

72. The method of claim 71, wherein said rAAV permissive cell is a packaging cell, wherein the genome of said packaging comprises a *cap* gene and a *rep* gene.

73. The method of claim 71, wherein either (a) said *rep* gene, said *cap* gene and said helper activity is provided in a single plasmid or (b) said *rep* gene and said *cap* gene is provided by a *rep/cap* plasmid and said helper activity is provided by a helper plasmid.

74. A method of obtaining an rAAV vector comprising the steps of (a) producing the rAAV using the method of anyone of claims 70-73 and (b) purifying the rAAV.