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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 August 2022 (11.08.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/169884 A2

(51) International Patent Classification:
A01K 67/027 (2006.01)

(21) International Application Number:
PCT/US2022/014956

(22) International Filing Date:
02 February 2022 (02.02.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/144,743 02 February 2021 (02.02.2021) US

(71) Applicant: **ALLEN INSTITUTE** [US/US]; 615 Westlake Avenue North, Seattle, Washington 98109 (US).

(72) Inventors: **GORE, Bryan**; 615 Westlake Avenue North, Seattle, Washington 98109 (US). **LEIN, Edward Sebastian**; 615 Westlake Avenue North, Seattle, Washington 98109 (US). **LEVI, Boaz P.**; 615 Westlake Avenue North, Seattle, Washington 98109 (US). **MACHEN, Deja**; 615 Westlake Avenue North, Seattle, Washington 98109 (US). **MARTINEZ, Refugio**; 615 Westlake Avenue North, Seattle, Washington 98109 (US). **MICH, John K.**; 615 Westlake Avenue North, Seattle, Washington 98109 (US). **TING, Jonathan**; 615 Westlake Avenue North, Seattle, Washington 98109 (US).

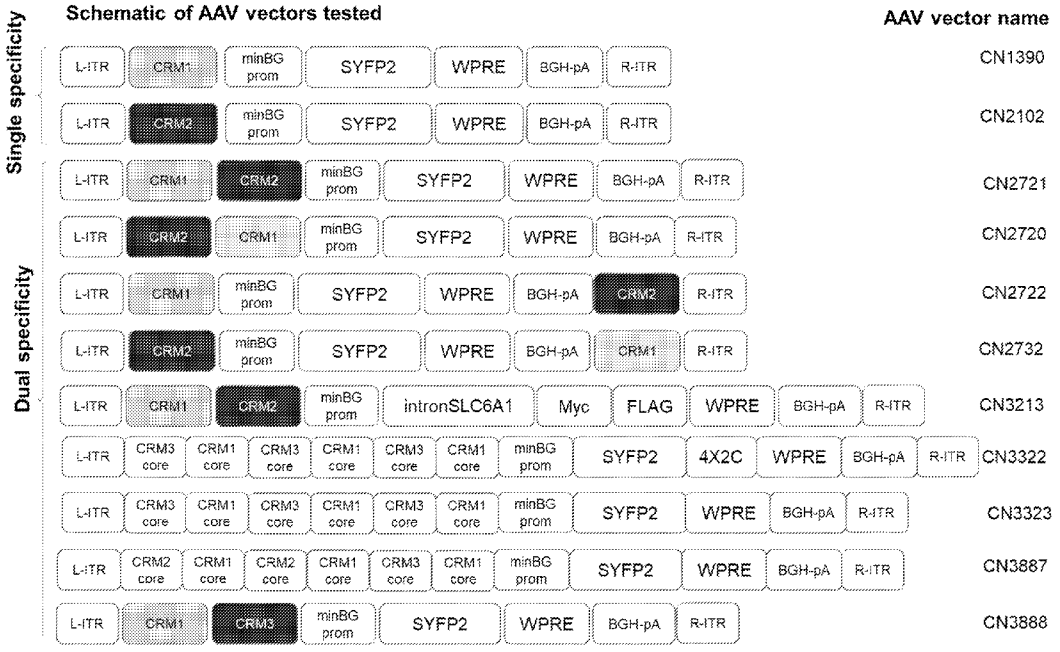
(74) Agent: **WINGER, C. Rachal**; Lee & Hayes PC, 601 W. Riverside Avenue Suite 1400, Spokane, Washington 99201 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

(54) Title: ARTIFICIAL EXPRESSION CONSTRUCTS FOR MODULATING GENE EXPRESSION IN GABAERGIC NEURONS AND ASTROCYTES

FIG. 1A



(57) Abstract: Artificial expression constructs for modulating gene expression in GABAergic neurons and astrocytes are described. The artificial expression constructs can be used to express SLC6A1 for the treatment of SLC6A1-associated disorders, among other uses.

[Continued on next page]

WO 2022/169884 A2

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

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

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

ARTIFICIAL EXPRESSION CONSTRUCTS FOR MODULATING GENE EXPRESSION IN GABAERGIC NEURONS AND ASTROCYTES

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to U.S. Provisional Patent Application No. 63/144,743 filed on February 2, 2021, which is incorporated herein by reference in its entirety as if fully set forth herein.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0002] This invention was made with government support under MH114126 awarded by the National Institutes of Health. The government has certain rights in the invention.

REFERENCE TO SEQUENCE LISTING

[0003] The Sequence Listing associated with this application is provided in text format in lieu of a paper copy and is hereby incorporated by reference into the specification. The name of the text file containing the Sequence Listing is A166-0029PCT_ST25.txt. The text file is 181 KB, was created on February 1, 2022, and is being submitted electronically via EFS-Web.

FIELD OF THE DISCLOSURE

[0004] The current disclosure provides artificial expression constructs for modulating gene expression in GABAergic neurons and astrocytes. The gene to be expressed can include SLC6A1 to treat SLC6A1-associated disorders, among other uses.

BACKGROUND OF THE DISCLOSURE

[0005] γ-Aminobutyric acid (GABA), an inhibitory neurotransmitter, is released from GABAergic neurons. GABA does not undergo enzymatic breakdown but is instead removed from the extracellular space through the action of GABA transporters. GABA transporters are expressed in different cell types, including inhibitory neurons and astrocytes, and belong to the solute carrier 6 (SLC6) family. The 6 types of GABA transporters include: A1/GAT1, A13/GAT2, A11/GAT3, A6/TauT, A8/CT1, and A12/BGT1 (Scimemi, Front Cell Neurosci. 2014; 8: 161).

[0006] The A1/GAT1 protein is encoded by the solute carrier family 6 member 1 (SLC6A1) gene. Gene mutations of SLC6A1 are characterized by mild-to-moderate intellectual disability, epilepsy, speech difficulties, behavioral problems (e.g. hyperactivity, attention deficit, aggressiveness, and autistic traits), and neurological signs (e.g. ataxia, hypotonia, tremor, and fine-motor impairment (Carvill, et al. Am J Hum Genet. 2015; 96:808–15; and Johannesen, et al. Epilepsia. 2018;

59:389–402).

SUMMARY OF THE DISCLOSURE

[0007] The current disclosure provides artificial expression constructs that drive gene expression in GABAergic neurons and astrocytes. The artificial expression constructs can be used to drive SLC6A1 gene expression to ameliorate disorders associated with SLC6A1 gene mutations, among other uses.

[0008] The artificial expression constructs include enhancer elements which drive gene expression in GABAergic neurons by including the I56i enhancer or a core thereof and drive expression in astrocytes by including one or more enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, or a core thereof (e.g., eHGT_387m(core2) or eHGT_390m(core2)).

[0009] In particular embodiments, the artificial enhancer elements include a concatenated core of an enhancer. Examples include a concatenated core of I56i. These artificial enhancer elements can provide higher levels and more rapid onset of transgene expression compared to a single full length original (native) enhancer.

[0010] In particular embodiments, the core of I56i (or I56i(core)) includes the sequence as set forth in any one of SEQ ID NOs: 4 and 5. In particular embodiments, these cores are concatenated and have 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core sequence. SEQ ID NOs: 6 and 7 provide three-copy concatemers of the selected enhancer cores.

[0011] In particular embodiments, the artificial expression constructs include a three-copy concatemer of the core of hI56i and a second enhancer selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, and eHGT_390m.

[0012] In particular embodiments, artificial enhancer elements include a combination concatenated enhancer. In particular embodiments, the combination concatenated enhancer includes a core of the enhancer selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m (e.g., eHGT_387m(core2)), and eHGT_390m (e.g., eHGT_390m(core2)) concatenated with the I56i(core). In particular embodiments, the core of eHGT_387m (eHGT_387m(core2)) includes the sequence as set forth in SEQ ID NO: 84. In particular embodiments, the core of eHGT_390m (eHGT_390m(core2)) includes the sequence as set forth in SEQ ID NO: 85.

[0013] In particular embodiments, a combination concatenated enhancer includes eHGT_387m(core2) and I56i(core) as set forth in SEQ ID NO: 88 (eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)). In particular

embodiments, a combination concatenated enhancer includes eHGT_390m(core2) and I56i(core) as set forth in SEQ ID NO: 86. In particular embodiments, the combination concatenated enhancer is concatenated to include 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the combination concatenated enhancer. In particular embodiments, SEQ ID NO: 89 provides a three-copy-concatemer of the eHGT_390m(core2)-I56i(core) combination concatenated enhancer.

BRIEF DESCRIPTION OF THE FIGURES

[0014] Some of the drawings submitted herein may be better understood in color. Applicant considers the color versions of the drawings as part of the original submission and reserves the right to present color images of the drawings in later proceedings.

[0015] FIGs. 1A, 1B. (FIG. 1A) Designs to generate two specificities of expression per AAV vector. Schematic design of single and dual specificity AAVs tested. (FIG. 1B) Additional exemplary schematics.

[0016] FIG. 2. Brain-wide expression of dual specificity vectors. Fluorescent images of sagittal sections of mouse brains transduced by the indicated viruses after intravenous delivery and packaged with PHP.eB capsid. White is SYFP expression. Images are montages.

[0017] FIG. 3. Visual cortex expression of dual specificity vectors. Fluorescent images of sagittal sections of mouse visual cortex (VISp) transduced by the indicated viruses after intravenous delivery and packaged with PHP.eB capsid. White is SYFP expression. Images are montages.

[0018] FIG. 4. Striatal expression of dual specificity vectors. Fluorescent images of sagittal sections of mouse striatum transduced by the indicated viruses after intravenous delivery and packaged with PHP.eB capsid. White is SYFP expression. Images are montages.

[0019] FIG. 5. Cerebellar expression of dual specificity vectors. Fluorescent images of sagittal sections of mouse cerebellum transduced by the indicated viruses after intravenous delivery and packaged with PHP.eB capsid. White is SYFP expression. Images are montages.

[0020] FIG. 6. Quantification of $SYFP^+$ cells transduced by CN1390 vector. Mouse visual cortex (VISp) transduced by CN1390 virus after intravenous delivery and packaged with PHP.eB capsid. Overlap of SYFP fluorescence GABAergic cell marker (Gad1) and astrocyte marker (Fgfr3) mRNA image by mFISH. Co-localization shown with different colored circles, and quantified. Images are montages.

[0021] FIG. 7. Quantification of $SYFP^+$ cells transduced by CN2102 vector. Mouse visual cortex (VISp) transduced by CN2102 virus after intravenous delivery and packaged with PHP.eB capsid. Overlap of SYFP fluorescence GABAergic cell marker (Gad1) and astrocyte marker (Fgfr3) mRNA image by mFISH. Co-localization shown by different colored circles, and quantified.

Images are montages.

[0022] FIG. 8. Quantification of $SYFP^+$ cells transduced by CN2102+CN1390 vectors. Mouse visual cortex (VISp) transduced by CN1390 and CN2102 viruses after intravenous delivery and packaged with PHP.eB capsid. Overlap of SYFP fluorescence GABAergic cell marker (Gad1) and astrocyte marker (Fgfr3) mRNA image by mFISH. Co-localization shown by different colored circles, and quantified. Images are montages.

[0023] FIG. 9. Quantification of $SYFP^+$ cells transduced by CN2721 vector. Mouse visual cortex (VISp) transduced by CN2721 virus after intravenous delivery and packaged with PHP.eB capsid. Overlap of SYFP fluorescence GABAergic cell marker (Gad1) and astrocyte marker (Fgfr3) mRNA image by mFISH. Co-localization shown by different colored circles, and quantified. Images are montages.

[0024] FIGs. 10A-10C. Codon optimization of SLC6A1. Characterization of SLC6A1 expression vectors with different codon optimization strategies, with and without an intron. Expression of SLC6A1 is shown by Western blot analysis (FIG. 10A) and input control is shown by staining for tubulin (FIG. 10B). (FIG. 10C) Quantification of SYFP expression normalized by loading control. In this experiment, HEK293 cells were transfected with 1µg DNA in triplicate in a 12-well plate for 96 hours. Cells were then lysed in RIPA. Coding sequences were under the control of a CMV promoter as follows: (1) No DNA, (2) CN2972: hSLC6A1_myc_ddk_native_CN2522GeneOpt1Splice (SEQ ID NO: 24), (3) CN2975: hSLC6A1_myc_ddk_native_Intron (SEQ ID NO: 33), (4) CN2976: hSLC6A1_myc_ddk_native_CN2522GeneOpt1Splice_Intron (SEQ ID NO: 36), (5) CN2974: hSLC6A1_myc_ddk_native_IDTCodonOptSplice1 (SEQ ID NO: 30), and (6) CN2973: hSLC6A1_myc_ddk_native_IDTCodonOptSplice1_Intron (SEQ ID NO: 27).

[0025] FIGs. 11A, 11B. (FIG. 11A) Epifluorescence micrograph image (inverted) showing native SYFP2 expression in mouse brain sagittal section 22 days after retro-orbital delivery of 1.0E12 viral genome copies of AAV vector #CN3323. Scale bar: 1 mm. (FIG. 11B) Higher magnification view of the thalamus region.

[0026] FIGs. 12A-12C. A mouse was injected by the intracerebroventricular (ICV) route on postnatal day 2 with 1E11vg of PHP.eB packaged CN3213 (pAAV-eHGT_3xhI56i(core)_eHGT_387m-minBG-intronSLC6A1-myc-flag-WPRE3-BGHpA) and was sacrificed on at postnatal day 21. (FIG. 12A) Sagittal section showing brain-wide expression of CN3213-expressed myc-tagged and codon-optimized human SLC6A1. (FIG. 12B and FIG. 12C) Magnification of cerebral cortex showing co-expression in astrocytes (arrowhead) and GABA-positive interneurons (arrow). Much of the of the SLC6A1 expressed in GABAergic cells is

trafficked into the dendrites.

[0027] FIGs. 13A-13C. Myc-tagged hSLC6A1 is trafficked into the GABAergic dendrites and appears as puncta throughout the neuropil. P2 aged neonatal animals were ICV injected with1E11vg PHP.eB serotype AAVs expressing myc-tagged hSLC6A1 with and astrocyte-only enhancer (FIG. 13A), a GABAergic inhibitory cell only enhancer (FIG. 13B), or a dual specificity enhancer pair (FIG. 13C). Brains were isolated at postnatal day 21, sectioned and stained for myc-tagged hSLC6A1 (white). Magnification, exposure times, and post-acquisition image adjustments are identical for each vector. Bottom row shows higher magnification of dendritic staining in the neuropil. Although astrocytes are clearly expressing myc-hSLC6A1 in (FIG. 13A) and (FIG. 13C) (arrowheads), and GABAergic cell bodies are only obvious in (FIG. 13B) (arrows), GABAergic cells are clearly expressing myc-hSLC6A1 in (FIG. 13C) since the dendrites are labeled throughout the neuropil as in (FIG. 13B) (asterisks). This staining is not seen in (FIG. 13A) where only astrocytes express myc-hSLC6A1.

[0028] FIG. 14. Sequences supporting the disclosure. Sequences include: hI56i – full length human hI56i enhancer (SEQ ID NO: 1), Murine I56i Enhancer (core is the same as human) (SEQ ID NO: 2), Zebrafish I46i Enhancer (SEQ ID NO: 3), hI56i core - human hI56i enhancer core (SEQ ID NO: 4), Core of the Zebrafish I46i Enhancer (SEQ ID NO: 5), 3xhI56i(core) (SEQ ID NO: 6), 3x Concatamerized Core of the Zebrafish I46i Enhancer (SEQ ID NO: 7), eHGT_375h (SEQ ID NO: 8), eHGT_376h (SEQ ID NO: 9), eHGT_390h (SEQ ID NO: 10), eHGT_373m (SEQ ID NO: 11), eHGT_375m (SEQ ID NO: 12), eHGT_386m (SEQ ID NO: 13), eHGT_387m (SEQ ID NO: 14), eHGT_387m(core2) (SEQ ID NO: 84), eHGT_390m (SEQ ID NO: 15), eHGT_390m(core2) (SEQ ID NO: 85), Combination Concatenated Enhancer (eHGT_387m(core2)-hI56i(core)) (SEQ ID NO: 95), Combination Concatenated Enhancer (eHGT_390m(core2)-hI56i(core)) (SEQ ID NO: 86), 3xhI56i(core)_eHGT_390m Enhancer (SEQ ID NO: 87), 3X Combination Concatenated Enhancer (eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)) (SEQ ID NO: 88), 3x Combination Concatenated Enhancer (eHGT_390m(core2)-hI56i(core)- eHGT_390m(core2)-hI56i(core)- eHGT_390m(core2)-hI56i(core)) (SEQ ID NO: 89), Beta-Globin Minimal Promoter (SEQ ID NO: 16), minCMV Promoter (SEQ ID NO: 17), Mutated minCMV Promoter (SEQ ID NO: 18), minRho Promoter (SEQ ID NO: 19), minRho* Promoter (SEQ ID NO: 20), Hsp68 minimal Promoter (proHsp68) (SEQ ID NO: 21), SLC6A1 encoding sequence from CN2972 (SEQ ID NO: 22), Myc-DDK tag sequence in CN2972 (SEQ ID NO: 23), CN2972 (SEQ ID NO: 24), SLC6A1 encoding sequence from CN2973 (SEQ ID NO: 25), Myc-DDK tag sequence in CN2973 (SEQ ID NO: 26), CN2973 (SEQ ID NO: 27), SLC6A1 encoding sequence from CN2974 (SEQ ID NO: 28), Myc-DDK tag sequence in CN2974

(SEQ ID NO: 26), CN2974 (SEQ ID NO: 30), SLC6A1 encoding sequence from CN2975 (SEQ ID NO: 31), Myc-DDK tag sequence in CN2975 (SEQ ID NO: 32), CN2975 (SEQ ID NO: 33), SLC6A1 encoding sequence from CN2976 (SEQ ID NO: 34), Myc-DDK tag sequence in CN2976 (SEQ ID NO: 23), CN2976 (SEQ ID NO: 36), CN2478 (SEQ ID NO: 37), SLC6A1 transcript variant (SEQ ID NO: 38), SLC6A1 transcript variant 2 (SEQ ID NO: 39), SLC6A1 transcript variant X5 (SEQ ID NO: 40), SLC6A1 transcript variant X1 (SEQ ID NO: 41), SLC6A1 transcript variant X2 (SEQ ID NO: 42), SLC6A1 transcript variant X6 (SEQ ID NO: 43), SLC6A1 transcript variant X3 (SEQ ID NO: 44), SLC6A1 transcript variant X4 (XM_017007072.2) (SEQ ID NO: 45), Sodium- and chloride-dependent GABA transporter 1 isoform a (SEQ ID NO: 46), Myc-DDK tag (SEQ ID NO: 47), SYFP2 (SEQ ID NO: 48), EGFP (SEQ ID NO: 49), Optimized Flp recombinase (FlpO) (SEQ ID NO: 50), Improved Cre recombinase (iCre) (SEQ ID NO: 51), tet-Transactivator version 2 (tTA2) (SEQ ID NO: 52), GCaMP6m (SEQ ID NO: 53), GCaMP6s (SEQ ID NO: 54), GCaMP6f (SEQ ID NO: 55), SP10 insulator (SP10ins) (SEQ ID NO: 56), 3xSP10ins (SEQ ID NO: 57), WPRE3 (SEQ ID NO: 58), WPRE (SEQ ID NO: 59), BGHpA (SEQ ID NO: 60), HGHpA (SEQ ID NO: 61), P2A (SEQ ID NO: 62), T2A (SEQ ID NO: 63), E2A (SEQ ID NO: 64), F2A (SEQ ID NO: 65), Exemplary Plasmid Backbone 1 – Left ITR (SEQ ID NO: 66), Exemplary Plasmid Backbone 1 – Right ITR (SEQ ID NO: 67), Exemplary Plasmid Backbone 2 – Left ITR (SEQ ID NO: 68), Exemplary Plasmid Backbone 2 – Right ITR (SEQ ID NO: 69), PHP.eB capsid (SEQ ID NO: 70), AAV9 VP1 capsid protein (SEQ ID NO: 71), CN1390 (SEQ ID NO: 72), CN2102 (SEQ ID NO: 73), CN2720 (SEQ ID NO: 74), CN2721 (SEQ ID NO: 75), CN2722 (SEQ ID NO: 76), CN2732 (SEQ ID NO: 77), CN3213 (SEQ ID NO: 90), CN3322 (SEQ ID NO: 91), CN3323 (SEQ ID NO: 92), CN3887 (SEQ ID NO: 93), and CN3888 (SEQ ID NO: 94).

DETAILED DESCRIPTION

[0029] The solute carrier 6 (SLC6) family of proteins includes transporters for neurotransmitters, amino acids, osmolytes, and energy metabolites. These proteins play an important role in neurotransmission and homeostasis.

[0030] γ-Aminobutyric acid (GABA), an inhibitory neurotransmitter released from GABAergic neurons, does not undergo enzymatic breakdown, and instead is transported back into cells following release through the action of GABA transporters. GABA transporters are expressed in different cell types, including inhibitory neurons and astrocytes. The 6 types of GABA transporters include: A1/GAT1, A13/GAT2, A11/GAT3, A6/TauT, A8/CT1, and A12/BGT1 (Scimemi, Front Cell Neurosci. 2014; 8: 161).

[0031] A1/GAT1 is expressed in GABAergic axon terminals and also present in astrocytes,

oligodendrocytes and microglia (Fattorini, et al., Glia 2017; 65:514-22). By moving sodium and chloride ions across the membrane in a fixed ratio with GABA, GAT1 generates a stoichiometric current (Lester, et al., Annu Rev Pharmacol Toxicol 1994; 34: 219-49) and forces the intracellular translocation of extracellular GABA within millliseconds of its release. Because GABA is removed so quickly, it is prevented from activating neighboring synapses (Isaacson, et al., Neuron 1993; 10: 165-75).

[0032] The A1/GAT1 protein (referred to hereafter as GAT1) is encoded by the solute carrier family 6 member 1 (SLC6A1) gene. The SLC6A1 gene on human chromosome 4 is also referred to as GAT1, GABATR, and GABATHG. The SLC6A1 gene has a nucleic acid sequence including sequences set forth in Accession NOs.: NM_003042.4 (SEQ ID NO: 38), NM_001348250.2 (SEQ ID NO: 39), XM 011534027.3 (SEQ ID NO: 40), XM 011534025.3 (SEQ ID NO: 41), XM 005265410.5 (SEQ ID NO: 42), XM 005265411.5 (SEQ ID NO: 43), SM_0170070071.2 (SEQ ID NO: 44) and XM_017007072.2 (SEQ ID NO: 45). SLC6A1 sequences, including codon optimized variants thereof are also provided as SEQ ID NOs: 22, 25, 28, 31, 34 within FIG. 14.

[0033] Gene mutations of SLC6A1 are characterized by a mild-to-moderate intellectual disability, epilepsy, speech difficulties, behavioral problems (e.g. hyperactivity, attention deficit, aggressiveness, and autistic traits), and neurological signs (e.g. ataxia, hypotonia, tremor, and fine-motor impairment (Carvill, et al. Am J Hum Genet. 2015; 96:808–15; and Johannesen, et al. Epilepsia. 2018; 59:389–402). For example, in a study of individuals with SLC6A1 mutations, most of which lead to GAT1 loss-of-function, the most common clinical features included: epilepsy (92/101, 91.1%), developmental delay and cognitive impairment (46/56, 82.1%) and autistic traits (20/92, 22.8%) (Goodspeed, et al., Brain Communications 2020; 2(2): fcaa170). Before this study, in 2015, pathogenic *SLC6A1* mutations were identified in 4% of individuals with a previously undiagnosed early-onset epilepsy with myoclonic atonic seizures and a 3p microdeletion in SLC6A1 and SLC6A11 was described in a patient with Doose Syndrome (Carvill, et al., Am J Hum Genet 2015; 96: 808-15). Additional studies identified autism spectrum disorder and developmental epileptic encephalopathy in patients with variants in *SLC6A1* (Rauch, et al., Lance 2012; 380: 1674-82; and Sanders, et al., Nature 2012; 485: 237-41). Further, an exome-wide trio sequencing study found an association between schizophrenia and *de novo* missense variants in *SLC6A1* (Rees, et al., Nat Neurosci 2020; 23: 179-84).

