

GENE THERAPY FOR NEURODEGENERATIVE DISORDERS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of priority under 35 U.S.C. §119(e) of U.S. Serial No. 62/942,059, filed November 29, 2019 and U.S. Serial No. 63/004,422, filed April 2, 2020, the entire contents of both are incorporated herein by reference in their entireties.

INCORPORATION OF SEQUENCE LISTING

[0002] The material in the accompanying sequence listing is hereby incorporated by reference into this application. The accompanying sequence listing text file, name APRES1110_2WO_Sequence_Listing.txt, was created on November 24, 2020, and is 105 kb. The file can be accessed using Microsoft Word on a computer that uses Windows OS

BACKGROUND OF THE INVENTION

TECHNICAL FIELD

[0003] The present disclosure relates generally to gene therapy for neurodegenerative disorders, and more specifically to polynucleotides and expression cassettes for delivery of therapeutic genes. In particular embodiments, a therapeutic gene is presenilin-1.

BACKGROUND INFORMATION

[0004] Alzheimer's disease (AD), also referred to as Alzheimer's, is a chronic neurodegenerative disease that is the cause of a majority of neurodegenerative dementia. Symptoms include difficulty with memory, problems with language, disorientation, mood swings, loss of motivation, and other behavioral problems such as withdrawal from family and society. Bodily functions are gradually lost, ultimately leading to death. Although the disease can last for more than ten years, the average life expectancy is three to nine years following diagnosis.

[0005] The disease is accompanied by a variety of neuropathologic features principal among which are the presence in the brain of amyloid plaques and the neurofibrillary degeneration of neurons. The etiology of this disease is complex, although in about 10% of AD cases it appears to be familial, being inherited as an autosomal dominant trait. Among these inherited forms of AD, there are at least four different genes, some of whose mutants confer inherited susceptibility to this disease. The $\epsilon 4$ (Cys112Arg) allelic polymorphism of the Apolipoprotein E (ApoE) gene has been associated with AD in a significant proportion of cases with onset late in life. A very small proportion of familial cases with onset before age 65 years have been associated with mutations in the β -amyloid precursor protein (APP) gene on chromosome 21. A third locus associated with a larger proportion of cases with early onset AD

has recently been mapped to chromosome 14q24.3. The majority (70-80%) of heritable, early-onset AD maps to chromosome 14 and appears to result from one of more than 20 different amino-acid substitutions within the protein presenilin-1 (PS1). A similar, although less common, AD-risk locus on chromosome 1 encodes a protein, presenilin-2 (PS-2, highly homologous to PS-1).

SUMMARY OF THE INVENTION

[0006] The present disclosure relates to polynucleotides and nucleic acid expression cassettes encoding presenilin-1 (PSEN-1) for the treatment of neurodegenerative disorders.

[0007] In some embodiments, the disclosure provides an isolated cDNA or a hybrid genomic/cDNA that encodes the naturally occurring human presenilin-1 amino acid sequence set forth in either SEQ ID NO: 12 (isoform X1) or SEQ ID NO:14 (isoform X2), wherein as compared to the cDNA corresponding to the naturally occurring PSEN-1 X1 isoform coding sequence (SEQ ID NO:15) or PSEN-1 X2 isoform coding sequence (SEQ ID NO: 13), the isolated cDNA or hybrid genomic/cDNA comprises codon optimization changes in at least 25% of the tolerant codons. In some aspects of these embodiments, no intolerant codons are altered in the PSEN-1 coding sequence in the isolated cDNA or hybrid genomic/cDNA. In some aspects of these embodiments, the isolated cDNA or hybrid genomic/cDNA comprises codon optimization changes in at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99% or all of the tolerant codons in the PSEN-1 coding sequence.

[0008] In some embodiments, the disclosure provides an isolated cDNA or hybrid genomic/cDNA that encodes the naturally occurring human presenilin-1 amino acid sequence set forth in either SEQ ID NO: 12 (isoform X1) or SEQ ID NO:14 (isoform X2), wherein the isolated cDNA or hybrid genomic/cDNA comprises 20 or less CpG dinucleotides. This is a reduction as compared to SEQ ID NO:1 or SEQ ID NO:13, each of which has 23 CpG dinucleotides in the PSEN1 open reading frame. It will be understood that in these embodiments, the replacement of any CpG dinucleotide present in SEQ ID NO:1 or SEQ ID NO:13 must be achieved by replacing either the cytosine or the guanine (or both) with another nucleotide that, due to the redundancy of the genetic code, does not alter the amino acid encoded by the codon containing the replaced nucleotide. In other words, any nucleotide substitution utilized to remove a CpG dinucleotide must preserve the amino acid sequence encoded by SEQ ID NO:1 or SEQ ID NO:13. In some aspects of these embodiments, the isolated cDNA or hybrid genomic/cDNA comprises less than 15, less than 12, less than 10, less than 9, less than 8, less than 7, less than 6, less than 5, less than 4, less than 3, one, or none of

the CpG dinucleotides present in SEQ ID NO:1 or SEQ ID NO:13. In some aspects of these embodiments, all intolerant codons present in SEQ ID NO:1 or SEQ ID NO:13 are preserved in the isolated cDNA or artificial gene that has a reduced number of CpG dinucleotides.

[0009] In some embodiments, the isolated cDNA or hybrid genomic/cDNA comprises codon optimization changes in at least 25% of the tolerant codons present in SEQ ID NO:1 or SEQ ID NO:13 and comprises 20 or less CpG dinucleotides. In some aspects of these embodiments, the isolated cDNA or hybrid genomic/cDNA comprises codon optimization changes in at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99% or all of the tolerant codons in the PSEN-1 coding sequence. In some aspects of these embodiments, the isolated cDNA or hybrid genomic/cDNA comprises less than 15, less than 12, less than 10, less than 9, less than 8, less than 7, less than 6, less than 5, less than 4, less than 3, one, or no CpG dinucleotides. In some aspects of these embodiments, all intolerant codons present in SEQ ID NO:1 or SEQ ID NO:13 are preserved in the isolated cDNA or artificial gene that has a reduced number of CpG dinucleotides.

[0010] In some embodiments, the disclosure provides a hybrid genomic/cDNA that comprises: 1) at least a portion or all of naturally occurring PSEN-1 exon 3 with two alternate splice donor sites as used to produce the cDNAs in SEQ ID NO:1 and SEQ ID NO: 13; 2) at least a portion of naturally occurring PSEN-1 intron 3, wherein the portion of intron 3 comprises a splice acceptor site; and 3) a nucleotide sequence capable of encoding upon expression both SEQ ID NO: 12 (isoform X1) and SEQ ID NO:14 (isoform X2) due to the use of the alternate splice donor sites, wherein the hybrid genomic/cDNA: a) includes less than 70% of naturally occurring PSEN-1 intron 3; b) includes less than 70% of naturally occurring PSEN-1 intron 4; c) lacks at least one of naturally occurring PSEN-1 introns 5, 6, 7, 8, or 9; and/or d) is less than 4.4 kb in length. In some aspects of these embodiments, the portion of the hybrid genomic/cDNA that encodes the naturally occurring human presenilin-1 amino acid sequence set forth in either SEQ ID NO: 12 (isoform X1) or SEQ ID NO:14 (isoform X2), comprises codon optimization changes in at least 25% of the tolerant codons wherein as compared to the cDNA corresponding to the naturally occurring PSEN-1 X1 isoform coding sequence (SEQ ID NO:15), or the PSEN-1 X2 isoform sequence (SEQ ID NO:13). In some aspects of these embodiments, the hybrid genomic/cDNA that encodes the naturally occurring human presenilin-1 amino acid sequence set forth in either SEQ ID NO: 12 (isoform X1) or SEQ ID NO:14 (isoform X2), comprises less than 50 CpG dinucleotides throughout the nucleotide sequence. In some embodiments, the hybrid genomic/cDNA comprises less than

20 CpG dinucleotides in the PSEN-1 coding sequence. In some embodiments, the hybrid genomic/cDNA comprises codon optimization changes in at least 30% of the tolerant codons in SEQ ID NO:15 or SEQ ID NO:13; less than 50 CpG dinucleotides throughout the nucleotide sequence; less than 20 CpG dinucleotides in the PSEN-1 coding sequence; and no changes in any intolerant codons in SEQ ID NO:1 or SEQ ID NO:13. In some more specific versions of any of the aspects set forth in this paragraph, the hybrid genomic/cDNA comprises less than 40, less than 30, less than 20, less than 15, less than 12, less than 10, less than 9, less than 8, less than 7, less than 6, less than 5, less than 4, less than 3, one, or no CpG dinucleotides throughout the nucleotide sequence. In some more specific versions of any of the aspects set forth in this paragraph, the hybrid genomic/cDNA comprises less than 15, less than 12, less than 10, less than 9, less than 8, less than 7, less than 6, less than 5, less than 4, less than 3, one, or no CpG dinucleotides in the PSEN-1 coding region. In some more specific versions of any of the aspects set forth in this paragraph, the hybrid genomic/cDNA comprises codon optimization changes in at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99% or all of the tolerant codons in the PSEN-1 coding sequence in SEQ ID NO:15 or SEQ ID NO:13.

[0011] Described herein, in some embodiments, are isolated polynucleotides set forth in SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8; or polynucleotides having at least 95% identity to SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8. In some embodiments, the isolated cDNA or hybrid genomic/cDNA is SEQ ID NO:6 (a cDNA), SEQ ID NO:7 (a cDNA), or SEQ ID NO:8 (a hybrid genomic/cDNA); or a polynucleotide having at least 95% identity to SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8 and encoding the same amino acid sequence as SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8, respectively. In some embodiments, the polynucleotide having at least 95% identity to SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8 and encoding the same amino acid sequence as SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8, respectively, maintains the intolerant codons present therein and either (1) maintains all optimized codons present therein; or (2) replaces one or more optimized codons therein with other codons that encode the same amino acid and are also optimized.

[0012] Described herein, in some embodiments, are isolated polynucleotides set forth in SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39; or polynucleotides having at least 95% identity to SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39 and encoding the same amino acid sequence as encoded by each of SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, and SEQ ID NO:39. In some aspects of these embodiments, the polynucleotide having at least 95% identity to SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38,

or SEQ ID NO:39 encodes the same amino acid sequence as SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, and SEQ ID NO:39, maintains the intolerant codons present therein and either (1) maintains all optimized codons present therein; or (2) replaces one or more optimized codons therein with other codons that encode the same amino acid and are also optimized. In some aspects of these embodiments, the isolated polynucleotide is SEQ ID NO:36. In some aspects of these embodiments, the isolated polynucleotide is SEQ ID NO:37. In alternate aspects of these embodiment, the isolated polynucleotide is SEQ ID NO:38. In alternate aspects of these embodiment, the isolated polynucleotide is SEQ ID NO:39.

[0013] In certain embodiments, the disclosure provides nucleic acid expression cassettes comprising any of the cDNA or hybrid genomic/cDNA polynucleotides encoding presenilin 1 set forth above.

[0014] In certain embodiments, nucleic acid expression cassette comprises sequences encoding a 5' AAV inverted terminal repeat sequence (ITR), a promoter with an optional enhancer, a polynucleotide encoding presenilin 1 and a 3' AAV ITR. In certain embodiments, a nucleic acid expression cassette comprises a full-length AAV 5' inverted terminal repeat (ITR) and a full-length 3' ITR. In certain embodiments, a nucleic acid expression cassette comprises a shortened version of the 5' ITR, termed Δ ITR, has been described in which the D-sequence and terminal resolution site (trs) are deleted (X. S. Wang, *et al.*, *J Mol Biol* 250:573–580, 1995; X. S. Wang, *et al.*, *J Virol* 70:1668–1677, 1996); C. Ling *et al.*, *J Virol*. Jan 2015, 89 (2) 952-961; DOI: 10.1128/JVI.02581-14). In certain embodiments, the ITRs are selected from a source which differs from the AAV source of the capsid. For example, AAV2 ITRs may be selected for use with an AAV capsid having a particular efficiency for a selected cellular receptor, target tissue or viral target. In certain embodiments, the AAV capsid is from AAV9. In one embodiment, the ITR sequences from AAV2, or the deleted version thereof (Δ ITR), however, ITRs from other AAV sources maybe selected. Where the source of the ITRs is from one AAV serotype and the AAV capsid is from another AAV serotype, the resulting vector may be termed pseudotyped. In certain embodiments, the ITRs and capsids are from AAV9. In certain embodiments, the ITRs are from single stranded or self-complementary AAV vectors. In certain embodiments, the ITRs may be part of the expression cassette, while in alternate embodiments, the ITRs may be part of the vector into which the expression cassette is cloned

[0015] In some embodiments, the one or more regulatory elements comprise a Kozak translation initiation signal such as a polynucleotide set forth in SEQ ID NO:5, or a nucleotide sequence having at least an 80% sequence identity to SEQ ID NO: 5.

[0016] In some embodiments, the one or more regulatory elements comprise a chromatin insulator sequence, such as the polynucleotide set forth in SEQ ID NO:4, or a nucleotide sequence having at least a 95% sequence identity to SEQ ID NO: 4.

[0017] In some embodiments, the one or more regulatory elements comprise promoter. In some aspects of these embodiments, the promoter is a neuron-specific promoter. A neuron-specific promoter can comprise (i) a polynucleotide set forth in SEQ ID NO:2; (ii) a polynucleotide set forth in SEQ ID NO:3; (iii) a functional fragment of SEQ ID NO:2 or SEQ ID NO:3; or (iv) polynucleotide with at least 95% identity to (i), (ii), or (iii). In alternate aspects of these embodiments, the promoter is selected from CAG (SEQ ID NO: 23), CBA (SEQ ID NO: 24), UBC (SEQ ID NO: 25), PGK (SEQ ID NO: 26), PKC, EF1a (SEQ ID NO: 27), GUSB, CMV (SEQ ID NO: 28), NSE (SEQ ID NO: 29), PDGF, desmin, MCK, MeCP2 (SEQ ID NO: 30), GFAP (SEQ ID NO: 31), CaMKII or MBP.

[0018] In some embodiments, the one or more regulatory elements comprise at least one mRNA stability element. The at least one mRNA stability element can comprise (i) a polynucleotide set forth in SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11; (ii) a functional variant of SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii). In some aspects of these embodiments, the nucleic acid expression cassette comprises a mRNA stability element located 5' of the open reading frame of the polynucleotide encoding PSEN1; and a mRNA stability element located 3' of the polyadenylation signal. In certain embodiments, the nucleic acid expression cassette comprises one or more polyadenylation enhancer elements, such as, for example Human growth hormone (hGH) polyadenylation signal sequences, rabbit beta-globin (rBG) polyadenylation signal sequences, SV40 polyadenylation signal sequences or bovine growth hormone (BGH) polyadenylation signal sequences.

[0019] In some embodiments, the one or more regulatory elements comprise one, two or three micro RNA ("miRNA" or "miR") binding sites to suppress expression of the encoded PSEN-1 in dorsal root ganglia. MicroRNAs are 19-25 nucleotide noncoding RNAs that bind to miRNA binding sites and down-regulate gene expression either by reducing nucleic acid molecule stability or by inhibiting translation. In some aspects of these embodiments, each miRNA binding site is independently selected from a binding site for any of the following miRNAs: miRNA-1914, miR1181, miR3918, miR939, miR324, miR650, MiR29C, or miR2277. In some aspects of these embodiments, the miRNA binding site(s) are located 3' to the coding sequence of the viral genome.

[0020] Other embodiments provide vectors comprising the nucleic acid expression cassettes provided herein. A vector can be a viral vector, such as an adeno-associated virus (AAV) vector, a retroviral vector, a lentiviral vector, or an adenoviral vector. An AAV vector can be AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVDJ, AAVrh10, AAV11, AAV12, AAV13, AAV14, AAV15, AAV16, AAV2/1, AAV2/5, AAV2/6, AAV2/7, AAV2/8, AAV2/9, AAV2/rh10, AAV2/11, or AAV2/12.

[0021] A vector as described herein can be a pseudotyped vector. Pseudotyping provides a mechanism for modulating a vector's target cell population. Pseudotyped vectors comprise the genome of one vector, e.g., the genome of one AAV serotype, in the capsid of a second vector, e.g., a second AAV serotype. A lentiviral vector may be pseudotyped with envelope glycoproteins derived from Rhabdovirus vesicular stomatitis virus (VSV) serotypes (Indiana and Chandipura strains), rabies virus (e.g., various Evelyn–Rokitnicki–Abelseth ERA strains and challenge virus standard (CVS)), Lyssavirus Mokola virus, a rabies-related virus, vesicular stomatitis virus (VSV), Mokola virus (MV), lymphocytic choriomeningitis virus (LCMV), rabies virus glycoprotein (RV-G), glycoprotein B type (FuG-B), a variant of FuG-B (FuG-B2) or Moloney murine leukemia virus (MuLV). A virus may be pseudotyped for transduction of one or more neurons or groups of cells.

[0022] Other embodiments provide nucleic acid expression cassettes comprising: (i) any of the cDNA or hybrid genomic/cDNA polynucleotides encoding presenilin 1 set forth above; (ii) a Kozak translation initiation signal; (iii) a neuron-specific promoter; (iv) a chromatin insulator sequence; (v) at least one mRNA stability element; or (v) any combination thereof. In some aspects of these embodiments, the nucleic acid expression cassettes comprises each of: (i) any of the cDNA or hybrid genomic/cDNA polynucleotides encoding presenilin 1 set forth above; (ii) a Kozak translation initiation signal; (iii) a neuron-specific promoter; (iv) a chromatin insulator sequence; and (v) at least one mRNA stability element. In more specific aspects of these embodiments, the Kozak translation initiation signal comprises a polynucleotide set forth in SEQ ID NO:5; the chromatin insulator sequence comprises a polynucleotide set forth in SEQ ID NO:4; the at least one mRNA stability element comprises a polynucleotide set forth in SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, or any combination thereof; and the neuron-specific promoter comprises a polynucleotide set forth in SEQ ID NO:2 or SEQ ID NO:3.

[0023] Described herein are methods of treating a neurodegenerative disease, disorder, or condition comprising administering to a subject in need thereof a nucleic acid expression cassette or vector. In some embodiments, the neurodegenerative disease, disorder, or condition

is Alzheimer's disease, posterior cortical atrophy (PCA), logopenic progressive aphasia (lvPPA), hippocampal sparing AD, frontotemporal dementia, frontotemporal lobar degeneration, Pick's disease, Lewy body dementia, aphasic variants of AD, behavioral-comportmental ("frontal") variant of AD, a dysexecutive variant, memory loss, cognitive impairment, or mild cognitive impairment.

[0024] Described herein are methods of producing presenilin 1 protein including transforming a host cell with an optimized polynucleotide set forth in SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39; or by a polynucleotide having at least 95% identity to SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39, and encoding the same polypeptide encoded by each of the foregoing, or with a vector encoding presenilin 1 optimized polynucleotide; and culturing the cell under conditions and for a time that allow expression of the presenilin 1 protein. In some aspects, the expression level of the presenilin 1 protein encoded by the optimized polynucleotide in the host cell is greater than a level of expression of presenilin 1 protein encoded by a wild-type polynucleotide in a host cell, thereby producing presenilin 1 protein.

[0025] It is to be understood that one, some, or all of the properties of the various embodiments described herein may be combined to form other embodiments of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] **Figure 1** is a bar graph showing the amount of presenilin-1 protein expressed from HEK293 cells harboring constructs comprising various codon-optimized versions of a presenilin-1 coding sequence under control of a CMV promoter, as well as from a wild-type presenilin-1 coding sequence. The asterisk "*" indicates a statistically significant difference ($p < 0.05$, One way ANOVA followed by Tukey's multiple comparison test).

[0027] **Figure 2** is a bar graph showing the amount of presenilin-1 protein expressed from HEK293 cells harboring constructs comprising various codon-optimized versions of a presenilin-1 coding sequence whose expression is driven by a CAG promoter (CAG-v1.5 containing PSEN1 coding sequence of SEQ ID NO:37; CAG v3.0 containing PSEN1 coding sequence of SEQ ID NO:39)), as well as from a wild-type presenilin-1 coding sequence (SEQ ID NO:15) driven by a CAG promoter. The asterisk "*" indicates a statistically significant difference compared to wild-type presenilin 1 ($p < 0.05$, One way ANOVA followed by Tukey's multiple comparison test).

[0028] **Figure 3** is a bar graph showing gamma secretase activity as measured by cleavage of NotchΔE to NICD in fibroblasts from familial Alzheimer's disease (FAD) patients harboring either a C410Y or G206A mutation in PSEN1. Fibroblasts were transformed with an empty vector ("NotchΔE+Empty"), or a vector containing SEQ ID NO:37 encoding PSEN1 ("NotchΔE+hPSENv1.5") in the presence or absence of the gamma secretase inhibitor DAPT and the levels of NICD measured. The asterisks "***" indicates a statistically significant difference compare to empty vector (p<0.01, One way ANOVA followed by Tukey's multiple comparison test).

[0029] **Figure 4** is a bar graph showing the level of Aβ40 production in fibroblasts from familial Alzheimer's disease (FAD) patients harboring a C410Y mutation in PSEN1 (C410Y) following transformation with either an empty vector ("Empty"), or a vector containing SEQ ID NO:37 encoding PSEN1 ("pAT028").

DETAILED DESCRIPTION

[0030] The present invention is based on the seminal discovery that optimized polynucleotides and expression cassettes encoding optimized therapeutic genes such as presenilin-1 can be used to increase expression levels of the therapeutic gene, as compared to a wild-type sequence, to deliver gene therapy for use in the treatment of neurodegenerative disorders.

[0031] Before the present compositions and methods are described, it is to be understood that this invention is not limited to particular compositions, methods, and experimental conditions described, as such compositions, methods, and conditions may vary. It is also to be understood that the terminology used herein is for purposes of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only in the appended claims.

[0032] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

Definitions

[0033] Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (e.g., in cell culture, molecular genetics, and biochemistry). Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, it will be understood that modifications and variations are encompassed within

the spirit and scope of the instant disclosure. The preferred methods and materials are now described.

[0034] As used herein, the term “about” in the context of a numerical value or range means $\pm 10\%$ of the numerical value or range recited or claimed, unless the context requires a more limited range. Further, the term “about” when used in connection with one or more numbers or numerical ranges, should be understood to refer to all such numbers, including all numbers in a range and modifies that range by extending the boundaries above and below the numerical values set forth. The recitation of numerical ranges by endpoints includes all numbers, e.g., whole integers, including fractions thereof, subsumed within that range (for example, the recitation of 1 to 5 includes 1, 2, 3, 4, and 5, as well as fractions thereof, e.g., 1.5, 2.25, 3.75, 4.1, and the like) and any range within that range.

[0035] As used herein, the singular forms “a”, “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms “including”, “includes”, “having”, “has”, “with”, or variants thereof are used in either the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term “comprising.” It must be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include the plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to a “protein” is a reference to one or more proteins, and includes equivalents thereof known to those skilled in the art and so forth.

[0036] As used herein, the terms “comprising,” “comprise” or “comprised,” and variations thereof, in reference to defined or described elements of an item, composition, apparatus, method, process, system, etc. are meant to be inclusive or open ended, permitting additional elements, thereby indicating that the defined or described item, composition, apparatus, method, process, system, etc. includes those specified elements--or, as appropriate, equivalents thereof--and that other elements can be included and still fall within the scope/definition of the defined item, composition, apparatus, method, process, system, etc.

[0037] By an “AAV vector” is meant a vector derived from an adeno-associated virus serotype, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV6, etc. AAV vectors can have one or more of the AAV wild-type genes deleted in whole or part, preferably the rep and/or cap genes, but retain functional flanking ITR sequences. Functional ITR sequences are necessary for the rescue, replication and packaging of the AAV virion. Thus, an AAV vector is defined herein to include at least those sequences required in *cis* for replication and packaging (e.g., functional ITRs) of the virus. ITRs don't need to be the wild-

type nucleotide sequences, and may be altered, e.g., by the insertion, deletion or substitution of nucleotides, so long as the sequences provide for functional rescue, replication and packaging.

[0038] The control elements are selected to be functional in a mammalian cell. The resulting construct which contains the operatively linked components is bounded (5' and 3') with functional AAV ITR sequences. By “adeno-associated virus inverted terminal repeats” or “AAV ITRs” is mean the art-recognized regions found at each end of the AAV genome which function together in *cis* as origins of DNA replication and as packaging signals for the virus. AAV ITRs, together with the AAV rep coding region, provide for the efficient excision and rescue from, and integration of a nucleotide sequence interposed between two flanking ITRs into a mammalian cell genome. The nucleotide sequences of AAV ITR regions are known. As used herein, an “AAV ITR” does not necessarily comprise the wild-type nucleotide sequence, but may be altered, e.g., by the insertion, deletion or substitution of nucleotides. Additionally, the AAV ITR may be derived from any of several AAV serotypes, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV6, etc. Furthermore, 5' and 3' ITRs which flank a selected nucleotide sequence in an AAV vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, i.e., to allow for excision and rescue of the sequence of interest from a host cell genome or vector, and to allow integration of the heterologous sequence into the recipient cell genome when AAV Rep gene products are present in the cell. Additionally, AAV ITRs may be derived from any of several AAV serotypes, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV 5, AAV6, etc. Furthermore, 5' and 3' ITRs which flank a selected nucleotide sequence in an AAV expression vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, i.e., to allow for excision and rescue of the sequence of interest from a host cell genome or vector, and to allow integration of the DNA molecule into the recipient cell genome when AAV Rep gene products are present in the cell.

[0039] The term “wild-type” and “native” are used herein interchangeably and refer to a form of a substance (e.g., a polynucleotide, a nucleotide sequence, a protein, etc.) that is found in nature. The term “wild-type presenilin-1 coding sequence” as used herein means the polynucleotide sequence set forth in SEQ ID NO:15.

[0040] The term “hybrid genomic/cDNA” as used herein means a non-naturally occurring nucleotide sequence that encodes a protein (e.g., a human presenilin-1), wherein the coding sequence for the protein is interrupted by one or more non-coding intronic sequences.

[0041] The term “intolerant codon” as used herein means a codon present in a reference nucleotide sequence that is not changed in a corresponding subject nucleotide sequence encoding the same amino acids sequence. Intolerant codon in SEQ ID NO: 15 are underlined.

[0042] The term “tolerant codon” as used herein means a codon present in a reference nucleotide sequence that is not an intolerant codon. A tolerant codon may be changed to a different codon encoding the same amino acid in a corresponding subject nucleotide sequence encoding the same amino acid sequence.

[0043] The term “optimized codon” means a codon set forth in Table 2 or Table 3. A codon in a subject nucleotide sequence is said to be “optimized” when the corresponding codon in a reference sequence is replaced with a different codon coding for the same amino acid and selected from a codon set forth in Table 1 or Table 2.

[0044] The term “codon optimization change” means the replacement of a tolerant codon in a reference sequence with a codon encoding the same amino acid selected from Table 2. For some amino acids, there exists more than one optimized codon (see Table 2). For the purpose of clarity, the term “codon optimization changes” includes replacing an optimized codon present in a reference sequence with a different optimized codon coding for the same amino acid set forth in Table 1.

[0045] The exons and introns of the PSEN1 gene can be identified with reference to GenBank reference sequence number NG_007386 as follows: Exon 1 consist of nucleotides 5037 to 5113; intron 1 consist of nucleotides 5,114 to 16,324; exon 2 consist of nucleotides 16,325 to 16,406; intron 2 consist of nucleotides 16,407 to 16,496; exon 3 consist of nucleotides 16,497 to 16,636; intron 3 consist of nucleotides 16,637 to 39,326; exon 4 consist of nucleotides 39,327 to 39,577; intron 4 consist of nucleotides 39,578 to 42,095; exon 5 consist of nucleotides 42,096 to 42,237; intron 5 consist of nucleotides 42,238 to 55,382; exon 6 consist of nucleotides 55,383 to 55,450; intron 6 consist of nucleotides 55,451 to 61,173; exon 7 consist of nucleotides 61,174 to 61,394; intron 7 consist of nucleotides 61,395 to 66,560; exon 8 consist of nucleotides 66,561 to 66,659; intron 8 consist of nucleotides 66,660 to 74,915; exon 9 consist of nucleotides 74,916 to 75,002; intron 9 consist of nucleotides 75,003 to 80,298; exon 10 consist of nucleotides 80,299 to 80,472; intron 10 consist of nucleotides 80,473 to 85,655; exon 11 consist of nucleotides 85,656 to 85,776; intron 11 consist of nucleotides 85,775 to 87,663; exon 12 consist of nucleotides 87,664 to 92,222.

[0046] The term “CpG dinucleotide” means any occurrence of the nucleotide sequence CG in a reference nucleotide sequence.

[0047] “Self-complementary AAV” refers to a construct in which a coding region carried by a recombinant AAV nucleic acid sequence has been designed to form an intra-molecular double-stranded DNA template. Upon infection, rather than waiting for cell mediated synthesis of the second strand, the two complementary halves of scAAV will associate to form one double stranded DNA (dsDNA) unit that is ready for immediate replication and transcription. See, e.g., D M McCarty et al, “Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis”, *Gene Therapy*, (August 2001), Vol 8, Number 16, Pages 1248-1254. Self-complementary AAVs are described in, e.g., U.S. Pat. Nos. 6,596,535; 7,125,717; and 7,456,683, each of which is incorporated herein by reference in its entirety.

[0048] As used herein, “operably linked,” “operable linkage,” “operatively linked,” or grammatical equivalents thereof refer to juxtaposition of genetic elements, e.g., a polynucleotide encoding a protein or RNA, a promoter, an enhancer, a polyadenylation sequence, etc., wherein the elements are in a relationship permitting them to operate in the expected manner. For example, a regulatory element, which can comprise promoter and/or enhancer sequences, is operatively linked to a coding region if the regulatory element helps initiate transcription of the coding sequence. There may be intervening residues between the regulatory element and coding region so long as this functional relationship is maintained.

[0049] As used in this specification and the appended claims, the term “or” is generally employed in its sense including “and/or” unless the content clearly dictates otherwise.

[0050] “Percent (%) identity” with respect to a reference polynucleotide or polypeptide sequence is defined as the percentage of nucleic acids or amino acids in a candidate sequence that are identical to the nucleic acids or amino acids in the reference polynucleotide or polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid or amino acid sequence identity can be achieved in various ways that are within the capabilities of one of skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, or Megalign software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For example, percent sequence identity values may be generated using the sequence comparison computer program BLAST. As an illustration, the percent sequence identity of a given nucleic acid or amino acid sequence, A, to, with, or against a given nucleic acid or amino acid sequence, B, (which can alternatively be phrased as a given nucleic acid or amino acid sequence, A that

has a certain percent sequence identity to, with, or against a given nucleic acid or amino acid sequence, B) is calculated as follows:

$$100 \text{ multiplied by (the fraction } X/Y)$$

where X is the number of nucleotides or amino acids scored as identical matches by a sequence alignment program (e.g., BLAST) in that program's alignment of A and B, and where Y is the total number of nucleic acids in B. It will be appreciated that where the length of nucleic acid or amino acid sequence A is not equal to the length of nucleic acid or amino acid sequence B, the percent sequence identity of A to B will not equal the percent sequence identity of B to A.