[0034] The current disclosure provides artificial expression constructs that drive gene expression in GABAergic neurons and astrocytes. The artificial expression constructs can be used to drive SLC6A1 gene expression to ameliorate disorders associated with SLC6A1 gene mutations, among other uses described herein. SLC6A1 gene expression can result in the expression of

functional GAT1 GABA transporters.

[0035] The artificial expression constructs disclosed herein drive gene expression in GABAergic neurons by including an I56i enhancer or a core thereof. In particular embodiments, the I56i enhancer core can be derived from, for example the human, murine, or zebrafish I56i enhancer (SEQ ID NOs: 1, 2, and 3 respectively). The selected cores of the I56i enhancer can include SEQ ID NO: 4 (core shared by human and mouse) or SEQ ID NO: 5 (zebrafish core). In particular embodiments, the cores are concatenated. For example, SEQ ID NO: 6 provides a three-copy concatemer of the selected human/murine I56i core while SEQ ID NO: 7 provides a three-copy concatemer of the selected zebrafish I56i core.

[0036] Of particular interest, the synthetic 3x human/murine core (referred to herein as the 3xhI56iCore; SEQ ID NO: 6) is shorter than the original full length enhancer sequence reported in Dimidschstein *et al.* (*Nat Neurosci* 19(12):1743-1749, 2016), despite being a 3x concatemer. Thus, this concatenated core provides more room for cargo genes linked to the enhancer, which is highly desirable. Moreover, the peak level of transgene expression driven by the 3xhI56iCore enhancer is much greater than simply three times the level of the original single full-length original enhancer.

[0037] The artificial expression constructs drive gene expression in astrocytes by including one or more astrocyte-specific enhancers. Examples of astrocyte-specific enhancers include eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof.

[0038] In particular embodiments, the artificial expression constructs include a combination concatenated enhancer. In particular embodiments, the combination concatenated enhancer includes a core of the enhancer selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, and eHGT_390m concatenated with the I56i(core). In particular embodiments, a combination concatenated enhancer includes eHGT_390m(core2) and I56i(core) as set forth in SEQ ID NO: 86. In particular embodiments, the combination concatenated enhancer is concatenated to include 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the combination concatenated enhancer. In particular embodiments, SEQ ID NO: 89 provides a three-copy-concatemer of the eHGT_390m(core2)-I56i(core) combination concatenated enhancer.

[0039] Particular embodiments provide artificial expression constructs including the features of vectors described herein including vectors: CN2720, CN2721, CN2722, CN2732, CN3213, CN3322, CN3323, CN3887, CN3888, CN2972, CN2973, CN2974, CN2975, or CN2976. In certain embodiments, the heterologous encoding sequence encoding SYFP2 in CN2720, CN2721,

CN2722, CN2732, CN3213, CN3322, CN3323, CN3887, and CN3888 is replaced or supplemented with an SLC6A1 gene sequence. The SLC6A1 encoding sequence can be codon optimized (see, e.g., FIGs. 10A-10C and 14). Particular embodiments provide artificial expression constructs including the features of vectors described herein including vectors ID10.01, ID10.02, ID10.03, ID10.04, ID10.05, ID10.06, ID10.07, ID10.08, ID10.09, ID10.10, ID10.11, ID10.12, ID10.13, ID10.14, ID10.15, ID10.16, ID10.17, ID10.18, ID10.19, ID10.20, ID10.21, ID10.22, ID10.23, ID10.24, ID10.25, ID10.26, ID10.27, ID10.28, ID10.29, ID10.30, ID10.31, ID10.32, ID11.01, ID11.02, ID11.03, ID11.04, ID11.05, ID11.06, ID11.07, ID11.08, ID11.09, ID11.10, ID11.11, ID11.12, ID11.13, ID11.14, ID11.15, ID11.16, ID12.01, ID12.02, ID12.03, ID12.04, ID12.05, ID12.06, ID12.07, ID12.08, ID12.09, ID12.10, ID12.11, ID12.12, ID12.13, ID12.14, ID12.15, ID12.16, ID13.01, ID13.02, ID13.03, ID13.04, ID13.05, ID13.06, ID13.07, ID13.08, ID13.09, ID13.10, ID13.11, ID13.12, ID13.13, ID13.14, ID13.15, ID13.16, ID14.01, ID14.02, ID14.03, ID14.04, ID14.05, ID14.06, ID14.07, ID14.08, ID14.09, ID14.10, ID14.11, ID14.12, ID14.13, ID14.14, ID14.15, ID14.16, ID15.01, ID15.02, ID15.03, ID15.04, ID15.05, ID15.06, ID15.07, ID15.08, ID15.09, ID15.10, ID15.11, ID15.12, ID15.13, ID15.14, ID15.15, ID15.16, ID16.01, ID16.02, ID16.03, ID16.04, ID16.05, ID16.06, ID16.07, ID16.08, ID16.09, ID16.10, ID16.11, ID16.12, ID16.13, ID16.14, ID16.15, ID16.16, ID17.01, ID17.02, ID17.03, ID17.04, ID17.05, ID17.06, ID17.07, ID17.08, ID17.09, ID17.10, ID17.11, ID17.12, ID17.13, ID17.14, ID17.15, ID17.16, ID18.01, ID18.02, ID18.03, ID18.04, ID18.05, ID18.06, ID18.07, ID18.08, ID18.09, ID18.10, ID18.11, ID18.12, ID18.13, ID18.14, ID18.15, ID18.16, ID19.01, ID19.02, ID19.03, ID19.04, ID19.05, ID19.06, ID19.07, ID19.08, ID19.09, ID19.10, ID19.11, ID19.12, ID19.13, ID19.14, ID19.15, ID19.16, ID20.01, ID20.02, ID20.03, ID20.04, ID20.05, ID20.06, ID20.07, ID20.08, ID20.09, ID20.10, ID20.11, ID20.12, ID20.13, ID20.14, ID20.15, ID20.16, ID21.01, ID21.02, ID21.03, ID21.04, ID21.05, ID21.06, ID21.07, ID21.08, ID21.09, ID21.10, ID21.11, ID21.12, ID21.13, ID21.14, ID21.15, ID21.16, ID22.01, ID22.02, ID22.03, ID22.04, ID22.05, ID22.06, ID22.07, ID22.08, ID22.09, ID22.10, ID22.11, ID22.12, ID22.13, ID22.14, ID22.15, ID22.16, ID23.01, ID23.02, ID23.03, ID23.04, ID23.05, ID23.06, ID23.07, ID23.08, ID23.09, ID23.10, ID23.11, ID23.12, ID23.13, ID23.14, ID23.15, ID23.16, ID24.01, ID24.02, ID24.03, ID24.04, ID24.05, ID24.06, ID24.07, ID24.08, ID24.09, ID24.10, ID24.11, ID24.12, ID24.13, ID24.14, ID24.15, ID24.16, ID25.01, ID25.02, ID25.03, ID25.04, ID25.05, ID25.06, ID25.07, ID25.08, ID25.09, ID25.10, ID25.11, ID25.12, ID25.13, ID25.14, ID25.15, ID25.16, ID26.01, ID26.02, ID26.03, ID26.04, ID26.05, ID26.06, ID26.07, ID26.08, ID26.09, ID26.10, ID26.11, ID26.12, ID26.13, ID26.14, ID26.15, ID26.16, ID27.01, ID27.02, ID27.03, ID27.04, ID27.05, ID27.06, ID27.07, ID27.08, ID27.09, ID27.10, ID27.11, ID27.12, ID27.13, ID27.14,

ID27.15, and ID27.16. Of note, vectors CN2972, CN2973, CN2974, CN2975, and CN2976 do not include enhancer sequences as disclosed herein and thus are not used to provide targeted gene expression in GABAergic neurons and astrocytes.

[0040] Aspects of the disclosure are now described with the following additional options and detail: (i) Artificial Expression Constructs & Vectors for Targeted Expression of Genes in Targeted Cell Types; (ii) Compositions for Administration (iii) Cell Lines Including Artificial Expression Constructs; (iv) Transgenic Animals; (v) Methods of Use; (vi) Kits and Commercial Packages; (vii) Exemplary Embodiments; and (viii) Closing Paragraphs. These headings are provided for organizational purposes only and do not limit the scope or interpretation of the disclosure.

[0041] (i) Artificial Expression Constructs & Vectors for Targeted Expression of Genes in Targeted Cell Types. Artificial expression constructs disclosed herein include (i) at least two enhancer sequences wherein at least one enhancer sequence leads to expression of a coding sequence within GABAergic neurons and at least one enhancer sequence leads to expression of a coding sequence within astrocytes, (ii) a coding sequence that is expressed, and (iii) a promoter. The artificial expression construct can also include other regulatory elements if necessary or beneficial.

[0042] "Enhancers" or "enhancer elements" increase the level of transcription associated with a promoter. In certain examples, enhancers are cis-acting sequences that and can function in either orientation relative to the promoter and the coding sequence that is to be transcribed and can be located upstream or downstream relative to the promoter or the coding sequence to be transcribed. There are art-recognized methods and techniques for measuring function(s) of enhancer elements. Particular examples of enhancer sequences utilized within artificial expression constructs disclosed herein include an I56i enhancer or a core thereof (e.g., hI56i core or 3XhI56i core) and one or more of the enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof.

[0043] Artificial expression constructs including at least two enhancer sequences can have the two enhancer sequences adjacent to each other or not adjacent to each other. The term "adjacent" refers to the position of two sequence segments relative to each other such that there is not an intervening functional sequence (e.g., promoter, enhancer, or heterologous coding sequence) between the two referenced sequence segments (e.g., enhancers). The term "not adjacent" refers to two sequence segments (e.g., enhancers) positioned such that there is an intervening functional sequence including a promoter, enhancer, and/or heterologous coding sequence between the two sequence segments. Enhancer sequences that are adjacent can have small

linking sequences between them, for example, residues appearing based on cloning strategies. These small linking segments are considered non-functional within the context of the current disclosure. In particular embodiments, an artificial expression construct including two enhancer sequences includes a first enhancer and a second enhancer. In particular embodiments, the first enhancer is adjacent to the second enhancer. In particular embodiments, the first enhancer is not adjacent to the second enhancer. In particular embodiments, the first enhancer is 5' of the second enhancer. In particular embodiments, the second enhancer is 5' of the first enhancer.

[0044] In particular embodiments, a targeted central nervous system cell type enhancer is an enhancer that is uniquely or predominantly utilized by the targeted central nervous system cell type. A targeted central nervous system cell type enhancer enhances expression of a gene in the targeted central nervous system. In certain embodiments, a targeted central nervous system cell type enhancer is also a selective targeted central nervous system type enhancer that enhances expression of a gene in the targeted central nervous system and does not substantially direct expression of genes in other non-targeted cell types, thus having cell type specific transcriptional activity.

[0045] When a heterologous coding sequence operatively linked to an enhancer disclosed herein leads to expression in a targeted cell type, it leads to expression of the administered heterologous coding sequence in the intended cell type.

[0046] When a heterologous coding sequence is selectively expressed in selected cells, it leads to expression of the administered heterologous coding sequence in the intended cell type, as explained din additional detail below. In particular embodiments, not substantially expressed in other cell types is less than 50% expression in a reference cell type as compared to a targeted cell type; less than 40% expression in a reference cell type as compared to a targeted cell type; less than 30% expression in a reference cell type as compared to a targeted cell type; less than 20% expression in a reference cell type as compared to a targeted cell type; or less than 10% expression in a reference cell type as compared to a targeted cell type. In particular embodiments, a reference cell type refers to non-targeted cells. The non-targeted cells can be within the same anatomical structure as the targeted cells and/or can project to a common anatomical area. In particular embodiments, a reference cell type is within an anatomical structure that is adjacent to an anatomical structure that includes the targeted cell type. In particular embodiments, a reference cell type is a non-targeted cell with a different gene expression profile than the targeted cells.

[0047] In particular embodiments, the product of the coding sequence may be expressed at low levels in non-selected cell types, for example at less than 1% or 1%, 2%, 3%, 5%, 10%, 15% or

20% of the levels at which the product is expressed in selected cells. In particular embodiments, the targeted central nervous system cell type is the only cell type that expresses the right combination of transcription factors that bind an enhancer disclosed herein to drive gene expression. Thus, in particular embodiments, expression occurs exclusively within the targeted cell type.

[0048] In particular embodiments, targeted cell types (e.g. neuronal, and/or non-neuronal) can be identified based on transcriptional profiles, such as those described in Tasic et al., Nature 563, 72-78 (2018) and Hodge et al., Nature 573, 61-68 (2019). For reference, the following description of cell types and distinguishing features is also provided:

[0049] Neocortical GABAergic neuron Subclasses:

- All: Express GABA synthesis genes Gad1/GAD1 and Gad2/GAD2.
- *Lamp5* and *Vip* GABAergic neurons: Developmentally derived from neuronal progenitors from the caudal ganglionic eminence (CGE) or preoptic area (POA).
- *Sst* and *Pvalb* GABAergic neurons: Developmentally derived from neuronal progenitors in the medial ganglionic eminence (MGE).
- *Lamp5* GABAergic neurons: Found in many neocortical layers, especially upper (L1-L2/3), and have mainly neurogliaform and single bouquet morphology.
- *Lamp5_Lhx6* GABAergic neurons: A subset of *Lamp5* GABAergic neurons that co-express *Lamp5* and *Lhx6*.
- Sncg GABAergic neurons: Found in many neocortical layers, and have molecular overlaps with *Lamp5* and *Vip* cells, but inconsistent expression of *Lamp5* or *Vip*, with more consistent expression of Sncg.
- Serpinf1 GABAergic neurons: Found in many neocortical layers, and have molecular overlaps with Sncg and *Vip* cells, but inconsistent expression of Sncg or *Vip,* with more consistent expression of Serpinf1.
- *Vip* GABAergic neurons: Found in many neocortical layers, but especially frequent in upper layers (L1-L4), and highly express the neurotransmitter vasoactive intestinal peptide (*Vip*).
- *Sst* GABAergic neurons: Found in many neocortical layers, but especially frequent in lower layers (L5-L6). They highly express the neurotransmitter somatostatin (*Sst)*, and frequently block dendritic inputs to postsynaptic neurons. Included in this subclass are sleep-active *Sst Chodl* neurons (which also express Nos1 and Tacr1) that are highly distinct from other *Sst* neurons but express some shared marker genes including *Sst*. In human, *SST* gene expression is often detected in layer 1 *LAMP5*+ GABAergic neuron

subtypes.

- *Pvalb* GABAergic neurons: Found in many neocortical layers, but especially frequent in lower layers (L5-L6). They highly express the calcium-binding protein parvalbumin (*Pvalb*), express neuropeptide Tac1, and frequently dampen the output of postsynaptic neurons. Most fast-spiking GABAergic neurons express *Pvalb* strongly. Included in this subclass are chandelier cells, which have distinct, chandelier-like morphology and express the markers *Cpne5* and *Vipr2* in mouse, and *NOG* and *UNC5B* in human.
- *Meis2*: A distinct subclass defined by a single type, only neocortical GABAergic neuron type that expresses Meis2 gene, and does not express some other genes that are expressed by all other neocortical GABAergic neuron types (for example, Thy1 and Scn2b). This type is found in L6b and subcortical white matter.

[0050] Neocortical glutamatergic neuron subclasses:

- All: Express glutamate transmitters Slc17a6 and/or Slc17a7. They all express Snap25 and lack expression of Gad1/Gad2.
- L2/3 IT glutamatergic neurons: Primarily reside in Layer 2/3 and have mainly intratelencephalic (cortico-cortical) projections.
- L4 IT glutamatergic neurons: Primarily reside in Layer 4 and mainly have either local or intratelencephalic (cortico-cortical) projections.
- L5 IT glutamatergic neurons: Primarily reside in Layer 5 and have mainly intratelencephalic (cortico-cortical) projections. Also called L5a.
- L5 PT glutamatergic neurons: Primarily reside in Layer 5 and have mainly cortico-subcortical (pyramidal tract or corticofugal) projections. Also called L5b or L5 CF (corticofugal) or L5 ET (extratelencephalic). This subclass includes cells that are located in the primary motor cortex and neighboring areas and are corticospinal projection neurons, which are associated with motor neuron/movement disorders, such as ALS. This subclass includes thick-tufted pyramidal neurons, including distinctive subtypes found only in specialized regions, e.g. Betz cells, Meynert cells, and von Economo cells.
- L5 NP glutamatergic neurons: Primarily reside in Layer 5 and have mainly nearby projections.
- L6 CT glutamatergic neurons: Primarily reside in Layer 6 and have mainly cortico-thalamic projections.
- L6 IT glutamatergic neurons: Primarily reside in Layer 6 and have mainly intratelencephalic (cortico-cortical) projections. Included in this subclass are L6 IT Car3

cells, which are highly similar to intracortical-projecting cells in the claustrum.

- L6b glutamatergic neurons: Primarily reside in the neocortical subplate (L6b), with local (near the cell body) projections and some cortico-cortical projections from VISp to anterior cingulate, and cortico-subcortical projections to the thalamus.
- CR neurons: A distinct subclass defined by a single type in L1, Cajal-Retzius cells express distinct molecular markers Lhx5 and Trp73.

[0051] Cerebellar Purkinje cells: large GABAergic neurons that are the only projection neurons and the sole output from the cerebellum. Their cell bodies form a single layer, so called 'Purkinje cell layer', and they express parvalbumin.

[0052] Deep cerebellar nucleus neurons: neurons located in the deep cerebellar nuclei structures. These include glutamatergic and GABAergic cells that express the gene Pvalb.

[0053] Non-neuronal Subclasses:

- Astrocytes: Neuroectoderm-derived glial cells which express the marker Aqp4 and often GFAP, but do not express neuronal marker SNAP25. They can have a distinct star-shaped morphology and are involved in metabolic support of other cells in the brain. Multiple astrocyte morphologies are observed in mouse and human
- Oligodendrocytes: Neuroectoderm-derived glial cells, which express the marker Sox10. This category includes oligodendrocyte precursor cells (OPCs). Oligodendrocytes are the subclass that is primarily responsible for myelination of neurons.
- VLMCs: Vascular leptomeningeal cells (VLMCs) are part of the meninges that surround the outer layer of the cortex and express the marker genes Lum and Col1a1.
- Pericytes: Blood vessel-associated cells that express the marker genes Kcnj8 and Abcc9. Pericytes wrap around endothelial cells and are important for regulation of capillary blood flow and are involved in blood-brain barrier permeability.
- SMCs: Specialized smooth-muscle cells which are blood vessel-associated cells that express the marker gene Acta2. SMCs cover arterioles in the brain and are involved in blood-brain barrier permeability.
- Endothelial cells: Cells that line blood vessels of the brain. Endothelial cells express the markers Tek and PDGF-B.
- Microglia: hematopoietic-derived immune cells, which are brain-resident macrophages, and perivascular macrophages (PVMs) that may be transitionally associated with brain tissue or included as a biproduct of brain dissection methods. Microglia are known to express Cx3cr1, Tmem119, and PTPRC (CD45).

[0054] In particular embodiments, a coding sequence is a heterologous coding sequence that

encodes GAT1. The heterologous coding sequence that encodes GAT1 can be a codon optimized SLC6A1 variant, for example as shown in FIG. 14 (SEQ ID NOs: 22, 25, 28, 31, 34).

[0055] In particular embodiments, a coding sequence is a heterologous coding sequence that encodes an effector element. An effector element is a sequence that is expressed to achieve, and that in fact achieves, an intended effect. Examples of effector elements include reporter genes/proteins and functional genes/proteins.

[0056] Exemplary reporter genes/proteins include those expressed by Addgene ID#s 83894 (pAAV-hDlx-Flex-dTomato-Fishell_7), 83895 (pAAV-hDlx-Flex-GFP-Fishell_6), 83896 (pAAV-hDlx-GiDREADD-dTomato-Fishell-5), 83898 (pAAV-mDlx-ChR2-mCherry-Fishell-3), 83899 (pAAV-mDlx-GCaMP6f-Fishell-2), 83900 (pAAV-mDlx-GFP-Fishell-1), and 89897 (pcDNA3-FLAG-mTET2 (N500)). Exemplary reporter genes particularly can include those which encode an expressible fluorescent protein, or expressible biotin; blue fluorescent proteins (*e.g.* eBFP, eBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire); cyan fluorescent proteins (*e.g.* eCFP, Cerulean, CyPet, AmCyanl, Midoriishi-Cyan, mTurquoise); green fluorescent proteins (*e.g.* GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green (mAzamigreen), CopGFP, AceGFP, avGFP, ZsGreenl, Oregon Green™(Thermo Fisher Scientific)); Luciferase; orange fluorescent proteins (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato, dTomato); red fluorescent proteins (mKate, mKate2, mPlum, DsRed monomer, mCherry, mRuby, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRedl, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred, Texas Red™ (Thermo Fisher Scientific)); far red fluorescent proteins (e.g., mPlum and mNeptune); yellow fluorescent proteins (*e.g.*, YFP, eYFP, Citrine, SYFP2, Venus, YPet, PhiYFP, ZsYellowl); and tandem conjugates.

[0057] GFP is composed of 238 amino acids (26.9 kDa), originally isolated from the jellyfish Aequorea victoria/Aequorea aequorea/Aequorea forskalea that fluoresces green when exposed to blue light. The GFP from A. victoria has a major excitation peak at a wavelength of 395 nm and a minor one at 475 nm. Its emission peak is at 509 nm which is in the lower green portion of the visible spectrum. The GFP from the sea pansy (Renilla reniformis) has a single major excitation peak at 498 nm. Due to the potential for widespread usage and the evolving needs of researchers, many different mutants of GFP have been engineered. The first major improvement was a single point mutation (S65T) reported in 1995 in Nature by Roger Tsien. This mutation dramatically improved the spectral characteristics of GFP, resulting in increased fluorescence, photostability and a shift of the major excitation peak to 488 nm with the peak emission kept at 509 nm. The addition of the 37°C folding efficiency (F64L) point mutant to this scaffold yielded enhanced GFP

(EGFP). EGFP has an extinction coefficient (denoted ε), also known as its optical cross section of 9.13X10-21 m^2/molecule, also quoted as 55,000 L/(mol•cm). Superfolder GFP, a series of mutations that allow GFP to rapidly fold and mature even when fused to poorly folding peptides, was reported in 2006.

[0058] The "yellow fluorescent protein" (YFP) is a genetic mutant of green fluorescent protein, derived from Aequorea victoria. Its excitation peak is 514 nm and its emission peak is 527 nm.

[0059] Exemplary functional molecules include functioning ion transporters, cellular trafficking proteins, enzymes, transcription factors, neurotransmitters, calcium reporters, channelrhodopsins, guide RNA, nucleases, microRNA, or designer receptors exclusively activated by designer drugs (DREADDs).

[0060] Ion transporters are transmembrane proteins that mediate transport of ions across cell membranes. These transporters are pervasive throughout most cell types and important for regulating cellular excitability and homeostasis. Ion transporters participate in numerous cellular processes such as action potentials, synaptic transmission, hormone secretion, and muscle contraction. Many important biological processes in living cells involve the translocation of cations, such as calcium (Ca2+), potassium (K+), and sodium (Na+) ions, through such ion channels. In particular embodiments, ion transporters include voltage gated sodium channels (e.g., SCN1A), potassium channels (e.g., KCNQ2), and calcium channels (e.g. CACNA1C)).

[0061] Exemplary enzymes, transcription factors, receptors, membrane proteins, cellular trafficking proteins, signaling molecules, and neurotransmitters include enzymes such as lactase, lipase, helicase, alpha-glucosidase, amylase; transcription factors such as SP1, AP-1, Heat shock factor protein 1, C/EBP (CCAA-T/enhancer binding protein), and Oct-1; receptors such as transforming growth factor receptor beta 1, platelet-derived growth factor receptor, epidermal growth factor receptor, vascular endothelial growth factor receptor, and interleukin 8 receptor alpha; membrane proteins, cellular trafficking proteins such as clathrin, dynamin, caveolin, Rab-4A, and Rab-11A; signaling molecules such as nerve growth factor (NGF), platelet-derived growth factor (PDGF), transforming growth factor β (TGFβ), epidermal growth factor (EGF), GTPase and HRas; and neurotransmitters such as cocaine and amphetamine regulated transcript, substance P, oxytocin, and somatostatin.