[0051] As used herein, the term “polynucleotide or gene expression” refers to the process by which a nucleic acid sequence or a polynucleotide is transcribed from a DNA template (such as into mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides may be collectively referred to as “polynucleotide or gene product.” If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell. The terms “polynucleotide or gene expression” and “expression” can be used interchangeably, unless context clearly indicates otherwise.

[0052] As used herein, the term “presenilin-1” denotes a protein encoded by the PSEN1 gene. Presenilin 1 is one of the four core proteins in the presenilin complex, which mediate the regulated proteolytic events of several proteins in the cell, including gamma secretase. Gamma-secretase is considered to play a strong role in generation of beta amyloid, accumulation of which is related to the onset of Alzheimer's disease, from the beta-amyloid precursor protein. There are two forms of presenilin-1 encoded by the PSEN1 gene based on alternate splicing. The predominant form in humans is the 467 amino acid isoform X1. The alternate form is the 463 amino acid isoform X2. Presenilin-1, presenilin 2 (PSEN2), and amyloid precursor protein (APP) are mostly associated with autosomal dominant forms of early onset Alzheimer's disease.

[0053] The term “promoter” as used herein is defined as a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a polynucleotide sequence. A “constitutive” promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a cell under most or all physiological conditions of the cell. An “inducible” promoter is a nucleotide sequence which, when operably linked with a polynucleotide which encodes or specifies a gene product, causes the gene product to be produced in a cell substantially only when an inducer which corresponds to the

promoter is present in the cell. A “tissue-specific” promoter is a nucleotide sequence which, when operably linked with a polynucleotide encodes or specified by a gene, causes the gene product to be produced in a cell substantially only if the cell is a cell of the tissue type corresponding to the promoter. Alternatively, heterologous control sequences can be employed. Useful heterologous control sequences generally include those derived from sequences encoding mammalian or viral genes. Examples include, but are not limited to, the phosphoglycerate kinase (PGK) promoter, CAG, neuronal promoters, promoter of Dopamine-1 receptor and Dopamine-2 receptor, the SV40 early promoter, mouse mammary tumor virus LTR promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), Rous sarcoma virus (RSV) promoter, synthetic promoters, hybrid promoters, and the like. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, will also find use herein. Such promoter sequences are commercially available from, e.g., Stratagene (San Diego, Calif.). For purposes of the present invention, both heterologous promoters and other control elements, such as CNS-specific and inducible promoters, enhancers and the like, will be of particular use. Examples of heterologous promoters include the CMV promoter. Examples of CNS specific promoters include those isolated from the genes of myelin basic protein (MBP), glial fibrillary acid protein (GFAP), and neuron specific enolase (NSE).

[0054] As used herein, the term “regulatory element” refers to a genetic element or polynucleotide that either alone or together with one or more additional regulatory elements influences or modulates expression of a polynucleotide or gene. A regulatory element can facilitate polynucleotide or gene expression, increase polynucleotide or gene expression, decrease polynucleotide or gene expression and/or confer selective polynucleotide or gene expression in a particular cell type or tissue. A regulatory element can influence or modulate polynucleotide or gene expression temporally and/or spatially. As used herein, the term “regulate polynucleotide or gene expression,” “influence polynucleotide or gene expression,” or “modulate polynucleotide or gene expression” refers to increasing polynucleotide or gene expression, decreasing polynucleotide or gene expression, and/or conferring selective polynucleotide or gene expression. “Regulating polynucleotide or gene expression,” “influencing polynucleotide or gene expression,” or “modulating polynucleotide or gene expression” can refer to temporal and/or spatial regulation.

[0055] A “transgene” is used herein to conveniently refer to a polynucleotide or a nucleic acid that is intended or has been introduced into a cell or organism. Transgenes include any nucleic acid, such as a gene that encodes a polypeptide or protein.

[0056] The term “variant,” when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to a wild type gene. This definition may also include, for example, “allelic,” “splice,” “species,” or “polymorphic” variants. A splice variant may have significant identity to a reference molecule but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. Of particular utility in the invention are variants of wild type gene products. Variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes that give rise to variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

[0057] Ranges: throughout this disclosure, various aspects of the invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.1, 2.2, 2.7, 3, 4, 5, 5.5, 5.75, 5.8, 5.85, 5.9, 5.95, 5.99, and 6. This applies regardless of the breadth of the range.

[0058] The present disclosure provides compositions and methods for treating subjects with Alzheimer's disease and other neurodegenerative diseases, disorders and conditions. In particular, aspect, the present disclosure contemplates gene therapy by providing a polynucleotide encoding a presenilin-1 (PSEN1) gene to a subject in need of treatment. Alzheimer's disease (AD) is the most common form of neurodegenerative disease of the brain. Pathological hallmarks of AD include intraneuronal accumulation of paired helical filaments composed of abnormal tau proteins and extracellular deposits of β -amyloid peptide (A β) in

neuritic plaques. Clinically, AD can be categorized into two phenotypes based on the ages of onset: early-onset AD (EOAD; <65 years) and late-onset AD (LOAD; >65 years), of which LOAD is the more common form worldwide. The proportion of EOAD in all AD cases is between 5% and 10%. Presenilin 1 (PSEN1), presenilin 2 (PSEN2), and amyloid precursor protein (APP) are mostly associated with autosomal dominant forms of EOAD. Apart from genetic factors, mutations are environmentally related. Genetic–environmental interactions may be caused by variation in the age of onset, neuropathological patterns, and disease duration.

[0059] PSEN1 and PSEN2 encode transmembrane proteins PS1 and PS2, respectively, that constitute the catalytic core of γ -secretase, the founding member of an emerging class of unconventional, Intramembrane-Cleaving Proteases (I-CLiPs). Active γ -secretase is a multiprotein complex composed of PS1 or PS2 together with nicastrin (NCT), the anterior pharynx-defective protein 1 (APH1), and the presenilin enhancer 2 (PEN2). Experimental evidence such as the binding of transition-state analogue γ -secretase inhibitors to PS1, as well as the abolishment of γ -secretase activity when PS1 lacks the aspartate residues critical for proteolysis have confirmed that presenilins harbor the active site of the enzymatic complex.

[0060] PS1 and PS2 play fundamental roles in cell signaling as part of the γ -secretase complex. The latter cleaves numerous type-I membrane proteins in their transmembrane domain releasing their corresponding intracellular domains, which are capable of influencing gene expression. The amyloid precursor protein (APP) is processed by the successive actions of β -secretase (BACE1) and γ -secretase, generating amyloid-beta peptides (A β) of different lengths, ranging from 37 to 46 amino acids. Cleavage of the APP C-terminal fragments (APP-CTFs) by γ -secretase also releases the APP intracellular domain (AICD), which has been recently involved in the regulation of brain ApoE expression, a major genetic determinant of AD, and in cholesterol metabolism. In addition, PS1 has been shown to interact with a growing list of proteins that modulate γ -secretase activity.

[0061] Nucleic Acid Compositions

[0062] Accordingly, the present disclosure provides an isolated cDNA or a hybrid genomic/cDNA that encodes the naturally occurring human presenilin-1 and characterized by one or more of: codon optimization only at some or all tolerant codons, reduction of CpG dinucleotides, or the presence of donor/acceptor splice sites to enable expression of both PSEN-1 isoforms; nucleic acid expression cassettes comprising the foregoing and additional regulatory elements; vectors comprising such expression cassettes; compositions comprising those vectors; and methods for gene therapy of neurodegenerative disorders such as

Alzheimer's disease that utilize any of the foregoing. Without being bound by theory, it is believed that the nucleic acid expression cassettes disclosed herein will result in increased and improved PSEN-1 expression as compared to native or mutated forms of PSEN-1 in patients in need thereof, e.g. Alzheimer's disease patients. Thus, PSEN-1 protein expression can be increased at a lower dose of the expression cassette or the vector comprising that expression cassette.

[0063] An embodiment provides nucleic acid expression cassettes comprising any of the cDNA or hybrid genomic/cDNA polynucleotides encoding presenilin 1 set forth above; and one or more regulatory elements operably linked to the polynucleotide encoding presenilin 1.

[0064] Any genetic element that modulates or influences polynucleotide or gene expression can be a regulatory element, including, for example, promoters, enhancers, chromatin insulators, translation initiation sequences such as strong and weak Kozak signal sequences and internal ribosomal entry sites, mRNA stability sequences, sequences that influence mRNA processing such as splicing and cleavage, sequences that influence mRNA export from the nucleus and/or mRNA retention, posttranslational response elements, non-coding sequences such as introns, poly A sequences, repressors, silencers, terminators, and others. Regulatory elements can function to modulate polynucleotide or gene expression at the transcriptional level, at the posttranscriptional level, at the translational level, or any combination thereof. Regulatory elements can increase the rate at which RNA transcripts are produced, increase the stability of RNA produced, increase the rate of protein synthesis from RNA transcripts, prevent RNA degradation and/or increase RNA stability to facilitate protein synthesis, for example.

[0065] An expression cassette as used herein may suitably comprise a promoter and poly A sequence. In certain preferred embodiments, an expression cassette may comprise a promoter, poly A sequence and mRNA stability element. A particularly preferred expression cassette may include a CAG promoter, Kozak, codon optimized PSEN1 (tolerant only), mRNA stability element and poly A. A specifically preferred expression cassette may include SEQ ID NO:23, SEQ ID NO:5, SEQ ID NO:6, SED ID NO:9, and SEQ ID NO:34. In certain embodiments, preferred vectors may comprise AAV surrounded by ITRs and packaged into an AAV9 or AAVrh10 capsid.

[0066] The nucleic acid expression cassettes described herein can comprise regulatory elements that regulate or modulate polynucleotide or gene expression at any step, including the transcriptional, posttranscriptional, and translational levels, for example. A regulatory element can regulate or modulate polynucleotide or gene expression at more than one level or function

in more than one way to regulate or modulate polynucleotide or gene expression. Thus, a regulatory element can have any function, or any combination of the functions described above. For example, a regulatory element can function as an mRNA stabilizing element and modulate, i.e., increase or decrease, translation. As yet another example, a regulatory element can modulate transcription initiation and modulate mRNA stability. A regulatory element can also have a predominant function by which it modulates polynucleotide or gene expression and have one or more additional functions that increase or decrease polynucleotide or gene expression. A regulatory element can comprise a sequence that is located within or overlaps with other regulatory elements that have the same or different functions in modulating polynucleotide or gene expression or that modulate polynucleotide or gene expression at the same or different steps.

[0067] Regulatory elements can be derived from coding or non-coding DNA sequences. Regulatory elements derived from non-coding DNA can be associated with genes, e.g., may be found in a gene, such as upstream sequences, introns, 3' and 5' untranslated regions (UTRs), and/or downstream regions. As used herein, the term “upstream” when referring to nucleic acid means 5' relative to another sequence and the term “downstream” means 3' relative to another sequence. The term “upstream” can be used interchangeably with the term “5'” when referring to location of sequences relative to each other, unless context clearly indicates otherwise. The term “downstream” can be used interchangeably with the term “3'” when referring to location of sequences relative to each other, unless context clearly indicates otherwise.

[0068] In some embodiments, regulatory elements derived from non-coding DNA sequences are not associated with a gene, e.g., may not be found in a gene. The genomic region from which a regulatory element is derived can be distinct from the genomic region from which an operably linked polynucleotide is derived. In some embodiments, a regulatory element is derived from a distal genomic region or location with respect to the genomic region or location from which the operably linked polynucleotide (such as a cDNA derived from an endogenous gene or an endogenous version of a heterologous gene, for example) is derived. In some embodiments, a regulatory element comprises intron sequences. Intron sequences can include sequences derived from any gene. In some embodiments, the intron sequences are derived from the genomic region from which an operatively linked polynucleotide is derived. For example, the nucleic acid expression cassettes described herein can include introns from an endogenous gene that corresponds to a polynucleotide or that gave rise to a polynucleotide in the form of a cDNA. As another example, the nucleic acid expression cassettes described herein can include introns from an endogenous gene that does not correspond to or gave rise to a polynucleotide.

[0069] In some embodiments, the one or more regulatory elements comprise a Kozak translation initiation signal such as a polynucleotide set forth in SEQ ID NO:5. In some embodiments, the one or more regulatory elements comprise a chromatin insulator sequence, such as the polynucleotide set forth in SEQ ID NO:4.

[0070] **Kozak Sequences:** A 5' UTR generally includes sequences that are recognized by the ribosome that allow the ribosome to bind and initiate translation. Exemplary sequences for translation initiation include Kozak initiation signal sequences. As used herein, the terms "Kozak initiation signal sequence," "Kozak consensus sequence," and "Kozak sequence" can be used interchangeably, unless context clearly indicates otherwise. A person of skill in the art will appreciate that a Kozak initiation signal sequence can be located in part in the 5' UTR and include the AUG translation initiation codon itself and the nucleotide immediately following or downstream of the AUG start codon, as described below.

[0071] Translation initiation of an mRNA typically occurs at an ATG codon that is recognized by a ribosome. The ATG codon at which translation begins may not be the first ATG start codon present in an mRNA sequence. A motif called a Kozak sequence can direct translation initiation to an ATG codon. The Kozak consensus sequence is defined as 5'-(gcc)gccRccAUGG-3, where the underlined AUG indicates the translation start codon; uppercase letters indicate conserved bases; "R" indicates the presence of a purine, with adenine more frequent; lowercase letters indicate the most common base at a position that can vary; and the sequence (gcc) is of uncertain significance. In addition to these features, other positions and features can contribute to translation initiation. Strong and weak Kozak consensus sequences have been described, with a strong Kozak consensus sequence including the features above that are considered optimal for translation initiation and a weak Kozak consensus sequence including features that deviate or differ from a strong Kozak consensus sequence. The amount of protein synthesized from an mRNA can depend on the strength of the Kozak sequence. For example, a CCACC (SEQ ID NO: 5) sequence immediately upstream of an AUG translation initiation codon can increase the rate of translation initiation compared to a sequence that differs from CCACC.

[0072] In some embodiments, the nucleic acid expression cassettes provided herein comprise a Kozak translation initiation signal. The Kozak translation initiation signal can be located immediately upstream or 5' of a translation initiation AUG codon. Any Kozak consensus sequence that is a strong Kozak sequence can be used. In some embodiments, the Kozak translation initiation signal comprises a sequence set forth in SEQ ID NO:5.

[0073] *Promoters:* Promoters are a major *cis*-acting element within the vector genome design that can dictate the overall strength of expression as well as cell-specificity. Accordingly, in certain embodiments, the promoter is a neuron-specific promoter. A neuron-specific promoter can provide selective expression of a polynucleotide or therapeutic gene in neuronal cells. Selective expression that is restricted or limited to a particular cell type can prevent or reduce off-target effects that are often undesirable and can result in side effects, for example. As used herein, “selective expression” refers to expression that is at least 1%, at least 2%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, and any number or range in between, higher in neurons as compared to non-neuronal cells. In some embodiments, there is no expression in non-neuronal cells.

[0074] In some embodiments, a neuron-specific promoter of the nucleic acid expression cassettes described herein provides for expression that is at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, and any number or range in between, higher as compared to expression provided by a promoter that can drive expression in any cell type. In some embodiments, a neuron-specific promoter of the nucleic acid expression cassettes described herein provides for expression that is at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, and any number or range in between, higher as compared to expression provided by a promoter that can drive expression in one or more non-neuronal cell types.

[0075] Any neuron-specific promoter can be used in the nucleic acid expression cassettes provided herein. Exemplary promoters include a somatostatin (SST; SEQ ID NO: 2) gene promoter, a neuropeptide Y (NPY; SEQ ID NO: 3) promoter, an alpha-calcium/calmodulin kinase 2A promoter, a synapsin I promoter (e.g., nucleotides 273-684 of SEQ ID NO:46), a neuron-specific enolase (NSE) (e.g., SEQ ID NO:29), a dopaminergic receptor 1 (Drd1a) promoter, a tubulin alpha I promoter, a GFAP promoter (e.g., SEQ ID NO:31) and known variations thereof (e.g., gfaABC(1)D) and others. Hybrid promoters can also be used. A hybrid promoter is a promoter that includes promoter sequences derived from more than one gene.

Promoters can be from any species, including human, rhesus macaque, mouse, rat, and chicken, for example. A neuron-specific promoter can comprise (i) a polynucleotide set forth in SEQ ID NO:2; (ii) a polynucleotide set forth in SEQ ID NO:3; (iii) a functional fragment of SEQ ID NO:2 or SEQ ID NO:3; or (iv) polynucleotide with at least 95% identity to (i), (ii), or (iii). In alternate aspects of these embodiments, the promoter comprises CAG, CBA, UBC, PKC, EF1 α , GUSB, CMV, NSE, PDGF, desmin, MCK, MeCP2, GFAP, CaMKII or MBP.

[0076] Constitutive promoters such as the human elongation factor 1 α -subunit (EF1 α) (e.g., SEQ ID NO:27 or nucleotides 237-1415 of SEQ ID NO:44), immediate-early cytomegalovirus (CMV) (e.g., SEQ ID NO:28), chicken β -actin (CBA) (e.g., SEQ ID NO:24 or nucleotides 237-890 of SEQ ID NO:43) and its derivative CAG (SEQ ID NO:23 or SEQ ID NO:40), the β glucuronidase (GUSB), ubiquitin C (UBC) (e.g., SEQ ID NO:25 or nucleotides 237-1323 of SEQ ID NO:42 or), phosphoglycerate kinase 1 (PGK) (e.g., SEQ ID NO:26), or even the native PSEN-1 promoter (e.g., nucleotides 237-1200 of SEQ ID NO:41) can be used to promote expression in most tissues. Generally, CBA and CAG promote the larger expression among the constitutive promoters; however, their size of ~1.7 kbs in comparison to CMV (~0.8 kbs) or EF1 α (~1.2 kbs) limits its use in vectors with packaging constraints such as AAV. The GUSB or UBC promoters can provide ubiquitous gene expression with a smaller size of 378 bps and 403 bps, respectively, but they are considerably weaker than the CMV or CBA promoter. Thus, modifications to constitutive promoters in order to reduce the size without affecting its expression have been pursued and examples such as the CBh (~800 bps) and the miniCBA (~800 bps) can promote expression comparable and even higher in selected tissues.

[0077] When expression is restricted to certain cell types within an organ, e.g. brain, central nervous system etc., promoters can be used to mediate this specificity. For example, within the nervous system promoters have been used to restrict expression to neurons, astrocytes, or oligodendrocytes. In neurons, the neuron-specific enolase (NSE) promoter drives stronger expression than ubiquitous promoters; however, its size of 2.2 kbs limits its use in smaller vectors. Additionally, the platelet-derived growth factor B-chain (PDGF- β), the synapsin (Syn), and the methyl-CpG binding protein 2 (MeCP2) (e.g., SEQ ID NO:30) promoters can drive neuron-specific expression at lower levels than NSE, but their sizes of 1.4 kbs, 470 bps and 229 bps, respectively, make them more suitable for vectors with limitations in size. In astrocytes, the 680 bps-long shortened version [gfaABC(1)D] of the glial fibrillary acidic protein (GFAP, 2.2 kbs) promoter can confer higher levels of expression with the same astrocyte-specificity as the GFAP promoter. Targeting oligodendrocytes can also be accomplished by the selection of the myelin basic protein (MBP) promoter, whose expression

is restricted to this glial cell (Gray SJ, *et al.*, Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. *Hum Gene Ther.* 2011;22:1143–1153).

[0078] Tissue specific promoters provide the advantage of limiting the expression to the desired cell or tissue. However, low levels of expression and/or large size may limit their use. To compensate for weak strength, the level of expression can be increased by adding enhancer elements such as from CMV.

[0079] **MicroRNA binding sites:** In some embodiments, the one or more regulatory elements comprise one, two or three micro RNA (“miRNA” or “miR”) binding sites to suppress expression of the encoded PSEN-1 in dorsal root ganglia. MicroRNAs are 19-25 nucleotide noncoding RNAs that bind to miRNA binding sites and down-regulate gene expression either by reducing nucleic acid molecule stability or by inhibiting translation. In some aspects of these embodiments, each miRNA binding site is independently selected from a binding site for any of the following miRNAs: miRNA-1914, miR1181, miR3918, miR939, miR324, miR650, MiR29C, or miR2277. In some aspects of these embodiments, the miRNA binding site(s) are located 3' to the mRNA stability element.

[0080] Endogenous miRNAs can ‘de-target’ or inhibit transgene expression when their exact complementary target sequences are engineered into an expression cassette. The level of repression, *in vitro*, correlates with the number of target sequences within the expression cassette. In an *in vivo* study, when an engineered lentiviral vector containing 4 copies of the neuronal-specific miR-124 target sequence was injected into mouse brain, PGK-driven transgene expression was de-targeted from neurons to only astrocytes (Colin A. *et al.*, Engineered lentiviral vector targeting astrocytes *in vivo*. *Glia*. 2009 Apr 15; 57(6):667-79). Endogenous miRNAs are a useful tool in obtaining transgene cell specificity because their respective binding sites are small, can be combined, and are robust in their ability to restrict expression.

[0081] **mRNA Stability Element:** Exemplary mRNA stability elements include a MALAT1 mRNA stability element, C-rich stability elements of HBA1, HBA2, lipoxigenase, alpha(I)-collagen, and tyrosine hydroxylase 3' UTRs, for example, AU-rich elements (AREs) of 3' UTRs, and others. An mRNA stability element can be, for example, an expression and nuclear retention element. An mRNA stability element can prevent or decrease degradation of mRNA. For example, degradation of mRNA can be decreased by about 5%, about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 95%, about 99%, and any number or range in between, when an mRNA stability element is

included as compared to a nucleic acid expression cassette that does not include an mRNA stability element. In an embodiment, there is no degradation of mRNA. Any sequence that prevents or decreases degradation of the mRNA can be an mRNA stability element. An mRNA stability element can be placed into any location in a nucleic acid expression cassette. For example, an mRNA stability element can be placed 3' to the open reading frame of a polynucleotide and before or 5' of a polyadenylation site. As another example, an mRNA stability element can be placed 3' to the open reading frame of a polynucleotide and 3' to a polyadenylation site. As yet another example, an mRNA stability element can be placed 5' to an open reading frame of a polynucleotide.

[0082] In some embodiments, an mRNA stability element comprises (i) a polynucleotide set forth in SEQ ID NO:9; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii). In some embodiments, an mRNA stability element comprises a polynucleotide with at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:9.

[0083] Accordingly, in certain embodiments, the at least one mRNA stability element comprises (i) a polynucleotide set forth in SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11; (ii) a functional variant of SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii).

[0084] In some embodiments, the nucleic acid expression cassettes embodied herein include an mRNA stability element comprising (i) a polynucleotide set forth in SEQ ID NO:10; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii). In some embodiments, the mRNA stability element comprises a polynucleotide with at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:10. In some embodiments, the mRNA stability element comprises (i) a polynucleotide set forth in SEQ ID NO:10; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii) is located 5' of an open reading frame of a polynucleotide encoding PSEN1 or other therapeutic gene.

[0085] In some embodiments, the nucleic acid expression cassettes described herein include an mRNA stability element comprising (i) a polynucleotide set forth in SEQ ID NO:11; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii). In some embodiments, the mRNA stability element comprises a polynucleotide with at least 80%, with at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:11.

[0086] In some embodiments, the mRNA stability element comprises a polynucleotide with at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:10. In some embodiments, the mRNA stability element comprising (i) a polynucleotide set forth in SEQ ID NO:10; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii) is located 5' of an open reading frame of a polynucleotide encoding PSEN1 or other therapeutic gene.

[0087] In some embodiments, the nucleic acid expression cassettes described herein include an mRNA stability element comprising (i) a polynucleotide set forth in SEQ ID NO:11; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii). In some embodiments, the mRNA stability element comprises a polynucleotide with at least 80%, with at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:11. In some embodiments, the mRNA stability element comprising (i) a polynucleotide set forth in SEQ ID NO:11; (ii) a functional variant thereof; or (iii) a polynucleotide with at least 95% sequence identity to (i) or (ii) is located 5' of an open reading frame of a PSEN1 nucleotide sequence.

[0088] In some aspects of these embodiments, the nucleic acid expression cassette comprises a mRNA stability element located 5' of the open reading frame of the polynucleotide encoding PSEN1; and a mRNA stability element located 3' of the polyadenylation signal.

[0089] *Polyadenylation Signal:* In certain embodiments, the nucleic acid expression cassette also comprises one or more polyadenylation enhancer elements, such as, for example,

human growth hormone (hGH; SEQ ID NO: 33; nucleotides 3330-3806 of SEQ ID NO:41) polyadenylation signal sequences, rabbit beta-globin (rBG; SEQ ID NO: 34 or 35; nucleotides 2139-2367 of SEQ ID NO:47) polyadenylation signal sequences, SV40 polyadenylation signal sequences or bovine growth hormone (BGH) polyadenylation signal sequences. The polyadenylation of a transcript is critical for nuclear export, translation, and mRNA stability. Therefore, the efficiency of transcript polyadenylation is important for transgene expression. The poly(A) tail contains binding sites for poly(A) binding proteins (PABPs). These proteins cooperate with other factors to affect the export, stability, decay, and translation of an mRNA. PABPs bound to the poly(A) tail may also interact with proteins, such as translation initiation factors, that are bound to the 5' cap of the mRNA. This interaction causes circularization of the transcript, which subsequently promotes translation initiation. Furthermore, it allows for efficient translation by causing recycling of ribosomes. While the presence of a poly(A) tail usually aids in triggering translation, the absence or removal of one often leads to exonuclease-mediated degradation of the mRNA. Polyadenylation itself is regulated by sequences within the 3'-UTR of the transcript. These sequences include cytoplasmic polyadenylation elements (CPEs), which are uridine-rich sequences that contribute to both polyadenylation activation and repression. CPE-binding protein (CPEB) binds to CPEs in conjunction with a variety of other proteins in order to elicit different responses.

[0090] *Chromatin Insulator Sequence:* In certain embodiments, a nucleic acid expression cassette can further comprise a chromatin insulator sequence. Packaging of genes into chromatin can render genes inaccessible to the transcription machinery of the cell, resulting in little or no gene expression. Chromatin insulators can protect a sequence from being packed into transcriptionally inactive chromatin. Including a chromatin insulator sequence in a nucleic acid expression cassette can keep a polynucleotide in an accessible state and allow transcription to occur. Any chromatin insulator can be used in the nucleic acid expression cassettes provided herein. Exemplary chromatin insulator sequences include a CTCF insulator, a gypsy insulator, and a β -globin locus. Chromatin insulator sequences from any species can be used, including mammals and non-mammals and vertebrates and non-vertebrates. As an example, a chromatin insulator sequence from human beta globin locus HS4 can be used. Other examples of chromatin insulator sequences include sequences from chicken and *Drosophila*. A chromatin insulator sequence can comprise a polynucleotide set forth in SEQ ID NO:4, a functional variant of SEQ ID NO:4, or a polynucleotide with at least 95% identity to SEQ ID NO:4. A chromatin insulator sequence can comprise at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:4 as long as the function of the reference sequence and the ability to protect a sequence with which it is associated from being packed into transcriptionally inactive chromatin is maintained.

[0091] *Transcription Termination Region*: A transcription termination region of a recombinant construct or expression cassette is a downstream regulatory region including a stop codon and a transcription terminator sequence. Transcription termination regions that can be used can be homologous to the transcriptional initiation region, can be homologous to the polynucleotide encoding a polypeptide of interest, or can be heterologous (i.e., derived from another source). A transcription termination region or can be naturally occurring, or wholly or partially synthetic. 3' non-coding sequences encoding transcription termination regions may be provided in a recombinant construct or expression construct and may be from the 3' region of the gene from which the initiation region was obtained or from a different gene. A large number of termination regions are known and function satisfactorily in a variety of hosts when utilized in both the same and different genera and species from which they were derived. Termination regions may also be derived from various genes native to the preferred hosts. The termination region is usually selected more for convenience rather than for any particular property.

[0092] Regulatory elements and polynucleotides of the nucleic acid expression cassettes provided herein can be combined in any fashion. In some embodiments, a nucleic acid expression cassette comprises a polynucleotide encoding presenilin 1, wherein the polynucleotide comprises any one of (I) a polynucleotide set forth in SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39; or (III) a polynucleotide having at least 95% identity to (I) or (II). In some embodiments, the polynucleotide comprises a sequence having at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, 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 99.5%, at least 99.9%, and any number or range in between, identity to SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39. In some embodiments, the nucleic acid expression cassette further comprises one or more regulatory elements operably linked to the polynucleotide encoding presenilin 1. In some embodiments, the one or more regulatory elements comprise a neuron-specific promoter.

[0093] In some embodiments, the nucleic acid expression cassette further comprises (i) a Kozak translation initiation signal; (ii) a chromatin insulator sequence; (iii) at least one mRNA stability element; or (iv) any combination thereof, wherein the one or more regulatory elements comprise a neuron-specific promoter. In some embodiments, the Kozak translation initiation signal comprises a polynucleotide set forth in SEQ ID NO:5; the chromatin insulator sequence comprises a polynucleotide set forth in SEQ ID NO:4; the at least one mRNA stability element comprises a polynucleotide set forth in SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, or any combination thereof; the neuron-specific promoter comprises a polynucleotide set forth in SEQ ID NO:2 or SEQ ID NO:3.

[0094] In some embodiments, the mRNA stability element comprising SEQ ID NO:9 is located 3' of an open reading frame of the polynucleotide encoding PSEN1 and 5' of a polyadenylation signal, the mRNA stability element comprising SEQ ID NO:10 is located 5' of an open reading frame of the polynucleotide encoding PSEN1, and the mRNA stability element comprising SEQ ID NO:11 is located 3' of an open reading frame of the polynucleotide encoding PSEN1.

[0095] In some embodiments, the nucleic acid expression cassettes provided herein comprise: (a) one or more regulatory elements operably linked to a polynucleotide encoding presenilin 1, wherein the polynucleotide comprises any one of (I) a polynucleotide set forth in SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39; (II) a polynucleotide having at least 95% identity to (I); and wherein the one or more regulatory elements comprise a neuron-specific promoter comprising a polynucleotide set forth in SEQ ID NO:2 or SEQ ID NO:3; (b) a Kozak translation initiation signal comprising a polynucleotide set forth in SEQ ID NO:5; (c) a chromatin insulator sequence comprising a polynucleotide set forth in SEQ ID NO:4; and (d) at least one mRNA stability element comprising a polynucleotide set forth in SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, or any combination thereof.

[0096] Other embodiments provide nucleic acid expression cassettes comprising: (i) any of the cDNA or hybrid genomic/cDNA polynucleotides encoding presenilin 1 set forth above; (ii) a Kozak translation initiation signal; (iii) a neuron-specific promoter; (iv) a chromatin insulator sequence; (v) at least one mRNA stability element; or (v) any combination thereof. The Kozak translation initiation signal can comprise a polynucleotide set forth in SEQ ID NO:5; the chromatin insulator sequence can comprise a polynucleotide set forth in SEQ ID NO:4; the at least one mRNA stability element can comprise a polynucleotide set forth in SEQ

ID NO:9, SEQ ID NO:10, SEQ ID NO:11, or any combination thereof; and the neuron-specific promoter can comprise a polynucleotide set forth in SEQ ID NO:2 or SEQ ID NO:3.

[0097] Codon-Optimization

[0098] Codon optimization can be utilized to enhance protein expression for heterologous gene expression. Codon optimization is a method of gene optimization, where in the synthetic gene sequence is modified to match the "codon usage pattern" for a particular organism. For example, in order to optimize expression of a particular amino acid sequence in a specific organism, one would select the "most frequently used codons" (from a list of degenerate codons for an amino acid), by that organism. See, Table 2 for a list of preferred codons used. Upon codon optimization, the encoded amino acid sequence remains the same but with the DNA sequence encoding the amino acid sequence is different, optimized for that organism. Accordingly, the disclosure provides a codon-optimized presenilin-1 (PSEN1)-encoding polynucleotide suitable for use in the compositions and methods described herein. The codon-optimized PSEN1 can include a full length hybrid genomic/cDNA (e.g. SEQ ID NO: 8), comprising one or more optimized codons set forth in Table 2. In certain embodiments, the PSEN-1 polynucleotide comprising SEQ ID NO: 1, comprises one or more optimized codons set forth in Table 2. In certain embodiments, a codon-optimized presenilin-1 (PSEN1)-encoding polynucleotide is set forth as SEQ ID NO: 6. In certain embodiments, a codon-optimized presenilin-1 (PSEN1)-encoding polynucleotide is set forth as SEQ ID NO: 36. In certain embodiments, a codon-optimized presenilin-1 (PSEN1)-encoding polynucleotide is set forth as SEQ ID NO: 37. In certain embodiments, a codon-optimized presenilin-1 (PSEN1)-encoding polynucleotide is set forth as SEQ ID NO: 38. In certain embodiments, a codon-optimized presenilin-1 (PSEN1)-encoding polynucleotide is set forth as SEQ ID NO: 39.

[0099] Vectors

[0100] A "vector" is a macromolecule or association of macromolecules that comprises or associates with a polynucleotide and which can be used to mediate delivery of the polynucleotide to a cell. Examples of vectors include plasmids, viral vectors, liposomes, and other gene delivery vehicles. A vector can comprise one or more elements for vector replication. A vector can be engineered to lack one or more elements for vector replication.

[0101] A vector can be an integrating or non-integrating vector, referring to the ability of the vector to integrate the nucleic acid expression cassette and/or polynucleotide into a genome of a cell. Either an integrating vector or a non-integrating vector can be used to deliver a nucleic acid expression cassette containing a polynucleotide. Examples of vectors include, but are not limited to, (a) non-viral vectors such as nucleic acid vectors including linear oligonucleotides

and circular plasmids; artificial chromosomes such as human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), and bacterial artificial chromosomes (BACs or PACs); episomal vectors; transposons (e.g., PiggyBac); and (b) viral vectors such as retroviral vectors, lentiviral vectors, adenoviral vectors, and AAV vectors. Viruses have several advantages for delivery of nucleic acids, including high infectivity and/or tropism for certain target cells or tissues. In some embodiments, a virus is used to deliver a nucleic acid molecule or nucleic acid expression cassette comprising one or more polynucleotide.

[0102] In some embodiments, the vector is a viral vector. In some embodiments, the viral vector is an adeno-associated virus (AAV) vector, a retroviral vector, a lentiviral vector, or an adenoviral vector. The vectors comprising the nucleic acid expression cassettes provided herein can be a viral vector, such as an adeno-associated virus (AAV) vector, a retroviral vector, a lentiviral vector, or an adenoviral vector.

[0103] In some embodiments, the AAV vector is AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVDJ, AAVrh10, AAV11, AAV12, AAV13, AAV14, AAV15, AAV16, AAV2/1, AAV2/5, AAV2/6, AAV2/7, AAV2/8, AAV2/9, AAV2/rh10, AAV2/11, or AAV2/12, single-stranded AAV (ssAAV) vector or self-complementary AAV (scAAV) vector. In some embodiments, the AAV vector is a hybrid or chimeric AAV serotype.

[0104] In some embodiments, the AAV vector comprises: a) promoter selected from a CAG promoter, a presenilin-1 promoter, a ubiquitin C promoter, a CBA promoter, a synapsin-1 promoter, a PGK promoter, and an EF1 α promoter, operatively linked to b) a presenilin-1 coding sequence selected from SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, or SEQ ID NO:39, or a polynucleotide having at least 95% identity to any of the foregoing PSEN-1 coding sequences and encoding a wild-type PSEN-1 amino acids sequence; and c) a polyadenylation sequence selected from a human growth hormone polyadenylation sequence and a rabbit β -globin polyadenylation sequence. In some aspects of these embodiments, the AAV vector additionally comprises, in between the promoter and the PSEN-1 coding sequence, an intron selected from a human beta globin intron (either wild-type or synthetic) or a minute virus of mice intron.

[0105] In some embodiments, the AAV vector comprises:

a. nucleotides 1-141 of SEQ ID NO:41 or a sequence having at least 95% identity thereto, nucleotides 237-1200 of SEQ ID NO:41 or a sequence having at least 95% identity thereto, nucleotides 1221-1786 of SEQ ID NO:41 or a sequence having at least 95% identity thereto, nucleotides 1899-3299 of SEQ ID NO:41 or a sequence having at least 95% identity thereto, nucleotides 3330-3806 of SEQ ID NO:41 or a sequence having at least 95% identity

thereto and nucleotides 4553-4693 of SEQ ID NO:41 or a sequence having at least 95% identity thereto;

b. Nucleotides 1-141 of SEQ ID NO:42 or a sequence having at least 95% identity thereto, nucleotides 237-1323 of SEQ ID NO:42 or a sequence having at least 95% identity thereto, nucleotides 1344-1909 of SEQ ID NO:42 or a sequence having at least 95% identity thereto, nucleotides 1983-3416 of SEQ ID NO:42 or a sequence having at least 95% identity thereto, nucleotides 3447-~~3923~~ of SEQ ID NO:42 or a sequence having at least 95% identity thereto, and nucleotides 4554-4694 of SEQ ID NO:42 or a sequence having at least 95% identity thereto;

c. Nucleotides 1-141 of SEQ ID NO:43 or a sequence having at least 95% identity thereto, nucleotides 237-890 of SEQ ID NO:43 or a sequence having at least 95% identity thereto, nucleotides 911-1476 of SEQ ID NO:43 or a sequence having at least 95% identity thereto, nucleotides 1550-2983 of SEQ ID NO:43 or a sequence having at least 95% identity thereto, nucleotides 3014-3490 of SEQ ID NO:43 or a sequence having at least 95% identity thereto, and nucleotides 4553-4694 or a sequence having at least 95% identity thereto of SEQ ID NO:43;

d. Nucleotides 1-141 or a sequence having at least 95% identity thereto, nucleotides 237-1415 or a sequence having at least 95% identity thereto, nucleotides 1436-2001 or a sequence having at least 95% identity thereto, nucleotides 2075-3508 or a sequence having at least 95% identity thereto, nucleotides 3539-4015 or a sequence having at least 95% identity thereto, and nucleotides 4500-4640 or a sequence having at least 95% identity thereto of SEQ ID NO:44 or a sequence having at least 95% identity thereto;

e. Nucleotides 1-141 of SEQ ID NO:45 or a sequence having at least 95% identity thereto, nucleotides 237-664 of SEQ ID NO:45 or a sequence having at least 95% identity thereto, nucleotides 684-1249 of SEQ ID NO:45 or a sequence having at least 95% identity thereto, nucleotides 1323-2756 of SEQ ID NO:45 or a sequence having at least 95% identity thereto, 2787-3263 of SEQ ID NO:45 or a sequence having at least 95% identity thereto, and nucleotides 4533-4673 of SEQ ID NO:45 or a sequence having at least 95% identity thereto;

f. Nucleotides 1-141 of SEQ ID NO:46 or a sequence having at least 95% identity thereto, nucleotides 237-684 of SEQ ID NO:46 or a sequence having at least 95% identity thereto, nucleotides 705-1270 of SEQ ID NO:46 or a sequence having at least 95% identity thereto, nucleotides 1344-2777 of SEQ ID NO:46 or a sequence having at least 95% identity thereto, nucleotides 2808-3284 of SEQ ID NO:46 or a sequence having at least 95% identity

thereto, and nucleotides 4554-4695 of SEQ ID NO:46 or a sequence having at least 95% identity thereto; or

g. Nucleotides 1-105 of SEQ ID NO:47 or a sequence having at least 95% identity thereto, nucleotides 113-766 of SEQ ID NO:47 or a sequence having at least 95% identity thereto, nucleotides 776-867 of SEQ ID NO:47 or a sequence having at least 95% identity thereto, nucleotides 881-2311 of SEQ ID NO:47 or a sequence having at least 95% identity thereto, nucleotides 2319-2367 of SEQ ID NO:47 or a sequence having at least 95% identity thereto, and nucleotides 2386-2526 of SEQ ID NO:47 or a sequence having at least 95% identity thereto.

[0106] In some embodiments, the AAV vector comprises: a) nucleotides 1-141, 237-1200, 1221-1786, 1899-3299, 3330-3806 and 4553-4693 of SEQ ID NO:41; b) nucleotides 1-141, 237-1323, 1344-1909, 1983-3416, 3447-3923, and 4554-4694 of SEQ ID NO:42; c) nucleotides 1-141, 237-890, 911-1476, 1550-2983, 3014-3490, and 4553-4694 of SEQ ID NO:43; d) nucleotides 1-141, 237-1415, 1436-2001, 2075-3508, 3539-4015, and 4500-4640 of SEQ ID NO:44; e) nucleotides 1-141, 237-664, 684-1249, 1323-2756, 2787-3263, and 4533-4673 of SEQ ID NO:45; e) nucleotides 1-141, 237-684, 705-1270, 1344-2777, 2808-3284, and 4554-4695 of SEQ ID NO:46; or f) nucleotides 1-105, 113-766, 776-867, 881-2311, 2319-2367, and 2386-2526 of SEQ ID NO:47.

[0107] In some embodiments, the AAV vector comprises a nucleotide sequence of any one of SEQ ID NOs:41-47, or a nucleotide sequence having 95% identity to any one of SEQ ID NOs:41-47.

[0108] It will be understood by one of skill in the art that in the above embodiments of AAV vectors wherein a sequence has less than 100% identity to a specified range of nucleotides, such sequence should provide a similar functionality (e.g., be a functional ITR pair; be a functional promoter that can drive expression of the PSEN-1 coding sequence; be a functional intron; encode the same amino acid sequence as wild-type PSEN-1; or be a functional polyadenylation sequence).

[0109] Techniques contemplated herein for gene therapy of somatic cells include delivery via a viral vector (e.g., retroviral, adenoviral, AAV, helper-dependent adenoviral systems, hybrid adenoviral systems, herpes simplex, pox virus, lentivirus, and Epstein-Barr virus), and non-viral systems, such as physical systems (naked DNA, DNA bombardment, electroporation, hydrodynamic, ultrasound, and magnetofection), and chemical systems (cationic lipids, different cationic polymers, and lipid polymers).

[0110] Viral gene therapy vectors or gene delivery vectors can have the ability to be reproducibly and/or stably propagated and purified to high titers; to mediate targeted delivery (e.g., to deliver the polynucleotide specifically to a tissue or organ of interest without widespread vector dissemination elsewhere or off-target delivery); and to mediate gene delivery and/or polynucleotide expression without inducing harmful side effects or off-target effects.

[0111] The term “AAV” is an abbreviation for adeno-associated virus, and may be used to refer to the virus itself or a derivative thereof. The term covers all serotypes, subtypes, and both naturally occurring and recombinant forms, except where required otherwise. The abbreviation “rAAV” refers to recombinant adeno-associated virus, also referred to as a recombinant AAV vector (or “rAAV vector”). The term “AAV” includes AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV 12, rh10, and hybrids thereof, avian AAV, bovine AAV, canine AAV, equine AAV, primate AAV, non-primate AAV, and ovine AAV. The genomic sequences of various serotypes of AAV, as well as the sequences of the native terminal repeats (TRs), Rep proteins, and capsid subunits are known in the art. Such sequences may be found in the literature or in public databases such as GenBank. An “rAAV vector” as used herein refers to an AAV vector comprising a polynucleotide sequence not of AAV origin (i.e., a polynucleotide heterologous to AAV), typically a sequence of interest for the genetic transformation of a cell. In general, the heterologous polynucleotide is flanked by at least one, and generally by two, AAV inverted terminal repeat sequences (ITRs). The term “rAAV vector” encompasses both rAAV vector particles and rAAV vector plasmids. An rAAV vector may either be single-stranded (ssAAV) or self-complementary (scAAV). An “AAV virus” or “AAV viral particle” or “rAAV vector particle” refers to a viral particle composed of at least one AAV capsid protein and an encapsidated polynucleotide rAAV vector. If the particle comprises a heterologous polynucleotide (i.e., a polynucleotide other than a wild-type AAV genome such as a polynucleotide or a nucleic acid expression cassette to be delivered to a mammalian cell), it is typically referred to as an “rAAV vector particle” or simply an “rAAV vector.” Thus, production of rAAV particle necessarily includes production of an rAAV vector, as such a vector is contained within an rAAV particle.

[0112] The cloning capacity of vectors or viral expression vectors can be a particular challenge for expression of large polynucleotides. For example, AAV vectors typically have a packaging capacity of ~4.8kb, lentiviruses typically have a capacity of ~8kb, adenoviruses typically have a capacity of ~7.5kb, and alphaviruses typically have a capacity of ~7.5 kb. Some viruses can have larger packaging capacities, for example herpesvirus can have a capacity of

>30kb and vaccinia a capacity of ~25kb. Advantages of using AAV for gene therapy include low pathogenicity, very low frequency of integration into the host genome, and the ability to infect dividing and non-dividing cells.

[0113] Several serotypes of AAV, non-pathogenic parvovirus, have been engineered for the purposes of gene delivery, some of which are known to have tropism for certain tissues or cell types. Viruses used for various gene-therapy applications can be engineered to be replication-deficient or to have low toxicity and low pathogenicity in a subject or a host. Such virus-based vectors can be obtained by deleting all, or some, of the coding regions from the viral genome, and leaving intact those sequences (e.g., inverted terminal repeat sequences) that are necessary for functions such as packaging the vector genome into the virus capsid or the integration of vector nucleic acid (e.g., DNA) into the host chromatin. A nucleic acid expression cassette comprising a polynucleotide, for example, can be cloned into a viral backbone such as a modified or engineered viral backbone lacking viral genes, and used in conjunction with additional vectors (e.g., packaging vectors), which can, for example, when co-transfected, produce recombinant viral vector particles.

[0114] In some cases, an AAV vector or an AAV viral particle, or virion, used to deliver a nucleic acid expression cassette into a cell, cell type, or tissue, in vivo or in vitro, is replication-deficient. In some cases, an AAV virus is engineered or genetically modified so that it can replicate and generate virions only in the presence of helper factors.

[0115] In some embodiments, a nucleic acid expression cassette is designed for delivery by an AAV or a recombinant AAV (rAAV). In some embodiments, a nucleic acid expression cassette is delivered using a lentivirus or a lentiviral vector. In some embodiments, larger polynucleotide, i.e., genes that exceed the cloning capacity of AAV, are preferably delivered using a lentivirus or a lentiviral vector.

[0116] The nucleic acid expression cassette can be designed for delivery by an optimized therapeutic retroviral vector, e.g., a lentiviral vector. The retroviral vector can be a lentiviral vector comprising a left (5') LTR; sequences which aid packaging and/or nuclear import of the virus, at least one regulatory element, optionally a lentiviral Rev response element (RRE); optionally a promoter or active portion thereof; a polynucleotide operably linked to one or more regulatory elements; optionally an insulator; and a right (3') retroviral LTR. A lentiviral vector can also include a posttranscriptional regulatory element, such as the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE). A lentiviral vector can be a self-inactivating (SIN) lentiviral vector. Any suitable packaging system can be used with a lentiviral vector, including second, third, and fourth generation packaging systems, for example. A

lentiviral vector can be pseudotyped. Any envelope glycoprotein can be used for pseudotyping, including, for example, a glycoprotein from vesicular stomatitis virus (VSV), rabies virus, Lyssavirus, Mokola virus, lymphocytic choriomeningitis virus (LCMV), Lassa fever virus (LFV), retroviruses, Moloney murine leukemia virus (MuLV), filoviruses, paramyxoviruses, measles virus, Nipah virus, orthomyxoviruses, and others. A lentiviral vector can be pseudotyped to alter tropism. Any cell type can be targeted by pseudotyping, including neuronal cells, for example.

[0117] Methods of Treatment

[0118] Methods of treating a neurodegenerative disease, disorder, or condition comprising administering to a subject in need thereof a nucleic acid expression cassette described herein. Any neurodegenerative disease, disorder, or condition can be treated with the nucleic acid expression cassettes provided herein. In some embodiments, the neurodegenerative disease, disorder, or condition is Alzheimer's disease, familial Alzheimer's disease, sporadic Alzheimer's disease, late-onset Alzheimer's disease, frontotemporal dementia, frontotemporal lobar degeneration, Pick's disease, Lewy body dementia, memory loss, cognitive impairment, or mild cognitive impairment. Other exemplary neurodegenerative diseases, disorders, or conditions include tauopathy, primary age-related tauopathy (PART), chronic traumatic encephalopathy (CTE), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17), amyotrophic lateral sclerosis-parkinsonism-dementia (ALS-PDC, Lytico-bodig disease), ganglioglioma, gangliocytoma, meningioangiomas, postencephalitic parkinsonism, subacute sclerosing panencephalitis (SSPE), lead encephalopathy, tuberous sclerosis, pantothenate kinase-associated neurodegeneration, synucleinopathy, Parkinson's disease, multiple system atrophy (MSA), neuroaxonal dystrophies, Parkinson's-like disease, Parkinsonism, prion diseases, motor neuron diseases, dementia, transmissible spongiform encephalopathies, systemic atrophies primarily affecting the central nervous system, trinucleotide repeat disorders, proteopathies, amyloidosis, neuronal ceroid lipofuscinoses, amyotrophic lateral sclerosis (ALS), Huntington's disease, traumatic brain injury, stroke, autism spectrum disorder (ASD), depression, anxiety, post-traumatic stress disorder (PTSD), schizophrenia, Attention-Deficit/Hyperactivity Disorder (ADHD), bipolar disorder, Obsessive-Compulsive Disorder (OCD), personality disorder, pain, and others.

[0119] Familial Alzheimer's disease (FAD) or early-onset familial Alzheimer's disease (EOFAD) is an uncommon form of Alzheimer's disease that usually strikes earlier in life, defined as before the age of 65 (usually between 50 and 65 years of age). FAD is inherited by

autosomal dominant mutation. Mutations in three different genes have been identified as responsible for the development of FAD, and other genes are being studied. As used here, “FAD” refers to an Alzheimer's disease caused by a mutation in any of those three genes, which code for presenilin 1 (PSEN-1), presenilin 2 (PSEN-2), and amyloid precursor protein (APP). “PSEN-1 mediated FAD” is meant to only refer to FAD caused by a mutation in the PSEN-1 gene.

[0120] As used herein, the terms “treat,” “treatment,” “therapy,” “therapeutic,” and the like refer to obtaining a desired pharmacologic and/or physiologic effect, including, but not limited to, alleviating, delaying or slowing the progression, reducing the effects or symptoms, inhibiting, ameliorating the onset of a disease or disorder, obtaining a beneficial or desired result with respect to a disease, disorder, or medical condition, such as a therapeutic benefit and/or a prophylactic benefit. “Treatment,” as used herein, covers any treatment of a disease in a mammal, particularly in a human, and includes: (a) inhibiting the disease, i.e., arresting its development; and (b) relieving the disease, i.e., causing regression of the disease. A therapeutic benefit includes eradication or amelioration of the underlying disorder being treated. Also, a therapeutic benefit is achieved with the eradication or amelioration of one or more of the physiological symptoms associated with the underlying disorder such that an improvement is observed in the subject, notwithstanding that the subject may still be afflicted with the underlying disorder. The methods of the present disclosure may be used with any mammal or other animal. In some cases, the treatment can result in a decrease or cessation of symptoms. A prophylactic effect includes delaying or eliminating the appearance of a disease or condition, delaying or eliminating the onset of symptoms of a disease or condition, slowing, halting, or reversing the progression of a disease or condition, or any combination thereof.

[0121] As used herein, the term “subject” refers to any individual or patient on which the methods disclosed herein are performed. The term “subject” can be used interchangeably with the term “individual” or “patient.” The subject can be a human, although the subject may be an animal, as will be appreciated by those in the art. Thus, other animals, including mammals such as rodents (including mice, rats, hamsters and guinea pigs), cats, dogs, rabbits, farm animals including cows, horses, goats, sheep, pigs, etc., and primates (including monkeys, chimpanzees, orangutans and gorillas) are included within the definition of subject.

[0122] The vectors provided herein can be administered in an amount effective to treat the neurodegenerative disease, disorder, or condition. The term “effective amount” or “therapeutically effective amount” refers to that amount of a composition described herein that is sufficient to affect the intended application, including but not limited to disease treatment,

as defined herein. The therapeutically effective amount may vary depending upon the intended treatment application (*in vivo*), or the subject and disease condition being treated, e.g., the weight and age of the subject, the severity of the disease condition, the manner of administration and the like, which can readily be determined by one of ordinary skill in the art. The term also applies to a dose that will induce a particular response in a target cell. The specific dose will vary depending on the particular composition chosen, the dosing regimen to be followed, whether it is administered in combination with other compounds, timing of administration, the tissue to which it is administered, and the physical delivery system in which it is carried. Exemplary AAV vector doses that can be administered include about 10^3 genome copies (GC)/kg, 10^4 GC/kg, 10^5 GC/kg, 10^6 GC/kg, 10^7 GC/kg, 10^8 GC/kg, 10^9 GC/kg, 10^{10} GC/kg, 10^{11} GC/kg, 10^{12} GC/kg, 10^{13} GC/kg, 10^{14} GC/kg, and any number or range in between, although higher or lower doses can be used.

[0123] Nucleic acid expression cassettes can be delivered by any suitable method or vectors. Exemplary methods include intracranial injection, stereotaxic injection, and intravenous injection. In some embodiments, nucleic acid expression cassettes are delivered as viral vectors.

[0124] The procedures described herein employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, genetics, immunology, cell biology, cell culture and transgenic biology, which are within the skill of the art. (See, e.g., Maniatis, et al., *Molecular Cloning*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1982); Sambrook, et al., (1989); Sambrook and Russell, *Molecular Cloning*, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001); Ausubel, et al., *Current Protocols in Molecular Biology*, John Wiley & Sons (including periodic updates) (1992); Glover, *DNA Cloning*, IRL Press, Oxford (1985); Russell, *Molecular biology of plants: a laboratory course manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984); Anand, *Techniques for the Analysis of Complex Genomes*, Academic Press, NY (1992); Guthrie and Fink, *Guide to Yeast Genetics and Molecular Biology*, Academic Press, NY (1991); Harlow and Lane, *Antibodies*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1988); *Nucleic Acid Hybridization*, B. D. Hames & S. J. Higgins eds. (1984); *Transcription And Translation*, B. D. Hames & S. J. Higgins eds. (1984); *Culture Of Animal Cells*, R. I. Freshney, A. R. Liss, Inc. (1987); *Immobilized Cells And Enzymes*, IRL Press (1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology*, Academic Press, Inc., NY; *Methods In Enzymology*, Vols. 154 and 155, Wu, et al., eds.; *Immunochemical Methods In Cell And Molecular Biology*, Mayer

and Walker, eds., Academic Press, London (1987); *Handbook Of Experimental Immunology*, Volumes I-IV, D. M. Weir and C. C. Blackwell, eds. (1986); Riott, *Essential Immunology*, 6th Edition, Blackwell Scientific Publications, Oxford (1988); Fire, et al., *RNA Interference Technology From Basic Science to Drug Development*, Cambridge University Press, Cambridge (2005); Schepers, *RNA Interference in Practice*, Wiley-VCH (2005); Engelke, *RNA Interference (RNAi): The Nuts & Bolts of siRNA Technology*, DNA Press (2003); Gott, *RNA Interference, Editing, and Modification: Methods and Protocols (Methods in Molecular Biology)*, Human Press, Totowa, N.J. (2004); and Sohail, *Gene Silencing by RNA Interference: Technology and Application*, CRC (2004)).

[0125] The compositions and methods are more particularly described below and the Examples set forth herein are intended as illustrative only, as numerous modifications and variations therein will be apparent to those skilled in the art.

[0126] The terms used in the specification generally have their ordinary meanings in the art, within the context of the compositions and methods described herein, and in the specific context where each term is used. Some terms have been more specifically defined below to provide additional guidance to the practitioner regarding the description of the compositions and methods.

[0127] All patents, patent applications, and other scientific or technical writings referred to anywhere herein are incorporated by reference herein in their entirety. The embodiments illustratively described herein suitably can be practiced in the absence of any element or elements, limitation or limitations that are specifically or not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising", "consisting essentially of", and "consisting of" may be replaced with either of the other two terms, while retaining their ordinary meanings. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by embodiments, optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the description and the appended claims.

[0128] Whenever a range is given in the specification, for example, a temperature range, a time range, or a composition or concentration range, all intermediate ranges and subranges, as

well as all individual values included in the ranges given are intended to be included in the disclosure. It will be understood that any subranges or individual values in a range or subrange that are included in the description herein can be excluded from the aspects herein. It will be understood that any elements or steps that are included in the description herein can be excluded from the claimed compositions or methods

[0129] In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group.

[0130] Presented below are examples describing therapeutic polynucleotides encoding presenilin 1, contemplated for the discussed applications. The following are provided for exemplification purposes only, to further illustrate the embodiments of the present invention, and are not intended to limit the scope of the invention described in broad terms above. While they are typical of those that might be used, other procedures, methodologies, or techniques known to those skilled in the art may alternatively be used.

EXAMPLES

EXAMPLE 1

[0131] This example describes modification of the human presenilin 1 (PSEN1) cDNA by codon-optimization.

[0132] The native cDNA sequence of the human presenilin 1 gene (PSEN1; GenBank Accession No. NM_000021.4, SEQ ID NO:1) is shown below (SEQUENCES section). The coding sequences are underlined in SEQ ID NO: 1. The coding sequence present in SEQ ID NO: 1 is repeated as SEQ ID NO: 15 which is broken up into codons and the intolerant codons are underlined. The open reading frame encoding the protein itself corresponds to nucleotides (nt) 213 through 1616. Notable elements of the mRNA are the long 5' untranslated sequence (212 nt) and long 3' untranslated sequence (4012 nt). The start codon at nt 213 is preceded by a weak Kozak translation initiation signal of CTCCA, missing the A residue at the -3 position relative to the A residue of the AUG start codon defined as +1.

[0133] The native PSEN1 cDNA (SEQ ID NO:1) was modified by codon optimization. Several methods of codon optimization can be used in which the DNA sequence encoding the protein is changed in ways that do not affect protein sequence. Codon optimization identifies preferred codons based on statistical surveys of codon usage or abundance of cognate tRNA level in cells.

[0134] SEQ ID NO:6 is a codon optimized PSEN1 cDNA that was generated by modified codon optimization with the additional constraint of allowing only tolerant synonymous codon changes. Tolerability was determined by comparison of DNA sequences encoding the same protein in related species. For example, a codon choice that is the same in all related species implies that changing this codon would not be tolerated. Thus, only codons that tolerate change are modified to preferred synonymous codons.

[0135] SEQ ID NO:6 was generated based on the open reading frame of n=11 primate cDNAs (Table 1). The cDNAs sequences were obtained from GenBank and aligned using CLUSTAL OMEGA facility (ebi.ac.uk/Tools/msa/clustalo/). Of the 467 codons in the human cDNA for PSEN1, 267 were invariant among all 11 species. These were designated as codons that could not tolerate change and were preserved in SEQ ID NO:6.

Table 1. Open Reading Frames of Primate PSEN1 cDNAs.

No.	GenBank Code	Order	Genus / Species	Common Name
1	NM_000021.4	Ape	<i>Homo sapiens</i>	Human
2	XM_024231237.1	Ape	<i>Pongo abelii</i>	Orang Utan
3		Ape	<i>Nomascus leucogenys</i>	Gibbon
4	XM_007987195.1	Old world	<i>Chlorocebus sabaeus</i>	African Green
5	XM_028850213.1	Old world	<i>Macaca mulatta</i>	Rhesus
6	XM_011983100.	Old world	<i>Mandrillus leucophaeus</i>	Drill
7		New World	<i>Aotus nancymae</i>	Ma's night monkey
8	XM_003924504.2	New World	<i>Saimiri boliviensis boliviensis</i>	Bolivian squirrel monkey
9	XM_021718361.1	Tasiers	<i>Carlito syrichta</i>	Philippine tarsier
10	XM_012656183.1	Prosimian	<i>Propithecus coquereli</i>	Coquerels sifaka
11	NM_001309945.1	Prosimian	<i>Microcebus murinus</i>	Gray mouse lemur

[0136] For the remaining codons where change is tolerated, synonymous codons were chosen in locations with bias of codon choice in human genes.

[0137] Step 1. Took human cDNA. Accept all 267 codons conserved across 11 primates. Changed tolerant codons according to rules in Table 2.

Table 2: Preferred Optimized Codons

Amino acid	Most preferred codon	Preferred codon	Non-preferred	Highly non-preferred
A (Alanine)	GCC	GCT, GCA		GCG
C (Cysteine)		TGT, TGC		
D (Aspartate)		GAT, GAC		
E (Glutamate)		GAA, GAG		
F (Phenylalanine)		TTT, TTC		
G (Glycine)	GGC	GGA, GGG	GGT	
H (Histidine)	CAC		CAT	
I (Isoleucine)	ATC	ATT	ATA	
K (Lysine)		AAA, AAG		
L (Leucine)	CTG	CTC	TTG, CTT	TTA, CTA
M (Methionine)		ATG		
N (Asparagine)		AAT, AAC		
P (Proline)	CCT	CCC, CCA		CCG
Q (Glutamine)	CAG		CAA	
R (Arginine)	AGA, AGG	CGC, CGG	CGA	CGT
S (Serine)	AGC	TCT, TCC, TCA, AGT		TCG
T (Threonine)	ACC	ACT, ACA		ACG
V (Valine)	GTG	GTC	GTT, GTA	
W (Tryptophan)		TGG		
Y (Tyrosine)		TAT, TAC		
STOP	TGA	TAA	TAG	

EXAMPLE 2

[0138] This example describes modification of the human presenilin 1 (PSEN1) cDNA by elimination of CpG dinucleotides.

[0139] In mammalian DNA, there is selective methylation of CpG dinucleotides. This affects the recruitment of chromatin proteins which in turn affects gene expression. In addition, there is an innate immune response to newly introduced DNA aimed at eliminating viral infection. This innate immune response is mediated by toll like receptor 9 (TLR9) which recognizes unmethylated CpG dinucleotides.

[0140] SEQ ID NO:7 uses the redundancy of the genetic code to completely eliminate CpG dinucleotides in the PSEN1 cDNA. The number of CpG dinucleotides was reduced from 24 in the native cDNA to zero. Elimination of CpG dinucleotides can reduce recognition of a viral vectors such as AAV, for example, and polynucleotides by antigen presenting cells and reduce immune responses to gene therapy, thereby prolonging polynucleotide expression and reducing the need for immunosuppressive therapies. However, it was not possible to both eliminate all CpG dinucleotides while also maintaining all intolerant codons. Thus, SEQ ID NO:7 includes changes to six intolerant codons as indicated by underlining in that sequence.