[0062] In particular embodiments, functional molecules include reporters of cell function and states such as calcium reporters. Intracellular calcium concentration is an important predictor of numerous cellular activities, which include neuronal activation, muscle cell contraction and second messenger signaling. A sensitive and convenient technique to monitor the intracellular calcium levels is through the genetically encoded calcium indicator (GECI). Among the GECIs,

green fluorescent protein (GFP) based calcium sensors named GCaMPs are efficient and widely used tools. The GCaMPs are formed by fusion of M13 and calmodulin protein to N- and C-termini of circularly permutated GFP. Some GCaMPs yield distinct fluorescence emission spectra (Zhao *et al.*,Science, 2011, 333(6051): 1888-1891). Exemplary GECIs with green fluorescence include GCaMP3, GCaMP5G, GCaMP6s, GCaMP6m, GCaMP6f, jGCaMP7s, jGCaMP7c, jGCaMP7b, and jGCaMP7f. Furthermore, GECIs with red fluorescence include jRGECO1a and jRGECO1b. AAV products containing GECIs are commercially available. For example, Vigene Biosciences provides AAV products including AAV8-CAG-GCaMP3 (Cat. No:BS4-CX3AAV8), AAV8-Syn-FLEX-GCaMP6s-WPRE (Cat. No:BS1-NXSAAV8), AAV8-Syn-FLEX-GCaMP6s-WPRE (Cat. No:BS1-NXSAAV8), AAV9-CAG-FLEX-GCaMP6m-WPRE (Cat. No:BS2-CXMAAV9), AAV9-Syn-FLEX-jGCaMP7s-WPRE (Cat. No:BS12-NXSAAV9), AAV9-CAG-FLEX-jGCaMP7f-WPRE (Cat. No:BS12-CXFAAV9), AAV9-Syn-FLEX-jGCaMP7b-WPRE (Cat. No:BS12-NXBAAV9), AAV9-Syn-FLEX-jGCaMP7c-WPRE (Cat. No:BS12-NXCAAV9), AAV9-Syn-FLEX-NES-jRGECO1a-WPRE (Cat. No:BS8-NXAAAV9), and AAV8-Syn-FLEX-NES-jRCaMP1b-WPRE (Cat. No:BS7-NXBAAV8).

[0063] In particular embodiments calcium reporters include the genetically encoded calcium indicators GECI, NTnC; Myosin light chain kinase, GFP, Calmodulin chimera; Calcium indicator TN-XXL; BRET-based auto-luminescent calcium indicator; and/or Calcium indicator protein OeNL(Ca2+)-18u).

[0064] In particular embodiments, functional molecules include modulators of neuronal activity like channelrhodopsins (e.g., channelrhodopsin-1, channelrhodopsin-2, and variants thereof). Channelrhodopsins are a subfamily of retinylidene proteins (rhodopsins) that function as light-gated ion channels. In addition to channelrhodopsin 1 (ChR1) and channelrhodopsin 2 (ChR2), several variants of channelrhodopsins have been developed. For example, Lin et al. (*Biophys J,* 2009, 96(5): 1803-14) describe making chimeras of the transmembrane domains of ChR1 and ChR2, combined with site-directed mutagenesis. Zhang et al. (*Nat Neurosci,* 2008, 11(6): 631-3) describe VChR1, which is a red-shifted channelrhodopsin variant. VChR1 has lower light sensitivity and poor membrane trafficking and expression. Other known channelrhodopsin variants include the ChR2 variant described in Nagel, et al., *Proc Natl Acad Sci USA,* 2003, 100(24): 13940-5), ChR2/H134R (Nagel, G., et al., *Curr Biol,* 2005, 15(24): 2279-84), and ChD/ChEF/ChIEF (Lin, J. Y., et al., *Biophys J,* 2009, 96(5): 1803-14), which are activated by blue light (470 nm) but show no sensitivity to orange/red light. Additional variants are described in Lin, *Experimental Physiology,* 2010, 96.1: 19-25 and Knopfel et al., *The Journal of Neuroscience,* 2010, 30(45): 14998-15004).

[0065] In particular embodiments, functional molecules include DNA and RNA editing tools such CRISPR/CAS (e.g., guide RNA and a nuclease, such as Cas, Cas9 or cpf1). Functional molecules can also include engineered Cpf1s such as those described in US 2018/0030425, US 2016/0208243, WO/2017/184768 and Zetsche *et al.* (2015) *Cell* 163: 759-771; single gRNA (see e.g., Jinek *et al.* (2012) Science 337:816-821; Jinek *et al.* (2013) eLife 2:e00471; Segal (2013) eLife 2:e00563) or editase, guide RNA molecules, microRNA, or homologous recombination donor cassettes.

[0066] As indicated, sequences are publicly-available. Further examples include, lactase (e.g., GenBank: EAX11622.1), lipase (e.g., GenBank: AAA60129.1), helicase (e.g., GenBank: AMD82207.1), amylase (e.g., GenBank: AAA51724.1), alpha-glucosidase (e.g., GenBank: ABI53718.1), transcription factor SP1 (e.g., UniProtKB/Swiss-Prot: P08047.3), transcription factor AP-1 (e.g., NP_002219.1), heat shock factor protein 1 (e.g., UniProtKB/Swiss-Prot: Q00613.1), CCAAT/enhancer-binding protein (C/EBP) beta isoform a (e.g., NP_005185.2), Oct-1 (e.g., UniProtKB/Swiss-Prot: P14859.2), TGFβ (e.g., GenBank: CAF02096.2), platelet-derived growth factor receptor (e.g., GenBank: AAA60049.1), epidermal growth factor receptor (e.g., GenBank: CAA25240.1), vascular endothelial growth factor receptor (e.g., GenBank: AAC16449.2), interleukin 8 receptor alpha (e.g., GenBank: AAB59436.1), caveolin (e.g., GenBank: CAA79476.1), dynamin (e.g., GenBank: AAA88025.1), clathrin heavy chain 1 isoform 1 (e.g., NP_004850.1), clathrin heavy chain 2 isoform 1 (e.g., NP_009029.3), clathrin light chain A isoform a (e.g., NP_001824.1), clathrin light chain B isoform a (e.g., NP_001825.1), ras-related protein Rab-4A isoform 1 (e.g., NP_004569.2), ras-related protein Rab-11A (e.g., UniProtKB/Swiss-Prot: P62491.3), platelet-derived growth factor (e.g., GenBank: AAA60552.1), transforming growth factor-beta3 (e.g., GenBank: AAA61161.1), nerve growth factor (e.g., GenBank: CAA37703.1), EGF (e.g., GenBank: CAA34902.2), cocaine and amphetamine regulated transcript (Chain A) (e.g., PDB: 1HY9_A), protachykinin-1 (e.g., UniProtKB - P20366), oxytocin-neurophysin 1 (e.g., UniProtKB - P01178), somatostatin (e.g., GenBank: AAH32625.1), genetically-encoded green calcium indicator NTnC (chain A) [synthetic construct] (e.g., PDB: 5MWC_A), calcium indicator TN-XXL [synthetic construct], (e.g., GenBank: ACF93133.1), BRET-based auto-luminescent calcium indicator [synthetic construct] (e.g., GenBank ADF42668.1), calcium indicator protein OeNL(Ca2+)-18u [synthetic construct], ((e.g., GenBank BBB18812.1), myosin light chain kinase, Green fluorescent protein, Calmodulin chimera (Chain A) [synthetic construct] ((e.g., PDB: 3EKJ_A), channelopsin *1* (e.g., *UniProt*KB - F8UVI5), channelopsin 1 (e.g., GenBank: AER58217.1), channelrhodopsin-2 ((e.g., UniProtKB - B4Y105), channel rhodopsin 2 [synthetic construct] ((e.g., GenBank: ABO64386.1), CRISPR-associated protein (Cas) (e.g., GenBank:

AKG27598.1), Cas9 [synthetic construct] (e.g., GenBank: AST09977.1), CRISPR-associated endonuclease Cpf1 (e.g., UniProtKB/Swiss-Prot: U2UMQ6.1), ribonuclease 4 or ribonuclease L (e.g., UniProtKB/Swiss-Prot: Q05823.2), deoxyribonuclease II beta (e.g., GenBank: AAF76893.1), sodium channel protein type 1 subunit alpha (e.g., UniProtKB - P35498), potassium voltage-gated channel subfamily KQT member 2 (e.g., UniProtKB - O43526), and voltage-dependent L-type calcium channel subunit alpha-1C (e.g., UniProtKB - Q13936).

[0067] Additional effector elements include Cre, iCre, dgCre, FlpO, and tTA2. iCre refers to a codon-improved Cre. dgCre refers to an enhanced GFP/Cre recombinase fusion gene with an N terminal fusion of the first 159 amino acids of the Escherichia coli K-12 strain chromosomal dihydrofolate reductase gene (DHFR or folA) harboring a G67S mutation and modified to also include the R12Y/Y100I destabilizing domain mutation. FlpO refers to a codon-optimized form of FLPe that greatly increases protein expression and FRT recombination efficiency in mouse cells. Like the Cre/LoxP system, the FLP/FRT system has been widely used for gene expression (and generating conditional knockout mice, mediated by the FLP/FRT system). tTA2 refers to tetracycline transactivator. 4x2C is a synthetic microRNA binding site element that allows silencing of virus mediated transgene expression in certain cell types. For example, it can be used to reduce or eliminate expression in many glutamatergic neuron populations in the brain. 4x2C is described in Sayeg et al., ACS Synth. Biol. 2015, 4, 7, 788–795.

[0068] Exemplary expressible elements are expression products that do not include effector elements, for example, a non-functioning or defective protein. In particular embodiments, expressible elements can provide methods to study the effects of their functioning counterparts. In particular embodiments, expressible elements are non-functioning or defective based on an engineered mutation that renders them non-functioning. In these aspects, non-expressible elements are as similar in structure as possible to their functioning counterparts.

[0069] Exemplary self-cleaving peptides include the 2A peptides which lead to the production of two proteins from one mRNA. The 2A sequences are short (e.g., 20 amino acids), allowing more use in size-limited constructs. Particular examples include P2A, T2A, E2A, and F2A. In particular embodiments, the artificial expression constructs include an internal ribosome entry site (IRES) sequence. IRES allow ribosomes to initiate translation at a second internal site on a mRNA molecule, leading to production of two proteins from one mRNA.

[0070] Coding sequences encoding molecules (e.g., RNA, proteins) described herein can be obtained from publicly available databases and publications. Coding sequences can further include various sequence polymorphisms, mutations, and/or sequence variants wherein such alterations do not affect the function of the encoded molecule. The term “encode” or “encoding”

refers to a property of sequences of nucleic acids, such as a vector, a plasmid, a gene, cDNA, mRNA, to serve as templates for synthesis of other molecules such as proteins.

[0071] The term “gene” may include not only coding sequences but also regulatory regions such as promoters, enhancers, insulators, and/or post-regulatory elements, such as termination regions. The term further can include all introns and other DNA sequences spliced from the mRNA transcript, along with variants resulting from alternative splice sites. The sequences can also include degenerate codons of a reference sequence or sequences that may be introduced to provide codon preference in a specific organism or cell type.

[0072] Promoters can include general promoters, tissue-specific promoters, cell-specific promoters, and/or promoters specific for the cytoplasm. Promoters may include strong promoters, weak promoters, constitutive expression promoters, and/or inducible promoters. Inducible promoters direct expression in response to certain conditions, signals or cellular events. For example, the promoter may be an inducible promoter that requires a particular ligand, small molecule, transcription factor or hormone protein in order to effect transcription from the promoter. Particular examples of promoters include minBglobin (or minBGprom), CMV, minCMV, minCMV* (minCMV* is minCMV with a SacI restriction site removed), minRho, minRho* (minRho* is minRho with a SacI restriction site removed), SV40 immediately early promoter, the Hsp68 minimal promoter (proHSP68), and the Rous Sarcoma Virus (RSV) long-terminal repeat (LTR) promoter. Minimal promoters have no activity to drive gene expression on their own but can be activated to drive gene expression when linked to a proximal enhancer element.

[0073] In particular embodiments, expression constructs are provided within vectors. The term vector refers to a nucleic acid molecule capable of transferring or transporting another nucleic acid molecule, such as an expression construct. The transferred nucleic acid is generally linked to, *e.g.*, inserted into, the vector nucleic acid molecule. A vector may include sequences that direct autonomous replication in a cell or may include sequences that permit integration into host cell DNA. Useful vectors include, for example, plasmids (*e.g.*, DNA plasmids or RNA plasmids), transposons, cosmids, bacterial artificial chromosomes, and viral vectors.

[0074] Viral vector is widely used to refer to a nucleic acid molecule that includes virus-derived components that facilitate transfer and expression of non-native nucleic acid molecules within a cell. The term adeno-associated viral vector refers to a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from AAV. The term "retroviral vector" refers to a viral vector or plasmid containing structural and functional genetic elements, or portions thereof, that are primarily derived from a retrovirus. The term "lentiviral vector" refers to a viral vector or plasmid containing structural and functional genetic

elements, or portions thereof, that are primarily derived from a lentivirus, and so on. The term "hybrid vector" refers to a vector including structural and/or functional genetic elements from more than one virus type.

[0075] Adenovirus vectors refer to those constructs containing adenovirus sequences sufficient to (a) support packaging of an artificial expression construct and (b) to express a coding sequence that has been cloned therein in a sense or antisense orientation. A recombinant Adenovirus vector includes a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb. In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification.

[0076] Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range, and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are cis elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression, and host cell shut-off. The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP is particularly efficient during the late phase of infection, and all the mRNAs issued from this promoter possess a 5'-tripartite leader (TPL) sequence which makes them preferred mRNAs for translation.

[0077] Other than the requirement that an adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of particular embodiments disclosed herein. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. In particular embodiments, adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in particular embodiments, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

[0078] As indicated, the typical vector is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical. The polynucleotide encoding the gene of interest may also be inserted in lieu of a deleted E3 region in E3 replacement vectors or in the E4 region where a helper cell line or helper virus complements the E4 defect.

[0079] Adeno-Associated Virus (AAV) is a parvovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replication is dependent on the presence of a helper virus, such as adenovirus. Various serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter.

[0080] The AAV DNA is 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs. There are two major genes in the AAV genome: rep and cap. The rep gene codes for proteins responsible for viral replications, whereas cap codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential cis components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three AAV viral promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins.

[0081] AAVs stand out for use within the current disclosure because of their superb safety profile and because their capsids and genomes can be tailored to allow expression in targeted cell populations. scAAV refers to a self-complementary AAV. pAAV refers to a plasmid adeno-associated virus. rAAV refers to a recombinant adeno-associated virus.

[0082] Other viral vectors may also be employed. For example, vectors derived from viruses such as vaccinia virus, polioviruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells.

[0083] Retroviruses are a common tool for gene delivery. "Retrovirus" refers to an RNA virus that reverse transcribes its genomic RNA into a linear double-stranded DNA copy and subsequently covalently integrates its genomic DNA into a host genome. Once the virus is integrated into the host genome, it is referred to as a "provirus." The provirus serves as a template for RNA

polymerase II and directs the expression of RNA molecules which encode the structural proteins and enzymes needed to produce new viral particles.

[0084] Illustrative retroviruses suitable for use in particular embodiments, include: Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV), feline leukemia virus (FLV), spumavirus, Friend murine leukemia virus, Murine Stem Cell Virus (MSCV), Rous Sarcoma Virus (RSV), and lentivirus.

[0085] "Lentivirus" refers to a group (or genus) of complex retroviruses. Illustrative lentiviruses include: HIV (human immunodeficiency virus; including HIV type 1, and HIV type 2); visna-maedi virus (VMV); the caprine arthritis-encephalitis virus (CAEV); equine infectious anemia virus (EIAV); feline immunodeficiency virus (FIV); bovine immune deficiency virus (BIV); and simian immunodeficiency virus (SIV). In particular embodiments, HIV based vector backbones (i.e., HIV cis-acting sequence elements) can be used.

[0086] A safety enhancement for the use of some vectors can be provided by replacing the U3 region of the 5' LTR with a heterologous promoter to drive transcription of the viral genome during production of viral particles. Examples of heterologous promoters which can be used for this purpose include, for example, viral simian virus 40 (SV40) (*e.g.*, early or late), cytomegalovirus (CMV) (*e.g.*, immediate early), Moloney murine leukemia virus (MoMLV), Rous sarcoma virus (RSV), and herpes simplex virus (HSV) (thymidine kinase) promoters. Typical promoters are able to drive high levels of transcription in a Tat-independent manner. This replacement reduces the possibility of recombination to generate replication-competent virus because there is no complete U3 sequence in the virus production system. In particular embodiments, the heterologous promoter has additional advantages in controlling the manner in which the viral genome is transcribed. For example, the heterologous promoter can be inducible, such that transcription of all or part of the viral genome will occur only when the induction factors are present. Induction factors include one or more chemical compounds or the physiological conditions such as temperature or pH, in which the host cells are cultured.

[0087] In particular embodiments, viral vectors include a TAR element. The term "TAR" refers to the "trans-activation response" genetic element located in the R region of lentiviral LTRs. This element interacts with the lentiviral trans-activator (tat) genetic element to enhance viral replication. However, this element is not required in embodiments wherein the U3 region of the 5' LTR is replaced by a heterologous promoter.

[0088] The "R region" refers to the region within retroviral LTRs beginning at the start of the capping group (i.e., the start of transcription) and ending immediately prior to the start of the

poly(A) tract. The R region is also defined as being flanked by the U3 and U5 regions. The R region plays a role during reverse transcription in permitting the transfer of nascent DNA from one end of the genome to the other.

[0089] In particular embodiments, expression of heterologous sequences in viral vectors is increased by incorporating posttranscriptional regulatory elements, efficient polyadenylation sites, and optionally, transcription termination signals into the vectors. A variety of posttranscriptional regulatory elements can increase expression of a heterologous nucleic acid. Examples include the woodchuck hepatitis virus posttranscriptional regulatory element (WPRE; Zufferey *et al.*, 1999, J. Virol., 73:2886); the posttranscriptional regulatory element present in hepatitis B virus (HPRE) (Smith *et al.*, Nucleic Acids Res. 26(21):4818-4827, 1998); and the like (Liu *et al.*, 1995, Genes Dev., 9:1766). In particular embodiments, vectors include a posttranscriptional regulatory element such as a WPRE or HPRE. In particular embodiments, vectors lack or do not include a posttranscriptional regulatory element such as a WPRE or HPRE.

[0090] Elements directing the efficient termination and polyadenylation of a heterologous nucleic acid transcript can increase heterologous gene expression. Transcription termination signals are generally found downstream of the polyadenylation signal. In particular embodiments, vectors include a polyadenylation signal 3' of a polynucleotide encoding a molecule (e.g., protein) to be expressed. The term "poly(A) site" or "poly(A) sequence" denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript by RNA polymerase II. Polyadenylation sequences can promote mRNA stability by addition of a poly(A) tail to the 3' end of the coding sequence and thus, contribute to increased translational efficiency. Particular embodiments may utilize BGHpA or SV40pA. In particular embodiments, a preferred embodiment of an expression construct includes a terminator element. These elements can serve to enhance transcript levels and to minimize read through from the construct into other plasmid sequences.

[0091] In particular embodiments, a viral vector further includes one or more insulator elements. Insulators elements may contribute to protecting viral vector-expressed sequences, *e.g.*, effector elements or expressible elements, from integration site effects, which may be mediated by cis-acting elements present in genomic DNA and lead to deregulated expression of transferred sequences (*i.e.*, position effect; see, *e.g.*, Burgess-Beusse *et al.*, *PNAS., USA*, 99:16433, 2002; and Zhan *et al.*, *Hum. Genet.*, 109:471, 2001). In particular embodiments, viral transfer vectors include one or more insulator elements at the 3' LTR and upon integration of the provirus into the host genome, the provirus includes the one or more insulators at both the 5' LTR and 3' LTR, by virtue of duplicating the 3' LTR. Suitable insulators for use in particular embodiments include the chicken β-globin insulator (see Chung *et al.*, *Cell* 74:505, 1993; Chung *et al.*, *PNAS USA* 94:575,

1997; and Bell *et al.*, *Cell* 98:387, 1999), SP10 insulator (Abhyankar et al., *JBC* 282:36143, 2007), or other small CTCF recognition sequences that function as enhancer blocking insulators (Liu et al., *Nature Biotechnology*, 33:198, 2015).

[0092] Beyond the foregoing description, a wide range of suitable expression vector types will be known to a person of ordinary skill in the art. These can include commercially available expression vectors designed for general recombinant procedures, for example plasmids that contain one or more reporter genes and regulatory elements required for expression of the reporter gene in cells. Numerous vectors are commercially available, e.g., from Invitrogen, Stratagene, Clontech, etc., and are described in numerous associated guides. In particular embodiments, suitable expression vectors include any plasmid, cosmid or phage construct that is capable of supporting expression of encoded genes in mammalian cell, such as pUC or Bluescript plasmid series.

[0093] Particular embodiments of vectors disclosed herein include:

Expression Construct Name	Features
CN2720	pAAV: eHGT_387m-3xhI56i(core)-minBGprom-SYFP2-WPRE3-BGHpA
CN2721	pAAV: 3xhI56i(core)-eHGT_387m-minBGprom-SYFP2-WPRE3-BGHpA
CN2722	pAAV: 3xhI56i(core)-minBGprom-SYFP2-WPRE3-eHGT_387m-BGHpA
CN2732	pAAV: eHGT_387m-minBGprom-SYFP2-WPRE3-3xhI56i(core)-BGHpA-
CN2972	pCDNA3.1-CMV_hSLC6A1_myc_native_CN2522GeneOpt1Splice_IRES2_SYFP2_BGHpA
CN2973	pCDNA3.1-CMV_hSLC6A1_myc_IDTcodonOptSplice1_Intron_IRES2_SYFP2_BGHpA
CN2974	pCDNA3.1-CMV_hSLC6A1_myc_IDTcodonOptSplice1_CN2522.str_IRES2_SYFP2_BGHpA
CN2975	pCDNA3.1-CMV_hSLC6A1_myc_native_Intron_IRES2_SYFP2_BGHpA
CN2976	pCDNA3.1-CMV_hSLC6A1_myc_native_CN2522GeneOpt1SpliceIntron_IRES2_SYFP2_BGHpA
CN3213	pAAV:3xhI56i(core)_eHGT_387m-minBGprom-intronSLC6A1-myc-flag-WPRE3-BGHpA
CN3322	rAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SYFP2-4X2C-WPRE3-BGHpA
CN3323	rAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SYFP2-WPRE3-BGHpA
CN3887	rAAV:eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SYFP2-WPRE3-BGHpA
CN3888	pAAV: eHGT_3xhI56i(core)_eHGT_390m-minBGprom-SYFP2-WPRE3-BGHpA
ID10.01	pAAV: 3xhI56i(core)-eHGT_375h-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.02	pAAV: 3xhI56i(core)-eHGT_376h-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.03	pAAV: 3xhI56i(core)-eHGT_390h-minBGprom-SLC6A1-WPRE3-BGHpA

ID10.04	pAAV: 3xhI56i(core)-eHGT_373m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.05	pAAV: 3xhI56i(core)-eHGT_375m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.06	pAAV: 3xhI56i(core)-eHGT_386m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.07	pAAV: 3xhI56i(core)-eHGT_387m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.08	pAAV: 3xhI56i(core)-eHGT_390m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.09	pAAV: hI56i-eHGT_375h-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.10	pAAV: hI56i-eHGT_376h-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.11	pAAV: hI56i-eHGT_390h-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.12	pAAV: hI56i-eHGT_373m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.13	pAAV: hI56i-eHGT_375m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.14	pAAV: hI56i-eHGT_386m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.15	pAAV: hI56i-eHGT_387m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.16	pAAV: hI56i-eHGT_390m-minBGprom-SLC6A1-WPRE3-BGHpA
ID10.17	pAAV: 3xhI56i(core)-eHGT_375h-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.18	pAAV: 3xhI56i(core)-eHGT_376h-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.19	pAAV: 3xhI56i(core)-eHGT_390h-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.20	pAAV: 3xhI56i(core)-eHGT_373m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.21	pAAV: 3xhI56i(core)-eHGT_375m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.22	pAAV: 3xhI56i(core)-eHGT_386m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.23	pAAV: 3xhI56i(core)-eHGT_387m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.24	pAAV: 3xhI56i(core)-eHGT_390m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.25	pAAV: hI56i-eHGT_375h-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.26	pAAV: hI56i-eHGT_376h-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.27	pAAV: hI56i-eHGT_390h-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.28	pAAV: hI56i-eHGT_373m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.29	pAAV: hI56i-eHGT_375m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.30	pAAV: hI56i-eHGT_386m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.31	pAAV: hI56i-eHGT_387m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID10.32	pAAV: hI56i-eHGT_390m-minBGprom-SLC6A1-Myc-FLAG-WPRE3-BGHpA
ID11.01	pAAV: eHGT_375h-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.02	pAAV: eHGT_376h-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.03	pAAV: eHGT_390h-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.04	pAAV: eHGT_373m-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.05	pAAV: eHGT_375m-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.06	pAAV: eHGT_386m-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.07	pAAV: eHGT_387m-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.08	pAAV: eHGT_390m-3xhI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.09	pAAV: eHGT_375h-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA

ID11.10	pAAV: eHGT_376h-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.11	pAAV: eHGT_390h-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.12	pAAV: eHGT_373m-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.13	pAAV: eHGT_375m-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.14	pAAV: eHGT_386m-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.15	pAAV: eHGT_387m-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID11.16	pAAV: eHGT_390m-hI56i-minBGprom-SLC6A1-WPRE3-BGHpA
ID12.01	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_375h-BGHpA
ID12.02	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_376h-BGHpA
ID12.03	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_390h-BGHpA
ID12.04	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_373m-BGHpA
ID12.05	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_375m-BGHpA
ID12.06	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_386m-BGHpA
ID12.07	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_387m-BGHpA
ID12.08	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-WPRE3-eHGT_390m-BGHpA
ID12.09	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_375h-BGHpA
ID12.10	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_376h-BGHpA
ID12.11	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_390h-BGHpA
ID12.12	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_373m-BGHpA
ID12.13	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_375m-BGHpA
ID12.14	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_386m-BGHpA
ID12.15	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_387m-BGHpA
ID12.16	pAAV: hI56i-minBGprom-SLC6A1-WPRE3-eHGT_390m-BGHpA
ID13.01	rAAV: eHGT_375h-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.02	rAAV: eHGT_376h-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.03	rAAV: eHGT_390h-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.04	rAAV: eHGT_373m-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.05	rAAV: eHGT_375m-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.06	rAAV: eHGT_386m-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.07	rAAV: eHGT_387m-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.08	rAAV: eHGT_390m-minBGprom-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID13.09	rAAV: eHGT_375h-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.10	rAAV: eHGT_376h-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.11	rAAV: eHGT_390h-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.12	rAAV: eHGT_373m-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.13	rAAV: eHGT_375m-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.14	rAAV: eHGT_386m-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.15	rAAV: eHGT_387m-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID13.16	rAAV: eHGT_390m-minBGprom-SLC6A1-WPRE3-hI56i-BGHpA
ID14.01	pAAV: hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.02	pAAV: hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.03	pAAV: hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.04	pAAV: hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-

	hI56i(core)-eHGT_373m(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.05	pAAV: hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.06	pAAV: hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.07	pAAV: hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.08	pAAV: hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-minBGprom-SLC6A1-WPRE3-BGHpA
ID14.09	pAAV: hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.10	pAAV: hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.11	pAAV: hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.12	pAAV: hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.13	pAAV: hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.14	pAAV: hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.15	pAAV: hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID14.16	pAAV: hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.01	pAAV: eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.02	pAAV: eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.03	pAAV: eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.04	pAAV: eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.05	pAAV: eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.06	pAAV: eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.07	pAAV: eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.08	pAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SLC6A1-WPRE3-BGHpA
ID15.09	pAAV: eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.10	pAAV: eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.11	pAAV: eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA

ID15.12	pAAV: eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.13	pAAV: eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.14	pAAV: eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.15	pAAV: eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID15.16	pAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SLC6A1-4X2C-WPRE3-BGHpA
ID16.01	pAAV: 3xhI56i(core)-eHGT_375h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.02	pAAV: 3xhI56i(core)-eHGT_376h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.03	pAAV: 3xhI56i(core)-eHGT_390h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.04	pAAV: 3xhI56i(core)-eHGT_373m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.05	pAAV: 3xhI56i(core)-eHGT_375m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.06	pAAV: 3xhI56i(core)-eHGT_386m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.07	pAAV: 3xhI56i(core)-eHGT_387m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.08	pAAV: 3xhI56i(core)-eHGT_390m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.09	pAAV: hI56i-eHGT_375h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.10	pAAV: hI56i-eHGT_376h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.11	pAAV: hI56i-eHGT_390h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.12	pAAV: hI56i-eHGT_373m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.13	pAAV: hI56i-eHGT_375m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.14	pAAV: hI56i-eHGT_386m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.15	pAAV: hI56i-eHGT_387m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID16.16	pAAV: hI56i-eHGT_390m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.01	pAAV: eHGT_375h-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.02	pAAV: eHGT_376h-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.03	pAAV: eHGT_390h-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.04	pAAV: eHGT_373m-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.05	pAAV: eHGT_375m-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.06	pAAV: eHGT_386m-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.07	pAAV: eHGT_387m-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-

	BGHpA
ID17.08	pAAV: eHGT_390m-3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.09	pAAV: eHGT_375h-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.10	pAAV: eHGT_376h-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.11	pAAV: eHGT_390h-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.12	pAAV: eHGT_373m-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.13	pAAV: eHGT_375m-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.14	pAAV: eHGT_386m-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.15	pAAV: eHGT_387m-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID17.16	pAAV: eHGT_390m-hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID18.01	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_375h-BGHpA
ID18.02	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_376h-BGHpA
ID18.03	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_390h-BGHpA
ID18.04	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_373m-BGHpA
ID18.05	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_375m-BGHpA
ID18.06	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_386m-BGHpA
ID18.07	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_387m-BGHpA
ID18.08	pAAV: 3xhI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_390m-BGHpA
ID18.09	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_375h-BGHpA
ID18.10	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_376h-BGHpA
ID18.11	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_390h-BGHpA
ID18.12	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_373m-BGHpA
ID18.13	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_375m-BGHpA
ID18.14	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_386m-BGHpA
ID18.15	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_387m-BGHpA
ID18.16	pAAV: hI56i-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-eHGT_390m-BGHpA
ID19.01	pAAV: eHGT_375h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.02	pAAV: eHGT_376h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.03	pAAV: eHGT_390h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.04	pAAV: eHGT_373m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.05	pAAV: eHGT_375m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.06	pAAV: eHGT_386m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA

ID19.07	pAAV: eHGT_387m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.08	pAAV: eHGT_390m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-3xhI56i(core)-BGHpA
ID19.09	pAAV: eHGT_375h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.10	pAAV: eHGT_376h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.11	pAAV: eHGT_390h-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.12	pAAV: eHGT_373m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.13	pAAV: eHGT_375m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.14	pAAV: eHGT_386m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.15	pAAV: eHGT_387m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID19.16	pAAV: eHGT_390m-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-hI56i-BGHpA
ID20.01	pAAV: 3xhI56i(core)-eHGT_375h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.02	pAAV: 3xhI56i(core)-eHGT_376h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.03	pAAV: 3xhI56i(core)-eHGT_390h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.04	pAAV: 3xhI56i(core)-eHGT_373m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.05	pAAV: 3xhI56i(core)-eHGT_375m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.06	pAAV: 3xhI56i(core)-eHGT_386m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.07	pAAV: 3xhI56i(core)-eHGT_387m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.08	pAAV: 3xhI56i(core)-eHGT_390m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.09	pAAV: hI56i-eHGT_375h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.10	pAAV: hI56i-eHGT_376h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.11	pAAV: hI56i-eHGT_390h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.12	pAAV: hI56i-eHGT_373m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.13	pAAV: hI56i-eHGT_375m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.14	pAAV: hI56i-eHGT_386m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.15	pAAV: hI56i-eHGT_387m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID20.16	pAAV: hI56i-eHGT_390m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.01	pAAV: eHGT_375h-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.02	pAAV: eHGT_376h-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.03	pAAV: eHGT_390h-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.04	pAAV: eHGT_373m-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.05	pAAV: eHGT_375m-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.06	pAAV: eHGT_386m-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-

	BGHpA
ID21.07	pAAV: eHGT_387m-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.08	pAAV: eHGT_390m-3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.09	pAAV: eHGT_375h-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.10	pAAV: eHGT_376h-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.11	pAAV: eHGT_390h-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.12	pAAV: eHGT_373m-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.13	pAAV: eHGT_375m-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.14	pAAV: eHGT_386m-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.15	pAAV: eHGT_387m-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID21.16	pAAV: eHGT_390m-hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID22.01	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_375h-BGHpA
ID22.02	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_376h-BGHpA
ID22.03	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_390h-BGHpA
ID22.04	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_373m-BGHpA
ID22.05	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_375m-BGHpA
ID22.06	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_386m-BGHpA
ID22.07	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_387m-BGHpA
ID22.08	pAAV: 3xhI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_390m-BGHpA
ID22.09	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_375h-BGHpA
ID22.10	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_376h-BGHpA
ID22.11	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_390h-BGHpA
ID22.12	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_373m-BGHpA
ID22.13	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_375m-BGHpA
ID22.14	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_386m-BGHpA
ID22.15	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_387m-BGHpA
ID22.16	pAAV: hI56i-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-eHGT_390m-BGHpA
ID23.01	pAAV: eHGT_375h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.02	pAAV: eHGT_376h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.03	pAAV: eHGT_390h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.04	pAAV: eHGT_373m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.05	pAAV: eHGT_375m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA

ID23.06	pAAV: eHGT_386m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.07	pAAV: eHGT_387m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.08	pAAV: eHGT_390m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-3xhI56i(core)-BGHpA
ID23.09	pAAV: eHGT_375h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.10	pAAV: eHGT_376h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.11	pAAV: eHGT_390h-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.12	pAAV: eHGT_373m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.13	pAAV: eHGT_375m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.14	pAAV: eHGT_386m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.15	pAAV: eHGT_387m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID23.16	pAAV: eHGT_390m-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-hI56i-BGHpA
ID24.01	pAAV: hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.02	pAAV: hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.03	pAAV: hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.04	pAAV: hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.05	pAAV: hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.06	pAAV: hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.07	pAAV: hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.08	pAAV: hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.09	pAAV: hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.10	pAAV: hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.11	pAAV: hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.12	pAAV: hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.13	pAAV: hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA

ID24.14	pAAV: hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.15	pAAV: hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID24.16	pAAV: hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.01	pAAV: eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.02	pAAV: eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.03	pAAV: eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.04	pAAV: eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.05	pAAV: eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.06	pAAV: eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.07	pAAV: eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.08	pAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.09	pAAV: eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.10	pAAV: eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.11	pAAV: eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.12	pAAV: eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.13	pAAV: eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-

	BGHpA
ID25.14	pAAV: eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.15	pAAV: eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID25.16	pAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SLC6A1-P2A-SYFP2-WPRE3-BGHpA
ID26.01	pAAV: hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.02	pAAV: hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.03	pAAV: hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.04	pAAV: hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.05	pAAV: hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.06	pAAV: hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.07	pAAV: hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core22)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.08	pAAV: hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.09	pAAV: hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.10	pAAV: hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.11	pAAV: hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.12	pAAV: hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.13	pAAV: hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.14	pAAV: hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.15	pAAV: hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-

	hI56i(core)-eHGT_387m(core2)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID26.16	pAAV: hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.01	pAAV: eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.02	pAAV: eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.03	pAAV: eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.04	pAAV: eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.05	pAAV: eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.06	pAAV: eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.07	pAAV: eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.08	pAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.09	pAAV: eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-eHGT_375h(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.10	pAAV: eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-eHGT_376h(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.11	pAAV: eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-eHGT_390h(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.12	pAAV: eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-eHGT_373m(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.13	pAAV: eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-eHGT_375m(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.14	pAAV: eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-eHGT_386m(core)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA

ID27.15	pAAV: eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-eHGT_387m(core2)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA
ID27.16	pAAV: eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-eHGT_390m(core2)-hI56i(core)-minBGprom-SYFP2-P2A-SLC6A1-WPRE3-BGHpA

[0094] Subcomponent sequences within the larger vector sequences can be readily identified by one of ordinary skill in the art and based on the contents of the current disclosure (see FIG. 14). Nucleotides between identifiable and enumerated subcomponents reflect restriction enzyme recognition sites used in assembly (cloning) of the constructs, and in some cases, additional nucleotides do not convey any identifiable function. These segments of complete vector sequences can be adjusted based on use of different cloning strategies and/or vectors. In general, short 6-nucleotide palindromic sequences reflect vector construction artifacts that are not important to vector function.

[0095] In particular embodiments vectors (e.g., AAV) with capsids that cross the blood-brain barrier (BBB) are selected. In particular embodiments, vectors are modified to include capsids that cross the BBB. Examples of AAV with viral capsids that cross the blood brain barrier include AAV9 (Gombash et al., Front Mol Neurosci. 2014; 7:81), AAVrh.10 (Yang, et al., Mol Ther. 2014; 22(7): 1299-1309), AAV1R6, AAV1R7 (Albright et al., Mol Ther. 2018; 26(2): 510), rAAVrh.8 (Yang, et al., supra), AAV-BR1 (Marchio et al., EMBO Mol Med. 2016; 8(6): 592), AAV-PHP.S (Chan et al., Nat Neurosci. 2017; 20(8): 1172), AAV-PHP.B (Deverman et al., Nat Biotechnol. 2016; 34(2): 204), AAV-PPS (Chen et al., Nat Med. 2009; 15: 1215), and PHP.eB. In particular embodiments, the PHP.eB capsid differs from AAV9 such that, using AAV9 as a reference, amino acids starting at residue 586: S-AQ-A (SEQ ID NO: 78) are changed to S-DGTLAVPFK-A (SEQ ID NO: 79). In particular embodiments, PHP.eb refers to SEQ ID NO: 70.

[0096] AAV9 is a naturally occurring AAV serotype that, unlike many other naturally occurring serotypes, can cross the BBB following intravenous injection. It transduces large sections of the central nervous system (CNS), thus permitting minimally invasive treatments (Naso et al., BioDrugs. 2017; 31(4): 317), for example, as described in relation to clinical trials for the treatment of spinal muscular atrophy (SMA) syndrome by AveXis (AVXS-101, NCT03505099) and the treatment of CLN3 gene-Related Neuronal Ceroid-Lipofuscinosis (NCT03770572).

[0097] AAVrh.10, was originally isolated from rhesus macaques and shows low seropositivity in humans when compared with other common serotypes used for gene delivery applications (Selot et al., Front Pharmacol. 2017; 8: 441) and has been evaluated in clinical trials LYS-SAF302, LYSOGENE, and NCT03612869.

[0098] AAV1R6 and AAV1R7, two variants isolated from a library of chimeric AAV vectors (AAV1 capsid domains swapped into AAVrh.10), retain the ability to cross the BBB and transduce the CNS while showing significantly reduced hepatic and vascular endothelial transduction.

[0099] rAAVrh.8, also isolated from rhesus macaques, shows a global transduction of glial and neuronal cell types in regions of clinical importance following peripheral administration and also displays reduced peripheral tissue tropism compared to other vectors.

[0100] AAV-BR1 is an AAV2 variant displaying the NRGTEWD (SEQ ID NO: 80) epitope that was isolated during in vivo screening of a random AAV display peptide library. It shows high specificity accompanied by high transgene expression in the brain with minimal off-target affinity (including for the liver) (Körbelin et al., EMBO Mol Med. 2016; 8(6): 609).

[0101] AAV-PHP.S (Addgene, Watertown, MA) is a variant of AAV9 generated with the CREATE method that encodes the 7-mer sequence QAVRTSL (SEQ ID NO: 81), transduces neurons in the enteric nervous system, and strongly transduces peripheral sensory afferents entering the spinal cord and brain stem.

[0102] AAV-PHP.B (Addgene, Watertown, MA) is a variant of AAV9 generated with the CREATE method that encodes the 7-mer sequence TLAVPFK (SEQ ID NO: 82). It transfers genes throughout the CNS with higher efficiency than AAV9 and transduces the majority of astrocytes and neurons across multiple CNS regions.

[0103] AAV-PPS, an AAV2 variant crated by insertion of the DSPAHPS (SEQ ID NO: 83) epitope into the capsid of AAV2, shows a dramatically improved brain tropism relative to AAV2.

[0104] For additional information regarding capsids that cross the blood brain barrier, see Chan et al., Nat. Neurosci. 2017 Aug: 20(8): 1172-1179.

[0105] (ii) Compositions for Administration. Artificial expression constructs and vectors of the present disclosure (referred to herein as physiologically active components) can be formulated with a carrier that is suitable for administration to a cell, tissue slice, animal (e.g., mouse, non-human primate), or human. Physiologically active components within compositions described herein can be prepared in neutral forms, as freebases, or as pharmacologically acceptable salts.

[0106] Pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like.

[0107] Carriers of physiologically active components can include solvents, dispersion media,

vehicles, coatings, diluents, isotonic and absorption delaying agents, buffers, solutions, suspensions, colloids, and the like. The use of such carriers for physiologically active components is well known in the art. Except insofar as any conventional media or agent is incompatible with the physiologically active components, it can be used with compositions as described herein.

[0108] The phrase "pharmaceutically-acceptable carriers" refer to carriers that do not produce an allergic or similar untoward reaction when administered to a human, and in particular embodiments, when administered intravenously (e.g. at the retro-orbital plexus).

[0109] In particular embodiments, compositions can be formulated for intravenous, intraparenchymal, intraocular, intravitreal, parenteral, subcutaneous, intracerebro-ventricular, intramuscular, intrathecal, intraspinal, intraperitoneal, oral or nasal inhalation, or by direct injection in or application to one or more cells, tissues, or organs.

[0110] Compositions may include liposomes, lipids, lipid complexes, microspheres, microparticles, nanospheres, and/or nanoparticles.

[0111] The formation and use of liposomes is generally known to those of skill in the art. Liposomes have been developed with improved serum stability and circulation half-times (see, for instance, U.S. Pat. No. 5,741,516). Further, various methods of liposome and liposome like preparations as potential drug carriers have been described (see, for instance U.S. Pat. Nos. 5,567,434; 5,552,157; 5,565,213; 5,738,868; and 5,795,587).

[0112] The disclosure also provides for pharmaceutically acceptable nanocapsule formulations of the physiologically active components. Nanocapsules can generally entrap compounds in a stable and reproducible way (Quintanar-Guerrero *et al.*, *Drug Dev Ind Pharm* 24(12):1113-1128, 1998; Quintanar-Guerrero *et al.*, *Pharm Res.* 15(7):1056-1062, 1998; Quintanar-Guerrero *et al.*, J*. Microencapsul*. 15(1):107-119, 1998; Douglas *et al.*, *Crit Rev Ther Drug Carrier Syst* 3(3):233-261, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles can be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present disclosure. Such particles can be easily made, as described in Couvreur *et al.*, *J Pharm Sci* 69(2):199-202, 1980; Couvreur *et al.*, *Crit Rev Ther Drug Carrier Syst*. 5(1)1-20, 1988; zur Muhlen *et al.*, *Eur J Pharm Biopharm*, 45(2):149-155, 1998; Zambaux *et al.*, *J Control Release* 50(1-3):31-40, 1998; and U.S. Pat. No. 5,145,684.

[0113] Injectable compositions can include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U.S. Pat. No. 5,466,468). For delivery via injection, the form is sterile and fluid to the extent that it can be delivered by syringe. In particular embodiments, it is stable under the conditions of

manufacture and storage, and optionally contains one or more preservative compounds against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion, and/or by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and/or antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In various embodiments, the preparation will include an isotonic agent(s), for example, sugar(s) or sodium chloride. Prolonged absorption of the injectable compositions can be accomplished by including in the compositions of agents that delay absorption, for example, aluminum monostearate and gelatin. Injectable compositions can be suitably buffered, if necessary, and the liquid diluent first rendered isotonic with sufficient saline or glucose.

[0114] Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. As indicated, under ordinary conditions of storage and use, these preparations can contain a preservative to prevent the growth of microorganisms.

[0115] Sterile compositions can be prepared by incorporating the physiologically active component in an appropriate amount of a solvent with other optional ingredients (e.g., as enumerated above), followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized physiologically active components into a sterile vehicle that contains the basic dispersion medium and the required other ingredients (e.g., from those enumerated above). In the case of sterile powders for the preparation of sterile injectable solutions, preferred methods of preparation can be vacuum-drying and freeze-drying techniques which yield a powder of the physiologically active components plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0116] Oral compositions may be in liquid form, for example, as solutions, syrups or suspensions, or may be presented as a drug product for reconstitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically

acceptable excipients such as binding agents (e.g., pregelatinized maize starch, polyvinyl pyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). Tablets may be coated by methods well-known in the art.

[0117] Inhalable compositions can be delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

[0118] Compositions can also include microchip devices (U.S. Pat. No. 5,797,898), ophthalmic formulations (Bourlais *et al*., *Prog Retin Eye Res*, 17(1):33-58, 1998), transdermal matrices (U.S. Pat. No. 5,770,219 and U.S. Pat. No. 5,783,208) and feedback-controlled delivery (U.S. Pat. No. 5,697,899).

[0119] Supplementary active ingredients can also be incorporated into the compositions.

[0120] Typically, compositions can include at least 0.1% of the physiologically active components or more, although the percentage of the physiologically active components may, of course, be varied and may conveniently be between 1 or 2% and 70% or 80% or more or 0.5-99% of the weight or volume of the total composition. Naturally, the amount of physiologically active components in each physiologically-useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of compositions and dosages may be desirable.

[0121] In particular embodiments, for administration to humans, compositions should meet sterility, pyrogenicity, and the general safety and purity standards as required by United States Food and Drug Administration (FDA) or other applicable regulatory agencies in other countries.

[0122] (iii) Cell Lines Including Artificial Expression Constructs. The present disclosure includes cells including an artificial expression construct described herein. A cell that has been transformed with an artificial expression construct can be used for many purposes, including in neuroanatomical studies, assessments of functioning and/or non-functioning proteins, and drug

screens that assess the regulatory properties of enhancers.

[0123] A variety of host cell lines can be used, but in particular embodiments, the cell is a mammalian cell. In particular embodiments, the artificial expression construct includes an I56i enhancer, a concatenated core thereof, or a concatenated core thereof and one or more of the enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof and the cell line can include, for example, human, primate, or murine cells. Cell lines which can be utilized for transgenesis in the present disclosure also include primary cell lines derived from living tissue such as rat or mouse brains and organotypic cell cultures, including brain slices from animals such as rats or mice. The PC12 cell line (available from the American Type Culture Collection, ATCC, Manassas, VA) has been shown to express a number of neuronal marker proteins in response to Neuronal Growth Factor (NGF). The PC12 cell line is considered to be a neuronal cell line and is applicable for use with this disclosure. JAR cells (available from ATCC) are a platelet derived cell-line that express some neuronal genes, such as the serotonin transporter gene, and may be used with embodiments described herein.

[0124] WO 91/13150 describes a variety of cell lines, including neuronal cell lines, and methods of producing them. Similarly, WO 97/39117 describes a neuronal cell line and methods of producing such cell lines. The neuronal cell lines disclosed in these patent applications are applicable for use in the present disclosure.

[0125] In particular embodiments, "neuronal" describes something that is of, related to, or includes, neuronal cells. Neuronal cells are defined by the presence of an axon and dendrites. The term "neuronal-specific" refers to something that is found, or an activity that occurs, in neuronal cells or cells derived from neuronal cells, but is not found in or occur in, or is not found substantially in or occur substantially in, non-neuronal cells or cells not derived from neuronal cells, for example glial cells such as astrocytes or oligodendrocytes.

[0126] In particular embodiments, non-neuronal cell lines may be used, including mouse embryonic stem cells. Cultured mouse embryonic stem cells can be used to analyze expression of genetic constructs using transient transfection with plasmid constructs. Mouse embryonic stem cells are pluripotent and undifferentiated. These cells can be maintained in this undifferentiated state by Leukemia Inhibitory Factor (LIF). Withdrawal of LIF induces differentiation of the embryonic stem cells. In culture, the stem cells form a variety of differentiated cell types. Differentiation is caused by the expression of tissue specific transcription factors, allowing the function of an enhancer sequence to be evaluated. (See for example Fiskerstrand *et al.*, *FEBS Lett* 458: 171-174, 1999).