[0141] SEQ ID NO:36 uses the redundancy of the genetic code to eliminate as many CpG dinucleotides in the PSEN1 cDNA as possible without altering any intolerant codons. In this construct, the number of CpG dinucleotides was reduced from 24 in the native cDNA to five.

EXAMPLE 3

[0142] This example describes modification of the human presenilin 1 (PSEN1) cDNA by inclusion of genomic sequences.

[0143] Previous gene therapy constructs have introduced exogenous or artificial introns into the expression cassette. For example, many AAV vectors used in clinical trials use the CAG promoter which includes an artificial hybrid intron of chicken beta actin and rabbit beta globin genes.

[0144] SEQ ID NO:8 is a hybrid genomic / cDNA PSEN1 gene sequence intended to direct pre-mRNA into the splicing apparatus and thereby enhance nuclear export and overall mRNA levels. SEQ ID NO:8 represents a shortened genomic version of PSEN1 that includes exons 2, intron 2, exon 3, intron 3, exon 4, intron 4 followed by the remainder of the protein coding gene in cDNA form. Introns 3 and 4 are too large to be inserted into an AAV gene transfer vector, for example, and are therefore internally shortened. Without being limited by theory, generally, splicing factors bind near the ends of introns and therefore internal deletions do not interfere with splicing.

[0145] Importantly, PSEN1 mRNA is found in two forms, one encoding the most abundant protein of length 467 amino acid and an alternate version (X2) encoding a 463 amino acid version of presenilin 1. There are two splice donors at the beginning of intron 3 separated by 12 nt. Alternate splicing at this location leads to different mRNAs encoding proteins that differ by a deletion of 4 amino acids. The significance of this alternative splicing is unknown but isoform X2 is seen across a wide range of primates (e.g., marmots: Gen Bank references XP_027787309.1 presenilin-1 isoform and XP_027787310.1 presenilin-1 isoform X2). This suggests some physiological significance.

[0146] SEQ ID NO:8 was designed with important features of intron 4 that allow for alternative splicing to produce isoforms X1 and X2 is enabled. Without being limited by theory, SEQ ID NO:8 will express both isoforms and therefore provide the full range of physiological effects that are provided by the native PSEN1 gene.

EXAMPLE 4

[0147] This example describes identification of neuron-specific promoter sequences.

[0148] PSEN1 expression should be specifically restricted to neurons to prevent A β accumulation in neurons. Previously reported AAV gene therapy vectors with neuron-specific expression included the neuron-specific elastase and synapsin 1 promoters.

[0149] The availability of RNA-Seq data from multiple cell types allowed an unbiased search for highly expressed neuron-specific genes (see web.stanford.edu/group/barres_lab/brain_rnaseq.html). Genes with high neuronal to endothelial expression ratio were identified and sorted by decreasing neuron expression level. Genes were then manually inspected to exclude candidates with potentially confounding factors (e.g., maternal expression / multiple transcription start sites) that might limit utility. Two novel highly expressed neuron-specific promoter sequences, SEQ ID NO:2 and SEQ ID NO:3, were identified by this method.

[0150] SEQ ID NO:2 includes a 480 base pair (bp) fragment of the human somatostatin gene (SST) from -407 to +73 relative to transcription start site. The use of the SST promoter from rhesus macaque has been previously reported with a fragment of ~ 300 bp that drives neuron-specific expression of a reporter polynucleotide in the context of lentiviral gene transfer.

[0151] SEQ ID NO:3 includes a 1000 bp segment from -952 to +48 relative to the mRNA start of the human neuropeptide Y (NPY) promoter. This will provide a highly specific expression pattern in brain.

EXAMPLE 5

[0152] This example describes regulatory elements to increase polynucleotide expression.

[0153] SEQ ID NO:4 from the human beta globin locus called HS4 can function as a chromatin insulator sequence. It has been used in the context of lentiviral gene transfer vectors to ensure ongoing expression of introduced polynucleotides.

[0154] SEQ ID NO:5 is a Kozak translation initiation signal. It can be used to replace the weak non-consensus Kozak signal in the native mRNA of the PSEN1 gene.

[0155] SEQ ID NO:4 and/or SEQ ID NO:5 can be used in nucleic acid expression cassettes in combination with any of the elements and features described herein. For example, SEQ ID NO:4 and/or SEQ ID NO:5 can be used in nucleic acid expression cassettes that include any one of the synthetic PSEN1 cDNA sequences set forth in SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8. The nucleic acid expression cassettes can further include any one of the neuron-specific promoters of SEQ ID NO:2 or SEQ ID NO:3. In addition, the nucleic acid expression

cassettes can include any one of the sequences set forth in SEQ ID NO:9, SEQ ID NO:10, or SEQ ID NO:11 described below (Example 6) that enhance mRNA expression by providing mRNA stability or enhancing mRNA transcription and processing, or any combination of SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:11.

EXAMPLE 6

[0156] This example describes regulatory sequences that enhance polynucleotide expression by conferring stability to mRNA or enhancing transcription and processing of mRNA.

[0157] SEQ ID NO:9 is an expression and nuclear retention element that confers mRNA stability. Expression and nuclear retention elements stabilize mRNAs by making complex secondary structures with the terminal polyadenylated sequence of the mRNA, thereby inhibiting 3' to 5' degradation. Without being limited by theory, the insertion of this sequence beyond the open reading frame and before polyadenylation site will provide promoter mRNA stability.

[0158] SEQ ID NO:10 corresponds to the 3' non-coding sequence of the native PSEN1 cDNA. Without being limited by theory, 3' untranslated sequences may contain important elements that enhance mRNA transcription and processing, thereby enhancing polynucleotide or gene expression. SEQ ID NO:10 in part or in its entirety can be appended to the 5' end of any presenilin coding sequence to enhance expression level.

[0159] SEQ ID NO:11 corresponds to the 5' non-coding sequence of the native PSEN1 cDNA. Without being limited by theory, 5' untranslated sequences can contain important elements that enhance mRNA stability, thereby enhancing polynucleotide or gene expression. SEQ ID NO:11 in part or in its entirety can be appended to the 3' end of any presenilin encoding sequence to enhance expression level.

[0160] SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:11 can be used in any combination in nucleic acid expression cassettes described herein. Nucleic acid expression cassettes that include SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11 or any combination of SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:11 can have any of the combination of elements and features described herein. For example, an expression cassette that includes SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11 or any combination of SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:11 can include any one of the synthetic PSEN1 cDNA sequences set forth in SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8. The nucleic acid expression cassettes can further include any one of the neuron-specific promoters of SEQ ID NO:2 or SEQ ID NO:3. Nucleic acid expression

cassettes can also include further regulatory elements that increase polynucleotide expression, such as SEQ ID NO:4, SEQ ID NO: 5, or both.

EXAMPLE 7

[0161] This example describes design of presenilin 1 (PSEN1) expression cassettes.

[0162] Any elements and features described herein, including sequences set forth in SEQ ID NOs:2-11, can be combined into presenilin 1 expression cassettes. For example, nucleic acid expression cassettes can include any one of the synthetic cDNA sequences set forth in SEQ ID NO:6, SEQ ID NO:7, or SEQ ID NO:8 that encode PSEN1. Expression of any one of the synthetic cDNAs can be driven by a neuron-specific promoter of SEQ ID NO:2 derived from the human somatostatin (SST) gene or SEQ ID NO:3 derived from the human neuropeptide Y (NPY) promoter.

[0163] A nucleic acid expression cassette that has any of the synthetic cDNA sequences and promoter sequences described above can further include any of the elements that increase polynucleotide expression, including, for example, a chromatin insulator sequence of SEQ ID NO:4, a Kozak consensus sequence of SEQ ID NO:5, an mRNA stability element of SEQ ID NO:9, a 3' non-coding sequence of SEQ ID NO:10 derived from the native PSEN1 cDNA, a 5' non-coding sequence of SEQ ID NO:11 derived from the native PSEN1 cDNA, or any combination of these elements. Selection of elements can be based on desired levels of expression, for example. For example, expression levels can vary with cell type or the brain region a neuron is found in, which can be used as a guide or criterion for inclusion or exclusion of regulatory elements that affect any step in gene expression, such as mRNA transcription, processing, stability, and/or translation, for example.

[0164] The nucleic acid expression cassettes can be included in a viral vector, for example. Any viral vector can be used, including adeno-associated virus (AAV) vectors, lentiviral vectors, retroviral vectors, and adenoviral vectors, for example.

EXAMPLE 8

[0165] This example describes the synthesis of two different codon-optimized PSEN-1 constructs and the expression of presenilin 1 protein from each of the constructs, as well as from a construct comprising wild-type PSEN-1 coding sequence.

[0166] Constructs encoding codon-optimized human presenilin 1 were designed by making changes to the cDNA sequence encoding wild-type PSEN-1 only at codons that are variable across primate sequences. The wild-type PSEN-1 cDNA sequence contains 267 codons that are conserved across 11 primate sequences (see underlined codons in SEQ ID NO:15). These

intolerant codons were left unchanged. The remaining 200 tolerant codons in the wild-type cDNA were considered for optimization.

[0167] For one construct (v2.0), conservative codons changes (34 codons in lower cases) were made to construct SEQ ID NO:38. Native wild-type PSEN-1 cDNA codons were conserved for codons translated into: phenylalanine, tyrosine, cysteine, histidine, asparagine, lysine, aspartic acid and glutamic acid because only two different codons of equal usage in primates encode these amino acids. For glutamine-encoding codons, CAG was preferred. For isoleucine-encoding codons, ATA codons were changed to ATC and either ATC or ATT were maintained if present in native sequence. Methionine (ATG) and tryptophan (TGG) encoding codons were unchanged. For proline-, threonine-, and alanine-encoding codons, every codon terminating with a guanine (G) was changed into a redundant codon terminating with a cytosine (C). For valine- and glycine-encoding codons, every codon terminating with a thymine (T) or adenine (A) was changed into a redundant codon terminating with a cytosine (C) or guanine (G), respectively. AGG, AGA, CGC, CGG, AGT, AGC, TCC, TCT, TCA, TTG, CTC and CTG codons were left unchanged. CGT codons were changed to CGC; CGA codons were changed to CGG; TCG codons were changed to TCC; TTA codons were changed to TTG; CTT codons were changed to CTC; and CTA codons were changed to CTG.

[0168] For another construct (v1.5) more codons changes (138 codons in lower cases) were made to construct SEQ ID NO:37 compared to native sequence. In this construct, native codons were conserved for codons translated into: tryptophan, cysteine and methionine. For glutamine-encoding codons, selected CAA codons were changed to CAG. For isoleucine-encoding codons, selected ATA and ATT codons were changed to ATC. For proline-encoding codons, selected codons were changed to CCC or CCT. For threonine-encoding codons, selected codons were changed to ACC or ACA. For alanine-encoding codons, selected codons were changed to GCC or GCT. Glycine-encoding codons not terminating with a cytosine (C) were changed into redundant codons terminating with a cytosine (C). For valine-encoding codons, GTG was preferred. AGC codons were left unchanged. For aspartic acid-encoding codons, selected codons were changed to GAT or GAC. For glutamic acid-encoding codons, GAA or GAG were preferred. For phenylalanine-encoding codons, TTT codons were changed to TTC. For histidine-encoding codons, CAT codons were changed to CAC. For lysine-encoding codons, selected AAA codons were changed to AAG. For leucine-encoding codons, most selected codons were changed to CTG. For asparagine-encoding codons, AAT codons were changed to AAC. For arginine-encoding codons, AGA was preferred, but AGG and CGG codons were also used. For serine-encoding codons, AGC was preferred. but TCC and TCT

were also used on selected codons. For tyrosine-encoding codons, TAT codons were changed to TAC.

[0169] Each of the PSEN-2.0, PSEN-1.5 and wild-type PSEN-1 coding sequences were separately cloned into cloning vector pCMV6-XL5 (Origene, Rockville, MD). The resulting constructs (WT (pAT001), v1.5 (pAT010) and v2.0 (pAT012)) were transfected into HEK293 cells to determine the effect of codon optimization on presenilin 1 expression. The 293 cells were harvested 48 hours post-transfection, lysed using 300 μ L of RIPA buffer (50mM Base / Tris-HCl, 150mM Sodium Chloride, 0.5% Sodium Deoxycholate, 0.1% Sodium Dodecyl Sulfate, 1% Nonidet P-40 substitute with added cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Sigma Aldrich), and the supernatant was collected. The total protein concentration of each sample was measured using the THERMO SCIENTIFIC™ PIERCE™ BCA™ Protein Assay according to the manufacturer's instructions.

[0170] ELISAs for detecting human presenilin 1 (PS1) protein in cell lysates were performed using the RayBio® Human Presenilin 1 ELISA Kit. Dilutional linearity (DL) and spike-in recovery (SR) was assessed using untransfected 293 cells to test the compatibility of cell lysates with this ELISA kit. Table 3 shows that cell lysates exhibit acceptable dilutional linearity (1:250-1:1000) and spike-in recovery.

Table 3. Dilutional linearity and spike-in recovery

Sample ID	Expected concentration (pg/mL)	Interpolated Values (pg/mL)		Accuracy (%RE)	Precision (%CV)	SR/DL %
Neat sample (cell lysate)	--	67.18	63.88	--	3.56%	--
Control Spike (assay diluent)	750	800.47	834.07	8.97%	2.91%	
Sample Spike (cell lysate)	750	808.80	1068.02	25.12%	19.53%	99%
1:250 dilution (cell lysate)	3000	2490.14	2654.88	14.25%	4.53%	83%
1:500 dilution (cell lysate)	1500	1305.48	1320.87	12.46%	0.83%	87%
1:1000 dilution (cell lysate)	750	747.87	766.82	0.98%	1.77%	100%

[0171] ELISAs for detecting human presenilin 1 (PS1) protein in transfected 293 cell lysates were performed using the RayBio® Human Presenilin 1 ELISA Kit. Technical duplicates were run at a 1:1000 dilution according to the manufacturer's instructions. Table 4 and Fig. 1 shows that transfection with plasmid pAT010 resulted in a 2.5-fold increase in PS1 expression when compared with the native sequence.

Table 4. PS1 expression in transfected 293 cell lysates

Sample ID	Interpolated Values (pg/mL)		Mean (pg/mL)	Precision (%CV)	Dilution	ug/mL total protein	pg PS1/ug total protein
WT pAT001 replicate 1	196.63	202.47	199.55	2%	1000	854.3	233.59
WT pAT001 replicate 2	225.13	227.79	226.46	1%	1000	820.4	276.03
V1.5 pAT010 replicate 1	486.76	499.90	493.33	2%	1000	675.4	730.39
V1.5 pAT010 replicate 2	418.15	413.65	415.90	1%	1000	767.1	542.14
V2.0 pAT012 replicate 1	209.48	240.41	224.95	10%	1000	824.1	272.95
V2.0 pAT012 replicate 2	215.98	223.97	219.98	3%	1000	645.9	340.60

EXAMPLE 9

[0172] This example describes another codon-optimized PSEN-1 coding sequence.

[0173] For construct v3.0 (SEQ ID NO:39) 140 tolerant codons (as indicated in lower case) were changed, as was the stop codon, compared to codons present in wild-type PSEN-1 coding sequence. Methionine (ATG) and tryptophan (TGG) encoding codons were unchanged. For glutamine (Q)-encoding codons, all tolerant codons were changed to CAG. For isoleucine (I)-encoding codons, all tolerant codons were changed to ATC. For proline (P)-encoding codons, all tolerant codons were changed to CCC. For threonine (T)-encoding codons, all tolerant codons were changed to ACC. For alanine (A)-encoding codons, all tolerant codons were changed to GCC. For glycine (G)-encoding codons, all tolerant codons were changed to GGC. For valine (V)-encoding codons, all tolerant codons were changed to GTG. For aspartic acid (D)-encoding codons, all tolerant codons were changed to GAC. For glutamic acid (E)-encoding codons, all tolerant codons were changed to GGC. For phenylalanine (F)-encoding codons, all tolerant codons were changed to TTC. For histidine (H)-encoding codons, all tolerant codons were changed to CAC. For lysine (K)-encoding codons, all tolerant codons were changed to AAG. For leucine (L)-encoding codons, all tolerant codons were changed to

CTG. For asparagine (N)-encoding codons, all tolerant codons were changed to ACC. For arginine (R)-encoding codons, all tolerant codons were changed to AGA, except for codon 307 that was changed from AGG to CGG. For serine (S)-encoding codons, all tolerant codons were changed to AGC. For tyrosine (Y)-encoding codons, all tolerant codons were changed to TAC.

EXAMPLE 10

[0174] This example describes the relative expression levels of PSEN1 driven by the CAG promoter from two different codon-optimized PSEN-1 coding sequence as compared to the wild-type PSEN1 coding sequence.

[0175] Plasmid pAAV-CAG-MCS (Vector Biolabs) was modified by replacing the ampicillin antibiotic resistance gene with a kanamycin resistance gene. The sequence of the CAG promoter therein is set forth in SEQ ID NO:40 and has 98% sequence identity to SEQ ID NO:23. We then inserted the PSEN-1 coding sequence of either SEQ ID NO:39 (“CAG-v3.0”; pAT029), SEQ ID NO:37 (“CAG-v1.5”; pAT024), or the wild-type PSEN-1 coding sequence (SEQ ID NO:15; “CAG-native”; pAT022) into the modified plasmid and used those plasmids to transfect HEK293 cells in order to determine the effect of codon optimization on presenilin 1 expression. The 293 cells were harvested 48 hours post-transfection, lysed using 300 μ L of RIPA buffer (50mM Base/Tris-HCl, 150mM NaCl, 0.5% Sodium Deoxycholate, 0.1% Sodium Dodecyl Sulfate, 1% Nonidet P-40 substitute with added cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Sigma Aldrich), and the supernatant was collected. The total protein concentration of each sample was measured using the THERMO SCIENTIFIC™ PIERCE™ BCA™ Protein Assay according to the manufacturer’s instructions.

[0176] ELISAs for detecting human presenilin 1 (PS1) protein in cell lysates were performed using the RayBio® Human Presenilin 1 ELISA Kit. Technical duplicates were run at a 1:40 dilution according to the manufacturer’s instructions. Data was analyzed using the two-tailed t test. Fig. 2 shows that transfection with plasmid comprising codon-optimized SEQ ID NOs:37 or 39 resulted in increased in PS1 expression when compared with the wild-type coding sequence and that the level of PSEN-1 expression from the SEQ ID NO:37-containing plasmid showed a statistically significant increase over that from the wild-type coding sequence (as indicated by the asterisk $p < 0.05$). PSEN-1 expression from the SEQ ID NO:39-containing plasmid showed a trend towards increased expression as compared to the wild-type coding sequence ($p = 0.100$).

[0177] In summary, synthetic cDNA sequences based on codon optimization, exclusion of CpG dinucleotides, and inclusion of genomic sequences, neuron-specific promoter sequences, and other regulatory elements that enhance any step in gene expression, such as mRNA

transcription, processing, stability, and/or translation, for example, can be combined according to desired expression levels in neurons. Combining multiple modes of enhancing polynucleotide expression by combining elements described above, some or all of which may have a relatively small effect on protein production depending on cell type, neuronal location, and other factors, can allow for a relatively large increase in expression from a nucleic acid expression cassette or vector molecule.

EXAMPLE 11

[0178] This example describes the effect of a codon-optimized PSEN-1 nucleotide sequence on the gamma-secretase activity of FAD patient fibroblasts.

[0179] Primary dermal fibroblasts from patients with Familial Alzheimer's disease (FAD) carrying C410Y or G206A mutations in PSEN1 were electroporated with plasmids encoding a Notch1 fragment (Notch1ΔE) with either an empty non-coding plasmid(pAAV-CAG-MCS-KanR) or human codon-optimized presenilin-1 (hPSEN1v1.5) of SEQ ID NO:37.

[0180] As described in Example 10, the cDNA plasmid pAAV-CAG-MCS (Vector Biolabs) was modified to replace the ampicillin resistance gene with a kanamycin resistance gene and create the resulting plasmid pAAV-CAG-MCS-KanR. The resulting plasmid pAAV-CAG-MCS-KanR was used as a control or further modified to contain the codon-optimized human presenilin 1 coding sequence (SEQ ID NO:37) to evaluate functional gamma-secretase activity. The Notch1ΔE plasmid encodes the transmembrane domain and a portion of the intracellular domain of human Notch1 but lacks the entire extracellular domain of Notch 1. Inside cells, the Notch1ΔE is cleaved by gamma -secretase and can be detected by an antibody specific (Cell Signaling, #4147) for the cleaved fragment, NICD. The gamma -secretase activity can be inhibited with a known gamma-secretase inhibitor DAPT (Sigma-Aldrich, D5942).

[0181] To determine whether the codon-optimized construct is functional, NICD was measured following treatment with hPSEN1v1.5 (3 μg) compared to a non-coding plasmid in FAD patient fibroblasts containing one of two PSEN1 pathogenic mutations (C410Y or G206A). Some fibroblasts were exposed to DAPT inhibitor 24 hours after electroporation and were harvested 48 hours post-transfection, lysed using 100 μL of RIPA buffer (50mM Base/Tris-HCl, 150mM NaCl, 0.5% Sodium Deoxycholate, 0.1% Sodium Dodecyl Sulfate, 1% Nonidet P-40 substitute with added cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Sigma Aldrich), and the supernatant was collected. The total protein concentration of each sample was measured using the THERMO SCIENTIFIC™ PIERCE™ BCA™ Protein Assay according to the manufacturer's instructions.

[0182] Total protein lysates from technical duplicates were electrotransferred to nitrocellulose membrane according to the manufacturer's instructions. Western blots for detecting cleaved Notch (NICD) in cell lysates were performed using recommended conditions. Fig. 3 shows that transfection with the plasmid containing SEQ ID NO:37 resulted in increased NICD when compared with the non-coding construct (Empty) in both C410Y and G206A PSEN1 mutant fibroblasts.

EXAMPLE 12

[0183] This example describes the effect of a codon-optimized PSEN-1 nucleotide sequence on the A β 40 levels in FAD patient fibroblasts.

[0184] Primary dermal fibroblasts from patients with Familial Alzheimer's disease (FAD) carrying the C410Y mutation in PSEN 1 were electroporated with a plasmid encoding the amyloid precursor protein (APP) C99 fragment and either a non-coding empty plasmid (pAAV-CAG-MCS-KanR) or a similar plasmid containing a codon-optimized human presenilin 1 coding sequence pAT028 (SEQ ID NO:37), each of which is described above, in order to evaluate functional gamma -secretase activity. Inside the cells, the C99 is sequentially cleaved by gamma -secretase to produce A β peptides of varying lengths, which are then released into the cell culture media.

[0185] To determine the functionality of the codon-optimized construct, A β 40 was measured in the cell culture media following electroporation of the cells with the above plasmids. The cell culture media was collected 48 hours post-transfection and analyzed for A β 40 via an MSD ELISA. Data was analyzed using the two-tailed t test. Fig 4 shows that transfection with the SEQ ID NO:37-containing plasmid resulted in increased A β 40 production when compared with the non-coding construct (Empty) in C410Y FAD fibroblasts, with a trend towards statistical significance (p=0.23).

EXAMPLE 13

[0186] This example describes the synthesis of various AAV vectors comprising a partially codon optimized PSEN-1 coding sequence of the invention.

[0187] In order to prepare AAV viral vectors containing the PSEN-1 coding sequences of the invention, we removed the CAG promoter from the pAAV-CAG-MCS-KanR plasmid described above and replaced it with varying expression cassettes. Each expression cassette comprised a different promoter operatively linked to the PSEN-1 coding sequence of SEQ ID NO:37 and a polyadenylation sequence. In addition, the cassettes further comprised a human beta globin intron between the promoter and the PSEN-1 coding sequence; an HA-tag in between the human beta globin intron and the PSEN-1 coding sequence, which may be

removed prior to use in subjects; and either human growth hormone or albumin genomic stuffer sequences following the polyadenylation sequence. The sequence of each of these expression cassettes and the AAV2 ITRs that flank them (which are already present in the modified pAAV-CAG-MCS-KanR plasmid) are set forth in SEQ ID Nos: 41-46.

[0188] For the production of a self-complementary AAV vector, the 5' AAV2 ITR in pAAV-CAG-MCS-KanR is modified prior to insertion of the expression cassette. The expression cassette for this construct comprised a CBA promoter, a minute virus of mice intron, an HA-tag, which may be removed prior to use in subjects, SEQ ID NO:37, and a rabbit β -globin polyadenylation sequence. The sequence of this expression cassette including the modified 5' and native 3' AAV2 ITRs from the modified pAAV-CAG-MCS-KanR plasmid into which it was inserted, is set forth in SEQ ID NO:47.

[0189] Each of the resulting vector genome plasmids containing the expression cassette were used to create recombinant AAV vectors using the triple plasmid transfection method (Xiao and Samulski, J Virol 72: 2224-2232, 1998). This method used an AAV serotype-specific rep and cap plasmid specific to the serotype of interest as well as the vector genome DNA plasmid, but eliminated the use of Ad infection by supplying the essential Ad genes on a third plasmid. Multipasmids transient transfection of adherent HEK293 cells is a widely used method for rAAV production (Grimm et al., Hum Gene Ther 9: 2745-2760, 1998; Matsushita et al., Gene Ther 5: 938-945, 1998) and can be used to create these recombinant AAV vectors. The AAV particles may be formulated in phosphate buffered saline (PBS) or in 10mM sodium phosphate, 180mM NaCl with 0.001% of pluronic acid (F-68) at a pH of about 7.4.

SEQUENCES

SEQ ID NO:1

>NM_000021.4 Homo sapiens presenilin 1 (PSEN1), transcript variant 1, mRNA. Coding sequence underlined

GGAAACAAAACAGCGGCTGGTCTGGAAGGAACCTGAGCTACGAGCCGCGGCGG
CAGCGGGGCGGCGGGGAAGCGTATACCTAATCTGGGAGCCTGCAAGTGACAACA
GCCTTTGCGGTCCTTAGACAGCTTGGCCTGGAGGAGAACACATGAAAGAAAGAA
CCTCAAGAGGCTTTGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCAATGAC
AGAGTTACCTGCACCGTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGACAAC
CACCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAACGGCAGGAGCAC
AACGACAGACGGAGCCTTGGCCACCCTGAGCCATTATCTAATGGACGACCCAG
GGTAACTCCCGGCAGGTGGTGGAGCAAGATGAGGAAGAAGATGAGGAGCTGAC
ATTGAAATATGGCGCCAAGCATGTGATCATGCTCTTTGTCCCTGTGACTCTCTGC
ATGGTGGTGGTTCGTGGCTACCATTAAAGTCAGTCAGCTTTTATACCCGGAAGGATG
GGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGACTGTGGGCCAGAGAG
CCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACT
ATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGC
TTATTATATCATCTCTATTGTTGCTGTTCTTTTTTCATTCATTTACTTGGGGGAAG
TGTTTAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGG

AATTTTGGTGTGGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTCGACTCC
AGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTTCATCAAGTA
CCTCCCTGAATGGACTGCGTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAG
TGGCTGTTTTGTGTCCGAAAGGTCCACTTCGTATGCTGGTTGAAACAGCTCAGGA
GAGAAATGAAACGCTTTTTCCAGCTCTCATTACTCCTCAACAATGGTGTGGTTG
TTGAATATGGCAGAAGGAGACCCGGAAGCTCAAAGGAGAGTATCCAAAAATTCC
AAGTATAATGCAGAAAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAA
TGATGATGGCGGGTTCAGTGAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGG
GCCTCATCGCTCTACACCTGAGTCACGAGCTGCTGTCCAGGAACCTTCCAGCAGT
ATCCTCGCTGGTGAAGACCCAGAGGAAAGGGGAGTAAAACCTTGGATTGGGAGAT
TTCATTTTCTACAGTGTTCTGGTTGGTAAAGCCTCAGCAACAGCCAGTGGAGACT
GGAACACAACCATAGCCTGTTTCGTAGCCATATTAATTGGTTTGTGCCTTACATT
ATTACTCCTTGCCATTTTCAAGAAAGCATTGCCAGCTCTTCCAATCTCCATCACCT
TTGGGCTTGTCTTCTACTTTGCCACAGATTATCTTGTACAGCCTTTTATGGACCAA
TTAGCATTCCATCAATTTTATATCTAGCATATTTGCGGTTAGAATCCCATGGATGT
TTCTTCTTTGACTATAACAAAATCTGGGGAGGACAAAGGTGATTTTCCTGTGTCC
ACATCTAACAAAGTCAAGATTCCCGGCTGGACTTTTGCAGCTTCCTTCCAAGTCT
TCCTGACCACCTTGCACTATTGGACTTTGGAAGGAGGTGCCTATAGAAAACGATT
TTGAACATACTTCATCGCAGTGGACTGTGTCCCTCGGTGCAGAACTACCAGATT
TGAGGGACGAGGTCAAGGAGATATGATAGGCCCGGAAGTTGCTGTGCCCCATCA
GCAGCTTGACGCGTGGTCACAGGACGATTTCACTGACACTGCGAACTCTCAGGA
CTACCGTTACCAAGAGGTTAGGTGAAGTGGTTTAAACCAAACGGAACCTCTTCATC
TTAAACTACACGTTGAAAATCAACCCAATAATTCTGTATTAACCTGAATTCTGAAC
TTTTCAGGAGGTACTGTGAGGAAGAGCAGGCACCAGCAGCAGAATGGGGAATGG
AGAGGTGGGCAGGGGTTCCAGCTTCCCTTTGATTTTTTGCTGCAGACTCATCCTTT
TTAAATGAGACTTGTTTTCCCTCTCTTTGAGTCAAGTCAAATATGTAGATTGCCT
TTGGCAATTCTTCTCTCAAGCACTGACACTCATTACCGTCTGTGATTGCCATTTCT
TTCCCAAGGCCAGTCTGAACCTGAGGTTGCTTTATCCTAAAAGTTTTAACCTCAG
GTTCCAAATTCAGTAAATTTTGGAAACAGTACAGCTATTTCTCATCAATTCTCTAT
CATGTTGAAGTCAAATTTGGATTTTCCACCAAATTCTGAATTTGTAGACATACTTG
TACGCTCACTTGCCCCAGATGCCTCCTCTGTCTCATTCTTCTCTCCACACAAGC
AGTCTTTTTCTACAGCCAGTAAGGCAGCTCTGTCTGTTAGCAGATGGTCCCATT
ATTCTAGGGTCTTACTCTTTGTATGATGAAAAGAATGTGTTATGAATCGGTGCTG
TCAGCCCTGCTGTGACAGCTTCTTCCACAGCAAATGAGATGTATGCCCAAAGACG
GTAGAATTAAAGAAGAGTAAAATGGCTGTTGAAGCACTTTCTGTCTGTTATTTT
GTTTTTGCTTTTGCCACACAGTAGCTCAGAATTTGAACAAATAGCCAAAAGCTGG
TGGTTGATGAATTATGAACTAGTTGTATCAACACAAAGCAAGAGTTGGGGAAAG
CCATATTTAACTTGGTGAGCTGTGGGAGAACCTGGTGGCAGAAGGAGAACCAAC
TGCCAAGGGGAAAGAGAAGGGGCCTCCAGCAGCGAAGGGGATACAGTGAGCTA
ATGATGTCAAGGAGGAGTTTCAGGTTATTCTCGTCAGCTCCACAAATGGGTGCTT
TGTGGTCTCTGCCCCGCTTACCTTTCTCTCAATGTACCTTTGTGTGAACTGGGCA
GTGGAGGTGCCTGCTGCAGTTACCATGGAGTTCAGGCTCTGGGCAGCTCAGTCAG
GCAAAACACACAAACAGCCATCAGCCTGTGTGGGCTCAGGGCACCTCTGGACAA
AGGCTTGTGGGGCATAACCTTCTTTACCACAGAGAGCCCTTAGCTATGCTGATCA
GACCGTAAGCGTTTATGAGAACTTAGTTTCCTCCTGTGGCTGAGGAGGGGCCAG
CTTTTTCTTTTTGCTGCTGTTTTCTCTCCCAATCTATGATATGATATGACCTGG
TTTGGGGCTGTCTTTGGTGTTTAGAATATTTGTTTTCTGTCCCAGGATATTTCTTAT
AAGAACCTAACTTCAAGAGTAGTGTGCGAGTACTGATCTGAATTTAAATTAATAAT
TGGCTTATATTAGGCAGTCACAGACAGGAAAAATAAGAGCTATGCAAAGAAAGG
GGGATTTAAAGTAGTAGGTTCTATCATCTCAATTCATTTTTTTCCATGAAATCCCT
TCTTCCAAGATTCATTCCCTCTCTCAGACATGTGCTAGCATGGGTATTATCATTGA