[0127] Methods to differentiate stem cells into neuronal cells include replacing a stem cell culture media with a media including basic fibroblast growth factor (bFGF) heparin, an N2 supplement (*e.g.,* transferrin, insulin, progesterone, putrescine, and selenite), laminin and polyornithine. A process to produce myelinating oligodendrocytes from stem cells is described in Hu, *et al.,* 2009, *Nat. Protoc.* 4:1614-22. Bibel, *et al.,* 2007, *Nat. Protoc.* 2:1034-43 describes a protocol to produce glutamatergic neurons from stem cells while Chatzi, *et al.,* 2009, *Exp. Neurol.* 217:407-16 describes a procedure to produce GABAergic neurons. This procedure includes exposing stem cells to all-trans-RA for three days. After subsequent culture in serum-free neuronal induction medium including Neurobasal medium supplemented with B27, bFGF and EGF, 95% GABA neurons develop

[0128] U.S. Publication No. 2012/0329714 describes use of prolactin to increase neural stem cell numbers while U.S. Publication No. 2012/0308530 describes a culture surface with amino groups that promotes neuronal differentiation into neurons, astrocytes and oligodendrocytes. Thus, the fate of neural stem cells can be controlled by a variety of extracellular factors. Commonly used factors include brain derived growth factor (BDNF; Shetty and Turner, 1998, *J. Neurobiol.* 35:395-425); fibroblast growth factor (bFGF; U.S. Pat. No.5,766,948; FGF-1, FGF-2); Neurotrophin-3 (NT-3) and Neurotrophin-4 (NT-4); Caldwell, *et al.,* 2001, *Nat. Biotechnol.* 1;19:475-9); ciliary neurotrophic factor (CNTF); BMP-2 (U.S. Pat. Nos. 5,948,428 and 6,001,654); isobutyl 3-methylxanthine; leukemia inhibitory growth factor (LIF; U.S. Patent No. 6,103,530); somatostatin; amphiregulin; neurotrophins (*e.g.,* cyclic adenosine monophosphate; epidermal growth factor (EGF); dexamethasone (glucocorticoid hormone); forskolin; GDNF family receptor ligands; potassium; retinoic acid (U.S. Patent No. 6,395,546); tetanus toxin; and transforming growth factor-α and TGF-β (U.S. Pat. Nos. 5,851,832 and 5,753,506).

[0129] In particular embodiments, yeast one-hybrid systems may also be used to identify compounds that inhibit specific protein/DNA interactions, such as transcription factors for the I56i enhancer, a core thereof, or a concatenated core thereof and one or more of the enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof.

[0130] Transgenic animals are described below. Cell lines may also be derived from such transgenic animals. For example, primary tissue culture from transgenic mice (e.g., also as described below) can provide cell lines with the artificial expression construct already integrated into the genome. (for an example see MacKenzie & Quinn, *Proc Natl Acad Sci USA* 96: 15251-15255, 1999).

[0131] (iv) Transgenic Animals. Another aspect of the disclosure includes transgenic animals, the

genome of which contains an artificial expression construct including an I56i enhancer, core thereof, or a concatenated core thereof and one or more of the enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof operatively linked to a heterologous coding sequence. In particular embodiments, the genome of a transgenic animal includes CN2720, CN2721, CN2722, CN2732, CN3213, CN3322, CN3323, CN3887, CN3888, CN2972, CN2973, CN2974, CN2975, CN2976, ID10.01, ID10.02, ID10.03, ID10.04, ID10.05, ID10.06, ID10.07, ID10.08, ID10.09, ID10.10, ID10.11, ID10.12, ID10.13, ID10.14, ID10.15, ID10.16, ID10.17, ID10.18, ID10.19, ID10.20, ID10.21, ID10.22, ID10.23, ID10.24, ID10.25, ID10.26, ID10.27, ID10.28, ID10.29, ID10.30, ID10.31, ID10.32, ID11.01, ID11.02, ID11.03, ID11.04, ID11.05, ID11.06, ID11.07, ID11.08, ID11.09, ID11.10, ID11.11, ID11.12, ID11.13, ID11.14, ID11.15, ID11.16, ID12.01, ID12.02, ID12.03, ID12.04, ID12.05, ID12.06, ID12.07, ID12.08, ID12.09, ID12.10, ID12.11, ID12.12, ID12.13, ID12.14, ID12.15, ID12.16, ID13.01, ID13.02, ID13.03, ID13.04, ID13.05, ID13.06, ID13.07, ID13.08, ID13.09, ID13.10, ID13.11, ID13.12, ID13.13, ID13.14, ID13.15, ID13.16, ID14.01, ID14.02, ID14.03, ID14.04, ID14.05, ID14.06, ID14.07, ID14.08, ID14.09, ID14.10, ID14.11, ID14.12, ID14.13, ID14.14, ID14.15, ID14.16, ID15.01, ID15.02, ID15.03, ID15.04, ID15.05, ID15.06, ID15.07, ID15.08, ID15.09, ID15.10, ID15.11, ID15.12, ID15.13, ID15.14, ID15.15, ID15.16, ID16.01, ID16.02, ID16.03, ID16.04, ID16.05, ID16.06, ID16.07, ID16.08, ID16.09, ID16.10, ID16.11, ID16.12, ID16.13, ID16.14, ID16.15, ID16.16, ID17.01, ID17.02, ID17.03, ID17.04, ID17.05, ID17.06, ID17.07, ID17.08, ID17.09, ID17.10, ID17.11, ID17.12, ID17.13, ID17.14, ID17.15, ID17.16, ID18.01, ID18.02, ID18.03, ID18.04, ID18.05, ID18.06, ID18.07, ID18.08, ID18.09, ID18.10, ID18.11, ID18.12, ID18.13, ID18.14, ID18.15, ID18.16, ID19.01, ID19.02, ID19.03, ID19.04, ID19.05, ID19.06, ID19.07, ID19.08, ID19.09, ID19.10, ID19.11, ID19.12, ID19.13, ID19.14, ID19.15, ID19.16, ID20.01, ID20.02, ID20.03, ID20.04, ID20.05, ID20.06, ID20.07, ID20.08, ID20.09, ID20.10, ID20.11, ID20.12, ID20.13, ID20.14, ID20.15, ID20.16, ID21.01, ID21.02, ID21.03, ID21.04, ID21.05, ID21.06, ID21.07, ID21.08, ID21.09, ID21.10, ID21.11, ID21.12, ID21.13, ID21.14, ID21.15, ID21.16, ID22.01, ID22.02, ID22.03, ID22.04, ID22.05, ID22.06, ID22.07, ID22.08, ID22.09, ID22.10, ID22.11, ID22.12, ID22.13, ID22.14, ID22.15, ID22.16, ID23.01, ID23.02, ID23.03, ID23.04, ID23.05, ID23.06, ID23.07, ID23.08, ID23.09, ID23.10, ID23.11, ID23.12, ID23.13, ID23.14, ID23.15, ID23.16, ID24.01, ID24.02, ID24.03, ID24.04, ID24.05, ID24.06, ID24.07, ID24.08, ID24.09, ID24.10, ID24.11, ID24.12, ID24.13, ID24.14, ID24.15, ID24.16, ID25.01, ID25.02, ID25.03, ID25.04, ID25.05, ID25.06, ID25.07, ID25.08, ID25.09, ID25.10, ID25.11, ID25.12, ID25.13, ID25.14, ID25.15, ID25.16, ID26.01, ID26.02, ID26.03, ID26.04, ID26.05, ID26.06, ID26.07,

ID26.08, ID26.09, ID26.10, ID26.11, ID26.12, ID26.13, ID26.14, ID26.15, ID26.16, ID27.01, ID27.02, ID27.03, ID27.04, ID27.05, ID27.06, ID27.07, ID27.08, ID27.09, ID27.10, ID27.11, ID27.12, ID27.13, ID27.14, ID27.15, or ID27.16. In particular embodiments, when a non-integrating vector is utilized, a transgenic animal includes an artificial expression construct including an I56i enhancer, a core thereof, or a concatenated core thereof and one or more of the enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof and/or CN2720, CN2721, CN2722, CN2732, CN3213, CN3322, CN3323, CN3887, CN3888, CN2972, CN2973, CN2974, CN2975, CN2976, ID10.01, ID10.02, ID10.03, ID10.04, ID10.05, ID10.06, ID10.07, ID10.08, ID10.09, ID10.10, ID10.11, ID10.12, ID10.13, ID10.14, ID10.15, ID10.16, ID10.17, ID10.18, ID10.19, ID10.20, ID10.21, ID10.22, ID10.23, ID10.24, ID10.25, ID10.26, ID10.27, ID10.28, ID10.29, ID10.30, ID10.31, ID10.32, ID11.01, ID11.02, ID11.03, ID11.04, ID11.05, ID11.06, ID11.07, ID11.08, ID11.09, ID11.10, ID11.11, ID11.12, ID11.13, ID11.14, ID11.15, ID11.16, ID12.01, ID12.02, ID12.03, ID12.04, ID12.05, ID12.06, ID12.07, ID12.08, ID12.09, ID12.10, ID12.11, ID12.12, ID12.13, ID12.14, ID12.15, ID12.16, ID13.01, ID13.02, ID13.03, ID13.04, ID13.05, ID13.06, ID13.07, ID13.08, ID13.09, ID13.10, ID13.11, ID13.12, ID13.13, ID13.14, ID13.15, ID13.16, ID14.01, ID14.02, ID14.03, ID14.04, ID14.05, ID14.06, ID14.07, ID14.08, ID14.09, ID14.10, ID14.11, ID14.12, ID14.13, ID14.14, ID14.15, ID14.16, ID15.01, ID15.02, ID15.03, ID15.04, ID15.05, ID15.06, ID15.07, ID15.08, ID15.09, ID15.10, ID15.11, ID15.12, ID15.13, ID15.14, ID15.15, ID15.16, ID16.01, ID16.02, ID16.03, ID16.04, ID16.05, ID16.06, ID16.07, ID16.08, ID16.09, ID16.10, ID16.11, ID16.12, ID16.13, ID16.14, ID16.15, ID16.16, ID17.01, ID17.02, ID17.03, ID17.04, ID17.05, ID17.06, ID17.07, ID17.08, ID17.09, ID17.10, ID17.11, ID17.12, ID17.13, ID17.14, ID17.15, ID17.16, ID18.01, ID18.02, ID18.03, ID18.04, ID18.05, ID18.06, ID18.07, ID18.08, ID18.09, ID18.10, ID18.11, ID18.12, ID18.13, ID18.14, ID18.15, ID18.16, ID19.01, ID19.02, ID19.03, ID19.04, ID19.05, ID19.06, ID19.07, ID19.08, ID19.09, ID19.10, ID19.11, ID19.12, ID19.13, ID19.14, ID19.15, ID19.16, ID20.01, ID20.02, ID20.03, ID20.04, ID20.05, ID20.06, ID20.07, ID20.08, ID20.09, ID20.10, ID20.11, ID20.12, ID20.13, ID20.14, ID20.15, ID20.16, ID21.01, ID21.02, ID21.03, ID21.04, ID21.05, ID21.06, ID21.07, ID21.08, ID21.09, ID21.10, ID21.11, ID21.12, ID21.13, ID21.14, ID21.15, ID21.16, ID22.01, ID22.02, ID22.03, ID22.04, ID22.05, ID22.06, ID22.07, ID22.08, ID22.09, ID22.10, ID22.11, ID22.12, ID22.13, ID22.14, ID22.15, ID22.16, ID23.01, ID23.02, ID23.03, ID23.04, ID23.05, ID23.06, ID23.07, ID23.08, ID23.09, ID23.10, ID23.11, ID23.12, ID23.13, ID23.14, ID23.15, ID23.16, ID24.01, ID24.02, ID24.03, ID24.04, ID24.05, ID24.06, ID24.07, ID24.08, ID24.09, ID24.10, ID24.11, ID24.12, ID24.13, ID24.14, ID24.15, ID24.16, ID25.01, ID25.02, ID25.03,

ID25.04, ID25.05, ID25.06, ID25.07, ID25.08, ID25.09, ID25.10, ID25.11, ID25.12, ID25.13, ID25.14, ID25.15, ID25.16, ID26.01, ID26.02, ID26.03, ID26.04, ID26.05, ID26.06, ID26.07, ID26.08, ID26.09, ID26.10, ID26.11, ID26.12, ID26.13, ID26.14, ID26.15, ID26.16, ID27.01, ID27.02, ID27.03, ID27.04, ID27.05, ID27.06, ID27.07, ID27.08, ID27.09, ID27.10, ID27.11, ID27.12, ID27.13, ID27.14, ID27.15, or ID27.16 within one or more of its cells.

[0132] Detailed methods for producing transgenic animals are described in U.S. Pat. No. 4,736,866. Transgenic animals may be of any nonhuman species, but preferably include nonhuman primates (NHPs), sheep, horses, cattle, pigs, goats, dogs, cats, rabbits, chickens, and rodents such as guinea pigs, hamsters, gerbils, rats, mice, and ferrets.

[0133] In particular embodiments, construction of a transgenic animal results in an organism that has an engineered construct present in all cells in the same genomic integration site. Thus, cell lines derived from such transgenic animals will be consistent in as much as the engineered construct will be in the same genomic integration site in all cells and hence will suffer the same position effect variegation. In contrast, introducing genes into cell lines or primary cell cultures can give rise to heterologous expression of the construct. A disadvantage of this approach is that the expression of the introduced DNA may be affected by the specific genetic background of the host animal.

[0134] As indicated above in relation to cell lines, the artificial expression constructs of this disclosure can be used to genetically modify mouse embryonic stem cells using techniques known in the art. Typically, the artificial expression construct is introduced into cultured murine embryonic stem cells. Transformed ES cells are then injected into a blastocyst from a host mother and the host embryo re-implanted into the mother. This results in a chimeric mouse whose tissues are composed of cells derived from both the embryonic stem cells present in the cultured cell line and the embryonic stem cells present in the host embryo. Usually the mice from which the cultured ES cells used for transgenesis are derived are chosen to have a different coat color from the host mouse into whose embryos the transformed cells are to be injected. Chimeric mice will then have a variegated coat color. As long as the germ-line tissue is derived, at least in part, from the genetically modified cells, then the chimeric mice crossed with an appropriate strain can produce offspring that will carry the transgene.

[0135] In addition to the methods of delivery described above, the following techniques are also contemplated as alternative methods of delivering artificial expression constructs to target cells or targeted tissues and organs of an animal, and in particular, to cells, organs, or tissues of a vertebrate mammal: sonophoresis (*e.g.,* ultrasound, as described in U.S. Pat. No. 5,656,016); intraosseous injection (U.S. Pat. No. 5,779,708); microchip devices (U.S. Pat. No. 5,797,898);

ophthalmic formulations (Bourlais *et al.*, *Prog Retin Eye Res*, 17(1):33-58, 1998); transdermal matrices (U.S. Pat. No. 5,770,219 and U.S. Pat. No. 5,783,208); feedback-controlled delivery (U.S. Pat. No. 5,697,899), and any other delivery method available and/or described elsewhere in the disclosure.

[0136] (v) Methods of Use. In particular embodiments, a composition including a physiologically active component described herein is administered to a subject to result in a physiological effect.

[0137] In particular embodiments, the disclosure includes the use of the artificial expression constructs described herein to modulate expression of a heterologous gene which is either partially or wholly encoded in a location downstream to that enhancer in an engineered sequence. Thus, there are provided herein methods of use of the disclosed artificial expression constructs in the research, study, and potential development of medicaments for preventing, treating or ameliorating the symptoms of a disease, dysfunction, or disorder.

[0138] Particular embodiments include methods of administering to a subject an artificial expression construct that includes an I56i enhancer, a core thereof, or a concatenated core thereof and one or more of the enhancers selected from eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, eHGT_390m, and cores thereof and/or CN2720, CN2721, CN2722, CN2732, CN3213, CN3322, CN3323, CN3887, CN3888, CN2972, CN2973, CN2974, CN2975, CN2976, ID10.01, ID10.02, ID10.03, ID10.04, ID10.05, ID10.06, ID10.07, ID10.08, ID10.09, ID10.10, ID10.11, ID10.12, ID10.13, ID10.14, ID10.15, ID10.16, ID10.17, ID10.18, ID10.19, ID10.20, ID10.21, ID10.22, ID10.23, ID10.24, ID10.25, ID10.26, ID10.27, ID10.28, ID10.29, ID10.30, ID10.31, ID10.32, ID11.01, ID11.02, ID11.03, ID11.04, ID11.05, ID11.06, ID11.07, ID11.08, ID11.09, ID11.10, ID11.11, ID11.12, ID11.13, ID11.14, ID11.15, ID11.16, ID12.01, ID12.02, ID12.03, ID12.04, ID12.05, ID12.06, ID12.07, ID12.08, ID12.09, ID12.10, ID12.11, ID12.12, ID12.13, ID12.14, ID12.15, ID12.16, ID13.01, ID13.02, ID13.03, ID13.04, ID13.05, ID13.06, ID13.07, ID13.08, ID13.09, ID13.10, ID13.11, ID13.12, ID13.13, ID13.14, ID13.15, ID13.16, ID14.01, ID14.02, ID14.03, ID14.04, ID14.05, ID14.06, ID14.07, ID14.08, ID14.09, ID14.10, ID14.11, ID14.12, ID14.13, ID14.14, ID14.15, ID14.16, ID15.01, ID15.02, ID15.03, ID15.04, ID15.05, ID15.06, ID15.07, ID15.08, ID15.09, ID15.10, ID15.11, ID15.12, ID15.13, ID15.14, ID15.15, ID15.16, ID16.01, ID16.02, ID16.03, ID16.04, ID16.05, ID16.06, ID16.07, ID16.08, ID16.09, ID16.10, ID16.11, ID16.12, ID16.13, ID16.14, ID16.15, ID16.16, ID17.01, ID17.02, ID17.03, ID17.04, ID17.05, ID17.06, ID17.07, ID17.08, ID17.09, ID17.10, ID17.11, ID17.12, ID17.13, ID17.14, ID17.15, ID17.16, ID18.01, ID18.02, ID18.03, ID18.04, ID18.05, ID18.06, ID18.07, ID18.08, ID18.09, ID18.10, ID18.11, ID18.12, ID18.13, ID18.14, ID18.15, ID18.16, ID19.01, ID19.02, ID19.03, ID19.04, ID19.05, ID19.06,

ID19.07, ID19.08, ID19.09, ID19.10, ID19.11, ID19.12, ID19.13, ID19.14, ID19.15, ID19.16, ID20.01, ID20.02, ID20.03, ID20.04, ID20.05, ID20.06, ID20.07, ID20.08, ID20.09, ID20.10, ID20.11, ID20.12, ID20.13, ID20.14, ID20.15, ID20.16, ID21.01, ID21.02, ID21.03, ID21.04, ID21.05, ID21.06, ID21.07, ID21.08, ID21.09, ID21.10, ID21.11, ID21.12, ID21.13, ID21.14, ID21.15, ID21.16, ID22.01, ID22.02, ID22.03, ID22.04, ID22.05, ID22.06, ID22.07, ID22.08, ID22.09, ID22.10, ID22.11, ID22.12, ID22.13, ID22.14, ID22.15, ID22.16, ID23.01, ID23.02, ID23.03, ID23.04, ID23.05, ID23.06, ID23.07, ID23.08, ID23.09, ID23.10, ID23.11, ID23.12, ID23.13, ID23.14, ID23.15, ID23.16, ID24.01, ID24.02, ID24.03, ID24.04, ID24.05, ID24.06, ID24.07, ID24.08, ID24.09, ID24.10, ID24.11, ID24.12, ID24.13, ID24.14, ID24.15, ID24.16, ID25.01, ID25.02, ID25.03, ID25.04, ID25.05, ID25.06, ID25.07, ID25.08, ID25.09, ID25.10, ID25.11, ID25.12, ID25.13, ID25.14, ID25.15, ID25.16, ID26.01, ID26.02, ID26.03, ID26.04, ID26.05, ID26.06, ID26.07, ID26.08, ID26.09, ID26.10, ID26.11, ID26.12, ID26.13, ID26.14, ID26.15, ID26.16, ID27.01, ID27.02, ID27.03, ID27.04, ID27.05, ID27.06, ID27.07, ID27.08, ID27.09, ID27.10, ID27.11, ID27.12, ID27.13, ID27.14, ID27.15, or ID27.16 as described herein to drive expression of a gene in GABAergic neurons and astrocytes. The subject can be an isolated cell, a network of cells, a tissue slice, an experimental animal, a veterinary animal, or a human.

[0139] As is well known in the medical arts, dosages for any one subject depends upon many factors, including the subject's size, surface area, age, the particular compound to be administered, sex, time and route of administration, general health, and other drugs being administered concurrently. Dosages for the compounds of the disclosure will vary, but, in particular embodiments, a dose could be from 10^5 to 10^{100} copies of an artificial expression construct of the disclosure. In particular embodiments, a patient receiving intravenous, intraparenchymal, intraspinal, retro-orbital, or intrathecal administration can be infused with from 10^6 to 10^{22} copies of the artificial expression construct.

[0140] Treating subjects includes delivering therapeutically effective amounts. Therapeutically effective amounts include those that provide effective amounts, prophylactic treatments and/or therapeutic treatments.

[0141] An "effective amount" is the amount of a composition necessary to result in a desired physiological change in the subject. Effective amounts are often administered for research purposes. Effective amounts disclosed herein can cause a statistically-significant effect in an animal model or in vitro assay relevant to the assessment of an SLC6A1-associated disorder's development, progression, and/or resolution.

[0142] A "prophylactic treatment" includes a treatment administered to a subject who does not

display signs or symptoms of an SLC6A1-associated disorder or displays only early signs or symptoms of an SLC6A1-associated disorder such that treatment is administered for the purpose of diminishing or decreasing the risk of developing the SLC6A1-associated disorder further. Thus, a prophylactic treatment functions as a preventative treatment against an SLC6A1-associated disorder. In particular embodiments, prophylactic treatments reduce, delay, or prevent the worsening of an SLC6A1-associated disorder.

[0143] A "therapeutic treatment" includes a treatment administered to a subject who displays symptoms or signs of an SLC6A1-associated disorder and is administered to the subject for the purpose of diminishing or eliminating those signs or symptoms of the SLC6A1-associated disorder. The therapeutic treatment can reduce, control, or eliminate the presence or activity of the SLC6A1-associated disorder and/or reduce control or eliminate side effects of the SLC6A1-associated disorder.

[0144] Function as an effective amount, prophylactic treatment or therapeutic treatment are not mutually exclusive, and in particular embodiments, administered dosages may accomplish more than one treatment type.

[0145] In particular embodiments, methods to determine the efficacy of the treatments using constructs disclosed herein will be measured before treatment, during the first year after treatment, and at other times. In particular embodiments, efficacy of the treatments using constructs disclosed herein will be determined to be effective if the evaluated measurements can be maintained at a normal or reduced from previous disorder levels.

[0146] Therapeutically effective amounts can be assessed using developmental tests for cognitive and motor function. One of ordinary skill in the art is aware of proper conditions under which to assess cognitive functioning, which can include various tests that are commonly employed. Representative tests include neuropsychological tests such as the Continuous Performance Test (CPT), Wisconsin Card Sorting Test, Trailmaking A+B, the Mini Mental State Exam (MMSE), List Learning (Verbal Memory), Digit Sequencing Task (Working Memory), Token Motor Task (Motor Speed), Category Instances (Semantic Fluency), Controlled Oral Word Association Test (Letter Fluency), Tower of London Test (Executive Function), Symbol Coding (Attention and Motor Speed), Affective Interference Test-Delayed Recognition Task, Stroop Test, the Brief Assessment of Cognition in Schizophrenia (BACS; includes a number of the tests above), tests included in the Measurement and Treatment Research to Improve Cognition in Schizophrenia battery (MATRICS), and the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-cog)

[0147] In particular embodiments, methods to determine the efficacy of the treatments in mild-to-

moderate intellectual disability can be measured by observing any positive change in the clinical symptoms of the patient. Classification of intellectual disability can be determined using the American Association on Intellectual and Developmental Disabilities (AAIDD) or the *Diagnostic and Statistical* Manual of Mental Disorders, 5th Edition (DSM-5), which is published by the American Psychiatric Association (Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders; Board on the Health of Select Populations; Board on Children, Youth, and Families; Institute of Medicine; Division of Behavioral and Social Sciences and Education; The National Academies of Sciences, Engineering, and Medicine; Boat TF, Wu JT, editors. Mental Disorders and Disabilities Among Low-Income Children. Washington (DC): National Academies Press (US); 2015 Oct 28. 9, Clinical Characteristics of Intellectual Disabilities). In particular embodiments, efficacy of the treatment of intellectual disability can be measured using IQ, severity based on daily skills, severity based on intensity of support needed, or SSI listings criteria. SSI listings do not specify severity levels, rather they describe the standards for meeting a listing level severity.

[0148] In particular embodiments, methods to determine the efficacy of the treatments in epilepsy can be the treatments effect on reducing or preventing seizures. In particular embodiments, the methods provided may reduce or prevent one or more different types of seizures. Ideally, the methods of the disclosure result in a total prevention of seizures. However, the disclosure also encompasses methods in which the instances of seizures are decreased by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80% or at least 90%.