GAAAGCACAGCTACAGCAAAGCCACCTGAATAGCAATTTGTGATTGGAAGCATT
CTTGAGGGATCCCTAATCTAGAGTAATTTATTTGTGTAAGGATCCCAAATGTGTT
GCACCTTTCATGATACATTTCTTCTCTGAAGAGGGTACGTGGGGTGTGTGATTTA
AATCCATCCTATGTATTACTGATTGTCCTGTGTAGAAAGATGGCAATTATTCTGTC
TCTTCTCCAAGTTTGAGCCACATCTCAGCCACATTGTTAGACAGTGTACAGAGA
ACCTATCTTTCCTTTTTTTTTTTTTTAAAGGACAGGATTTTGCTGTGTTGCCAGGCT
AGACTTGAACCTCCTGGGCTCAAGTAATCCACCTCAGCCTGAGTAGACTGAGACTAC
AGCCCATCTTATTTCTTTAAATCATTTCATCTCAGGCAGAGAACTTTTCCCTCAAAC
ATTCTTTTTAGAAATTAGTTCAGTCAATTCCTAAAACATCCAAATGCTAGTCTTCCAC
CATGAAAAATAGATTGTCACTGGAAAGAACAGTAGCAATTTCCATAAGGATGTG
CCTTCACTCACACGGGACAGGCGGTGGTTATAGAGTCGGGCAAAACCAGCAGTA
GAGTATGACCAGCCAAGCCAATCTGCTTAATAAAAAGATGGAAGACAGTAAGGA
AGGAAAGTAGCCACTAAGAGTCTGAGTCTGACTGGGCTACAGAATAAAGGGTAT
TTATGGACAGAATGTCATTACATGCCTATGGGAATACCAATCATATTTGGAAGAT
TTGCAGATTTTTTTTTCAGAGAGGAAAGACTCACCTTCCTGTTTTTGGTTCTCAGTA
GGTTCGTGTGTGTTCTAGAATCACAGCTCTGACTCCAAATGACTCAATTTCTCA
ATTAGAAAAAGTAGAAGCTTTCTAAGCAACTTGGAAGAAAACAGTCATAAGTAA
GCAATTTGTTGATTTTACTACAGAAGCAACAACCTGAAGAGGCAGTGTTTTACTT
TCAGACTCCGGGATTCCCATTCTGTAGTCTCTCTGCTTTTAAAAACCCTCCTTTTG
CAATAGATGCCCAAACAGATGATGTTTATTACTTGTTATTTACGTGGCCTCAGAC
AGTGTATGTATTCTCGATATAACTTGTAGAGTGTGAAATATAAGTTTAACTACCA
ATAAGGTCTCCAGGGTTAGATGACTGCGGGAAGCCTTTGATCCCAACCCCCA
AGGCTTTGTATATTTGATCATTGTGTGATCTAACCTGGAAGAAAAAGAGCTCAGA
AACCCTATGAAAAAATTTGTTTCAGTGTCTTCTGTGTTCCCGTAGGTTCTGGAGTC
TGAGGATGCAAAGATGAATAAGATAAATTCTCAGAATGTAGTTATAATCTCTTGT
TTTCTGGTATATGCCATCTTTCTTTAACTTCTCTAAAATATTGGGTATTTGTCAA
TAACCACTTTTAAACAGTTACCATTACTGAGGGCTTATACATTGGTGTATATAAAG
TGACTTGATTGAGAAATCAATCCATTGAGTAAAGTACTCCTTCTCTAAATTTGCTG
TTATGTCTATAAGGAACAGTTTGACCTGCCCTTCTCCTCACCTCCTCACCTGCCTT
CCAACATTGAATTTGGAAGGAGACGTGAAAATTGGACATTTGGTTTTGCCCTTGG
GCTGGAAACTATCATATAATCATAAAGTTTGAGCCTAGAAAGTATCCTTGTGATCT
TCTCACCTCTTTAAATTCCCACAACACAAGAGATTAAAAACAGAGGTTTCAGCTC
TTCATAGTGCGTTGTGAAATGGCTGGCCAGAGTGTACCAACAAAGCTGTATCGG
GCTCACAGCTCAGAGACATCTGCATGTGATCATCTGCATAGTCCCTCCTCTAAC
GGGAAACACCTCAGATTTGCATATAAAAAAGCACCTGGTGCTGAAATGAACCC
CTTTCTTGAACATCAAAGCTGTCTCCACAGCCTTGGGCAGCAGGGTGCCTCTTA
GTGGATGTGCTGGGTCCACCCTGAGCCCTGACATGTGGTGGCAGCATTGCCAGTT
GGTCTGTGTGTCTGTGTAGCAGGGACGATTTCCAGAAAGCAATTTTCTTTTGA
AATACGTAATTGTTGAGACTAGGCAGTTTCAAAGTCAGCTGCATATAGTAGCAAG
TACAGGACTGTCTTGTTTTTGGTGTCTTGGAGGTGCTGGGGTGAGGGTTTCAGT
GGGATCATTACTCTCACATGTTGTCTGCCTTCTGCTTCTGTGGACACTGCTTTGT
ACTTAATTCAGACAGACTGTGAATACACCTTTTTTATAAATACCTTTCAAATTCTT
GGTAAGATATAATTTGATAGCTGATTGCAGATTTTCTGTATTTGTCAGATTAATA
AAGACTGCATGAATCCA

SEQ ID NO:2

> Human SST promoter

acactaaaatgtagagtatgatgacagatggagttgtctgggtacatttgtgtgcatttaagggatagatgtatttgcctttaagagctg
agtgtttgagcctctgtttgtgttaattgagtgcatgtgtgggagtgaaattgtggaatgtgtatgctcatagcactgagtgaaataaa
agattgtataaatctggggcatgtggaattgtgtgcctgtgcgtgtgcagtatttttttaagtaagccacttagatctgtcacct
cccctgtctctgtgattgatttgcgaggctaattgtgcgtaaaagggctggtgagatctggggcgccctctagcctgacgtcagag

agagagtttaaaacagaggagacgggttgagagcacacaagccgcttaggagcgagggtcggagccatcgctgctgctgctgac
cgcgctagagtttgaccagcc

SEQ ID NO:3

> Human NPY promoter

ttttgccaggggatgtggcttgactggagagaaaggagataaggatgtaaacacatgtagggcatacccccctattttattctct
gaatcctaaccctcagaataagttcttattcttgagaatcaatgacattatcttaagctaaattaatcaagcctccacagtggtctctctcaa
tagtggtgtggccttcctagaagtaattttccaaattcagtgatactttaagttcagatttaattgatatgaatctgtgatacactctaa
aataagattattttattgaaaagtgactgtaactttccctttatctaggaagagctctaagttagaagatgtttgcactttaccgaaggctg
tgtcttgaagcaccgccgagcaactctgagagccttgattttgtgtcctcagcatatgtttgtgtaatacagaaagagaagcagttgcc
agtgaagggtgatgttgctccaaaattatagttgatccacaaacacacacatacatgcaaaggattgtttgcttcacggtttttg
atatattaattcaatgctgttggaacagcacaaaaactaagtgtcagtttaacagaatcactgtccttttagcattaaaaatacatggaactta
atgctttaatttcccaacatgectttttatttagaaagattcagactttttatcttagaataaaaatgccattttatttagaaagatacaggag
cattcattcacggaactttcagatctcagtcactgcataaaatcttgatcctgtaataatagtttctgtatcttgcatttcattcaacagggtt
aacgcgatgagcaaatatgttcacgtttttaacatgttctgtttaaatacagaacccacattctcaacgttaattgaacgtacataggacta
tacaagggttagtaataagacagaaactgttgcctatttaaccacgcctcactttgga

SEQ ID NO:4

> HS4 CHROMATIN INSULATOR

cagcctaaagctttttcccgtatccccaggtgtctgcaggctcaaagagcagcgagaaagcgttcagaggaaagcgatcccgtgcc
accttccccgtgcccgggctgtcccgacgctgccgggctcggggatgcggggggagcgccggaccggagcggagccccgggg
ggctcgtcgtgccccctagcgggggagggacgtaattacatccctgggggctttgggggggggctgtccccgtgagctc

SEQ ID NO:5

> KOZAK CONSENSUS SEQUENCE

ccacc

SEQ ID NO:6

> Human PSEN1 with all tolerant, non-preferred codons changed to highly preferred synonymous codon. Changed codons in lower case:

ATGACAGAGTTACCTGCacctTTGTCCTACTTCCAGAATGCACAGATGTCTGAGGA
CAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAACGGCAGGA
GCACAACGACAGACGGAGCctgGGCCACCCTGAGCCActgTCTAATGGAaagCCCCA
GGGTAACCTCCCGGCAGGTGGTGGAGcagGATGAGGAAGAAGATGAGGAGCTGAC
ActgAAATATGGCGCCAAGcacGTGATCATGCTCTTTGTCCCTGTGACTCTCTGCATG
GTGGTGGTTCGTGGCTACCATTAAGTCAGTCAGCTTTTATACCCGGAAGGATGGGC
AGCTAATCTATACCCCATTCACAGAAGATACCGAGACTGTGGGCCAGAGAGCCC
TGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCATTGTTGTCATGACTATC
CTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGctgATT
ATATCATCTctgTTGctgCTGTTCTTTTTTTCATTTCATTACctgGGGGAAGTGTTTAA
ACCTATAACGTTGCTGTGGACTACATTACTGTGCACTCCTGATCTGGAATTTTggc
GTGGTGGGAATGATTTCCATTCACTGGAAAggcCCActgagaCTCCAGCAGGCATATC
TCATTATGATTAGTGCCCTCATGGCCCTGGTGTATTATCAAGTACCTCCCTGAATGG
ACTgccTGGCTCATCTTGGCTGTGATTTCACTGTATGATTTAGTGGCTGTTctgTGTcc
tAAAGGTCCActgCGTATGCTGgtgGAAACAGCTCAGGAGAGAAATGAAaccctgTTTC
CAGCTCTCATTTACTCCTCAACAATGGTGTGGctgGTGAATATGGCAGAAGGAGAC
cctGAAGCTCAAAGGAGAgTgTCCAAAAATTCCAAGTATAATGCAGAAAGCACAGA
AAGGGAGTCAcagGACACTGTTGCAGAGAATGATGATGGCGGGTTCAGTGAGGAA
TGGGAAGCCCAGAGGGACAGTcacctgGGGCCTcacCGCTCTACACCTGAGTCAagaG
CTGCTGTCCAGGAActgTCCAGCAGTATCCTCGCTggcGAAGACCCAGAGGAAAGG
GGAGTAAAACTTGATTGGGAGATTTCAATTTCTACAGTGTTCTGGTTggcAAAGC

CTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCatcT
 TAATTggcctgTGCCTTACActgctgCTCctgGCCATTTTCAAGAAAGCActgCCAGCTctgC
 CAATCTCCATCACCTTTGGGCTTGTCTTCTACTTTGCCACAGATTATctggtgCAGCC
 TTTTATGGACcagctgGCATTcaccagTTTTATATCtaa

SEQ ID NO:7

> SYNTHETIC PSEN1 CDNA with all CpG dinucleotides removed, all tolerant codons changed to highly preferred synonymous codon, and some intolerant codon changed as a consequence of eliminating CpG dinucleotides. Changed codons in lower case; changed intolerant codons in lower case and underlined.

ATGACAGAGTTACCTGCACcaTTGTCTACTTCCAGAATGCACAGATGTCTGAGGA
 CAACCACCTGAGCAATACTGTAggaAGCCAGAATGACAATAGAGAAgaCAGGAGC
 ACaatGACAGAggaAGCCTTGGCCACCCTGAGCCATTATCTAATGGAgaCCCCAGGG
 TAACTCCagaCAGGTGGTGGAGCAAGATGAGGAAGAAGATGAGGAGCTGACATTG
 AAATATggtGCCAAGCATGTGATCATGCTCTTTGTCCCTGTGACTCTCTGCATGGTG
 GTGgtGTGGCTACCATTAAGTCAGTCAGCTTTTATACCagaAAGGATGGGCAGCTA
 ATCTATACCCCATTCACAGAAGATactGAGACTGTGGGCCAGAGAGCCCTGCACTC
 AATTCTGAATGCTGCCATCATGATCAGTGTcATTGTTGTCATGACTATCCTCCTGG
 TGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTTATTATATC
 ATCTCTATTGTTGCTGTTctTTTTTCATTCATTTACTTGGGGGAAGTGTTTAAACC
 TATaatGTTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGAATTTTGGTGTG
 GTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTagaCTCCAGCAGGCATATCT
 CATTATGATTAGTGCCTCATGGCCCTGGTGTtTATCAAGTACCTCCCTGAATGG
 ACTgccTGGCTCATCTTGGCTGTGATTTCAGTATATGATTTAGTGGCTGTTTTGTGT
 cccAAAGGTCCACTTagaATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAaccCTT
 TTTCCAGCTCTCATTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAAG
 GAGACccaGAAGCTCAAAGGAGAGTATCCAAAAATTCCAAGTATAATGCAGAAAG
 CACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAATGATGATggtGGGTTCAGT
 GAGGAATGGGAAGCCCAGAGGGACAGTCATCTAGGGCCTCATaggTCTACACCTG
 AGTCAagaGCTGCTGTCCAGGAACTTTCCAGCAGTATCctgGCTGGTGAAGACCCAG
 AGGAAAGGGGAGTAAAACCTTGGATTGGGAGATTTCAATTTCTACAGTGTCTGGT
 TGGTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTttTG
 TAGCCATATTAATTGGTTTGTGCCTTACATTATTACTCCTTGCCATTTTCAAGAAA
 GCATTGCCAGCTCTTCCAATCTCCATCACCTTTGGGCTTGTCTTCTACTTTGCCAC
 AGATTATCTTGTACAGCCTTTTATGGACCAATTAGCATTCCATCAATTTTATATCT
 AG

SEQ ID NO:8

> GENOMIC / CDNA HYBRID PSEN1 GENE

cccAGATCTgacaacagccttgcggtccttagacagcttggcctggaggagaacacatgaaagaaaggttgttctgcttaatg
 taatctatgaaagtgttttataacagtataattgtagtcacaaagttctgttttcttccctttcagaacctcaagaggcttgttctgtg
 aaacagtattctatacagttgccaccatgacagagttacctgcaccgtgtcctacttcagaatgcacagatgtctgaggacaaccac
 ctgagcaatactgtacgtagccaggtacagtgctcagtcctgaactgccttggcagactggattcactatcatctcccctcacctctga
 gaaatgctgaggggtaggcaggccttctctacttttagaactcatagtgacgggtctgtgttaatcccagggttaaccgttaccttga
 ttctgctgagaatctgatttactgaaatgttttctgtgcttatagaatgacaatagagaacggcaggagcacaacgacagacggagc
 ctggccaccctgagccattatctaatggacgacccagggttaactcccggcaggtggtggagcaagatgaggaagaagatgagga
 gctgacattgaaatatggcgccaagcatgtgatcatgctcttgtccctgtgactctctgcatggtggtggtcgtgctaccattaaagtcag
 tcagctttatacccggaaggtatggcgagctgtacgtatgagtttgtttatttctcaagccagtggtgctttctttacagcatgtcatc
 atcacctgaagccctctgattgaaggggcatgactgaattaagaaaaaagaaattctgttggaggtggttaattgtggttggtgatct
 cattaacactgacctaggccttgtgttgtttattgtagaatctatacccccattcacagaagataccgagactgtgggcccagagagcc
 ctgcactcaattctgaatgctgccatcatgatcagtgctgctgactatcctcctggtggttctgtataataacaggtgctataag

tcacatcgccctggcttattatcatctctattgttgctgttctttttcattcatttacttgggggaagtgtttaaaccataacgttgctgtg
gactacattactgttgcaactcctgatctggaattttggtgtggtgggaatgattccattcactggaaaggtccacttcgactccagcaggc
atatctcattatgattagtgccctcatggccctggtgtttatcaagtacctccctgaatggactgctggtcctatcttggtgtgatttcagta
tatgatttagtggtctgtttgtgtccgaaaggtccacttcgatgctggttgaacagctcaggagagaaatgaaacgcttttccagctctc
atttactcctcaacaatggtgtggttggtgaatatggcagaaggagacccggaagctcaaaggagagtatccaaaattccaagtataa
tgcagaaagcacagaaagggagtcacaagacactgttgacagagaatgatgatggcgggttcagtgggaatgggaagcccagagg
gacagtcatctagggcctcatcgtctacacctgagtcacgagctgctgtcagggaactttccagcagatcctcgtgtggaagacc
agaggaaaggggagtaaaacttgattgggagatttcattttctacagtgttctggttggttaaagcctcagcaacagccagtgaggact
ggaacacaaccatagcctgtttcgtagccatattaattggtttgtccttacatttactccttgccattttcaagaaagcattgccagctctt
ccaatctccatcaccttgggctgttttctactttgccacagattatctgtacagcctttatggaccaatagcattccatcaattttatatct
agcataGTCGACc

SEQ ID NO:9

>MALAT1 MRNA STABILITY ELEMENT

tagggtcatgaaggtttttcttctgagaaaacaacagctattgtttctcaggttttgccttttctagcttaaaaaaaaaaaaa
gcaaaagatgctggttggttgcaactcctggttccaggacggggtcaaatccctgcggcgtctttgctttgact

SEQ ID NO:10

> 3' FLANKING SEQUENCE FROM NATIVE PSEN1 RNA

GGAAACAAAACAGCGGCTGGTCTGGAAGGAACCTGAGCTACGAGCCGCGGCGG
CAGCGGGGCGGCGGGGAAGCGTATACCTAATCTGGGAGCCTGCAAGTGACAACA
GCCTTTGCGGTTCCTTAGACAGCTTGGCCTGGAGGAGAACACATGAAAAGAAAGAA
CCTCAAGAGGCTTTGTTTTCTGTGAAACAGTATTTCTATACAGTTGCTCCA

SEQ ID NO: 11

> 5' FLANKING SEQUENCE FROM NATIVE PSEN1 MRNA

CATATTTGCGGTTAGAATCCCATGGATGTTTCTTCTTTGACTATAACAAAATCTGG
GGAGGACAAAGGTGATTTTCTGTGTCCACATCTAACAAAGTCAAGATTCCCAGC
TGGAGTTTTCAGCTTCTTCCAGTCTTCTGACCACCTTGCACTATTGGACTTT
GGAAGGAGGTGCCTATAGAAAACGATTTTGAACATACTTCATCGCAGTGGAAGT
TGTCCTCGGTGCAGAACTACCAGATTTGAGGGACGAGGTCAAGGAGATATGA
TAGGCCCCGGAAGTTGCTGTGCCCCATCAGCAGCTTGACGCGTGGTTCACAGGACG
ATTTCACTGACACTGCGAACTCTCAGGACTACCGTTACCAAGAGGTTAGGTGAAG
TGGTTTAAACCAAACGGAACCTTTCATCTTAACTACACGTTGAAAATCAACCCA
ATAATTCTGTATTAAGTGAATTCTGAAGTTTTCAGGAGGTACTGTGAGGAAGAGC
AGGCACCAGCAGCAGAATGGGGAATGGAGAGGTGGGCAGGGGTTCCAGCTTCCC
TTTGATTTTTTGTGCTGCAGACTCATCCTTTTTAAATGAGACTTGTTTTCCCCTCTCTT
TGAGTCAAGTCAAATATGTAGATTGCCTTTGGCAATTCTTCTTCTCAAGCACTGA
CACTCATTACCGTCTGTGATTGCCATTTCTTCCCAAGGCCAGTCTGAACCTGAGG
TTGCTTTATCCTAAAAGTTTAAACCTCAGGTTCCAAATTCAGTAAATTTTGAAAC
AGTACAGCTATTTCTCATCAATTCTCTATCATGTTGAAGTCAAATTTGGATTTTCC
ACCAAATTCTGAATTTGTAGACATACTTGACGCTCACTTGCCCCAGATGCCTCC
TCTGTCTCATTCTTCTCTCCCACACAAGCAGTCTTTTTCTACAGCCAGTAAGGCA
GCTCTGTCTGGTAGCAGATGGTCCCATTATTCTAGGGTCTTACTCTTTGTATGAT
GAAAAGAATGTGTTATGAATCGGTGCTGTGAGCCCTGCTGTCAGACCTTCTTCCA
CAGCAATGAGATGTATGCCCAAAGACGGTAGAATTAAAGAAGAGTAAATGG
CTGTTGAAGCACTTTCTGTCTGGTATTTTGTGTTTTGCTTTTGCCACACAGTAGCT
CAGAATTTGAACAAATAGCCAAAAGCTGGTGGTTGATGAATTATGAAGTAGTTGT
ATCAACACAAAGCAAGAGTTGGGGAAAGCCATATTTAACTTGGTGAGCTGTGGG
AGAACCTGGTGGCAGAAGGAGAACCAACTGCCAAGGGGAAAGAGAAGGGGCCT
CCAGCAGCGAAGGGGATACAGTGAGCTAATGATGTCAAGGAGGAGTTTCAGGTT

ATTCTCGTCAGCTCCACAAATGGGTGCTTTGTGGTCTCTGCCCCGCTTACCTTTCC
TCTCAATGTACCTTTGTGTGAACTGGGCAGTGGAGGTGCCTGCTGCAGTTACCAT
GGAGTTCAGGCTCTGGGCAGCTCAGTCAGGCCAAAACACACAAACAGCCATCAGC
CTGTGTGGGCTCAGGGCACCTCTGGACAAAGGCTTGTGGGGCATAACCTTCTTTA
CCACAGAGAGCCCTTAGCTATGCTGATCAGACCGTAAGCGTTTATGAGAACTTA
GTTTCCTCCTGTGGCTGAGGAGGGGCCAGCTTTTTCTTCTTTTGCCTGCTGTTTCT
CTCCCAATCTATGATATGATATGACCTGGTTTGGGGCTGTCTTTGGTGTTAGAAT
ATTTGTTTTCTGTCCCAGGATATTTCTTATAAGAACCTAACTTCAAGAGTAGTG
CGAGTACTGATCTGAATTTAAATTAATAATTGGCTTATATTAGGCAGTCACAGACA
GGAAAAATAAGAGCTATGCAAAGAAAGGGGGATTAAAGTAGTAGGTTCTATCA
TCTCAATTCATTTTTTTTCCATGAAATCCCTTCTTCCAAGATTCATTCCCTCTCTCAG
ACATGTGCTAGCATGGGTATTATCATTGAGAAAGCACAGCTACAGCAAAGCCAC
CTGAATAGCAATTTGTGATTGGAAGCATTCTTGAGGGATCCCTAATCTAGAGTAA
TTTATTTGTGTAAGGATCCCAAATGTGTTGCACCTTTCATGATACATTTCTTCTCT
GAAGAGGGGTACGTGGGGTGTGTGTATTTAAATCCATCCTATGTATTACTGATTGT
CCTGTGTAGAAAGATGGCAATTATTCTGTCTCTTCTCCAAGTTTGAGCCACATCT
CAGCCACATTGTTAGACAGTGTACAGAGAACCTATCTTTCCTTTTTTTTTTTTAA
AGGACAGGATTTTGCTGTGTTGCCAGGCTAGACTTGAACCTCCTGGGCTCAAGTA
ATCCACCTCAGCCTGAGTAGCTGAGACTACAGCCCATCTTATTTCTTTAAATCATT
CATCTCAGGCAGAGAACTTTTCCCTCAAACATTCTTTTTAGAAATTAGTTCAGTCAT
TCCTAAAACATCCAAATGCTAGTCTTCCACCATGAAAAATAGATTGTCACTGGAA
AGAACAGTAGCAATTTCCATAAGGATGTGCCTTCACTCACACGGGACAGGCGGT
GGTTATAGAGTCGGGCAAAACCAGCAGTAGAGTATGACCAGCCAAGCCAATCTG
CTTAATAAAAAGATGGAAGACAGTAAGGAAGGAAAGTAGCCACTAAGAGTCTG
AGTCTGACTGGGCTACAGAATAAAGGGTATTTATGGACAGAATGTCATTACATGC
CTATGGGAATACCAATCATATTTGGAAGATTTGCAGATTTTTTTTTCAGAGAGGAA
AGACTCACCTTCCTGTTTTTGGTTCTCAGTAGGTTTCGTGTGTGTTCTAGAAATCAC
AGCTCTGACTCCAAATGACTCAATTTCTCAATTAGAAAAAGTAGAAGCTTTCTAA
GCAACTTGGAAGAAAACAGTCATAAGTAAGCAATTTGTTGATTTTACTACAGAA
GCAACACTGAAGAGGCAGTGTTTTTACTTTTCAGACTCCGGGATTCCCATCTGT
AGTCTCTCTGCTTTTAAAAACCCTCCTTTTGCAATAGATGCCCAAACAGATGATG
TTATTACTTGTTATTTACGTGGCCTCAGACAGTGTATGTATTCTCGATATAACTT
GTAGAGTGTGAAATATAAGTTTAACTACCAAATAAGGTCTCCCAGGGTTAGATG
ACTGCGGGAAGCCTTTGATCCCAACCCCCAAGGCTTTGTATATTTGATCATTGT
GATCTAACCCTGGAAGAAAAAGAGCTCAGAAACCACTATGAAAAATTTGTTCA
GTGTTTTCTGTGTTCCCGTAGGTTCTGGAGTCTGAGGATGCAAAGATGAATAAGA
TAAATTCTCAGAATGTAGTTATAATCTCTTGTCTTCTGGTATATGCCATCTTTCTTT
AACTTCTCTAAAATATTGGGTATTTGTCAAATAACCACTTTTAACAGTTACCATTA
CTGAGGGCTTATACATTGGTGTTATAAAAGTGACTTGATTTCAGAAATCAATCCAT
TCAGTAAAGTACTCCTTCTCTAAATTTGCTGTTATGTCTATAAGGAACAGTTTGAC
CTGCCCTTCTCCTCACCTCCTCACCTGCCTTCCAACATTGAATTTGGAAGGAGAC
GTGAAAATTGGACATTTGGTTTTGCCCTTGGGCTGGAAACTATCATATAATCATA
AGTTTGAGCCTAGAAGTGATCCTTGTGATCTTCTCACCTCTTTAAATTCCCACAAC
ACAAGAGATTAAAAACAGAGGTTTCAGCTCTTCATAGTGCGTTGTGAAATGGCTG
GCCAGAGTGTACCAACAAAGCTGTCATCGGGCTCACAGCTCAGAGACATCTGCA
TGTGATCATCTGCATAGTCCTCTCCTCTAACGGGAAACACCTCAGATTTGCATAT
AAAAAAGCACCTTGGTGCTGAAATGAACCCCTTTCTTGAACATCAAAGCTGTCTC
CCACAGCCTTGGGCAGCAGGGTGCCTCTTAGTGAGTGTGCTGGGTCCACCCTGAG
CCCTGACATGTGGTGGCAGCATTGCCAGTTGGTCTGTGTGTCTGTGTAGCAGGGA
CGATTTCCAGAAAGCAATTTTCTTTTGAAATACGTAATTGTTGAGACTAGGCA
GTTTCAAAGTCAGCTGCATATAGTAGCAAGTACAGGACTGTCTTGTTTTTGGTGT

CCTTGGAGGTGCTGGGGTGAGGGTTTCAGTGGGATCATTACTCTCACATGTTGT
CTGCCTTCTGCTTCTGTGGACACTGCTTTGTACTTAATTCAGACAGACTGTGAATA
CACCTTTTTTATAAATACCTTTCAAATTCCTGGTAAGATATAATTTTGATAGCTGA
TTGCAGATTTTCTGTATTTGTCAGATTAATAAAGACTGCATGAATCCA

SEQ ID NO: 12

>XP_011535274.1 human presenilin-1 isoform X1 [467 amino acids]

MTELPAPLSYFQNAQMSEDNHLSTNVRSDNDRRQEHNDRRSLGHPEPLSNGRPQ
GNSRQVVEQDEEDEELTKYGAKHVIMLFVPVTLCMVVVATIKSVSFYTRKDGQ
LIYTPFTEDTETVGQRALHSILNAAIMISVIVMTILLVVLKYRCYKVIHAWLISSLL
LLFFFSFIYLGEVFKTYNVAVDYITVALLIWNFGVVGMIHKGPLRLQQAAYLIMIS
ALMALVFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALI
YSSTMVWLVNMAEGDPEAQRVSKNSKYNAESTERESQDTVAENDDGGFSEWEA
QRDShLGPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATA
SGDWNTTIACFVAILIGLCLTLLLAIFKKALPALPISITFGLVFYFATDYLVPFMDQL
AFHQFYI

SEQ ID NO: 13

>Gene (cDNA) for human PSEN1 463 amino acid isoform X2 (GenBank NM_007318.3)

atgacagagttacctgcaccgtgtcctacttccagaatgcacagatgtctgaggacaaccacctgagcaataactaatgacaatagaga
acggcaggagcacaacgacagacggagccttggccacctgagccattatctaatggacgacccagggaactcccgagcaggtg
gtggagcaagatgaggaagaagatgaggagctgacattgaaatattggcgccaagcatgtgatcatgctcttctgctcctgtgactctctg
catggtggtggtcgtggtcctaccattaagtcagtcagctttataccgggaaggatgggcagctaatctatacccccattcacagaagatac
cgagactgtggccagagagccctgactcaattctgaatgctgccatcatgatcagtgctattgtgcatgactatcctcctggtggtt
ctgtataatacaggtgctataaggtcatccatgctggtctattatcatctctattgtgtgctgtctttttcattcatttactgggggaagt
gtttaaaccataacgtgctgtggactacattactgtgactcctgatctggaattttggtggtgggaatgattccattcactggaaa
ggctcacttcgactccagcaggcagatctcattatgattagtgccctcatggccctggtgtttatcaagtacccctgaatggactgctg
gctcatcttggtgtgattcagtatatgatttagtggtgtgttctggtcgaagggtccacttcgtatgctggttgaaacagctcaggagag
aatgaacgcttttccagctctcatttactcctcaacaatggtgtggttggtgaatatggcagaaggagacccggaagctcaaggag
agtatccaaaaattccaagtataatgcagaaagcacagaaaggagtcacaagacactgttcagagaatgatgatggcgggttcagt
gaggaaagggaagccagaggacagtcactagggcctcatcgtctacacctgagtcacgagctgctgtccaggaaacttccagc
agtatcctcgctggtgaagaccagaggaaagggagtaaaactggattgggagatttcatttctacagtgcttctggttggttaaagcct
cagcaacagccagtgagactggacacaacatagcctgttctgtagccatataattggttgccttacattattactccttgccattt
tcaagaaagcattgccagctctccaatctcatcaccttgggctgtttctacttggccagattatctgtacagcctttatggacca
ttagcattccatcaattttatctag