[0149] Generally, a seizure can include convulsions, repetitive movements, unusual sensations, and combinations thereof. Seizures can be categorized as focal seizures (also referred to as partial seizures) and generalized seizures. Focal seizures affect only one side of the brain, while generalized seizures affect both sides of the brain. Specific types of focal seizures include simple focal seizures, complex focal seizures, and secondarily generalized seizures. Simple focal seizures can be restricted or focused on a particular lobe (*e.g.*, temporal lobe, frontal lobe, parietal lobe, or occipital lobe). Complex focal seizures generally affect a larger part of one hemisphere than simple focal seizures, but commonly originate in the temporal lobe or the frontal lobe. When a focal seizure spreads from one side (hemisphere) to both sides of the brain, the seizure is referred to as a secondarily generalized seizure. Specific types of generalized seizures include absences (also referred to as petit mal seizures), tonic seizures, atonic seizures, myoclonic seizures, tonic clonic seizures (also referred to as grand mal seizures), and clonic seizures.

[0150] In particular embodiments, methods described herein may reduce the frequency of

seizures, reduce the severity of seizures, change the type of seizures (*e.g.*, from a more severe type to a less severe type), or a combination thereof in a patient after treatment compared to the absence of treatment (*e.g.*, before treatment), or compared to treatment with an alternative conventional treatment.

[0151] In particular embodiments, methods to determine the efficacy of the treatments on speech difficulties can include the use of speech assessments. In particular embodiments, a speech-language pathologist assesses the verbal expression of the patient. Assessment tools include the diadochokinetic (DDK) rate to measure the repetitions of sounds within a set period of time; the Motor ABC test, the Beery-Buktenica Developmental Test of Visual-Motor Coordination (Beery VMI); MRI, CT, or EMG tests; Voice Handicap Index (VHI); Frenchay Dysarthria Assessment (FDA); Radbound Dysarthria Assessment (RDA); oral-motor examinations; and other speech and language examinations.

[0152] In particular embodiments, methods to determine the efficacy of the treatments on behavioral problems can include amelioration of at least one clinical symptom and/or at least one physical parameter associated with the behavioral problem. In particular embodiments, the behavioral problem can include attention deficit disorder (ADD) or attention deficit hyperactivity disorder (ADHD). An effective treatment results in an improvement in the patient's ADHD rating scale IV (ARS-IV), ADHD self-report scale (ASRS), clinical global impression (CGI), and/or cognitive functions.

[0153] ADHD rating scale IV (ARS-IV) rates the following behaviors: 1. Fails to give close attention to details or makes careless mistakes in work. 2. Fidgets with hands or feet or squirms in seat. 3. Has difficulty sustaining attention in tasks or play activities. 4. Leaves seat in situations in which remaining seated is expected. 5. Does not seem to listen when spoken to directly. 6. Runs about or climbs excessively in situations in which it is inappropriate. 7. Does not follow through on instructions and fails to finish work. 8. Has difficulty playing or engaging in leisure activities quietly. 9. Has difficulty organizing tasks and activities. 10. Is "on the go" or acts as if "driven by a motor." 11. Avoids tasks that require sustained mental effort. 12. Talks excessively. 13. Loses things necessary for tasks or activities. 14. Blurts out answers before questions have been completed. 15. Is easily distracted. 16. Has difficulty awaiting turn. 17. Is forgetful in daily activities. 18. Interrupts or intrudes on others.

[0154] Kessler et al (*Psychological Medicine,* 35:245-256, 2005) report the WHO adult ADHD self-report scale (ASRS), for use in the general population. The ASRS Symptom Checklist is a self-reported questionnaire used to assist in the diagnosis of adult ADHD.

[0155] The clinical global impression-improvement scale (CGI-I) is a 7 point scale that requires

the clinician to assess how much the patient's illness has improved or worsened relative to a baseline state at the beginning of the intervention, and rated as: 1, very much improved; 2, much improved; 3, minimally improved; 4, no change; 5, minimally worse; 6, much worse; or 7, very much worse.

[0156] In particular embodiments, treatment efficacy in autism spectrum disorder may be assessed. A variety of standardized evaluation schemes are available for monitoring the course, severity, and spectrum of functional impairments in patients with autism spectrum disorder or suspected to be at risk for autism-spectrum disorder. Such schemes also may be used to assess the evolution of autism symptoms over time or in response to treatment. Of these, the Autism Diagnostic Observation Schedule (ADOS-2, in its most current iteration, described in Gotham et al. J Autism Dev Disord. 2007 Apr;37(4):613-27) is uniquely useful for patients of wide age ranges as it has a variety of modules that account for the developmental level and age of the patient. It includes a standardized administration of interactive activities introduced by the examiner which are designed to elicit social interactions, communication, and repetitive behaviors for the purpose of diagnosing an autism spectrum disorder, with procedures optimized for patients from less than 48 months through adulthood. Also useful for evaluating communication impairment in autism spectrum disorder is the Expressive One Word Picture Vocabulary Test (EOWPVT), which assesses verbal expression and the ability to name and generate words (described in Chapman et al. Early Hum Dev. 2015 Jun; 91(6): 373-379.) Additional metrics that may be used to gauge improvement of ASD patients include the caregiver-administered Aberrant Behavior Checklist (ABC, see for e.g. Kaat et al. J Autism Dev Disord. 2014 May;44(5): l 103-16.) and Autism Treatment Evaluation Checklist (ATEC, see for e.g. Geier et al. Mental Health Research in Intellectual Disabilities 2013; 6: 255-67). Additionally, a modified version of the Clinical Global Impressions scale may be used to judge patient progress.

[0157] In particular embodiments, methods to determine the efficacy of the treatments on neurological signs can include treatments that improve a patient's symptoms or otherwise reduces, alleviates, or minimizes adverse conditions. In particular embodiments, the neurological signs can include ataxia, hypotonia, and other movement disorders. In particular embodiments, changes in muscle tone, strength, reflexes, hyperflexibility, posture, endurance, MRI, CT, EMG, or EEG scans can be assessed to measure effective treatment of hypotonia. In particular embodiments, an assessment of writing and eating skills, eye movements, gait, balance and coordination, speech, MRI, or CT scans can be used to measure effective treatment of ataxia.

[0158] Numerous movement disorders affect infants and children and a number of motor and developmental tests can be utilized to assess therapeutically effective amounts within this context.

Examples include the the Peabody Developmental Motor Scale (PDMS-II), Alberta Infant Motor Scale (AIMS), Bayley Scales of Infant and Toddler Development®-Third Edition (Bayley-III), or the Comprehensive Developmental Inventory for Infants and Toddlers (CDIIT).

[0159] PDMS-II is a skill-based measure of gross and fine motor development for infants and children from birth through 5 years of age. This tool separates motor development into gross and fine motor skills. Through a combination of the composite scores for the gross and fine motor skills, the examiner has a reliable estimate of the child's motor skills. It includes 4 gross motor and 2 fine motor subtests, as follows: Reflexes (gross motor); Stationary (gross motor); Locomotion (gross motor); Object Manipulation (gross motor); Grasping (fine motor); and Visual-Motor Integration (fine motor).

[0160] Scoring the PDMS-II relies on raw scores, percentiles, standard scores, and age equivalents for the subtests, and quotients for the composites. Raw scores are total points accumulated by a child on a subtest. Developmental ages are often used to convey information to parents of young children. Age equivalents for PDMS-II are called "motor ages" which convey to parents that their child is "passing" on items that a child of a certain chronological age would typically pass. Age equivalents for PDMS-II subtests are generated from Table C.1 in the PDMS-II manual or by PDMS-II software scoring and report systems.

[0161] AIMS is a 58-item observational measure of infant motor performance for use from birth through the age of independent walking (18 months). It assesses the sequential development of motor milestones in terms of progressive development and integration of antigravity muscle control. The test assesses infant movement in 4 positions: prone, supine, sitting, and standing. The AIMS total score is calculated by summing the scores for the 58 items with a range of scores between 0 and 58. Higher scores indicate more mature motor development. The infant's score can then be converted to a percentile and compared with age-equivalent peers from the normative sample.

[0162] Bayley-III offers a standardized assessment of cognitive and motor development for children between 1 and 42 months of age. The assessment measures cognitive, communication, physical, social/emotional, and adaptive areas of development to identify children with developmental delays. The test includes 5 scales of development: Cognitive Scale, Language Scale, Motor Scale, Social Emotional Scale, and Adaptive Behavior Scale. It is possible to present results for developmental age corresponding to each subscale vs chronological age.

[0163] The diagnostic test of the CDIIT is one of the child developmental tests covering 5 developmental subtests used for children aged 3 to 72 months.

[0164] The amount of expression constructs and time of administration of such compositions will

be within the purview of the skilled artisan having benefit of the present teachings. It is likely, however, that the administration of effective amounts of the disclosed compositions may be achieved by a single administration, such as for example, a single injection of sufficient numbers of infectious particles to provide an effect in the subject. Alternatively, in some circumstances, it may be desirable to provide multiple, or successive administrations of the artificial expression construct compositions or other genetic constructs, either over a relatively short, or a relatively prolonged period of time, as may be determined by the individual overseeing the administration of such compositions. For example, the number of infectious particles administered to a mammal may be 10^7, 10^8, 10^9, 10^{10}, 10^{11}, 10^{12}, 10^{13}, or even higher, infectious particles/ml given either as a single dose or divided into two or more administrations as may be required to achieve an intended effect. In fact, in certain embodiments, it may be desirable to administer two or more different expression constructs in combination to achieve a desired effect.

[0165] In certain circumstances it will be desirable to deliver the artificial expression construct in suitably formulated compositions disclosed herein either by pipette, retro-orbital injection, subcutaneously, intraocularly, intravitreally, parenterally, subcutaneously, intravenously, intraparenchymally, intracerebro-ventricularly, intramuscularly, intrathecally, intraspinally, intraperitoneally, by oral or nasal inhalation, or by direct application or injection to one or more cells, tissues, or organs. The methods of administration may also include those modalities as described in U.S. Pat. No. 5,543,158; U.S. Pat. No. 5,641,515 and U.S. Pat. No. 5,399,363.

[0166] (vi) Kits and Commercial Packages. Kits and commercial packages contain an artificial expression construct described herein. The artificial expression construct can be isolated. In particular embodiments, the components of an expression product can be isolated from each other. In particular embodiments, the expression product can be within a vector, within a viral vector, within a cell, within a tissue slice or sample, and/or within a transgenic animal. Such kits may further include one or more reagents, restriction enzymes, peptides, therapeutics, pharmaceutical compounds, or means for delivery of the compositions such as syringes, injectables, and the like.

[0167] Embodiments of a kit or commercial package will also contain instructions regarding use of the included components, for example, in basic research, electrophysiological research, neuroanatomical research, and/or the research and/or treatment of a disorder, disease or condition.

[0168] The Exemplary Embodiments and Experimental Examples below are included to demonstrate particular embodiments of the disclosure. Those of ordinary skill in the art should recognize in light of the present disclosure that many changes can be made to the specific

embodiments disclosed herein and still obtain a like or similar result without departing from the spirit and scope of the disclosure.

[0169] (vii) Exemplary Embodiments.

1. An artificial expression construct including (i) a first enhancer including a core of an I56i enhancer; (ii) a second enhancer including one or more of eHGT_387m, eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_390m, or a core thereof; (iii) a promoter; and (iv) a heterologous encoding sequence.
2. The artificial expression construct of embodiment 1, wherein the first enhancer is adjacent to the second enhancer.
3. The artificial expression construct of embodiment 1, wherein the first enhancer is not adjacent to the second enhancer.
4. The artificial expression construct of any of embodiments 1-3, wherein the first enhancer is 5' of the second enhancer.
5. The artificial expression construct of any of embodiments 1-3, wherein the second enhancer is 5' of the first enhancer.
6. The artificial expression construct of any of embodiments 1-5, wherein the core of the I56i enhancer is a I56i human core or a I56i zebrafish core.
7. The artificial expression construct of embodiment 6, wherein the core includes the sequence as set forth in SEQ ID NO: 4 or 5 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 4 or 5.
8. The artificial expression construct of any of embodiments 1-7, wherein the core of the I56i enhancer is concatenated.
9. The artificial expression construct of embodiment 8, wherein the concatenated core of the I56i enhancer includes 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the I56i human core or the I56i zebrafish core.
10. The artificial expression construct of embodiment 9, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence set forth in SEQ ID NO: 4 and/or SEQ ID NO: 5 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 4 or 5.
11. The artificial expression construct of embodiment 9, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence set forth in SEQ ID NO: 4 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 4.
12. The artificial expression construct of embodiment 9, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence set forth in SEQ ID NO: 5 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 5.

13. The artificial expression construct of embodiment 9, including 3 copies of SEQ ID NO: 4 or 3 copies of a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 4.

14. The artificial expression construct of embodiment 9, including 3 copies of SEQ ID NO: 5 or 3 copies of a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 5.

15. The artificial expression construct of embodiment 8, wherein the concatenated core of the I56i enhancer has the sequence as set forth in SEQ ID NO: 6 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 6.

16. The artificial expression construct of embodiment 8, wherein the concatenated core of the I56i enhancer has the sequence as set forth in SEQ ID NO: 7 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 7.

17. The artificial expression construct of embodiment 8, wherein the second enhancer includes eHGT_387m or eHGT_390m or a sequence having at least 90% sequence identity to the sequence as set forth for eHGT_387m or eHGT_390m.

18. The artificial expression construct of any of embodiments 1-17, wherein the second enhancer is a core of an enhancer selected from eHGT_387m, eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, or eHGT_390m.

19. The artificial expression construct of any of embodiments 1-18, wherein the second enhancer is a core of eHGT_387m or eHGT_390m or a sequence having at least 90% sequence identity to the sequence as set forth for a core of eHGT_387m or eHGT_390m.

20. The artificial expression construct of embodiment 19, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence set forth in SEQ ID NO: 84 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 84.

21. The artificial expression construct of embodiment 19 or 20, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence set forth in SEQ ID NO: 85 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 85.

22. The artificial expression construct of any of embodiments 1-21, wherein the second enhancer core is concatenated with the core of the I56i enhancer to create a combination concatenated enhancer.

23. The artificial expression construct of embodiment 22, wherein the combination concatenated enhancer includes the sequence as set forth in SEQ ID NO: 95, or SEQ ID NO: 86 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 95 or SEQ ID NO: 86.

24. The artificial expression construct of embodiment 22 or 23, wherein the combination concatenated enhancer includes 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core of the second enhancer and 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core of the I56i enhancer.

25. The artificial expression construct of any of embodiments 22-24, including 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the combination concatenated enhancer.

26. The artificial expression construct of any of embodiments 22-25, wherein the combination concatenated enhancer includes 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of eHGT_387m(core2); 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of eHGT_390m(core2); and/or 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core of the I56i enhancer.

27. The artificial expression construct of any of embodiments 22-26, wherein the combination concatenated enhancer includes 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence set forth in SEQ ID NO: 95 or SEQ ID NO: 86.

28. The artificial expression construct of any of embodiments 22-27, wherein the combination concatenated enhancer includes 3 copies of the sequence set forth in SEQ ID NO: 95 or SEQ ID NO: 86 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 95 or SEQ ID NO: 86.

29. The artificial expression construct of any of embodiments 22-28, wherein the combination concatenated enhancer includes the sequence set forth in SEQ ID NO: 88 or SEQ ID NO: 89 or a sequence having at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 88 or SEQ ID NO: 89.

30. The artificial expression construct of any of embodiments 1-29, wherein the heterologous encoding sequence encodes GAT1.

31. The artificial expression construct of embodiment 30, wherein the heterologous encoding sequence is a codon-optimized SLC6A1 gene.

32. The artificial expression construct of embodiment 31, wherein heterologous encoding sequence has at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 22, 25, 28, 31, 34, 38, 39, 40, 41, 42, 43, 44, or 45.

33. The artificial expression construct of embodiment 31 or 32, wherein heterologous encoding sequence has the sequence as set forth in SEQ ID NO: 22, 25, 28, 31, 34, 38, 39, 40, 41, 42, 43, 44, or 45.

34. The artificial expression construct any of embodiments 1-33, wherein the heterologous encoding sequence encodes an effector element, or an expressible element.

35. The artificial expression construct of embodiment 34, wherein the effector element includes a reporter protein or a functional molecule.

36. The artificial expression construct of embodiment 35, wherein the reporter protein includes a fluorescent protein.

37. The artificial expression construct of embodiment 35, wherein the functional molecule includes a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, microRNA, homologous recombination donor cassette, or a designer receptor exclusively activated by designer drug (DREADD).

38. The artificial expression construct of embodiment 34, wherein the expressible element includes a non-functional molecule.

39. The artificial expression construct of embodiment 38, wherein the non-functional molecule includes a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, microRNA, homologous recombination donor cassette, or DREADD.

40. The artificial expression construct of any of embodiments 1-39, wherein the artificial expression construct is associated with a capsid that crosses the blood brain barrier.

41. The artificial expression construct of embodiment 40, wherein the capsid includes PHP.eB, AAV-BR1, AAV-PHP.S, AAV-PHP.B, or AAV-PPS.

42. The artificial expression construct of any of embodiments 1-41, wherein the artificial expression construct includes or encodes a skipping element.

43. The artificial expression construct of embodiment 42, wherein the skipping element includes a 2A peptide or an internal ribosome entry site (IRES).

44. The artificial expression construct of embodiment 43, wherein the 2A peptide includes T2A, P2A, E2A, or F2A.

45. The artificial expression construct of any of embodiments 1-44, wherein the artificial expression construct includes or encodes a set of features selected from: a concatenated core of an I56i enhancer, eHGT_387m, eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_390m, eHGT_387m(core2), eHGT_375h(core), eHGT_376h(core), eHGT_390h(core), eHGT_373m(core), eHGT_375m(core), eHGT_386m(core), eHGT_390m(core2), AAV, scAAV, rAAv, minBglobin, CMV, minCMV, minRho, minRho*, fluorescent protein, codon-optimized SLC6A1, 4X2C, Cre, iCre, dgCre, FlpO, tTA2, SP10, WPRE, WPRE3, hGHpA, and/or BGHpA.

46. The artificial expression construct of any of embodiments 1-45, wherein the artificial expression construct includes the features of: CN2721, CN3213, CN2720, CN2722, CN2732,

CN3322, CN3323, CN3887, CN3888, CN2972, CN2973, CN2974, CN2975, CN2976, ID10.01, ID10.02, ID10.03, ID10.04, ID10.05, ID10.06, ID10.07, ID10.08, ID10.09, ID10.10, ID10.11, ID10.12, ID10.13, ID10.14, ID10.15, ID10.16, ID10.17, ID10.18, ID10.19, ID10.20, ID10.21, ID10.22, ID10.23, ID10.24, ID10.25, ID10.26, ID10.27, ID10.28, ID10.29, ID10.30, ID10.31, ID10.32, ID11.01, ID11.02, ID11.03, ID11.04, ID11.05, ID11.06, ID11.07, ID11.08, ID11.09, ID11.10, ID11.11, ID11.12, ID11.13, ID11.14, ID11.15, ID11.16, ID12.01, ID12.02, ID12.03, ID12.04, ID12.05, ID12.06, ID12.07, ID12.08, ID12.09, ID12.10, ID12.11, ID12.12, ID12.13, ID12.14, ID12.15, ID12.16, ID13.01, ID13.02, ID13.03, ID13.04, ID13.05, ID13.06, ID13.07, ID13.08, ID13.09, ID13.10, ID13.11, ID13.12, ID13.13, ID13.14, ID13.15, ID13.16, ID14.01, ID14.02, ID14.03, ID14.04, ID14.05, ID14.06, ID14.07, ID14.08, ID14.09, ID14.10, ID14.11, ID14.12, ID14.13, ID14.14, ID14.15, ID14.16, ID15.01, ID15.02, ID15.03, ID15.04, ID15.05, ID15.06, ID15.07, ID15.08, ID15.09, ID15.10, ID15.11, ID15.12, ID15.13, ID15.14, ID15.15, ID15.16, ID16.01, ID16.02, ID16.03, ID16.04, ID16.05, ID16.06, ID16.07, ID16.08, ID16.09, ID16.10, ID16.11, ID16.12, ID16.13, ID16.14, ID16.15, ID16.16, ID17.01, ID17.02, ID17.03, ID17.04, ID17.05, ID17.06, ID17.07, ID17.08, ID17.09, ID17.10, ID17.11, ID17.12, ID17.13, ID17.14, ID17.15, ID17.16, ID18.01, ID18.02, ID18.03, ID18.04, ID18.05, ID18.06, ID18.07, ID18.08, ID18.09, ID18.10, ID18.11, ID18.12, ID18.13, ID18.14, ID18.15, ID18.16, ID19.01, ID19.02, ID19.03, ID19.04, ID19.05, ID19.06, ID19.07, ID19.08, ID19.09, ID19.10, ID19.11, ID19.12, ID19.13, ID19.14, ID19.15, ID19.16, ID20.01, ID20.02, ID20.03, ID20.04, ID20.05, ID20.06, ID20.07, ID20.08, ID20.09, ID20.10, ID20.11, ID20.12, ID20.13, ID20.14, ID20.15, ID20.16, ID21.01, ID21.02, ID21.03, ID21.04, ID21.05, ID21.06, ID21.07, ID21.08, ID21.09, ID21.10, ID21.11, ID21.12, ID21.13, ID21.14, ID21.15, ID21.16, ID22.01, ID22.02, ID22.03, ID22.04, ID22.05, ID22.06, ID22.07, ID22.08, ID22.09, ID22.10, ID22.11, ID22.12, ID22.13, ID22.14, ID22.15, ID22.16, ID23.01, ID23.02, ID23.03, ID23.04, ID23.05, ID23.06, ID23.07, ID23.08, ID23.09, ID23.10, ID23.11, ID23.12, ID23.13, ID23.14, ID23.15, ID23.16, ID24.01, ID24.02, ID24.03, ID24.04, ID24.05, ID24.06, ID24.07, ID24.08, ID24.09, ID24.10, ID24.11, ID24.12, ID24.13, ID24.14, ID24.15, ID24.16, ID25.01, ID25.02, ID25.03, ID25.04, ID25.05, ID25.06, ID25.07, ID25.08, ID25.09, ID25.10, ID25.11, ID25.12, ID25.13, ID25.14, ID25.15, ID25.16, ID26.01, ID26.02, ID26.03, ID26.04, ID26.05, ID26.06, ID26.07, ID26.08, ID26.09, ID26.10, ID26.11, ID26.12, ID26.13, ID26.14, ID26.15, ID26.16, ID27.01, ID27.02, ID27.03, ID27.04, ID27.05, ID27.06, ID27.07, ID27.08, ID27.09, ID27.10, ID27.11, ID27.12, ID27.13, ID27.14, ID27.15, or ID27.16.

47. A vector including an artificial expression construct of any of embodiments 1-46.

48. The vector of embodiment 47, wherein the vector includes a viral vector.

49. The vector of embodiment 48, wherein the viral vector includes a recombinant adeno-associated viral (AAV) vector.

50. A transgenic cell including an artificial expression construct of any of embodiments 1-46 and/or a vector of embodiments 47-49.

51. The transgenic cell of embodiment 50, wherein the transgenic cell is a GABAergic neuron or an astrocyte.

52. The transgenic cell of embodiment 50 or 51, wherein the transgenic cell is murine, human, or non-human primate.

53. A non-human transgenic animal including an artificial expression construct of any of embodiments 1-46 and/or a vector of embodiments 47-49 and/or a transgenic cell of embodiment 51 or 52.

54. The non-human transgenic animal of embodiment 53, wherein the non-human transgenic animal is a mouse or a non-human primate.

55. An administrable composition including an artificial expression construct of any of embodiments 1-46 and/or a vector of embodiments 47-49 and/or a transgenic cell of embodiment 51 or 52.

56. A kit including an artificial expression construct of any of embodiments 1-46 and/or a vector of embodiments 47-49 and/or a transgenic cell of embodiment 51 or 52 and/or a non-human transgenic animal of embodiment 53 or 54 and/or an administrable composition of embodiment 55.

57. A method for expressing a gene within a population of cells in vivo or in vitro, the method including providing the administrable composition of embodiment 55 in a sufficient dosage and for a sufficient time to a sample or subject including the population of cells thereby expressing the gene within the population of cells.

58. The method of embodiment 57, wherein the gene encodes GAT1, an effector element, or an expressible element.

59. The method of embodiment 58, wherein the gene is a codon-optimized SLC6A1 gene.

60. The method of embodiment 57 or 58, wherein the gene has at least 90% sequence identity to a sequence as set forth in SEQ ID NO: 22, 25, 28, 31, or 34.