SEQ ID NO: 14

>presenilin-1 isoform X2 [463 amino acids]

MTELPAPLSYFQNAQMSEDNHLSTNDRNDRRQEHNDRRSLGHPEPLSNGRPQGN
SRQVVEQDEEDEELTKYGAKHVIMLFVPVTLCMVVVATIKSVSFYTRKDGQLIYTPF
TEDTETVGQRALHSILNAAIMISVIVMTILLVVLKYRCYKVIHAWLISSLLLLFFFS
FIYLGEVFKTYNVAVDYITVALLIWNFGVVGMIHKGPLRLQQAAYLIMISALMAL
VFIKYLPEWTAWLILAVISVYDLVAVLCPKGPLRMLVETAQERNETLFPALIYSSTMV
WLNVMAEGDPEAQRVSKNSKYNAESTERESQDTVAENDDGGFSEWEAQRDShL
GPHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFYSVLVGKASATASGDWNT
TIACFVAILIGLCLTLLLAIFKKALPALPISITFGLVFYFATDYLVPFMDQLAFHQFYI

SEQ ID NO: 15

>presenilin-1 isoform X1 coding sequence with intolerant codons underlined

ATG ACA GAG TTA CCT GCA CCG TTG TCC TAC TTC CAG AAT GCA CAG ATG
TCT GAG GAC AAC CAC CTG AGC AAT ACT GTA CGT AGC CAG AAT GAC AAT

AGA GAA CGG CAG GAG CAC AAC GAC AGA CGG AGC CTT GGC CAC CCT GAG
CCA TTA TCT AAT GGA CGA CCC CAG GGT AAC TCC CGG CAG GTG GTG GAG
CAA GAT GAG GAA GAA GAT GAG GAG CTG ACA TTG AAA TAT GGC GCC AAG
CAT GTG ATC ATG CTC TTT GTC CCT GTG ACT CTC TGC ATG GTG GTG GTC GTG
GCT ACC ATT AAG TCA GTC AGC TTT TAT ACC CGG AAG GAT GGG CAG CTA
ATC TAT ACC CCA TTC ACA GAA GAT ACC GAG ACT GTG GGC CAG AGA GCC
CTG CAC TCA ATT CTG AAT GCT GCC ATC ATG ATC AGT GTC ATT GTT GTC ATG
ACT ATC CTC CTG GTG GTT CTG TAT AAA TAC AGG TGC TAT AAG GTC ATC CAT
GCC TGG CTT ATT ATA TCA TCT CTA TTG TTG CTG TTC TTT TTT TCA TAT
ATT TAC TTG GGG GAA GTG TTT AAA ACC TAT AAC GTT GCT GTG GAC TAC ATT
ACT GTT GCA CTC CTG ATC TGG AAT TTT GGT GTG GTG GGA ATG ATT TCC ATT
CAC TGG AAA GGT CCA CTT CGA CTC CAG CAG GCA TAT CTC ATT ATG ATT
AGT GCC CTC ATG GCC CTG GTG TTT ATC AAG TAC CTC CCT GAA TGG ACT
GCG TGG CTC ATC TTG GCT GTG ATT TCA GTA TAT GAT TTA GTG GCT GTT TTG
TGT CCG AAA GGT CCA CTT CGT ATG CTG GTT GAA ACA GCT CAG GAG AGA
AAT GAA ACG CTT TTT CCA GCT CTC ATT TAC TCC TCA ACA ATG GTG TGG TTG
GTG AAT ATG GCA GAA GGA GAC CCG GAA GCT CAA AGG AGA GTA TCC AAA
AAT TCC AAG TAT AAT GCA GAA AGC ACA GAA AGG GAG TCA CAA GAC ACT
GTT GCA GAG AAT GAT GAT GGC GGG TTC AGT GAG GAA TGG GAA GCC CAG
AGG GAC AGT CAT CTA GGG CCT CAT CGC TCT ACA CCT GAG TCA CGA GCT
GCT GTC CAG GAA CTT TCC AGC AGT ATC CTC GCT GGT GAA GAC CCA
GAG GAA AGG GGA GTA AAA CTT GGA TTG GGA GAT TTC ATT TTC TAC
AGT GTT CTG GTT GGT AAA GCC TCA GCA ACA GCC AGT GGA GAC TGG
AAC ACA ACC ATA GCC TGT TTC GTA GCC ATA TTA ATT GGT TTG TGC CTT ACA
TTA TTA CTC CTT GCC ATT TTC AAG AAA GCA TTG CCA GCT CTT CCA ATC TCC
ATC ACC TTT GGG CTT GTT TTC TAC TTT GCC ACA GAT TAT CTT GTA CAG CCT
TTT ATG GAC CAA TTA GCA TTC CAT CAA TTT TAT ATC TAG

Inverted terminal repeats (ITR) for single-stranded AAV2

SEQ ID NO: 16

nc_001401.2 nt 1-145 and 4535-4679

>left ITR

ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg
 ggccgacctca gtgagcgagc gagcgcgagc agaggagtg gccactcca tcactagggg ttctt

SEQ ID NO: 17

>right ITR

aggaac ccctagtgt gagttggcc actccctctc tgcgcgctcg ctgcctcact gaggccgggc gaccaaaggt
 cgcccgacgc ccgggctttg cccgggcggc ctactgagc gagcgagcgc gcagagaggg agtgcccaa

Inverted terminal repeats for self-complementary AAV2

SEQ ID NO: 18

>left ITR

ctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgctggcgacctttggtcgcccgccctcagtgagcgagc
 gagcgcgagagagggagtg

SEQ ID NO: 19

>right ITR

aggaacccctagtgtgagttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgc
 ccgggctttgcccgggcggccctcagtgagcgagcgagcgcgagc

Introns**SEQ ID NO: 20**

>MVM NC_001510.1 minute virus of mice intron nt 2312-2403

aagaggtaa gggtttaagg gatggttggt tgggtgggta ttaatgtta attacctgtt ttacaggcct gaaatcactt ggttttaggt
tgg

SEQ ID NO: 21

>hybrid adenovirus SD/IgG SA; ac_000008.1 adenovirus type 5 tripartite leader, nt 573-758
and nc_000078.6 mus musculus strain c57bl/6j nt 115158736-115158695

tctgccac ggaggtgtta ttaccgaaga aatggccgcc agtcttttgg accagctgat cgaagaggta ctggctgata
atctccacc tcttagccat ttgaaccac ctacccttca cgaactgt t gatttagacg tgacggcccc cgaagatccc
aacgaggagg cggtttcga gatttttc tgacatc cactttgcct ttctctccac aggtgtccac tcccaaa

SEQ ID NO: 22

>hBG intron chimeric cmv promoter 601-734, cmv intron1 1263-1294, followed by the intron
ii of the beta-globin gene. the 3' end of the intron is fused to the first 20 nucleotides of exon 3
of the beta globin gene.

tcagatcgctggagacgcatccacgctgtttgacctccatagaagacacccgggaccgatccagcctccgcggttcgaatcccg
gccgggaacgggtgcatggaaacgcggttccccgtgccaagagtgacgtaagtaccgcctatagagctataggcccaaaaaatg
cttctcttttaataatactttttgttatcttatttctaatacttccctaactcttcttcttcagggaataatgatacaatgtatcatgcctctttgc
accattctaaagaataacagtgataatttctgggttaaggcaatagcaatatttctgcataataatatttctgcataataattgtaactgatga
agagggttcatattgctaataagcagctacaatccagctaccattctgctttattttatggttgggataaggctggattattctgagccaagc
taggcccttttctaatactgttcatacctcttattctctccacagctcctgggaacgtgctggtctgtgctgccccatcatttggc
aaagaatt

Promoters**SEQ ID NO: 23**

>CAG It727518.1, nt 3074 to 4750

gacattg attattgact agttattaat agtaatcaat tacgggggtca ttagtccata gcccatatat ggagtccgc gttacataac
ttacggtaaa tggcccgccct ggctgaccgc ccaacgacct cgcgccattg acgtcaataa tgacgtatgt tcccatagta
acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgccac ttggcagtac atcaagtgtg
tcatatgcca agtacgcccc ctattgacgt caatgacgggt aaatggcccg cctggcatta tggccagtac atgacctat
gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc atggtcgagg tgagccccac gttctgcttc
actctcccca tctccccccc cteccacccc ccaattttgt attatttat ttttaatta tttgtgcag cgaagggggc gggggggggg
ggggggcccc cccagggcgg ggccggggcgg ggcgaggggc ggggcggggc gaggcggaaa ggtgcggcgg
cagccaatca gagcggcgcg ctccgaaagt ttcttttat ggcgaggcgg cggcgggcggc ggccctataa aaagcgaagc
gcgcggcggg cgggagtcgt tgcgcgtgc ctccccccg tgcgccgctc cgcgcggcc tcgcgcggcc cgcgccggct
ctgactgacc gcgttactcc cacaggtgag cggcggggac ggcccttctc ctccgggctg taattagcgc ttggttaat
gacggcttgt ttctttctg tggctgcgtg aaagcctga ggggctccgg gagggccctt tgtgcggggg gagcggctcg
gggggtgctg gcgtgtgtgt gtgcgtgggg agcgcggcgt gcggctccgc gctgcccggc ggctgtgagc gctgcggcg
cggcgcgggg ctttgtcgc tccgagtggt gcgcgagggg agcgcggcgg ggggcgggtgc cccgcggtgc
gggggggggct gcgaggggaa caaaggctgc gtgcgggggtg tgtgcgtggg ggggtgagca gggggtgtgg gcgcgtcgt
cgggctgcaa cccccctgc accccctcc ccgagttgct gagcacggcc cggcttcggg tgcggggctc cgtacggggc
gtggcgcggg gctcgcgtg cggggcgggg ggtggcgga ggtgggggtg ccggcgggg cggggccgcc
tcggcgccgg gagggctcgg gggaaggggc gcgcggggcc ccggagcgcc ggcggctgtc gaggcgcggc
gagccgacg cattgccctt tatgtaatc gtgcgagagg gcgcagggac ttctttgtc ccaaatctgt gcggagccga
aatctgggag gcgcggccgc accccctcta gcggcgcggg gggaagcgg tgcggcgcc gcaggaagga
aatggcggg gagggccttc gtgcgtgcc gcgcggcgt ccccttctc ctctccagcc tcggggctgt ccggggggg
acggctgct tcggggggga cggggcaggg cggggttcgg ctctggcgt gtgaccggcg gctctagagc ctctgtaac
catgtcatg cttctctt ttctacag

SEQ ID NO: 24

>CBA nc_006273.2, nt 175625–175294 and GenBank accession no. x00182.1, nt 280–540.

cgcgctgacattgattattgactagtattataatagtaataatcaggggcattagtcacagccatataatggagttccggttacataact
tacggtaaatggcccgctggtgacgccccaacgacccccgccattgacgtcaataatgacgtatgtcccatagtaacgccaata
gggactttccattgacgtcaatgggtggagtatttacggtaaaactgccacttggcagtagacatcaagtgtatcatatgccaaagtagccc
cctattgacgtcaatgacggtaaatggcccgctggcattatgccagtagacgttatgggactttcctacttggcagtagacatctacg
tattagtcacgtcattaccatgtcagggccacgttctgcttactctccccatctccccccctccccacccccattttgtattttattttt
taattattttgtgcagcgatggggcgggggggggggcgcgccagggcgggcgggcgggcgagggcgggcgggcgggcg
cgagggcgagaggtgcggcgagccaatcagagcgcgcgctccgaaagtcttctttatggcgagggcgggcgggcgggcgggcg
cctataaaaagcgaagcgcgcgggcg

SEQ ID NO: 25

>UBC d63791 nt 3553–4729

gggtgcagcgccctccgcgcgggttttggcgccctcccgggcgccccctccacggcgagcgctgccacgtcagacgaagg
gcgcgaggagcgctcctgaccttccgcccggacgctcaggacagcgcccgctgctcataagactcggccttagaacccagtagtca
gcagaaggacatttaggacgggacttgggtgactctaggcgactgttttcttccagagagcggaacagggcgaggaagtagtcc
cttctcgcgattctcgggagggatctccgtggggcggtgaacgccgatgattatataaggacgcccgggtgtggcacagctagttc
cgctgcagccgggatttgggtcgcggttctgttggatcgctgtgacgtcacttggtagtgcgggctgctgggctggccgggg
cttctgtggccgcccggcgctcggtgggacgggaagcggtgtggagagaccgccaagggtgtagctgggtccgcgagcaagggt
gccctgaactgggggttgggggagcgacaaaatggcggtgttcccgagcttgaatggaagacgcttgaaggcgggctgtga
ggctgtgaaacaagggtggggggcatggtggcggaagaaccaaggctttagggccttcgtaatgcgggaaagctcttattcgg
gtgagatgggctggggcaccatctggggaccctgacgtgaagtgtgactgactggagaactcgggttctgctgtctggttcggggg
cggcagttatgcgggtcgcttggggcagtgacccgtaccttggggagcgcgccctcgtcgtgacgtcaccgttctgttggct
tataatgcagggtggggccacctcgggtaggtgtcggttaggcttttctccgtcgagagacgaggggtcgggcctagggttaggct
ctcctgaatcgacagggcgccggacctctggtaggggaggggataagtagggcgtagtttcttggctgggtttatgtacctatcttta
agtagctgaagctccggtttgaactatgcgtcgggttggcgagtggtttgtgaagtttttaggcacctttgaaatgtaatcattgg
gtcaatatgtaattttcagtgtagactagtaaa

SEQ ID NO: 26

>PGK nc_000086.7 106186725–106187235

ttctaccgggtaggggagggcgcttttcccaaggcagctctggagcatgcgcttttagcagccccgctgggcacttggcgctacacaagt
gcctctggcctcgacacattccacatccaccggtaggcgccaaccggctccgttcttgggtggcccttcgcgccaccttactactctc
ccctagtcaggaagtcccccccgcccgagctcgcgtcgtcagggacgtgacaatatggaagtagcacgtctcactagctcgtcgtc
agatggacagcaccgctgagcaatgggaagcggttaggcctttggggcagcgcccaatagcagcttctccttctgcttctgggtc
gaggtcgggaaggggtgggtccggggcgggctcagggggcgggctcagggggcgggcgggcgcccgaaggtcctccggagg
ccggcattctgcacgctcaaaaagcgacgtctgccgcgtgttctcctctcctcactctccggcccttctgacct

SEQ ID NO: 27

>Ef1a j04617.1 nt 379–1560

gctccggtgcccgtagtgggcagagcgacatcgccacagtccccgagaagtggggggagggggtcggaattgaaccgggtgc
ctagagaagggtggcggggtaactgggaaagtgtgtgtgtactggctccgcttttccgagggtgggggagaaccgtatata
agtgcagtagtcggtgaacgttcttttcgcaacgggtttgcccgccagaacacaggtgaagtccgtgtgtgttccgcccgtgg
cctctttacgggttatggcccttgcgtgccttgaattacttccacgcccttggctgcagtagctgattcttgatcccagcttcgggttga
agtgggtgggagaggttcaggccttgcgttaaggagccccctcgcctcgtgcttgagttagggccttggcctggcgctggggccgc
cgctgtgcaatctgggtggcaccttcgcgcctgtctcgtcttgcataagtctctagccatttaaaattttgatgacctgctgcgacgctt
ttttctggcaagatagcttctgtaaatcgggccaagatctgcacactgttatttcgggttttggggccgcccggcgacggggccggt
gctccccagcgcatgttcggcgagggcgggcgctgcgagcgggccaccgagaatcgagcgggggtagcttcaagctggccgg
cctgctctgggtccttggcctcgcgcggcgtgtatcgcggcccttggcgggcaaggctggccggtcggcaccagttgctgtgagc
ggaaaagtggccgcttcccgccctgctgcaggagagctcaaaatggagagcgcgcgctcgggagagcgggcggggtgagtcacc
cacacaaaggaaaaggcccttccgtcctcagccgtcgttcatgtactccagggagtagccggcgccgtccaggcacctcgatta
gttctcgagcttttgagtagctgtctttaggttggggggagggggtttatgcagtaggagtttccacactgagtggttgagactga

agttaggccagcttggcacttgatgtaattctccttggaaattgccctttttgagtttggatcttggttcattctcaagcctcagacagtggttc
aaagttttttcttccatttcaggtgctgga

SEQ ID NO: 28

>CMV

gtcgacattgattattgactagttattataatgaatcaattacggggcattagttcatagcccatatatggagttccgcgttacataacttac
ggtaaatggcccgcttggtgacgcccacgacccccgccattgacgtcaataatgacgtatgttcccatagtaacgccaataggg
actttccattgacgtcaatgggtggagttttacggtaaacgtgcccacttggcagtagatcaaggtatcatatgccaagtagccccctat
tgacgtcaatgacggtaaatggcccgcttggtgacgtacacgtacgttatgggactttctacttggcagtagacatctacgtatta
gtcatcgctattaccaggtgatgagggttttggcagtagacatcaatgggctggtgtagcggttgactcacggggatttccaaagtctccacc
ccattgacgtcaatgggtggttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgcccattgacgcaaatgg
gcggtagcgctgtacgtgggaggtctatataagcagagctctctggtaactagagaaccactgcttactggcttatcgaaattaata
cgactcactataggagacccaagctggctagcggttaactt

SEQ ID NO: 29>NSE similar to *Rattus norvegicus* neuron specific enolase gene ab038993.1 nt 1023-2715

agctctgagctcctctctgctgcgccaatccttccaacccccctatggtggtatggctgacacagaaaaatgtctgctctgtatgggacat
ttgccctcttctccaaatataagacaggatgaggcctagcttttgcgtctccaaagttttaaaagaacacattgcacggcatttagggact
ctaaagggtggaggaggaatgagggaattgcatcatccaaggctggtcctcatccatcactgcttccaggggccagagtggttcca
ggaagtattcttcaaaaggaagccgcatctgtagctaacactcagagccattttctgcgttaacccctccgacctatatacaggag
taacatgatcagtgacctgggggagctggccaaactgcgggacctgcccaagctgaggggccttgggtgctgctggacaacccctgtgc
cgatgagactgactaccgccagggagggcctggtgcagatggcacacctaagagcgctagacaaaagtagtactatgaggacgaggac
cgggcagaaagctgaggagatccgacagaggctgaaggaggaaacaggagcaagaactcgaacccggaccaagacatggaaccgta
cctcccgccaaacttagtggctctctagcctgcagggacagtaaaagtgatggcaggaaaggcagccccggaggtcaaaaggctgg
gcacgcgaggaggagaggccagagtcagaggctgcgggtatctcagatatgaaggaaagatgagagaggctcaggaaggagtaag
aaaagacacaagagaccagagaaggagagaagaattagagaggaggcagaggaccgctgtctctacagacatagctgtgtagaga
ctgggaggaaggagatgaacctgagcgcatgaagggaaggaggtggctgggttatggaggtatgtagctggccaggggaaaa
gatcctgcactaaaaatctgaagctaaaaataacaggacacgggtggagaggcgaaaggaggcagagtgaggcagagagact
gagaggcctggggatgtgggcattccggtagggcacacagttcacttgtcttcttttccaggaggccaaagatgctgacgtcaaga
atcataataacccagtgggggaccaccgcatcatagccctgttacaagaagtgggagatgttcttttgtccagactggaatccgtt
acatccgagggtcaggttctgtggtggtcatctgtgtggttctgttggcctacctaagtcctaagcacagctctcaagcagat
ccgaggcgactaagatgtagtaggggtgtctggagagaagagccgaggaggtgggctgtgtagatcagttcagtttcaataa
aaaggcggttttatactgtgtcaggttcgtgaacccctgtgtgggcttctcatctgtctgggttagtacctgccactatactggaataa
ggggacgcctgttccctcaggttggtggacaagggtatgagcatccgtgtacttatgggggttgcagcttggtcctggatcgcccg
gcccctccccaccggttcggttcccaccaccaccccgctcgtacgtgcgttccgctgcagctcttgactcatcggggcccccg
ggtcatatgcgctcgctcgctctataggcgccgccccctgccaccccccgccgctgggagccgcagcccgccgacctct
gctctctctgcgccg

SEQ ID NO: 30

>MeCP2 nc_000086.7 nt -677to+56 mouse mecp2 methyl cpg binding protein 2

tgccattataaacgtctgcaaaagcaaggtttgatatgttgattttactgtcagccttaagagtgcgacatctgctaatttagtgaataat
acaatcagtagaccctttaaacaagtcccttggcttgaacaacgcaggctcctcaacaggcaactttgctacttctacagaaaatga
taataaagaatgctggtgaagtcaaatgcttatcacaatggtgaactactcagcaggagggtctaataggcgccaagagcctagact
tccttaagcgccagagtcacaaagggccagttaatcctcaacattcaaatgctgcccaaaaaccagcccctctgtgccctagccgc
ctctttttccaaagtacagtagaactccaccaatccgcagctgaatgggtccgctcttttccctgctaaacagacaggaactcctgc
caattgagggcgctaccgctaaggctccgccccagcctgggctccacaaccaatgaagggtaatctcgacaaagagcaaggggtg
gggcgcgggcgcgaggtgcagcagcacacaggctggtcgggagggcgggcgcgacgtctgccgtgcggggtcccgcatc
ggttgcgcgcgcgtccctcctcctcgagagagggctgtgtaaaacccgtccggaaaatggccgcccgtgccg
ccaccgccgcccggccgcccgcggcggagcggagaggag

SEQ ID NO: 31

>GFAP nc_000017.11 nt -1991-0 *Homo sapiens* glial fibrillary acidic protein (GFAP)

ggcaacatggcaagaccctatctctacaaaaaaagttaaaaaatcagccacgtgtggtgacacacacctgtagtcacagctattcagg
aggctgagggtgaggggatcacttaaggctgggaggttgaggctgcagtgagtcgtgtggtgcccactgcactccagcctgggcaac
agtgagaccctgtctcaaaagacaaaaaaaaaaaaaaaaaaaaaagaacatatcctggtgtggagtaggggacgtgctctgacag
aggctcgggggacctgagctggctctgtgagctgggagggagggcagacagccaggccttctctgcaagcagacctggcagcattgg
gctggccgccccagggcctcctctcatgcccagtgaaatgactcaccttggcacagacacaaatgttcggggtgggcacagtgctt
gctcccgcgcacccagccccctcaaatgccttccgagaagcccattgagcagggggccttgacattgcacccagcctgacagcc
tggcatcttgggataaaagcagcacagccccctaggggctgccccttgcctgtgtgtggcaccggcggtggagaacaaggcttattc
agcctgtgcccaggaaagggatcaggggatgccagggcatggacagtggggtggcaggggggagaggagggctgtctgcttcc
cagaagtccaagacacaaatgggtgaggggactgggcagggttctgacctgtgggaccagagtgaggggcgtagatggacctg
aagttccagggaacagggccaggtctcaggctcctagttgggcccagtggtccagcgtttccaaacccatccatcccagag
gttcttccatctctccagcgtgatgtgtgggaactcagggaataaataatctccagtgggagacggagggggtggcagggaacgggg
cgctgcaggaataaagacgagccagcacagccagctcatgtgtaacggcttgggagctgtcaaggcctggtctctgggagagag
gcacagggaggccagacaaggaagggtgacctggaggacagatccaggggctaaagtcctgataaggcaagagagtgccgg
ccccctctgccctatcaggacctccactgccacatagaggccatgattgaccttagacaaagggtggtgtccaatcccagcccc
agccccagaactcagggaatgaatgggcagagagcaggaaatgtggacatctgtgtcaagggaaggactccaggagtgctgtgg
gaatgaggcctagtaggaatgaggtggcccttgagggtacagaacagggtcattcttcgcaaattccagaccttgacggcactta
cagctgagtgagataatgcctgggttatgaatcaaaaagttggaaagcaggtcagaggtcatctgttacagcccttctcctttttt
tttttttttgtagacaaggtctctctgttggccaggctggagtggcgcaaacacagctcactgcagcctcaacctactgggctcaag
caatcctcagcctcagcctccccaaagtgtggtggattacaagcatgagccacccactcagcccttctccttttaattgatgcataat
aattgtaagtattcatcatggtccaaccaacctttctgaccaccttctagagagagggtccttctgcttcagcggtcagggccccag
accatggtctggtcctcagggtaccacctgcctcatgcaggagttggcgtgccagggaagctctgcctctgggcacagtgacctcagt
gggtgaggggagctctcccatagctgggtgcggcccaacccacccctcaggctatgccagggggtgttgcagggggaccc
gggcatcgccagctctagccactcctcataaaagccctcgcatccaggagcga gcagagccagagcagg

SEQ ID NO: 32

Tag

>HA

tatccttatgacgtgcctgactatgcc

Polyadenylation signals**SEQ ID NO: 33**

>hGH ng_011676.1 nt 6537-7013

gggtggcatccctgtgacctcccccagtgccctctctggccctggaagttgccactccagtcccaccagccttgcctaataaaatta
agttgcatcattttgtctgactaggtgtcctctataatattatggggtggaggggggtggtatggagcaaggggcaagttgggaagaca
acctgtaggcctgcggggtctattgggaaccaagctggagtgagtgccacaatcttgctcactgcaatctccgctcctgggtca
agcgattctcctgcctcagcctccccagttgttgggattccaggcatgcatgaccaggctcagctaattttgttttttgtagagcggg
gtttaccatattggcaggctggtctcaactcctaactcaggtgatctaccaccttggcctcccaattgctgggattacaggcgtg
aaccactgtcccttccctgtcctt

SEQ ID NO: 34

>rBG ah001222.2, nt 1623-1749

gatcttttccctctgccaaaattatggggacatcatgaagccccttgagcatctgacttctggctaataaaggaaattttttcattgcaa
tagtgtgttggaattttttgtgtctctcactcg

SEQ ID NO: 35

>rBG ah001222.2, nt 1690-1745

aataaaggaaattttttcattgcaatagtggttggaattttttgtgtctctca

SEQ ID NO: 36

SYNTHETIC PSEN1 CDNA with maximum number of CpG dinucleotides removed by altering only tolerant codons. Changed codon are indicated in lower case.

ATGACAGAGTTACCTGCACcaTTGTCTACTTCCAGAATGCACAGATGTCTGAGGACAACCA
CCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAAagaCAGGAGCACAatGACAGAC
GGAGCCTTGCCACCCTGAGCCATTATCTAATGGAAgaCCCCAGGGTAACTCCagaCAGGTG
GTGGAGCAAGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATggtGCCAAGCATGTGAT
CATGCTCTTTGTCCCTGTGACTCTCTGCATGGTGGTGgttGTGGCTACCATTAAGTCAGTCA
GCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGAAGATACCGAGACT
GTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATCAGTGTCAATTGTTGT
CATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAGGTCATCCATGCCTGGC
TTATTATATCATCTCTATTGTTGCTGTTCTTTTTTTTTCATTCAATTTACTTGGGGGAAGTGT
AAAACCTATAaattGTGCTGTGGACTACATTACTGTTGCACTCCTGATCTGGAATTTTGGTGT
GGTGGGAATGATTTCCATTCACTGGAAAGGTCCACTTagaCTCCAGCAGGCATATCTCATT
TGATTAGTGCCCTCATGGCCCTGGTGTATCAAGTACCTCCCTGAATGGACTgccTGGCTC
ATCTTGGCTGTGATTTTCAGTATATGATTTAGTGGCTGTTTTGTGTcccAAAGGTCCACTTCG
TATGCTGGTTGAAACAGCTCAGGAGAGAAATGAAaccTTTTTCCAGCTCTCATTTACTCCT
CAACAATGGTGTGGTTGGTGAATATGGCAGAAGGAGACccaGAAGCTCAAAGGAGAGTATCC
AAAAATTCCAAGTATAATGCAGAAAGCACAGAAAGGGAGTCACAAGACACTGTTGCAGAGAA
TGATGATggtGGGTTGAGTGAAGATGGGAAGCCAGAGGGACAGTCATCTAGGGCCTCATa
ggTCTACACCTGAGTCAagaGCTGCTGTCCAGGAACTTTCCAGCAGTATCctgGCTGGTGAA
GACCCAGAGGAAAGGGGAGTAAACCTTGGATTGGGAGATTTCAATTTCTACAGTGTCTGTGT
TGTTAAAGCCTCAGCAACAGCCAGTGGAGACTGGAACACAACCATAGCCTGTTTTGTAGCCA
TATTAATTTGGTTTGTGCTTACATTATTACTCTTCCATTTTCAAGAAAGCATTGCCAGCT
CTTCCAATCTCCATCACCTTTGGGCTTGTCTTCTACTTTGCCACAGATTATCTTGTACAGCC
TTTTATGGACCAATTAGCATTCATCAATTTTATATCTAG

SEQ ID NO:37

hPSEN1v1.5. PSEN1 with certain tolerant codons changed to highly preferred synonymous codons. Changed codons are in lower case.

ATGACAGaaTTACCTgccccTTGagcTACTTCCAGAATGCACAGATGagcGAGGACAAC
CACCTGAGCAATACTGTACGTAGCCAGAATGACaacAGAGAACGGCAGgaaCACAAACGAC
aggCGGAGCctgGGCCACCCTGAGccctgtCTAATGGAagaCCCCAGGGTAAcagcaga
CAGGTGGTGgaaCAAGATGAGGAAGaggacGAGGAGCTGaccctgaagtacGGCGCCAAG
cacGTGATCATGCTCcttctgtgcccGTGACTCTCTGCATGGTGGTGgtgGTGGCTacaatc
AAGagcGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGAA
gacACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAatcCTGAATgccGCCATCATGATC
agcGTCATTGTTGTTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAAG
GTCATCCATGCCTGGctgatcATATCATCTctgtTGGctgCTGTTCTTTTTTAgcTTCATT
TACctgggcGAAGTGTTTAAACCTATAACGTTgccGTGGACTACATTACTGTTgccCTC
CTGATCTGGaacttcggcGTGGTGggcATGATTTCCATTCACTGGAAAgccccctgaga
ctgCAGCAGGCAtacCTCATTATGatctcccGCCCTCATGGCCCTGGTgttcATCAAGTAC
ctgcccagagTGGACTgctTGGCTCATCTTGGCTGTGatctccgtgTATGATTTAGTGGCT
GTTctgTGTcctAAAGGTCCAActgCGTATGCTGgtgGAAACAGCTCAGgaaAGAAATGAA
acactgTTTcctGCTctgATTTACTCCTCAACAATGGTGTGGctcGTGAATATGgccGAA
GGAGACcctGAAGccCAAcggAGAggtGCCAAAaacTCCAAGTATAacgccagagAGCACA
GAAAGGGAGagccaggatacaGTTgccGAGAATgacGATGGCggcTTCAGTGAGGAATGG
GAAGCCCAGAGGGACagccacctgGGGCCTcacagaagcaccCTGAGtctagagccGCT
GTCCAGGAAActgTCCAGCTccATCctggccggcGAAGACcccgaagAAAGGGGAGTAAAA
CTTGAActgGGAGATTTcatcTTCTACAGTGTtctcGTTggcAAAGCCagcGCAACAgct
agcGGAGACTGGAACACAacaATAGCCTGTTTCGTAGCCatcTTAATTggcctgTGCCCT
ACActtctgCTCctgGCCatcTTCAAGaaggccctgCCAgccctgctATCagcATCACC

ttcGGGCTTGTTTTCTACTTTGCCaccGATTATctggtgCAGcccttcATGGACcagctg
gccTTCaccagTTTtacATCTAG

SEQ ID NO: 38

hPSEN1v2.0. PSEN1 with certain tolerant codons changed to highly preferred synonymous codons. Changed codons are in lower case.

ATGACAGAGTTACCTGCaccTTGTCTACTTCCAGAATGCACAGATGTCTGAGGACAAC
CACCTGAGCAATACTGTACGTAGCCAGAATGACAATAGAGAACGGCAGGAGCACAACGAC
AGACGGAGCcctcGGCCACCCCTGAGCCAttgTCTAATGGAcggCCCCAGGGTAACCTCCCGG
CAGGTGGTGGAGcagGATGAGGAAGAAGATGAGGAGCTGACATTGAAATATGGCGCCAAG
CATGTGATCATGCTCTTTGTCCCTGTGACTCTCTGCATGGTGGTGGTCTGGCTACCATT
AAGTCAGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGAA
GATACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATTCTGAATGCTGCCATCATGATC
GTCATCCATGCCCTGGctcATTATATCATCTctgTTGTTGCTGTTCTTTTTTTCATTCAATT
TACTTGGGGGAAGTGTTTAAAACCTATAACGTTGCTGTGGACTACATTACTGTTGCACTC
CTGATCTGGAATTTTggcGTGGTGgggATGATTTCCATTCACTGGAAAggcCCActcggg
CTCCAGCAGGCATATCTCATTATGATTAGTGCCCTCATGGCCCTGGTGTATCAAGTAC
CTCCCTGAATGGACTgcccTGGCTCATCTTGGCTGTGATTTCAgtgTATGATTTAGTGGCT
GTTTTGTGTcccAAAGGTCCActcCGTATGCTGgtcGAAACAGCTCAGGAGAGAAATGAA
accctcTTTCAGCTCTCATTTACTCCTCAACAATGGTGTGGTTGGTGAATATGGCAGAA
GGAGACcccGAAGCTCAAAGGAGAgTgTCCAAAAATTCCAAGTATAATGCAGAAAGCACA
GAAAGGGAGTCAcagGACACTGTTGCAGAGAATGATGATGGCGGGTTCAGTGAGGAATGG
GAAGCCCAAGAGGACAGTCACTctgGGGCCTCATCGCTCTACACCTGAGTCAcggGCTGCT
GTCCAGGAAActcTCCAGCAGTATCCTCGCTggcGAAGACCCAGAGGAAAGGGGAGTAAAA
CTTGAGATTGGGAGATTTCATTTTCTACAGTGTCTGGTTggcAAAGCCTCAGCAACAGCC
AGTGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCatcTTAATTggcTTGTGCCTT
ACAttgttgCTCctcGCCATTTTCAAGAAAGCATTGCCAGCTctcCCAATCTCCATCACC
TTTGGGCTTGTTTTCTACTTTGCCACAGATTATctcgtgCAGCCTTTTATGGACcagttg
GCATTCCATCAATTTTATATCTAG

SEQ ID NO:39

hPSEN1 v3.0. PSEN1 with tolerant codons changed to highly preferred synonymous codons. Changed codons are in lower case.

ATGACAGAGTTACCTgccccTTGagcTACTTCCAGAATGCACAGATGagcGAGGACAACCA
CCTGAGCAATACTGTACGTAGCCAGAATGACaacAGAGAACGGCAGGAGCACAACGACAGAC
GGAGCcctgGGCCACCCCTGAGccccctgagcAATGGAAgaCCCCAGGGTAACagcCGGCAGGTG
GTGGAGcagGATGAGGAagaggacGAGGAGCTGaccctgaagtacGGCGCCAAGcacGTGAT
CATGCTCctcgtgcccGTGACTCTCTGCATGGTGGTGGTggtgGTGGCTACCatcAAGagcGTCA
GCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGAAgacACCGAGACT
GTGGGCCAGAGAGCCCTGCACTCAatcCTGAATgccGCCATCATGATCagcGTCATTGTTGT
CATGACTATCCTCCTGGTGGTTCTGTATAAAATACAGGTGCTATAAGGTCATCCATGCCTGGc
tgatcATATCATCTctgTTGctgCTGTTCTTTTTTtagcTTCATTTACctgggcGAAGTGTTT
AAAACCTATAACGTTgcccGTGGACTACATTACTGTTgccCTCCTGATCTGGaacttcggcGT
GGTGggcATGATTAgcATTCACTGGAAAgccccctgagactgCAGCAGGCAtacCTCatcA
TGatcagcGCCCCCATGGCCCT
GGTGttcATCAAGTACctgccccagTGGACTgccTGGCTCATCTTGGCTGTGatcagcgtgT
ATGATTTAGTGGCTGTTctgTGTcccAAAGGTCCActgCGTATGCTGgtggagACAGCTCAG
GAGAGAAATGAAaccctgTTTcccGCTctgATTTACTCCTCAaccATGGTGTGGctgGTGAA
TATGgccGAAGGAGACcccGAAgcccAAcggAGAgTgagcAAAaacagcAAGTATAacgccg
agAGCACAGAAAGGGAGagccagGACaccGTTgccGAGAATgacGATGGCggcTTCAGTGAG
gagTGGGAAGCCAGAGGGACagccacctgGGGccccacagaagcaccctcGAGagcagagc
cGCTGTCCAGGAActgTCCAGCagcATCctggccggcGAAGACcccGAGGAAAGGGGAGTAa

agCTTGGActgGGAGATTTCatcTTCTACAGTGTCTGGTTggcAAAGCCagcGCAACAGCC
 agcGGAGACTGGAACACAACCATAGCCTGTTTCGTAGCCatcTTAATTggcctgTGCCTTac
 cctgctgCTCctgGCCatcTTCAAGaagggccctgCCAgccctgcccATCagcATCACcttcG
 GGCTTGTTCCTACTTTGCCaccGATTATctggtgCAGcccttcATGGACcagctggccTTC
 caccagTTTtacATCTGA

SEQ ID NO:40

CAG promoter used in pAT029, pAT024, and pAT022

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
 TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCTGGC
 TGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAG
 TAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAAC
 TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC
 GTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTATGGG
 ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGCA
 GGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAA
 TTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGG
 GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGC
 GGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTA
 TGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCG
 GGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCC
 CGCCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCC
 CTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAAATGACGGCTCGTTTCTTTCTG
 TGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAG
 CGGCTCGGGGGGTGCGTGCGTGTGTGTGCGTGCGGAGCGCCGCGTGCGGCC
 GCGCTGCCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGGGCTTTGTGCGCTC
 CGCGTGTGCGCGAGGGGAGCGCGGCGGGGGCGGTGCCCGCGGTGCGGGGGG
 GCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGCGGGGGGTGAGCA
 GGGGGTGTGGGCGCGGCGGTGCGGTGTAACCCCCCTGCACCCCCCTCCCCG
 AGTTGCTGAGCACGGCCCGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCG
 GGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGG
 CGGGGCCGCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGA
 GCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATC
 GTGCGAGAGGGCGCAGGGACTTCTTTGTCCCAAATCTGGCGGAGCCGAAATCT
 GGGAGGCGCCCGCGCACCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCTTCGTGCGTCGCCGCGCCGCCGTCCCC
 TTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGGG
 ACGGGGAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTC
 TGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAG

SEQ ID NO: 41

Expression cassette for AAV production containing AAV2 ITRs (1-141, 4553-4693),
Presenilin 1 promoter (underlined 237-1200), a synthetic human beta-globin intron (1221-
 1786), HA-tag (1866-1898), codon-optimized human presenilin 1v1.5 (uppercase SEQ ID NO:
 37, 1899-3299), human growth hormone polyadenylation sequence (3330-3806) and albumin
 genomic sequence (3810-4522).

cctgcaggcagctgcgcgctcgtcgtcactgagccgcccgggcaagcccggcgctcgggcgacctttggctgcccggcctc
 agtgagcgagcgagcgcgagagggagtgcccaactccatcactaggggtcctcgcgccaatcagtcgataactataacggt
 cctaagtagcgatttaaatacgcgtctcttaagtagccccgggacgcgtcaattgagatctcgggggtggaattttaaagaaacttta
gaagaatgtaacttcccagataccatgtaccgttaatttcattttcggtttttgaatacccatgtttgacatttccgttcaccttgattaat
aaggtagtattcatttttagtttagcttttgatataatgttaagtgtggtatgctgtctaataaattagacattggtactgtctttacaaaaac

tggacaaagagcagcagatgcaaaaatcaagtgacccagcaaacagacacattttctgctctcagctagcttgcacctagaaga
ctggttctcaaaaggggagtcgaagaatcgccggaggatgtttaaatagcagtttctcaggttctgccaccaccagaagtgttgattcat
tgagtggtagggagagggcagagataatttcggaatttaacagcattctcttgattgtgatgcagctggttcgcaaataggtagccctaaagaa
atgacaggtgttaaatfaggatggccatcgcttgatgccgggagagcacacgctggcccaatttatatagggccttctcctcag
ctcgagcagcctcagaaccccgacaacccacgccagcgtctggcggttccgtcaggtgggaaggccaagtgagctctggg
ttctccccgaatcgttttccagggccggagcccgcccccttctcctgctcctccccctcctcgtggccggccgccaacgacg
ccagagccggaaatgacgacaacggtgagggttctcgggcgggcgctgggacaggcagctccggggtccgcggttcacatcgg
aaacaaaacagcgcgctgctctggaaggaaacctgagctacgagccgcgcgcgagcgggcgcgcggggaaagcggtatgtcctgat
ggggagctccggcgaagccaggaaggcaccgggacatggcggaagcttctgttagtaaccgtcagatcgctggagacgccat
ccacgctgttttgacctccatagaagacaccgggaccgatccagcctccggttcgaatcccgccgggaacggtgacgttggaaac
gcggtattccccgtgccaaagtgacgtaagtaccgctatagagctatagggcccaaaaaatgctttctcttttaatactattttgtt
atcttttcttaatactttccctaattctcttcttcaggcgaataatgatacaatgtatcatgccttttgaccattctaaagaataacagt
aatttctgggttaaggcaatagcaatatttctgcatataaatttctgcatataaattgtaactgatgaagggttcatattgctaatagca
gctacaatccagctaccattctgctttatgttggtgggataaggctggatttctgagtcgaagctaggcccttttgtaactggttca
tacctttatctctcccacagctcctgggcaacgtgctggtctgtgctggcccatcatttggcaagaattaccggtctcctgggc
aacgtgctggttattgtgctgtctcatatttggcaagaattcacgcccagagccgcccaccATGgcctaccatacagatgtccag
attacgtACAGAATTACCTGCCCCCTTGAGCTACTTCCAGAATGCACAGATGAGCGA
GGACAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAACAGAGAACGGCA
GGAACACAACGACAGGCGGAGCCTGGGCCACCCTGAGCCCCCTGTCTAATGGAAG
ACCCCAGGGTAACAGCAGACAGGTGGTGGAAACAAGATGAGGAAGAGGACGAGG
AGCTGACCCTGAAGTACGGCGCCAAGCACGTGATCATGCTCTTCGTGCCCGTGAC
TCTCTGCATGGTGGTGGTGGTGGCTACAATCAAGAGCGTCAGCTTTTATACCCGG
AAGGATGGGCAGCTAATCTATACCCCATTACAGAAGACACCGAGACTGTGGGC
CAGAGAGCCCTGCACTCAATCCTGAATGCCGCCATCATGATCAGCGTCATTGTTG
TCATGACTATCCTCCTGGTGGTTCGTGTATAAATACAGGTGCTATAAGGTCATCCA
TGCCTGGCTGATCATATCATCTCTGTTGCTGCTGTTCTTTTTTAGCTTCATTTACCT
GGGCGAAGTGTTTAAACCTATAACGTTGCCGTGGACTACATTACTGTTGCCCTC
CTGATCTGGAACCTTCGGCGTGGTGGGCATGATTTCCATTCACTGGAAAGGCCCCC
TGAGACTGCAGCAGGCATACCTCATTATGATCTCCGCCCTCATGGCCCTGGTGTT
CATCAAGTACCTGCCCCAGTGGACTGCTTGGCTCATCTTGGCTGTGATCTCCGTG
TATGATTTAGTGGCTGTTCTGTGTCTTAAGGTCCACTGCGTATGCTGTGTGAAA
CAGCTCAGGAAAGAAATGAAACACTGTTTCTGCTCTGATTTACTCTCAACAAT
GGTGTGGCTCGTGAATATGGCCGAAGGAGACCCTGAAGCCCAACGGAGAGTGTC
CAAAAACCTCCAAGTATAACGCCGAGAGCACAGAAAGGGAGAGCCAGGATACAG
TTGCCGAGAATGACGATGGCGGCTTCAGTGAGGAATGGGAAGCCCAGAGGGACA
GCCACCTGGGGCCTCACAGAAGCACCCCTGAGTCTAGAGCCGCTGTCCAGGAAC
TGTCCAGCTCCATCCTGGCCGGCGAAGACCCCGAAGAAAGGGGAGTAAACTTG
GACTGGGAGATTTTCATCTTCTACAGTGTTCTCGTTGGCAAAGCCAGCGCAACAGC
TAGCGGAGACTGGAACACAACAATAGCCTGTTTCGTAGCCATCTTAATTGGCCTG
TGCCTTACACTTCTGCTCCTGGCCATCTTCAAGAAGGCCCTGCCAGCCCTGCCTAT
CAGCATCACCTTCGGGCTTGTTTTCTACTTTGCCACCGATTATCTGGTGCAGCCCT
TCATGGACCAGCTGGCCTTCCACCAAGTTTACATCTAGtaagcgccgacctaggagctcctg
aggggggtgcatccctgtgacctccccagtgccctcctggccctggaggtgccaactcagtgcccaccagcctgtcctaataaa
attaagttgcatcatttctgctgactaggtgtccttctataatattatgggggtgaggggggtggtatggagcaaggggcaaggggggaa
gacaacctgtagggcctcggggtctattgggaaccaagctggagtgagtgccacaatcttgctcactgcaatctccgctcctgg
gttcaagcgattctcctcagcctccgagttgttgggattccagcatgcatgaccaggctcagtaattttgttttggtagaga
cgggggttcacataattggccaggctggtctccccctctaattctcaggtgacttaccaccttggcctcccaattgctgggattacagg
cgtgaaccactgctcccttccctgtccttggcctagggttcttgagacctctacaagagggggagttgacacttgggggtactttctggtgt
aacgaactaatagcctgaaaaaagaagtcagtgttttcagcaaggcaagaactgtctaatactagtagataaaacagagaacacttg
gccggaatcaactaagatgtgctatgttccattcatattatctccatctgcagagtagtggttagtgagggtgaaacattctcc
tgaacaactagttaaacttggctttagttccacctgtaccacttgcataatctgggaaagtgaattgcctaattcagtgacattaataaatt

SEQ ID NO: 42

ctctgagcgagctgcgcgtctgctctactgagcgccggcgcaagccggcgctggggcgaccttggctgcccgcctc
agttagcgagcgagcgcgagagagggagtgggcaactccatcactaggggttcctgcggccaattcagtgataactataacggt
cctaaggtagcgatttaaatacgcgtctcttaaggtagccccgggacgcgtcaattgagatctcggtcagcggcctccgcgcgggg
ttttggcgcctcccgccggcgccccctctcacggcgagcgtgccacgtcagacgaagggcgacgagcgttctgatacttcc
gccccgacgtcagacagcgccccgtctctcataagactcgcccttagaacccagtatcagcagaaggacattttagacggga
cttggtgactctagggcactgtgttttcttcagagagcggaacagcgagggaaaagttagtcccttctcgccgattctcgcgagggat
ctccgtggggcggtgaacgccgatgattatataaggacgcgccgggtgtgtgcacagctagttccgtgcgacggggatttgggtcgc
ggttcttgtttgtgatcgtgtgatcgtcattgtgtgaagttcgccggtctgtgggtctgccggggcttctgtgcccgccggcgccgtcg
gtgggacgggaagcgtgtgtgagagacgcccaagggtctgtatgtctgggtccgcgagcaaggttgccttgaactgggggttgggggg
agcgcacaaaaatggcgctgttcccaggtcttgaatggaagacgttgtgaagcgggctgtgaggtcgttgaacaaggttgggggg
catggtggcgcgcaagaacccaaggcttctgagcccttcgctaattcgccggaaaagctcttattcgggtgagatgggtcggggcaccatc
tggggacctgacgtgaagtttgcactgactcgagaactcgggttttgcgtcgtgttgcggggcgccgacgttatcggttgcgttgg
cagtgcacccgtacctttgggagcgcgccctcgtcgtgtcgtgacgtcacccgttctgttggccttataatgcaggggtggggccacctg
ccggtaggtgtgcggtagcgttttctccgtcgcaggacgcagggttcgggcttagggtagcctctctgaatcgacagcgccggac
ctctgtgtaggggagggataagtgagcgctcagtttcttggtcggttttatgtacctatctttaaagtacgtgaagctccggttttgaa
atgactagtaaaaagcttcgtttagtgaacctgcagatcgccctggagacgccatccacgtgtttgacctcataagaagacacggga
cgatccagcctccgcggattcgaaatccgcggccgggaacgggtgacgttggaggtcccttgcgaagagtgcagtgaattgacc
gccataagagtcatataggccacaaaatgcttcttctttaaataactttgtttgattcttattctaacttccctaactcttctttcagg
gcaataatgatacgtatgatcatgcctctttgcacatttctaagaataaacagtgataatttctgggttaaggcaataagcaatttctgcata
taaatatttctgcataataaattgtaactgatgtaagaggttcatattgtctaatagcactacaatccagctaccattctgctttattttgtggt
gggataaggctggattattctgagtccaagctaggcccttttgctaactcgtgttcataaccttattcttctcccacagctcctgggcaacgt
gctggtctgtgtctggcccatcatttggcaagaattaccggtgccaacgtgctggttattgtgtgtctcatcatttggcaagaatt
cacgccccagagccgccaccATGgcctaccatacagatgttccagattacgtACAGAATTACCTGCCCCCTTG
AGCTACTTCCAGAATGCACAGATGAGCGAGGACAACCACCTGAGCAATACTGTA
CGTAGCCAGAATGACAACAGAGAACGGCAGGAACACAACGACAGGCGGAGCCT
GGGCCACCCTGAGCCCCTGTCTAATGGAAGACCCCAGGGTAACAGCAGACAGGT
GGTGGAAACAAGATGAGGAAGAGGACGAGGAGCTGACCCTGAAGTACGGCGCCA
AGCACGTGATCATGCTCTTCGTGCCCCGTGACTCTCTGCATGGTGGTGGTGGTGGC
TACAATCAAGAGCGTCAGCTTTTTATACCCGGAAGGATGGGCAGCTAATCTATACC
CCATTACAGAAGACACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATCCTG
AATGCCGCCATCATGATCAGCGTCATTGTTGTCATGACTATCCTCCTGGTGGTTCT
GTATAAATACAGGTGCTATAAAGGTACATCCATGCCTGGCTGATCATATCATCTCTG
TTGCTGCTGTTCTTTTTAGCTTCATTTACCTGGGCGAAGTGTTTAAACCTATAA
CGTTGCCGTGGACTACATTACTGTTGCCCTCCTGACTCTGGAACCTCGGCCTGGTG
GGCATGATTTCCATTACTGGAAGGCCCCCTGAGACTGCAGCAGGCATACCTCA
TTATGATCTCCGCCCTCATGGCCCTGGTGTTTCATCAAGTACCTGCCGAGTGAC
TGCTTGGCTCATCTTGGCTGTGATCTCCGTGTATGATTTAGTGGCTGTTCTGTGTC

CTAAAGGTCCACTGCGTATGCTGGTGGAAACAGCTCAGGAAAGAAATGAAACAC
 TGTTTCCTGCTCTGATTTACTCCTCAACAATGGTGTGGCTCGTGAATATGGCCGA
 AGGAGACCCTGAAGCCCAACGGAGAGTGTCCAAAACTCCAAGTATAACGCCGA
 GAGCACAGAAAGGGAGAGCCAGGATACAGTTGCCGAGAATGACGATGGCGGCT
 TCAGTGAGGAATGGGAAGCCCAGAGGGACAGCCACCTGGGGCCTCACAGAAGC
 ACCCTGAGTCTAGAGCCGCTGTCCAGGAAGTGTCCAGCTCCATCCTGGCCGGCG
 AAGACCCGAAGAAAGGGGAGTAAACTTGGACTGGGAGATTTCATCTCTAC
 GTGTTCTCGTTGGCAAAGCCAGCGCAACAGCTAGCGGAGACTGGAACACAACAA
 TAGCCTGTTTCGTAGCCATCTTAATTGGCCTGTGCCTTACACTTCTGCTCCTGGCC
 ATCTTCAAGAAGGCCCTGCCAGCCCTGCCTATCAGCATCACCTTCGGGCTTGT
 TCTACTTTGCCACCGATTATCTGGTGCAGCCCTTCATGGACCAGCTGGCCTTCCAC
 CAGTTTTACATCTAGtaagcgccgccctaggagctcctcaggggggtggcatcctgtgacccctccccagtgct
 ctctggccctggaagttgccactccagtgcccaccagcctgtcctaataaataagttgcatcatctttgtctgactagtgctctctat
 aatattatgggtggaggggggtggatggagcaaggggcaaggggggaagacaacctgtagggcctgccccgtctaltgggaac
 caagctggagtgagtgccacatcttggtcactgcaatctccgctcctgggtcaagcattctcctcagcctccccagttgt
 tgggattccagcatgcatgaccagctcagtaattttgtttttgttagagacggggttcaccatattggccaggtgctctcccc
 tctaatactcaggtgatctaccacctggcctccaaattgctgggattacaggcgtgaaccactgctcccttccctgctctctgattta
 aaataactataccagcaggagagcgtcagacacagcataggctacctggccatgcccaaccgggtgggacatttgagttgcttg
 gcactgtcctctcatgctgtgggtccactcagtagatgctgttggaattcctggcctaggcgtgtgccagctgctcgtccccacatt
 ctggcttctctcctccatcttagctgtttcctcatgagaatgtccaaattcgaatttctatttaaccattatatattactgtttgctat
 tatctctgccccagtagattgttagctccagaagagaagagatcatgtctttgtcttagatagatgccatctgctggtacaatctctg
 gcacatgttacaggcaacaactactgtggaattggtaagcatgaatagaagaatgagtgaaatgaatgaatagacaataggcagaa
 atccagcctcaaagagcttacagcttggttaagaggaataaaatgtctgcaaatagccacaggacaggtcaaggaaggaggggctat
 ttccagctgagggcacccatcaggaagcaccacagcttcttagggataacagggtaatggcgcgggccgaggaaccctag
 tgatggagttggcactcctctctgctgctcgtcgtcactgagggcgggcgaccaaaggctcgccgacgccccgggctttgcc
 gggcgccctcagtgagcgagcgagcgagctgcctgcagg

SEQ ID NO: 43

Expression cassette for AAV production containing AAV2 ITRs (1-141, 4554-4694), CBA promoter (underlined, 237-890), a synthetic human beta-globin intron (911-1476), HA-tag (1556-1582), codon-optimized human presenilin 1v1.5 (uppercase SEQ ID NO: 37, 1550-2983), human growth hormone polyadenylation (3014-3490) and human growth hormone genomic sequence (3014-3491).

cctgcaggcagctgcgcgctcgtcgtcactgaggccgcccgggcaagcccggcgctcggcgacctttggtcgcccgccctc
 agtgagcgagcgagcgcgagagaggagtgccaaactccactaggggtcctgcggccaaatcagtcgataactatacgggt
 cctaaggtagcgatttaatacgcgctctcctaaggtagccccgggacgcgtcaattgagatctcgcacattgattattgactagtatttaa
 tagtaataattacgggtcatttagttcatagcccatataggagtccgcgttacataaactacggtaaatggccccgctggtgaccgc
 ccaacgacccccgccattgacgtcaataatgacgtatgttcccatagtaacgccaataggacatttccattgacgtcaatgggtggagt
 atttacggtaaactgccacttggcagtagcatcaagtgtatcatatgccaaagtagccccctattgacgtcaatgaggttaaatggccg
 cctggcattatgccagtagcatgacatttgggacttctacttggcagtagcatctacgtatttagtcatcgctattaccatgtcagggcca
 cgttctgcttactctccccatctccccccctccccaccccaattttgtatttttttttaattttttgtcagcgatggggcgggg
 gggggggcgcgccagcgcgggcgggcgggcgggcgagggcgggcgggcgagcgagagtgcgcgcgccagcca
 atcagagcgcgcgctccgaaagtcttctttatggcgagcgcgcgcgcgcgcgccctataaaaagcgaaagcgcgcgcgggcg
 gggagcaagctcgttttagtaaccgtcagatcgctggagacgcatccacgctgtttgacctccatagaagacacgggaccgat
 ccagcctccgcgattcgaatccccggcggaacgggtgattggaacgcggattccccgtgccaagagtgacgttaagtaccgcctat
 agagctctataggccacaaaaatgcttctcttttaatactttttgtttatcttatttctaatactttccctaactcttcttccagggaata
 atgatacaatgtatcatgctcttgcaccatttcaagaataacagtataatttctgggttaaggcaatagcaatatttctgcataataata
 ttctgcataataattgtaactgatgtaagaggtttcatattgctaatagcagctacaatccagctaccattctgctttattttgtggtggg
 aaggctggattattctgagtcgaagtagccctttgtcaatcgtgttcatacctcttattctctccacagctcctgggcaacgtgctgtg
 tctgtgtgctggccatcatttgcaagaattaccgggtgcaacgtgctgtgtgtctctcatattttgcaagaattacgc
 cccagagccgccaccATGgcctaccatacagatgttcagattacgctACAGAATTACCTGCCCCCTTGAGC

TACTTCCAGAATGCACAGATGAGCGAGGACAACCACTGAGCAATACTGTACGT
AGCCAGAATGACAACAGAGAACGGCAGGAACACAACGACAGGCGGAGCCTGGG
CCACCCTGAGCCCCTGTCTAATGGAAGACCCCAGGGTAACAGCAGACAGGTGGT
GGAACAAGATGAGGAAGAGGACGAGGAGCTGACCCTGAAGTACGGCGCCAAGC
ACGTGATCATGCTCTTCGTGCCCCGTGACTCTCTGCATGGTGGTGGTGGTGGCTAC
AATCAAGAGCGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCC
ATTCACAGAAGACACCGAGACTGTGGGCCAGAGAGCCCTGCACCTAATCTGAA
TGCCGCCATCATGATCAGCGTCATTGTTGTCATGACTATCCTCCTGGTGGTTCTGT
ATAAATACAGGTGCTATAAGGTCATCCATGCCTGGCTGATCATATCATCTCTGTT
GCTGCTGTTCTTTTATAGCTTCATTTACCTGGGCGAAGTGTTTAAACCTATAACG
TTGCCGTGGACTACATTACTGTTGCCCTCCTGATCTGGAACCTCGGCGTGGTGGG
CATGATTTCCATTCACTGGAAAGGCCCCCTGAGACTGCAGCAGGCATACCTCATT
ATGATCTCCGCCCTCATGGCCCTGGTGTTCATCAAGTACCTGCCCGAGTGGACTG
CTTGGCTCATCTTGGCTGTGATCTCCGTGTATGATTTAGTGGCTGTTCTGTGTCCT
AAAGGTCCACTGCGTATGCTGGTGGAAACAGCTCAGGAAAGAAATGAAACACTG
TTTCTGCTCTGATTTACTCCTCAACAATGGTGTGGCTCGTGAATATGGCCGAAG
GAGACCCTGAAGCCCAACGGAGAGTGTCCAAAACTCCAAGTATAACGCCGAGA
GCACAGAAAGGGAGAGCCAGGATACAGTTGCCGAGAATGACGATGGCGGCTTC
AGTGAGGAATGGGAAGCCCAGAGGGACAGCCACCTGGGGCCTCACAGAAGCAC
CCCTGAGTCTAGAGCCGCTGTCCAGGAAGTGTCCAGCTCCATCCTGGCCGCGCAA
GACCCCGAAGAAAGGGGAGTAAACTTTGACTGGGAGATTTCATCTTCTACAGT
GTTCTCGTTGGCAAAGCCAGCGCAACAGCTAGCGGAGACTGGAACACAACAATA
GCCTGTTTCGTAGCCATCTTAATTGGCCTGTGCCTTACACTTCTGCTCCTGGCCAT
CTTCAAGAAGGCCCTGCCAGCCCTGCCTATCAGCATCACCTTCGGGCTTGTTTTCT
ACTTTGCCACCGATTATCTGGTGCAGCCCTTCATGGACCAGCTGGCCTTCCACCA
GTTTTACATCTAGTaaagcgccgccctaggagctcctcgagggggtggcaccctgtgacccctccccagtgccctcc
tgccctggaagtgccactccagtgccaccagcctgtgctcaataaaatgaagtgcatcattttgtctgactaggtgtcctctataat
tatgggggtgaggggggtgtatggagcaagggggaagacaacctgtagggcctgagggtgtattgggaaccaag
ctggagtgagtgccacatcttggtcactgcaatccgcctcctgggtcaagcgattctcctgacctcagctcccgagtggtggg
attccaggcatgcatgaccaggctcagctaattttgtttttgtagagacggggttcacatattggccaggctggtctccccctcta
atctcaggtgatctaccaccttgccctccaaattgctgggattacaggcggtgaaccactgctcccttcctgtccttctgattttaaata
actataccagcaggagacgtccagacacagcataggctacctggccatgcccacccggtgggacatttgagttgcttctggcact
gtcctctcatgttggttcactcagtagatgcctgttgaaattcctgggcttagggctgtgcccagctgcctcgtcccgctacacctctggc
ttctctcctcccatatcttagctgtttctcctagagaatgttccaaattcgaaatttctatttaaccattatatttactgtttgcttattatc
tgccccagtagattgttagctccagaagagaaggatcatgtcttttcttatctagatatgcccactgctggtacatctctggcaca
tgttacaggcaacaactactgtggaattggtgaatgcatgaatagaagaatgagtgaaatgaatgaatagacaataggcagaaatccag
cctcaagagcttacagctctggttaagaggaataaaatgtctgcaaatagccacaggacaggtcaaagggaaggagggtatttccag
ctgaggggaccccatcaggaaagcaccacagacttctacaactactagacacatctcgatgcttttcaacttctatcaatggatgct
ccctggagaataatccccaaagtgaattacttagcacgtccagttaggtagatccttgtgacttctggttgttcagagatcatcaacca
gtgcaacaatcccccatcaatacacagcagtgccctgccccctcctcccccagggtctccgaggccctcctccgtgctgaacccc
ctggacatatcatatggcaactgaagtgccaacgagatataggaagtgaacacgatgtacactgaaacgtgcaatacaaatatgca
gcatgaagtgctcgttactaaccgagctacgtgggtgcttcttctaccacttcttaatgcctatggacacctcattctgtggt
gaagttcctgtgttcaatagggataacagggtaatggcgccggcgaggaaaccctagtgatggagttggccactccctctctgcg
cgctcgtcgtcactgaggccggcgaccaaaggctgcccagcggcggttggccggcgccctcagtgagcgagcgagc
gcgcagctgcctgcagg