61. The method of any of embodiments 57-60, wherein the gene has the sequence as set forth in SEQ ID NO: 22, 25, 28, 31, or 34.

62. The method of embodiment 57, wherein the population of cells includes GABAergic neurons and astrocytes.

63. The method of embodiment 57, wherein the providing includes pipetting.

64. The method of embodiment 63, wherein the pipetting is to a brain slice.
65. The method of embodiment 64, wherein the brain slice includes GABAergic neurons and astrocytes.
66. The method of embodiment 64 or 65, wherein the brain slice is murine, human, or non-human primate.
67. The method of embodiment 57, wherein the providing includes administering to a living subject.
68. The method of embodiment 67, wherein the living subject is a human, non-human primate, or a mouse.
69. The method of embodiment 67, wherein the living subject has an SL6CA1-associated disorder.
70. The method of embodiment 69, wherein the SL6CA1-associated disorder includes impaired cognitive function, impaired motor function, mild-to-moderate intellectual disability, epilepsy, speech difficulty, attention deficit disorder, attention deficit hyperactivity disorder, or an autism spectrum disorder.
71. The method of any of embodiments 67-70, wherein the administering to a living subject is through injection.
72. The method of embodiment 71, wherein the injection includes intravenous injection, intraparenchymal injection into brain tissue, intracerebroventricular (ICV) injection, intra-cisterna magna (ICM) injection, or intrathecal injection.
73. An artificial expression construct having a sequence with at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 90, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 93, or SEQ ID NO: 94.
74. An artificial expression construct having the sequence as set forth in SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 90, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 93, or SEQ ID NO: 94.

[0170] (viii) Closing Paragraphs. Variants of the sequences disclosed and referenced herein are also included. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological activity can be found using computer programs well known in the art, such as DNASTAR™ (Madison, Wisconsin) software. Preferably, amino acid changes in the protein variants disclosed herein are conservative amino acid changes, i.e., substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains.

[0171] In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and generally can be made without altering a biological activity of a resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g., Watson et al. *Molecular Biology of the Gene,* 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p. 224). Naturally occurring amino acids are generally divided into conservative substitution families as follows: Group 1: Alanine (Ala), Glycine (Gly), Serine (Ser), and Threonine (Thr); Group 2: (acidic): Aspartic acid (Asp), and Glutamic acid (Glu); Group 3: (acidic; also classified as polar, negatively charged residues and their amides): Asparagine (Asn), Glutamine (Gln), Asp, and Glu; Group 4: Gln and Asn; Group 5: (basic; also classified as polar, positively charged residues): Arginine (Arg), Lysine (Lys), and Histidine (His); Group 6 (large aliphatic, nonpolar residues): Isoleucine (Ile), Leucine (Leu), Methionine (Met), Valine (Val) and Cysteine (Cys); Group 7 (uncharged polar): Tyrosine (Tyr), Gly, Asn, Gln, Cys, Ser, and Thr; Group 8 (large aromatic residues): Phenylalanine (Phe), Tryptophan (Trp), and Tyr; Group 9 (non-polar): Proline (Pro), Ala, Val, Leu, Ile, Phe, Met, and Trp; Group 11 (aliphatic): Gly, Ala, Val, Leu, and Ile; Group 10 (small aliphatic, nonpolar or slightly polar residues): Ala, Ser, Thr, Pro, and Gly; and Group 12 (sulfur-containing): Met and Cys. Additional information can be found in Creighton (1984) Proteins, W.H. Freeman and Company.

[0172] In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, J. Mol. Biol. 157(1), 105-32). Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: Ile (+4.5); Val (+4.2); Leu (+3.8); Phe (+2.8); Cys (+2.5); Met (+1.9); Ala (+1.8); Gly (−0.4); Thr (−0.7); Ser (−0.8); Trp (−0.9); Tyr (−1.3); Pro (−1.6); His (−3.2); Glutamate (−3.5); Gln (−3.5); aspartate (−3.5); Asn (−3.5); Lys (−3.9); and Arg (−4.5).

[0173] It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, i.e., still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity.

[0174] As detailed in U.S. Pat. No. 4,554,101, the following hydrophilicity values have been

assigned to amino acid residues: Arg (+3.0); Lys (+3.0); aspartate (+3.0±1); glutamate (+3.0±1); Ser (+0.3); Asn (+0.2); Gln (+0.2); Gly (0); Thr (−0.4); Pro (−0.5±1); Ala (−0.5); His (−0.5); Cys (−1.0); Met (−1.3); Val (−1.5); Leu (−1.8); Ile (−1.8); Tyr (−2.3); Phe (−2.5); Trp (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

[0175] As outlined above, amino acid substitutions may be based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like.

[0176] As indicated elsewhere, variants of gene sequences can include codon optimized variants, sequence polymorphisms, splice variants, and/or mutations that do not affect the function of an encoded product to a statistically-significant degree.

[0177] Variants of the protein, nucleic acid, and gene sequences disclosed herein also include sequences with at least 70% sequence identity, 80% sequence identity, 85% sequence, 90% sequence identity, 95% sequence identity, 96% sequence identity, 97% sequence identity, 98% sequence identity, or 99% sequence identity to the protein, nucleic acid, or gene sequences disclosed herein.

[0178] "% sequence identity" refers to a relationship between two or more sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between protein, nucleic acid, or gene sequences as determined by the match between strings of such sequences. "Identity" (often referred to as "similarity") can be readily calculated by known methods, including those described in: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, NY (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, NY (1994); Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, NJ (1994); Sequence Analysis in Molecular Biology (Von Heijne, G., ed.) Academic Press (1987); and Sequence Analysis Primer (Gribskov, M. and Devereux, J., eds.) Oxford University Press, NY (1992). Preferred methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations may be performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR, Inc., Madison, Wisconsin). Multiple alignment of the sequences can also be performed using the Clustal method

of alignment (Higgins and Sharp CABIOS, 5, 151-153 (1989) with default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Relevant programs also include the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, Wisconsin); BLASTP, BLASTN, BLASTX (Altschul, et al., J. Mol. Biol. 215:403-410 (1990); DNASTAR (DNASTAR, Inc., Madison, Wisconsin); and the FASTA program incorporating the Smith-Waterman algorithm (Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Publisher: Plenum, New York, N.Y.. Within the context of this disclosure it will be understood that where sequence analysis software is used for analysis, the results of the analysis are based on the "default values" of the program referenced. As used herein "default values" will mean any set of values or parameters, which originally load with the software when first initialized.

[0179] Variants also include nucleic acid molecules that hybridizes under stringent hybridization conditions to a sequence disclosed herein and provide the same function as the reference sequence. Exemplary stringent hybridization conditions include an overnight incubation at 42 °C in a solution including 50% formamide, 5XSSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5XDenhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1XSSC at 50 °C. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, moderately high stringency conditions include an overnight incubation at 37°C in a solution including 6XSSPE (20XSSPE=3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50 °C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5XSSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

[0180] The term concatenate is broadly used to describe linking together into a chain or series. It is used to describe the linking together of nucleotide or amino acid sequences into a single nucleotide or amino acid sequence, respectively. The term “concatamerize” should be interpreted to recite: “concatenate.”

[0181] As will be understood by one of ordinary skill in the art, each embodiment disclosed herein can comprise, consist essentially of or consist of its particular stated element, step, ingredient or component. Thus, the terms "include" or "including" should be interpreted to recite: "comprise, consist of, or consist essentially of." The transition term "comprise" or "comprises" means has, but is not limited to, and allows for the inclusion of unspecified elements, steps, ingredients, or components, even in major amounts. The transitional phrase "consisting of" excludes any element, step, ingredient or component not specified. The transition phrase "consisting essentially of" limits the scope of the embodiment to the specified elements, steps, ingredients or components and to those that do not materially affect the embodiment. A material effect would cause a statistically significant reduction in targeted expression in GABAergic neurons and astrocytes utilizing an artificial expression construct disclosed herein.

[0182] In particular embodiments, artificial means not naturally occurring.

[0183] Unless otherwise indicated, all numbers expressing quantities of ingredients, properties such as molecular weight, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless indicated to the contrary, the numerical parameters set forth in the specification and attached claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits and by applying ordinary rounding techniques. When further clarity is required, the term "about" has the meaning reasonably ascribed to it by a person skilled in the art when used in conjunction with a stated numerical value or range, i.e. denoting somewhat more or somewhat less than the stated value or range, to within a range of ±20% of the stated value; ±19% of the stated value; ±18% of the stated value; ±17% of the stated value; ±16% of the stated value; ±15% of the stated value; ±14% of the stated value; ±13% of the stated value; ±12% of the stated value; ±11% of the stated value; ±10% of the stated value; ±9% of the stated value; ±8% of the stated value; ±7% of the stated value; ±6% of the stated value; ±5% of the stated value; ±4% of the stated value; ±3% of the stated value; ±2% of the stated value; or ±1% of the stated value.

[0184] Notwithstanding that the numerical ranges and parameters setting forth the broad scope of the invention are approximations, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors necessarily resulting from the standard deviation found in their respective testing measurements.

[0185] The terms "a," "an," "the" and similar referents used in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. Recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., "such as") provided herein is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the invention.

[0186] Groupings of alternative elements or embodiments of the invention disclosed herein are not to be construed as limitations. Each group member may be referred to and claimed individually or in any combination with other members of the group or other elements found herein. It is anticipated that one or more members of a group may be included in, or deleted from, a group for reasons of convenience and/or patentability. When any such inclusion or deletion occurs, the specification is deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

[0187] Certain embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Of course, variations on these described embodiments will become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventor expects skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

[0188] Furthermore, numerous references have been made to patents, printed publications, journal articles and other written text throughout this specification (referenced materials herein). Each of the referenced materials are individually incorporated herein by reference in their entirety for their referenced teaching.

[0189] In closing, it is to be understood that the embodiments of the invention disclosed herein are illustrative of the principles of the present invention. Other modifications that may be employed

are within the scope of the invention. Thus, by way of example, but not of limitation, alternative configurations of the present invention may be utilized in accordance with the teachings herein. Accordingly, the present invention is not limited to that precisely as shown and described.

[0190] The particulars shown herein are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of various embodiments of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for the fundamental understanding of the invention, the description taken with the drawings and/or examples making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

[0191] Definitions and explanations used in the present disclosure are meant and intended to be controlling in any future construction unless clearly and unambiguously modified in the following examples or when application of the meaning renders any construction meaningless or essentially meaningless. In cases where the construction of the term would render it meaningless or essentially meaningless, the definition should be taken from Webster's Dictionary, 3rd Edition or a dictionary known to those of ordinary skill in the art, such as the Oxford Dictionary of Biochemistry and Molecular Biology (Ed. Anthony Smith, Oxford University Press, Oxford, 2004).

CLAIMS

What is claimed is:

1. An artificial expression construct comprising a concatenated hI56i(core) enhancer having the sequence as set forth in SEQ ID NO: 6, an eHGT_387m enhancer having the sequence as set forth in SEQ ID NO: 14, a promoter, and a GAT1 protein encoding sequence.
2. An artificial expression construct comprising (i) a first enhancer having a core of an I56i enhancer; (ii) a second enhancer comprising one or more of eHGT_387m, eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_390m, or a core thereof; (iii) a promoter; and (iv) a heterologous encoding sequence.
3. The artificial expression construct of claim 2, wherein the first enhancer is adjacent to the second enhancer.
4. The artificial expression construct of claim 2, wherein the first enhancer is not adjacent to the second enhancer.
5. The artificial expression construct of claim 2, wherein the first enhancer is 5' of the second enhancer.
6. The artificial expression construct of claim 2, wherein the second enhancer is 5' of the first enhancer.
7. The artificial expression construct of claim 2, wherein the core of the I56i enhancer is a I56i human core or a I56i zebrafish core.
8. The artificial expression construct of claim 7, wherein the core comprises the sequence as set forth in SEQ ID NOs: 4 or 5.
9. The artificial expression construct of claim 2, wherein the core of the I56i enhancer is concatenated.
10. The artificial expression construct of claim 9, wherein the concatenated core of the I56i enhancer has 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the I56i human core or the I56i zebrafish core.
11. The artificial expression construct of claim 10, having 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence as set forth in SEQ ID NO: 4 and/or SEQ ID NO: 5.
12. The artificial expression construct of claim 10, having 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence as set forth in SEQ ID NO: 4.
13. The artificial expression construct of claim 10, having 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence as set forth in SEQ ID NO: 5.
14. The artificial expression construct of claim 10, having 3 copies of SEQ ID NO: 4.
15. The artificial expression construct of claim 10, having 3 copies of SEQ ID NO: 5.

16. The artificial expression construct of claim 9, wherein the concatenated core of the I56i enhancer has the sequence as set forth in SEQ ID NO: 6.
17. The artificial expression construct of claim 9, wherein the concatenated core of the I56i enhancer has the sequence as set forth in SEQ ID NO: 7.
18. The artificial expression construct of claim 9, wherein the second enhancer comprises eHGT_387m or eHGT_390m.
19. The artificial expression construct of claim 2, wherein the second enhancer is a core of an enhancer selected from eHGT_387m, eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, or eHGT_390m.
20. The artificial expression construct of claim 2, wherein the second enhancer is a core of eHGT_387m or eHGT_390m.
21. The artificial expression construct of claim 20, having 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence as set forth in SEQ ID NO: 84.
22. The artificial expression construct of claim 20, having 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence as set forth in SEQ ID NO: 85.
23. The artificial expression construct of claim 2, wherein the second enhancer core is concatenated with the core of the I56i enhancer to create a combination concatenated enhancer.
24. The artificial expression construct of claim 23, wherein the combination concatenated enhancer comprises the sequence as set forth in SEQ ID NO: 95, or SEQ ID NO: 86.
25. The artificial expression construct of claim 23, wherein the combination concatenated enhancer has 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core of the second enhancer and 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core of the I56i enhancer.
26. The artificial expression construct of claim 23, having 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the combination concatenated enhancer.
27. The artificial expression construct of claim 23, wherein the combination concatenated enhancer has 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of eHGT_387 m(core2); 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of eHGT_390m(core2); and/or 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the core of the I56i enhancer.
28. The artificial expression construct of claim 23, wherein the combination concatenated enhancer has 2, 3, 4, 5, 6, 7, 8, 9, or 10 copies of the sequence as set forth in SEQ ID NO: 95 or SEQ ID NO: 86.
29. The artificial expression construct of claim 23, wherein the combination concatenated enhancer has 3 copies of the sequence as set forth in SEQ ID NO: 95 or SEQ ID NO: 86.
30. The artificial expression construct of claim 23, wherein the combination concatenated

enhancer has the sequence as set forth in SEQ ID NO: 88 or SEQ ID NO: 89.

31. The artificial expression construct of claim 2, wherein the heterologous encoding sequence encodes GAT1.

32. The artificial expression construct of claim 31, wherein the heterologous encoding sequence is a codon-optimized SLC6A1 gene.

33. The artificial expression construct of claim 32, wherein heterologous encoding sequence has the sequence as set forth in SEQ ID NOs: 22, 25, 28, 31, 34, or 38-45.

34. The artificial expression construct of claim 2, wherein the heterologous encoding sequence encodes an effector element, or an expressible element.

35. The artificial expression construct of claim 34, wherein the effector element comprises a reporter protein or a functional molecule.

36. The artificial expression construct of claim 35, wherein the reporter protein comprises a fluorescent protein.

37. The artificial expression construct of claim 35, wherein the functional molecule comprises a functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, microRNA, homologous recombination donor cassette, or a designer receptor exclusively activated by designer drug (DREADD).

38. The artificial expression construct of claim 34, wherein the expressible element comprises a non-functional molecule.

39. The artificial expression construct of claim 38, wherein the non-functional molecule comprises a non-functional ion transporter, enzyme, transcription factor, receptor, membrane protein, cellular trafficking protein, signaling molecule, neurotransmitter, calcium reporter, channelrhodopsin, CRISPR/CAS molecule, editase, guide RNA molecule, microRNA, homologous recombination donor cassette, or DREADD.

40. The artificial expression construct of claim 2, wherein the artificial expression construct is associated with a capsid that crosses the blood brain barrier.

41. The artificial expression construct of claim 40, wherein the capsid comprises PHP.eB, AAV-BR1, AAV-PHP.S, AAV-PHP.B, or AAV-PPS.

42. The artificial expression construct of claim 2, wherein the artificial expression construct comprises or encodes a skipping element.

43. The artificial expression construct of claim 42, wherein the skipping element comprises a 2A peptide or an internal ribosome entry site (IRES).

44. The artificial expression construct of claim 43, wherein the 2A peptide comprises T2A, P2A,

E2A, or F2A.

45. The artificial expression construct of claim 2, wherein the artificial expression construct comprises or encodes a set of features selected from: a concatenated core of an I56i enhancer, eHGT_387m, eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_390m, eHGT_387m(core2), eHGT_375h(core), eHGT_376h(core), eHGT_390h(core), eHGT_373m(core), eHGT_375m(core), eHGT_386m(core), eHGT_390m(core2), AAV, scAAV, rAAv, minBglobin, CMV, minCMV, minRho, minRho*, fluorescent protein, codon-optimized SLC6A1, 4X2C, Cre, iCre, dgCre, FlpO, tTA2, SP10, WPRE, WPRE3, hGHpA, and/or BGHpA.

46. The artificial expression construct of claim 2, wherein the artificial expression construct comprises the features of: CN2721, CN3213, CN2720, CN2722, CN2732, CN3322, CN3323, CN3887, CN3888, CN2972, CN2973, CN2974, CN2975, CN2976, ID10.01, ID10.02, ID10.03, ID10.04, ID10.05, ID10.06, ID10.07, ID10.08, ID10.09, ID10.10, ID10.11, ID10.12, ID10.13, ID10.14, ID10.15, ID10.16, ID10.17, ID10.18, ID10.19, ID10.20, ID10.21, ID10.22, ID10.23, ID10.24, ID10.25, ID10.26, ID10.27, ID10.28, ID10.29, ID10.30, ID10.31, ID10.32, ID11.01, ID11.02, ID11.03, ID11.04, ID11.05, ID11.06, ID11.07, ID11.08, ID11.09, ID11.10, ID11.11, ID11.12, ID11.13, ID11.14, ID11.15, ID11.16, ID12.01, ID12.02, ID12.03, ID12.04, ID12.05, ID12.06, ID12.07, ID12.08, ID12.09, ID12.10, ID12.11, ID12.12, ID12.13, ID12.14, ID12.15, ID12.16, ID13.01, ID13.02, ID13.03, ID13.04, ID13.05, ID13.06, ID13.07, ID13.08, ID13.09, ID13.10, ID13.11, ID13.12, ID13.13, ID13.14, ID13.15, ID13.16, ID14.01, ID14.02, ID14.03, ID14.04, ID14.05, ID14.06, ID14.07, ID14.08, ID14.09, ID14.10, ID14.11, ID14.12, ID14.13, ID14.14, ID14.15, ID14.16, ID15.01, ID15.02, ID15.03, ID15.04, ID15.05, ID15.06, ID15.07, ID15.08, ID15.09, ID15.10, ID15.11, ID15.12, ID15.13, ID15.14, ID15.15, ID15.16, ID16.01, ID16.02, ID16.03, ID16.04, ID16.05, ID16.06, ID16.07, ID16.08, ID16.09, ID16.10, ID16.11, ID16.12, ID16.13, ID16.14, ID16.15, ID16.16, ID17.01, ID17.02, ID17.03, ID17.04, ID17.05, ID17.06, ID17.07, ID17.08, ID17.09, ID17.10, ID17.11, ID17.12, ID17.13, ID17.14, ID17.15, ID17.16, ID18.01, ID18.02, ID18.03, ID18.04, ID18.05, ID18.06, ID18.07, ID18.08, ID18.09, ID18.10, ID18.11, ID18.12, ID18.13, ID18.14, ID18.15, ID18.16, ID19.01, ID19.02, ID19.03, ID19.04, ID19.05, ID19.06, ID19.07, ID19.08, ID19.09, ID19.10, ID19.11, ID19.12, ID19.13, ID19.14, ID19.15, ID19.16, ID20.01, ID20.02, ID20.03, ID20.04, ID20.05, ID20.06, ID20.07, ID20.08, ID20.09, ID20.10, ID20.11, ID20.12, ID20.13, ID20.14, ID20.15, ID20.16, ID21.01, ID21.02, ID21.03, ID21.04, ID21.05, ID21.06, ID21.07, ID21.08, ID21.09, ID21.10, ID21.11, ID21.12, ID21.13, ID21.14, ID21.15, ID21.16, ID22.01, ID22.02, ID22.03, ID22.04, ID22.05, ID22.06, ID22.07, ID22.08, ID22.09, ID22.10, ID22.11, ID22.12, ID22.13, ID22.14, ID22.15, ID22.16, ID23.01, ID23.02, ID23.03, ID23.04, ID23.05, ID23.06, ID23.07, ID23.08, ID23.09,

ID23.10, ID23.11, ID23.12, ID23.13, ID23.14, ID23.15, ID23.16, ID24.01, ID24.02, ID24.03, ID24.04, ID24.05, ID24.06, ID24.07, ID24.08, ID24.09, ID24.10, ID24.11, ID24.12, ID24.13, ID24.14, ID24.15, ID24.16, ID25.01, ID25.02, ID25.03, ID25.04, ID25.05, ID25.06, ID25.07, ID25.08, ID25.09, ID25.10, ID25.11, ID25.12, ID25.13, ID25.14, ID25.15, ID25.16, ID26.01, ID26.02, ID26.03, ID26.04, ID26.05, ID26.06, ID26.07, ID26.08, ID26.09, ID26.10, ID26.11, ID26.12, ID26.13, ID26.14, ID26.15, ID26.16, ID27.01, ID27.02, ID27.03, ID27.04, ID27.05, ID27.06, ID27.07, ID27.08, ID27.09, ID27.10, ID27.11, ID27.12, ID27.13, ID27.14, ID27.15, or ID27.16.

47. A vector comprising an artificial expression construct of claim 2.

48. The vector of claim 47, wherein the vector comprises a viral vector.

49. The vector of claim 48, wherein the viral vector comprises a recombinant adeno-associated viral (AAV) vector.

50. A transgenic cell comprising an artificial expression construct of claim 2.

51. The transgenic cell of claim 50, wherein the transgenic cell is a GABAergic neuron or an astrocyte.

52. The transgenic cell of claim 50, wherein the transgenic cell is murine, human, or non-human primate.

53. A non-human transgenic animal comprising an artificial expression construct of claim 2.

54. The non-human transgenic animal of claim 53, wherein the non-human transgenic animal is a mouse or a non-human primate.

55. An administrable composition comprising an artificial expression construct of claim 2.

56. A kit comprising an artificial expression construct of claim 2.

57. A method for expressing a gene within a population of cells in vivo or in vitro, the method comprising providing the administrable composition of claim 55 in a sufficient dosage and for a sufficient time to a sample or subject comprising the population of cells thereby expressing the gene within the population of cells.

58. The method of claim 57, wherein the gene encodes GAT1, an effector element, or an expressible element.

59. The method of claim 58, wherein the gene is a codon-optimized SLC6A1 gene.

60. The method of claim 57, wherein the population of cells comprises GABAergic neurons and astrocytes.

61. The method of claim 57, wherein the providing comprises pipetting.

62. The method of claim 61, wherein the pipetting is to a brain slice.

63. The method of claim 62, wherein the brain slice comprises GABAergic neurons and

astrocytes.

64. The method of claim 62, wherein the brain slice is murine, human, or non-human primate.

65. The method of claim 57, wherein the providing comprises administering to a living subject.

66. The method of claim 65, wherein the living subject is a human, non-human primate, or a mouse.

67. The method of claim 65, wherein the living subject has an SL6CA1-associated disorder.

68. The method of claim 67, wherein the SL6CA1-associated disorder comprises impaired cognitive function, impaired motor function, mild-to-moderate intellectual disability, epilepsy, speech difficulty, attention deficit disorder, attention deficit hyperactivity disorder, or an autism spectrum disorder.

69. The method of claim 65, wherein the administering to a living subject is through injection.

70. The method of claim 69, wherein the injection comprises intravenous injection, intraparenchymal injection into brain tissue, intracerebroventricular (ICV) injection, intra-cisterna magna (ICM) injection, or intrathecal injection.

71. An artificial expression construct having a sequence with at least 90% sequence identity to the sequence as set forth in SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 90, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 93, or SEQ ID NO: 94.