SEQ ID NO: 44

Expression cassette for AAV production containing AAV2 ITRs (1-141, 4500-4640), EF-1alpha promoter (underlined, 237-1415), a synthetic human beta-globin intron (1436-2001), HA-tag (2081-2107), codon-optimized human presenilin 1v1.5 (uppercase SEQ ID NO: 37,

2075-3508), human growth hormone polyadenylation (3539-4015) and human growth hormone genomic sequence (4016-4469).

cctgcaggcagctgcgcgtcgtcgtcactgagccgcccgggcaaagcccgggctcgggagacatttggtcgcggcctc
agtgcgagcgcgcgcagagaggagtgccaaactccatcactaggggtcctgcggccaattcagtcgataactataacggt
cctaaggtagcgatttaataacgcgtctcttaaggtagccccgggacgcgtcaattgagatctcggctccgggtccccgtcagtgggc
agagccacatcggccacagtcggcagagggggggggaggggtcgccaattgaaccggtccctagagaaggtggcgccggg
taaacgggaaagtgcgtgactggtccgctcttcccgagggtgggggagaaccgtatataagtcagtagtcgcccgtgaac
gtctttttcgcaacgggtttgcccgcagaaacacaggtaaagtgcgtgtgtgttcccgccggcctggccttttacgggttatggcctt
gcgtgccttgaattactccacctggctgcagtagctgattcttgcacccgagcttcgggttggaagtgggtgggagagttcgagccct
ggccttaaggagcccttcgctcgtgcttgaagtggcctggcgtggcgctggggcccgccgtgcgaatctgtgtgcaccttc
ggcctgtctcgtgcttgcgataagtcctagccatttaaaattttgatgacctgctgcgacgcttttttggcaagatagcttcta
ggggccaagatcgcacactggtatttgggtttggggccggggcgacggggcccgctgcgtcccgccgacatgttcggc
gagcgggggcctgcgagcgccgcccaccgagaatcgagcggggtagtctcaagctggccggcctgctctggtgctgtcgcg
ccgctgtatcggccgcccctggcgcgcaaggctggcccggtcgccaccagtgcgtgagcggaagatggccgctcccgcc
ctgctgcaggagctcaaatggaggacggcgccgctgggagagcgggcggtgagtcacccacacaaaagaaaaggcccttc
cgtctcagccctcgttcatgtactccacggagtagccggcgccgtccagggcacctcgattagtctcgcgttttgagtagctgt
ctttaggttgggggaggggtttatgcgatggagtttcccacactgagtggttgagactgaagtagccagcttggcactgtatgt
aattctccttgaatttgcctttttgatttggatcttgttcattctcaagcctcagacagtggttcaaagttttttctcatttcaggtgtcg
tgaagcttctgttagtgaaccgtcagatcgctggagacgcatccacgctgtttgacctcatagaagacaccgggaccgatccag
cctccgcgattcgaatcccgccgggaacggtgattggaacgcggttccccgtgccaagagtgacgtaagtaccgctataga
gtctataggcccaaaaaatgcttctctttaaatacttttggatcttatttctaatacttccctaattcttcttcagggaataatg
atacaatgtatcatgctcttgcaccattctaaagaataacagtgataatttctgggttaaggcaatagcaatatttctgcatataatatttc
tgcataataattgtaactgatgaagaggttcatattgctaataagcagctacaatccagctaccattctgctttatttttggttgggataag
gctggattattctgagtcgaagctaggccctttgctaactggttcatacctcttatcttctccacagctcctggcaacgtgctggtct
gtgtgctggcccatcatttggcaagaattaccggtggcaacgtgctggttattgtgctgtctcatcatttggcaagaattcacgcc
cagagccgccaccATGgcctaccatacgtatgtccagattacgtACAGAATTACCTGCCCCCTTGAGCT
ACTTCCAGAATGCACAGATGAGCGAGGACAACACCTGAGCAATACTGTACGTA
GCCAGAATGACAACAGAGAACGGCAGGAACACAACGACAGGCGGAGCCTGGGC
CACCCTGAGCCCCCTGTCTAATGGAAGACCCCAGGGTAACAGCAGACAGGTGGTG
GAACAAGATGAGGAAGAGGACGAGGAGCTGACCCTGAAGTACGGCGCCAAGCA
CGTAGATCATGCTCTTTCGTGCCCCGTGACTCTCTGCATGGTGGTGGTGGGTACA
ATCAAGAGCGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCAT
TCACAGAAGACACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATCCTGAATG
CCGCCATCATGATCAGCGTCATTGTTGTCATGACTATCCTCCTGGTGGTCTGTAT
AAATACAGGTGCTATAAGGTCATCCATGCCTGGCTGATCATATCATCTCTGTTGC
TGCTGTTCTTTTTTAGCTTCATTTACCTGGGCGAAGTGTTTAAACCTATAACGTT
GCCGTGGACTACATTACTGTTGCCCTCCTGATCTGGAACCTCGGCGTGGTGGGCA
TGATTTCCATTCATCTGGAAGGCCCCCTGAGACTGCAGCAGGCATACCTCATTAT
GATCTCCGCCCTCATGGCCCTGGTGTTCATCAAGTACCTGCCCAGTGGACTGCT
TGGCTCATCTTGGCTGTGATCTCCGTGTATGATTTAGTGGCTGTTCTGTGTCTAA
AGGTCCACTGCGTATGCTGGTGGAAACAGCTCAGGAAAGAAATGAAACACTGTT
TCCTGCTCTGATTTACTCCTCAACAATGGTGTGGCTCGTGAATATGGCCGAAGGA
GACCCTGAAGCCCAACGGAGAGTGTCCAAAAACTCCAAGTATAACGCCGAGAGC
ACAGAAAGGGAGAGCCAGGATACAGTTGCCGAGAATGACGATGGCGGCTTCAGT
GAGGAATGGGAAGCCAGAGGGACAGCCACCTGGGGCCTCACAGAAGCACCCC
TGAGTCTAGAGCCGCTGTCCAGGAAGTGTCCAGCTCCATCCTGGCCGGCGAAGA
CCCCGAAGAAAGGGGAGTAAACTTGGACTGGGAGATTTTCATCTTCTACAGTGT
TCTCGTTGGCAAAGCCAGCGCAACAGCTAGCGGAGACTGGAACACAACAATAGC
CTGTTTCGTAGCCATCTTAATTGGCCTGTGCCTTACACTTCTGCTCCTGGCCATCT
TCAAGAAGGCCCTGCCAGCCCTGCCTATCAGCATCACCTTCGGGCTTGTTTTCTA
CTTGGCCACCGATTATCTGGTGCAGCCCTTCATGGACCAGCTGGCCTTCCACCAG

TTTACATCTAGtaagcgccgcccctagggagctcctcgagggggtggcatccctgtgacccctccccagtgccctcctg
 gccctggaagtggccactccagtgcccaccagcctgtcctaataaaattaagttgcatcattttgtctgactaggtgtccttataatatta
 tgggggtggaggggggtggtatggagcaagggggcaaggggggaagacaacctgtagggcctgcggggtctattgggaaccaagct
 ggagtgacagtggcacaatcttggtcactgcaatctccgctcctgggtcaagcgattctcctgcctcagcctcccgagttgtgggatt
 ccaggcatgcatgaccaggctcagctaattttgtttttgtagagacggggttcaccatattggccaggctggtctccccctcta
 ctgaggtgatctaccaccttggcctccaaattgctgggattacaggcgtgaaccactgtcccttccctgtccttctggcctagg
 ctgtgccagctgcctcgtcccgctcaccttctggtcttctcctccatcttagctgttttctcatgagaatgtccaaattcgaaatttc
 tatttaaccattatatttactgtttgtctattatctctgccccagtagattgttagctccagaagagaagagcatgtcttttgcctatctag
 atatgcccattctgcttgtaacaatctctggcacatgttacaggcaacaactactgtggaattggtgaatgcataagaagaatgag
 gaattgaatgaatagacaatagggcagaaatccagcctcaaaagagcttacagctgtggaagaggaataaaatgtctgcaaatagccacag
 gacagggtcaaaaggaagggggtatttccagctgaggggcacccatcaggaaagcaccacagacttcttagggataaacagggt
 atggcgccggcgccaggaacccctagtgatggagttggccactcctctctgcgcgctcgtcgtcactgaggccggcgacc
 aggtcgccgacgcccgggctttgccggggcgcctcagtgagcgagcgagcgcgcagctgcctgcagg

SEQ ID NO: 45

Expression cassette for AAV production containing AAV2 ITRs (1-141, 4533-4673), PGK promoter (underlined, 237-663), a synthetic human beta-globin intron (684-1249), HA-tag (1329-1355), codon-optimized human presenilin 1v1.5 (uppercase SEQ ID NO: 37, 1323-2756), human growth hormone polyadenylation (2787-3263) and human growth hormone genomic sequence (3264-4502).

cctgcaggcagctgcgcgctcgtcgtcactgaggccgcccgggcaagccccgggctcgggcgacctttggtcggccgacctc
 agtgagcgagcgagcgcgagagaggagtgggcaactccatcactaggggtcctgcggccaattcagtcgataactataacggt
 cctaaggtagcgatttaataacgctccttaaggtagccccgggacgctcaattgagatctcgcctcgaattccacgggggtgggg
ttgcgcctttccaaggcagccctgggttgcgcaggacgcggctgctctgggcgtggttccgggaacgcagcgccgcccaccc
gggtctcgcacattcttcacgtcgttgcgcagcgtacccggatcttcgccgtacccttctgggccccccggcgacgcttctgctcc
gccctaagtcgggaaggttcttgcgttcggcgctgccggacgtgacaaaagggaagccgcagctcactagtaccctcgcag
acggacagcgccaggagcaatggcacgcgcggcgaccgagtggtgtgccaatagcggctgctcagcagggcgccggag
agcagcgccgggaagggcggtgcccgaagcggtgtggcggtagtgtggcccaagcttcgtttagtgaaccgtcagat
 cgcttgagacgccatccacgtgtttgacctccatagaagacacgggacccgatccagcctccgcggttcgaatccggcg
 gaacggtgcattggaacgcggattccccgtgccaagagtgacgtaagtaccgctatagagtcataaggccacaaaaatgctttt
 cttaataatactttttgttatcttatttaatactttccctaactcttcttcttcagggcaataatgatacaatgtatcatgcctcttgaccatt
 cttaattgctaataagcagctacaatccagctaccattctgctttattttgtggtgggaaggtggtatttctgagtcgaagctagg
 ccttttctaactgtgttcatacctcttatcttctccacagctcctgggcaacgtgctggtctgtgtgctggccatcactttggcaaga
 attaccggtggcaacgtgctggtattgtgctgtctcatcattttggcaagaattcacgccccagagccgcccaccATGgcctacca
 tacgatgttccagattacgtACAGAATTACCTGCCCCCTTGAGCTACTTCCAGAATGCACAG
 ATGAGCGAGGACAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAACAGA
 GAACGGCAGGAACACAACGACAGGCGGAGCCTGGGCCACCCTGAGCCCCCTGTCT
 AATGGAAGACCCCAGGGTAACAGCAGACAGGTGGTGGGAACAAGATGAGGAAGA
 GGACGAGGAGCTGACCCTGAAGTACGGCGCCAAGCACGTGATCATGCTCTTCGT
 GCCCCGTGACTCTCTGCATGGTGGTGGTGGTGGCTACAATCAAGAGCGTCAGCTTT
 TATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGAAGACACCGAG
 ACTGTGGGCCAGAGAGCCCTGCACTCAATCCTGAATGCCGCCATCATGATCAGC
 GTCATTGTTGTCATGACTATCCTCCTGGTGGTTCTGTATAAATACAGGTGCTATAA
 GGTCATCCATGCCTGGCTGATCATATCATCTCTGTTGCTGCTGTTCTTTTTTAGCTT
 CATTACCTGGGCGAAGTGTTTAAACCTATAACGTTGCCGTGGACTACATTACT
 GTTGCCCTCCTGATCTGGAACCTCGGCGTGGTGGGCATGATTTCCATTCACTGGA
 AAGGCCCCCTGAGACTGCAGCAGGCATACCTCATTATGATCTCCGCCCTCATGGC
 CCTGGTGTTCATCAAGTACCTGCCCCAGTGGACTGCTTGGCTCATCTTGGCTGTG
 ATCTCCGTGTATGATTTAGTGGCTGTTCTGTGTCTAAAGGTCCACTGCGTATGCT

GGTGGAAACAGCTCAGGAAAGAAATGAAACACTGTTTCCTGCTCTGATTTACTCC
TCAACAATGGTGTGGCTCGTGAATATGGCCGAAGGAGACCCTGAAGCCCAACGG
AGAGTGTCCAAAACTCCAAGTATAACGCCGAGAGCACAGAAAGGGAGAGCCA
GGATACAGTTGCCGAGAATGACGATGGCGGCTTCAGTGAGGAATGGGAAGCCCA
GAGGGACAGCCACCTGGGGCCTCACAGAAGCACCCCTGAGTCTAGAGCCGCTGT
CCAGGAACCTGTCCAGCTCCATCTTGGCCGGCGAAGACCCCGAAGAAAGGGGAGT
AAAACCTGGAGCTGGGAGATTTTCATCTTCTACAGTGTTCCTGTTGGCAAAGCCAGC
GCAACAGCTAGCGGAGACTGGAACACAACAATAGCCTGTTTCGTAGCCATCTTA
ATTGGCCTGTGCCTTACACTTCTGCTCCTGGCCATCTTCAAGAAGGCCCTGCCAG
CCCTGCCTATCAGCATCACCTTCGGGCTTGTCTTCTACTTTGCCACCGATTATCTG
GTGCAGCCCTTCATGGACCAGCTGGCCTTCCACCAGTTTTACATCTAGTaaagcgccgc
cctagggagctcctcgagggggtggcatcctgtgacctccccagtgectctcctggccctggaagtggcactccagtggccacc
agcctgtgcttaataaaatgaattgcatcattttgtctgactaggtgctctctataattatgggggtggaggggggtgtatggagcaa
gggggcaaggggggaagacaacctgtagggcctgcggggctattgggaaccaagctggagtgcagtgccacaactctggctcactg
caatctccgctcctgggtcaagcgattctcctgcctcagcctcccgagttgttgggattccaggcatgcatgaccaggctcagtaatt
ttgttttttggtagagacggggttcaccatattggccaggtggtctccccctcctaattcaggtgatctaccacettggcctccaa
attgctgggattacaggcgtgaaccactgctccctcctgctcttctgattttaaataactataaccagcaggaggacgtccagacaca
gcataggctacctggccatgcccacccgggtgggacattgagttgctgcttggcactgtcctctcatgcttgggtccactcagtagat
gcctgttgaaattcctgggctagggctgtgccagctgctcgtcccgacattctggcttctctcctccatattcttagctgtttctc
atgagaatgttccaaattcgaaattctatttaaccattatatttactgttggctattatctctgccccagtagattgttagctccagaaga
gaaaggatcatgtctttgttatctagatatgcccatctgctggtacaatctctggcacatgttacaggcaacaactactgtggaattgg
tgaatgcatgaatagaagaatgagtgaaatgaatgaatagacaataggcagaaatccagcctcaaagagcttacagctgtgtaaggga
ataaaatgtctgcaaatagccacaggacagggtcaaaaggagggggctatftccagctgagggcaccccatcaggaaagcacccc
agacttctacaaactactagacacatctcgatgctttcacttctctatcaatggatcgtctccctggagaataatccccaaagtgaattac
ttagcacgtccagtttagtagatccttgtgtacttctgtgttctagagatcatcaaccagtgcacaaatcccccatcaatacacagca
gtgcttgcctctcctccccgaggtcttcgagggccttctcctgctgaacccctggacatatcatatggcaactgaagtgtcc
aacgagatataggaagtgaacacgatgtacactgaacgtgcaatacaaatatgcagcatgaagtgcctcggttactaacccgagc
tacgctgggtgcttcttttaccactttcctaatgctatggacacctcattctgtgctgaagtcttctgttcaattcccccatcttcat
gaacatcctgtgtaggacctcaccctgtcctgtagctttgactgaggcaagttctgtccatgcctagtagtgcaccaccttacta
gatgagacttctaaagagcttgcatggaaggaaagccccggggccttggaaagccatcacttagaaactgggagagctccaggcaa
gccacctcatcttagggataacagggtatggcgcgggcgccaggaaacccctagtgatggagtggccactcctctctgcgcgctc
gctcgctcactgagggcgggcgaccaaaggtcgcccgacgcccgggcttggccggcgccctcagtgagcgagcgagcgcgca
ctgcctgcagg

SEQ ID NO: 46

Expression cassette for AAV production containing AAV2 ITRs (1-141, 4554-4694), *synapsin 1* promoter (underlined, 237-684), a synthetic human beta-globin intron (705-1270), HA-tag (1350-1376), codon-optimized human presenilin 1v1.5 (uppercase SEQ ID NO: 37, 1344-2777), human growth hormone polyadenylation (2808-3284) and human growth hormone genomic sequence (3285-4523).

cctgcaggcagctgcgcgctcgtcgtcactgagggcgccgggcaaagccggcgctgggagacctttgtgccccggcctc
agtgcagcgagcgagcgcgagagaggagtgcccaactccatcactaggggtcctgcggccaattcagtcgataactataacggt
cctaaggtagcgatttaataacgcgctctcttaaggtagccccgggacgcgtcaattgagatctcagtgcaagtgggttttagaccag
gatgagggcggtgggggtgcctacctgacgacccgacccactggacaagcaccacccccattccccaaattgcgcat
cccctatcagagaggggggaggggaaacaggatgcggcgaggcgctgcgactgcagcttcagcaccgagacagtgcttgc
ccccgcctggcgcgcgccaccgccctcagcactgaaggcgcgctgacgtcactgcgggtccccgcacaaactccccctt
cgggccacttggctgcgtccgcgccgccggcgccagccggaccgcaccacgcagggcgcgagataggggggacagggcg
cgaccatctgcgctcgcgccggcgactcagcgctgcctcagctcgggtggcgagcgaggaagtcgtctgcctcagagcg
cagaagcttctgttagtgaaaccgtcagatgcctggagacgcatccacgctgtttgacctccatagaagacacggggacgcgatca
gctccgcggttcgaatccggccgggaacggtgcattggaacgggattccccgtgccaagagtgcagtaagtaaccgcctatag

agctctatagggccacaaaaatgctttcttctttaataactttttgtttatcttatttctaatactttccctaatactttctttcagggcaataat
gatacaatgtatcatgacctctttgaccattctaaagaataacagtataatttctgggttaaggcaatagcaatatttctgcatataaatattt
ctgcatataaattgtaactgatgtaagagggtttcatattgctaatagcagctacaatccagctaccattctgctttattttgtggtgggataa
ggctggattattctgagccaagctaggcccttttgtaatcgtgttcatacctcttatcttctccacagctcctgggcaacgtgctggtc
tgtgtgctggcccatcactttggcaagaattaccggtggcaacgtgctggtattgtgctgtctcatcattttggcaagaattcacgcc
cagagccgccaccATGgcctaccatacagatgtccagattacgtACAGAATTACCTGCCCCCTTGAGCT
ACTTCCAGAATTGCACAGATGAGCGAGGACAACCACCTGAGCAATACTGTACGTA
GCCAGAATTGACAACAGAGAACGGCAGGAACACAACGACAGGCGGAGCCTGGGC
CACCTGAGCCCCCTGTCTAATGGAAGACCCCAGGGTAACAGCAGACAGGTGGTG
GAACAAGATGAGGAAGAGGACGAGGAGCTGACCCTGAAGTACGGCGCCAAGCA
CGTGATCATGCTCTTCGTGCCCCGTGACTCTCTGCATGGTGGTGGTGGTGGCTACA
ATCAAGAGCGTCAGCTTTTATACCCGGAAGGATGGGCAGCTAATCTATACCCAT
TCACAGAAGACACCGAGACTGTGGGCCAGAGAGCCCTGCACTCAATCCTGAATG
CCGCCATCATGATCAGCGTCATTGTTGTCTGACTATCCTCCTGGTGGTTCTGTAT
AAATACAGGTGCTATAAGGTCATCCATGCCTGGCTGATCATATCATCTCTGTTGC
TGCTGTTCTTTTTTAGCTTCATTTACCTGGGCGAAGTGTTTAAACCTATAACGTT
GCCGTGGACTACATTACTGTTGCCCTCCTGATCTGGAACCTCGGCGTGGTGGGCA
TGATTTCCATTCACTGGAAGGCCCCCTGAGACTGCAGCAGGCATACCTCATTAT
GATCTCCGCCCTCATGGCCCTGGTGTTCATCAAGTACCTGCCCAGTGGACTGCT
TGGCTCATCTTGGCTGTGATCTCCGTGTATGATTTAGTGGCTGTTCTGTGTCCTAA
AGGTCCACTGCGTATGCTGGTGGAAACAGCTCAGGAAAGAAATGAAACACTGTT
TCCTGCTCTGATTTACTCCTCAACAATGGTGTGGCTCGTGAATATGGCCGAAGGA
GACCCTGAAGCCCAACGGAGAGTGTCCAAAACTCCAAGTATAACGCCGAGAGC
ACAGAAAGGGAGAGCCAGGATACAGTTGCCGAGAATGACGATGGCGGCTTCAGT
GAGGAATGGGAAGCCAGAGGGACAGCCACCTGGGGCCTCACAGAAGCACCCC
TGAGTCTAGAGCCGCTGTCCAGGAAGTGTCCAGCTCCATCCTGGCCGGCGAAGA
CCCCGAAGAAAGGGGAGTAAACTTTGGACTGGGAGATTTTCATCTTCTACAGTGT
TCTCGTTGGCAAAGCCAGCGCAACAGCTAGCGGAGACTGGAACACAACAATAGC
CTGTTTCGTAGCCATCTTAATTGGCCTGTGCCTTACACTTCTGCTCCTGGCCATCT
TCAAGAAGGCCCTGCCAGCCCTGCCTATCAGCATCACCTTCGGGCTTGTTTTCTA
CTTTGCCACCGATTATCTGGTGCAGCCCTTCATGGACCAGCTGGCCCTTCCACCAG
TTTTACATCTAGtaagcggccgacctaggagctcctcagggggtggcatccctgtgacccctccccagtgccctcctg
gccctggaagtggccactccagtgcccaccagcctgtgctaataaaattagttgcatcattttgtctgactagggtgctcttataata
tgggggtggagggggtggtatggagcaaggggcaagggggaagacaacctgtagggcctgcgggtctattgggaaccaagct
ggagtgcagtgggcacaactctggctcactgcaatctcgccctcctgggttcaagcgattctcctgcctcagcctcccgagtggtgggatt
ccaggcatgcatgaccaggctcagtaattttgtttttggtagagacggggttcaccatattggccaggctggtctccccctcta
ctcagggtgatctaccaccttggcctccaaattgctgggattacaggcgtgaaccactgctccctcctgctcttctgattttaaataa
ctataccagcaggaggacgtccagacacagcataggctacctggccatgcccaccgggtgggacatttgagttgcttcttggcactg
tcctctcatgcttgggtccactcagtagatgctgttgaaattcctgggcctagggtgtgccagctgcctcgtcccgctacacttctggtt
cttctcctccatacttagctgttttctcatgagaatgttccaaattcgaaatttctatlaaccattatatttacttgttcttattatctt
gccccagtagattgttagctccagaagagaaaggatcatgtctttgcttatctagatatgccatctgcctggtacaatctctggcacat
gttacaggcaacaactacttgtggaattggtgaatgcatgaatagaagaatgagtgaaatgaatagaacaataggcagaaatccag
cctcaaaagagcttacagctgtgtaagaggaataaaatgtctgcaaatagccacaggacaggtcaaagggaaggagggtctattccag
ctgagggcaccatcaggaaagcaccacagacttctacaactactagacacatctcgatgcttttacttctatcaatgatcgtct
ccctggagaataatccccaaagtgaattacttagcacgtccagttaggtagatccttgtgacttctggttgttcagagatcatcaacca
gtgcaaaatcccccatcaatacacagcagtgccctgccccctcccccgagggtctccgaggcccttccctgctgctgaacccc
ctggacatatcatatggcaactgaagtgtccaacagatataggaagtgaacacgatgtacactgaacgtgcaatacaaatatgca
gcatgaagtgcctcggttactaacccagctacgctgggtgcttcttttaccactttcctaatgcctatggacacctcattctgtggt
gaagtgcctgtgttaattcccccatcttcaatgaacatcctgtgtagggacctcaccctgtcctgctagcttgcactgaggcaagttc
tgtccatgcttagtgccaccactttactagatgagacttcaagagcttggcatggaaggaaagcccgggggccttgaagcc
atcacttagaaactgggagagctccaggcaagccacctcatcttagggataacagggtaatggcggggcccaggaacccctagt

gatggaggttggccactccctctctgcgcgctcgtcgtcactgaggccgggcgaccaaaggtcgcccgaccccggttggcc
ggggcgccctcagtgagcgagcgagcgcgagctgcctgcagg

SEQ ID NO: 47

Expression cassette for production of a self-complementary AAV containing native and modified AAV2 ITRs (1-105, 2386-2526), CBA promoter (underlined, 113-766), minute virus of mice intron (776-867), HA-tag (884-910), codon-optimized human presenilin 1v1.5 (uppercase SEQ ID NO: 37, 881-2311), rabbit beta-globin polyadenylation sequence (2319-2367).

ggcgcgctcgctgctcactgagccgcgcggcgcaagccgcggcgctggcgccgaccttggcgcccgccctcagtgagcgagc
gagcgcgagagaggagtgagatctccgcacattgattattgactatttaataagtaataaattacggggctcattagtctatagcccat
atatggaggttccgcggttacataacttacggtaaatggcccgccctggctgaccgcccacgacccccggccattgacgtcaataatgac
gtatgttccccatagtaacgccaatagggactttccattgacgtcaatgggtggagtaftttacgggtaaactccccacttggcagttacatcaa
gtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcattatgccagttacatgaccttatgggac
tttctacttggcagttacatctacgtattatgcatcgctattaccatgtcgcaggccacgttctgttctacttccccatctccccccccctcccc
accccccaattttgtattttatttttttaattatttttgcagcgatggggcgggggggggggggcgcgccgagcgggggcgggcg
ggcgagggggcgggcgggcgagcgcgagaggtgcgcgcgcagccaatcagagcgcgcgctccgaaagtttctctttatgg
cgagcgcgcgcgcgcgccctataaaaaagcgaagcgcgcgcgggcggggagcaagcttcgtaagaggtgaagggttaagg
gatggtgtgtgtgggtattaatgtttaattacctgttttacaggcctgaaatcactgtgttttaggttggtgtagccgccaccATGta
cccatacagatgttccagattacgctACAGAATTACCTGCCCCCTTGAGCTACTTCCAGAATGCAC
AGATGAGCGAGGACAACCACCTGAGCAATACTGTACGTAGCCAGAATGACAACA
GAGAACGGCAGGAACACAACGACAGGCGGAGCCTGGGCCACCCTGAGCCCCCTGT
CTAATGGAAGACCCCAGGGTAACAGCAGACAGGTGGTGGAACAAGATGAGGAA
GAGGACGAGGAGCTGACCCTGAAGTACGGCGCCAAGCACGTGATCATGCTCTTC
GTGCCCCGTGACTCTCTGCTATGGTGGTGGTGGTGGCTACAATCAAGAGCGTCAGCT
TTTATACCCGGAAGGATGGGCAGCTAATCTATACCCCATTCACAGAAGACACCG
AGACTGTGGGCCAGAGAGCCCTGCACTCAATCCTGAATGCCGCCATCATGATCA
GCGTCATTGTTGTTCATGACTATCCTCCTGGTGGTTCGTGTATAAATACAGGTGCTAT
AAGGTCATCCATGCCTGGCTGATCATATCATCTCTGTTGCTGCTGTTCTTTTTTAG
CTTCATTTACCTGGGCGAAGTGTTTAAACCTATAACGTTGCCGTGGACTACATT
ACTGTTGCCCTCCTGATCTGGAACCTTCGGCGTGGTGGGCATGATTTCATTCACT
GGAAAGGCCCCCTGAGACTGCGAGCAGGCATACCTCATTATGATCTCCGCCCTCAT
GGCCCTGGTGTTCATCAAGTACCTGCCCGAGTGGAAGTCTTGCTCATCTTGGCT
GTGATCTCCGTGTATGATTTAGTGGCTGTTCTGTGTCCTAAAGGTCCACTGCGTAT
GCTGGTGGAACAGCTCAGGAAAGAAATGAAACACTGTTTCTGCTCTGATTAC
TCCTCAACAATGGTGTGGCTCGTGAATATGGCCGAAGGAGACCCTGAAGCCCCA
CGGAGAGTGTCCAAAACTCCAAGTATAACGCCGAGAGCACAGAAAGGGAGAG
CCAGGATACAGTTGCCGAGAAATGACGATGGCGGCTTCACTGAGGAATGGGAAGC
CCAGAGGGGACAGCCACCTGGGGCTCACAGAAGCACCCCTGAGCTAGAGCCGC
TGTCCAGGAACACTGTCCAGCTCCATCCTGGCCGGCGAAGACCCCGAAGAAAGGGG
AGTAAAACTTGACTGAGGAGATTTTCATCTTCTACAGTGTTCTCGTTGGCAAAGCC
AGCGCAACAGCTAGCGGAGACTGGAACACAACAATAGCCTGTTTCGTAGCCATC
TTAATTGGCCTGTGCCCTACACTTCTGCTCCTGGCCATCTTCAAGAAGGCCCTGCC
AGCCCTGCCTATCAGCATACCTTTCGGGCTTGTTTTCTACTTTGCCACCGATTATC
TGGTGCAGCCCTTCATGGACCAGCTGGCCTTCCACCAGTTTTTACATCTAGcgcgccga
ataaaagatctttattttcattagatctgtgtgtgtgtttttgtgtgtagggataaacagggttaataggaacccctagtgtgaggttggccac
tcctctctgtcgcgctcgtcgtcactgaggccggcgaccaaaggtgcgccgacgcccgggctftggccgggcggcctcagtga
cgagcgagcgcgagctgcctgcagg

[0190] Any and all references and citations to other documents, such as patents, patent applications, patent publications, journals, books, papers, web contents, that have been made throughout this disclosure are hereby incorporated herein by reference in their entirety for all purposes.

[0191] Although the present invention has been described with reference to specific details of certain embodiments thereof in the above examples, it will be understood that modifications and variations are encompassed within the spirit and scope of the invention. Accordingly, the invention is limited only by the following claims.