72. An artificial expression construct having the sequence as set forth in SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 74, SEQ ID NO: 75, SEQ ID NO: 76, SEQ ID NO: 77, SEQ ID NO: 90, SEQ ID NO: 91, SEQ ID NO: 92, SEQ ID NO: 93, or SEQ ID NO: 94.

FIG. 1A

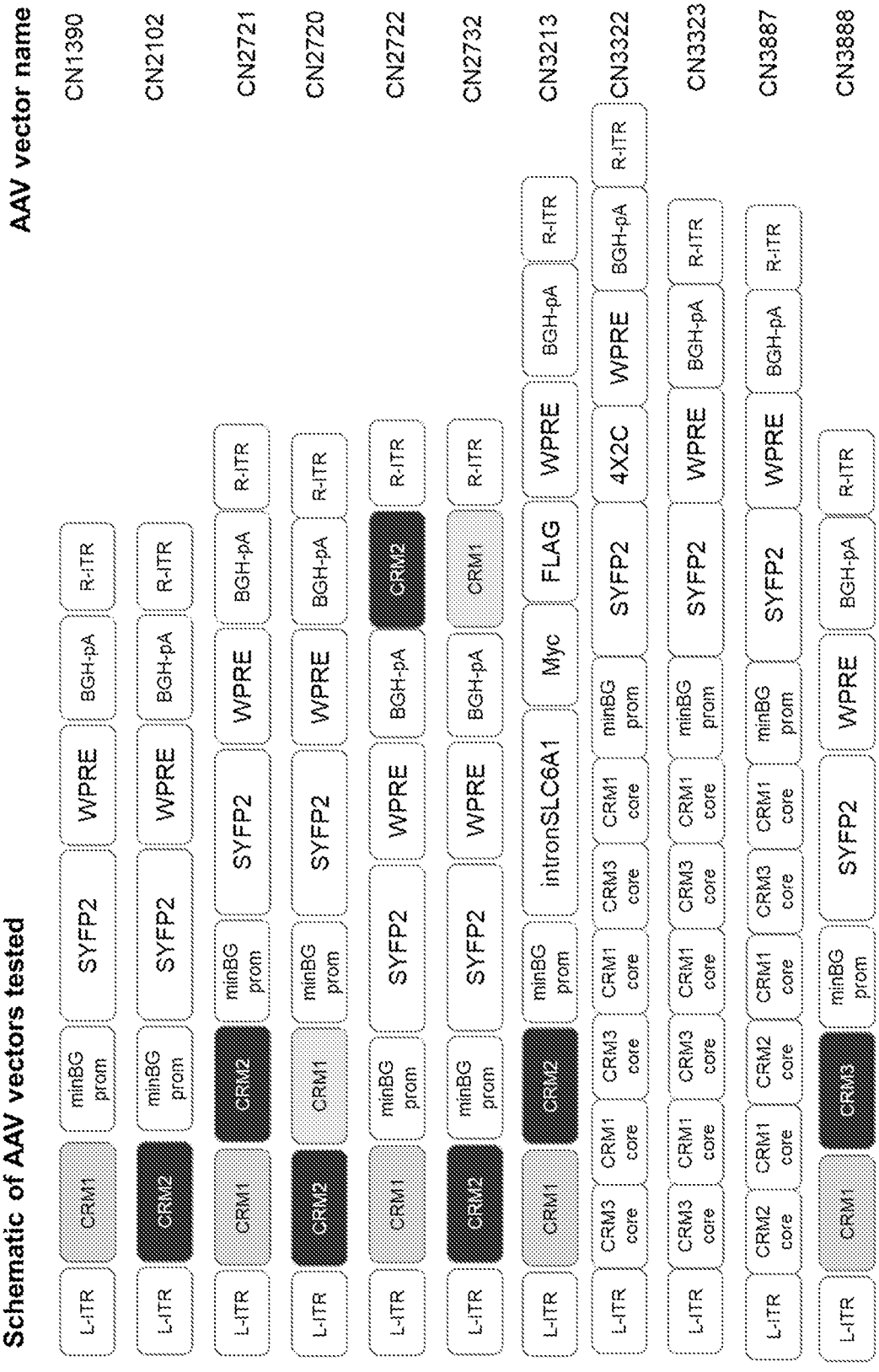

Schematic of AAV vectors tested
AAV vector name
Single specificity
Dual specificity
CN1390
CN2102
CN2721
CN2720
CN2722
CN2732
CN3213
CN3322
CN3323
CN3887
CN3888
L-ITR
CRM1
CRM2
CRM3
CRM8
minBG prom
SYFP2
WPRE
BGH-pA
R-ITR
core
intronSLC6A1
Myc
FLAG
4X2C

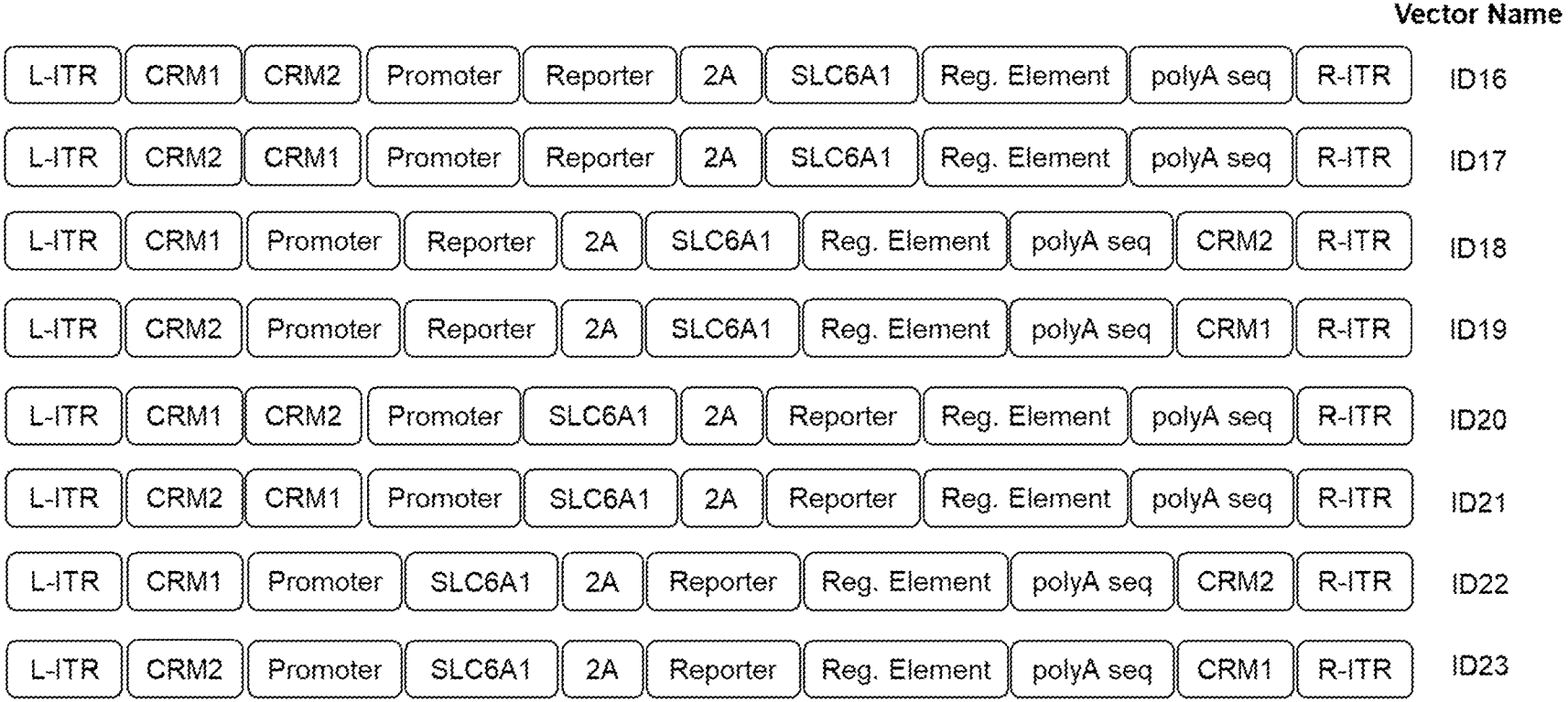
FIG. 1A cont'd

Enhancers tested
CRM1
CRM2
CRM3
3xhl56icore
eHGT_387m
eHGT_390m
CRM1 core
CRM31 core
CRM3 core
hl56i(core)
eHGT_387m(core2)
eHGT_390m(core2)

FIG. 1B

CRM1: I56i or 3xI56icore
CRM1 core: I56i(core)
CRM2: eHGT_375h, eHGT_376h, eHGT_390h, eHGT_373m, eHGT_375m, eHGT_386m, eHGT_387m, or eHGT_390m
CRM2 core: core of eHGT_375h, core of eHGT_376h, core of eHGT_390h, core of eHGT_373m, core of eHGT_375m, core of eHGT_386m, core of eHGT_387m (e.g., eHGT_387m(core2)), or core of eHGT_390m (e.g., eHGT_390m(core2)).

FIG. 1B cont'd



CRM1: I56i or 3xI56icore
CRM2: eHGT_375h, eHGT_376h, eHGT_ 390h, eHGT_ 373m, eHGT_ 375m, eHGT_ 386m, eHGT_ 387m, or eHGT_ 390m.
2A: P2A, T2A, E2A, or F2A

FIG. 1B cont'd

L-ITR
Promoter
SLC6A1
2A
Reporter
Reg. Element
polyA seq
R-ITR
ID24
L-ITR
Promoter
SLC6A1
2A
Reporter
Reg. Element
polyA seq
R-ITR
ID25
L-ITR
Promoter
Reporter
2A
SLC6A1
Reg. Element
polyA seq
R-ITR
ID26
L-ITR
Promoter
Reporter
2A
SLC6A1
Reg. Element
polyA seq
R-ITR
ID27
CRM1 core: I56i(core)
CRM2 core: core of eHGT_375h, core of eHGT_376h, core of eHGT_390h, core of eHGT_373m, core of eHGT_375m, core of eHGT_386m, core of eHGT_387m (e.g., eHGT_387m(core2)), or core of eHGT_390m (e.g., eHGT_390m(core2))
2A: P2A, T2A, E2A, or F2A

FIG. 2

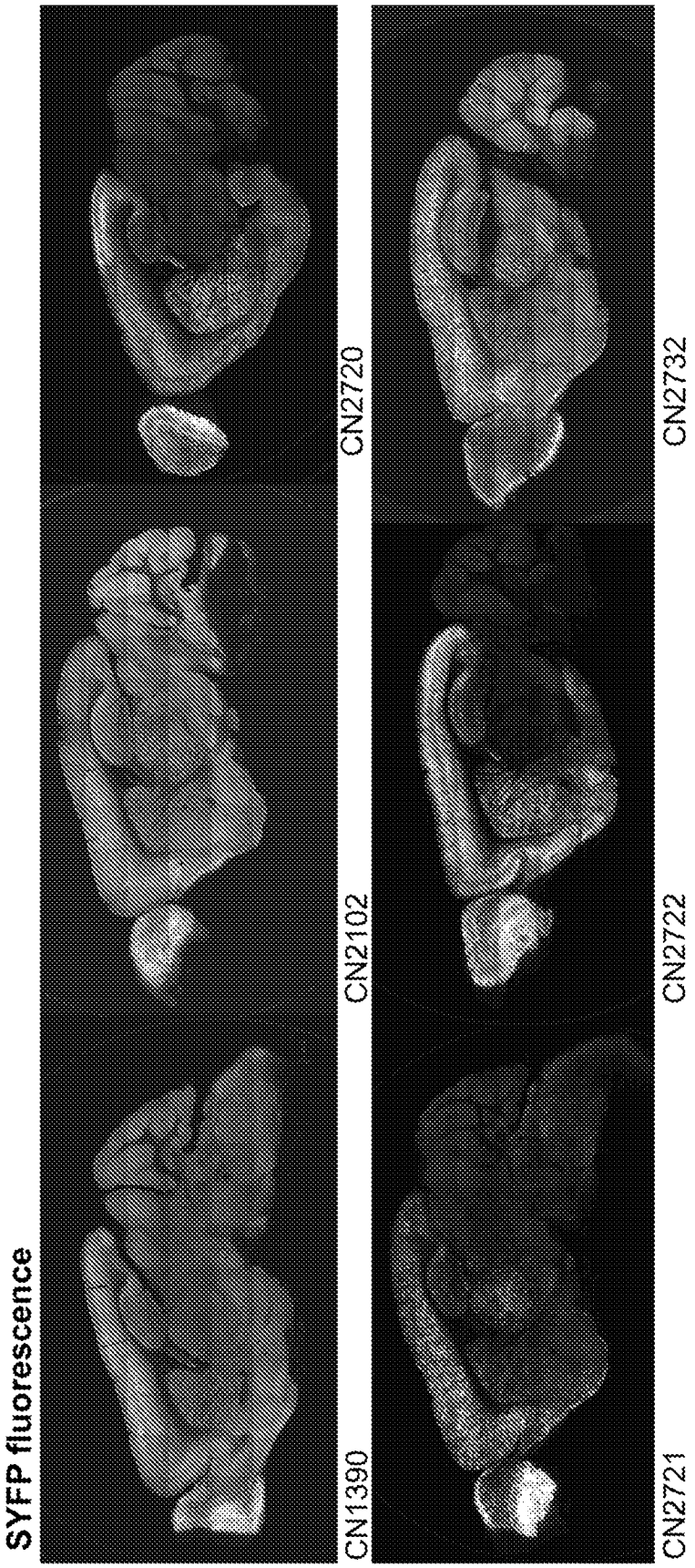

SYFP fluorescence
CN1390
CN2102
CN2720
CN2721
CN2722
CN2732

FIG. 3

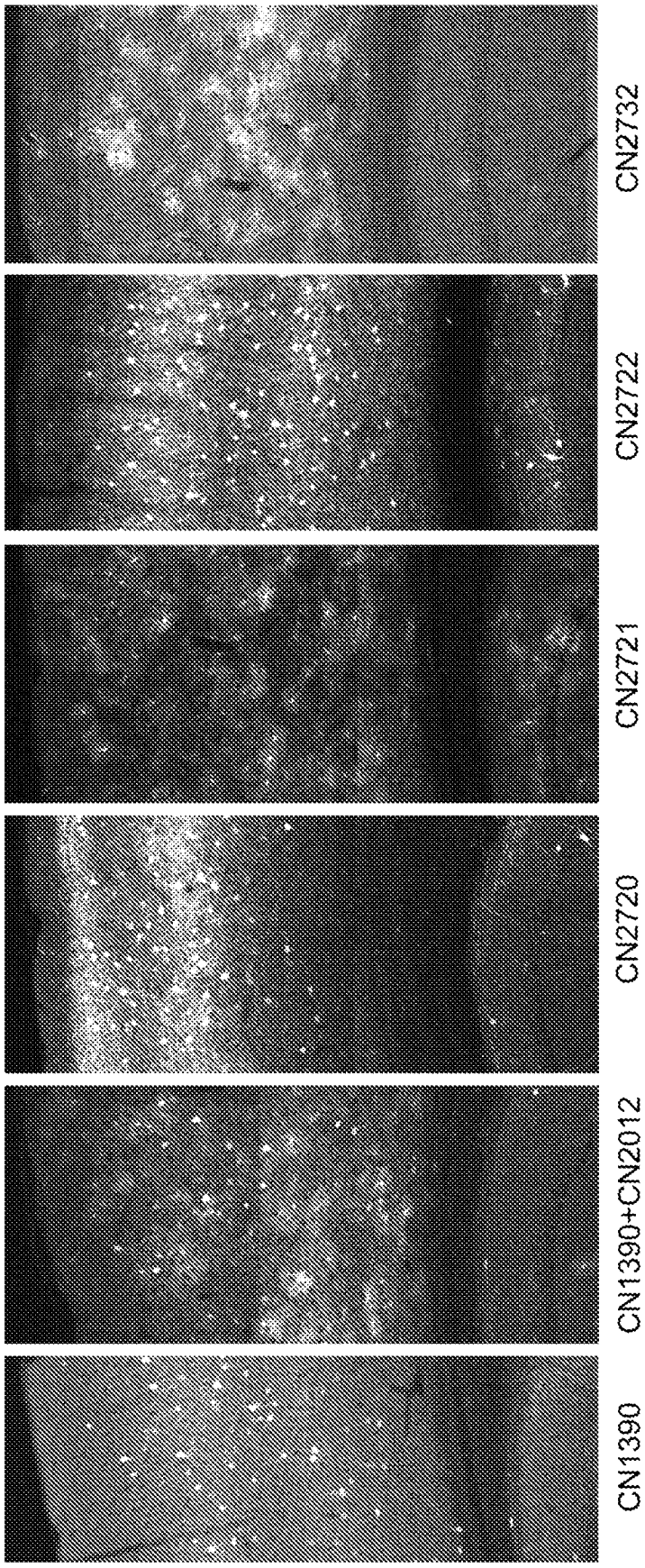


FIG. 4

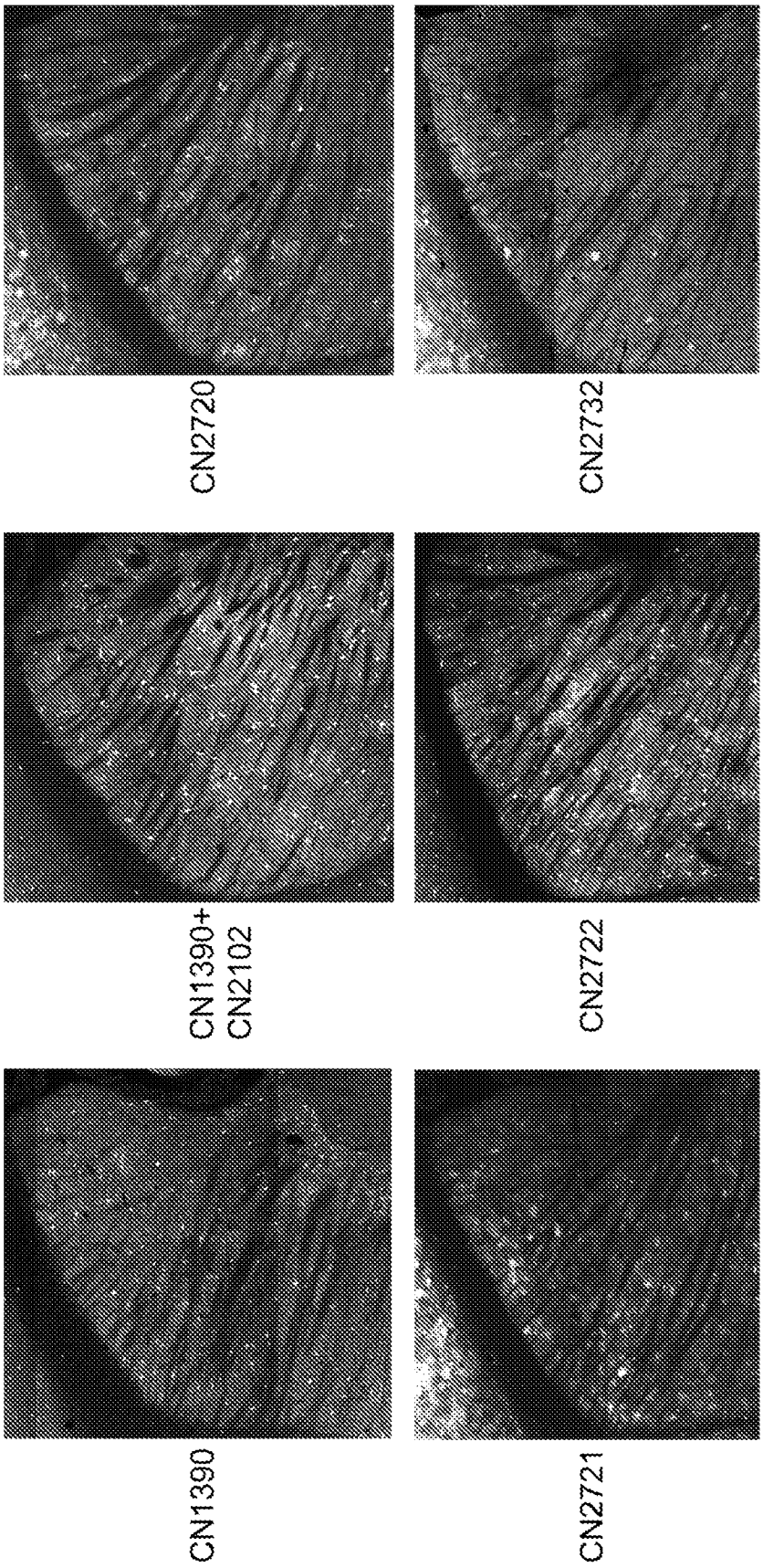


FIG. 5

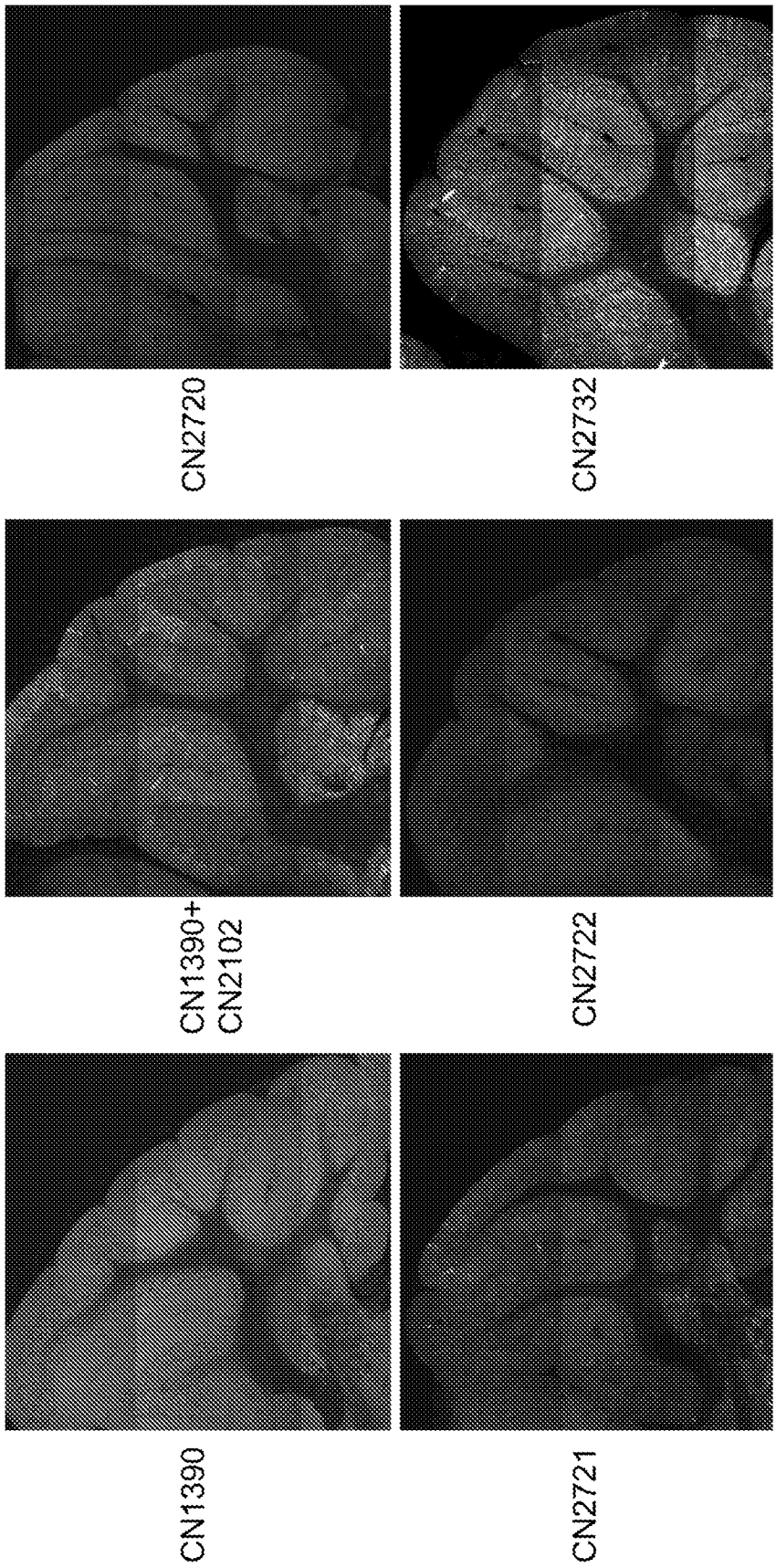


FIG. 6

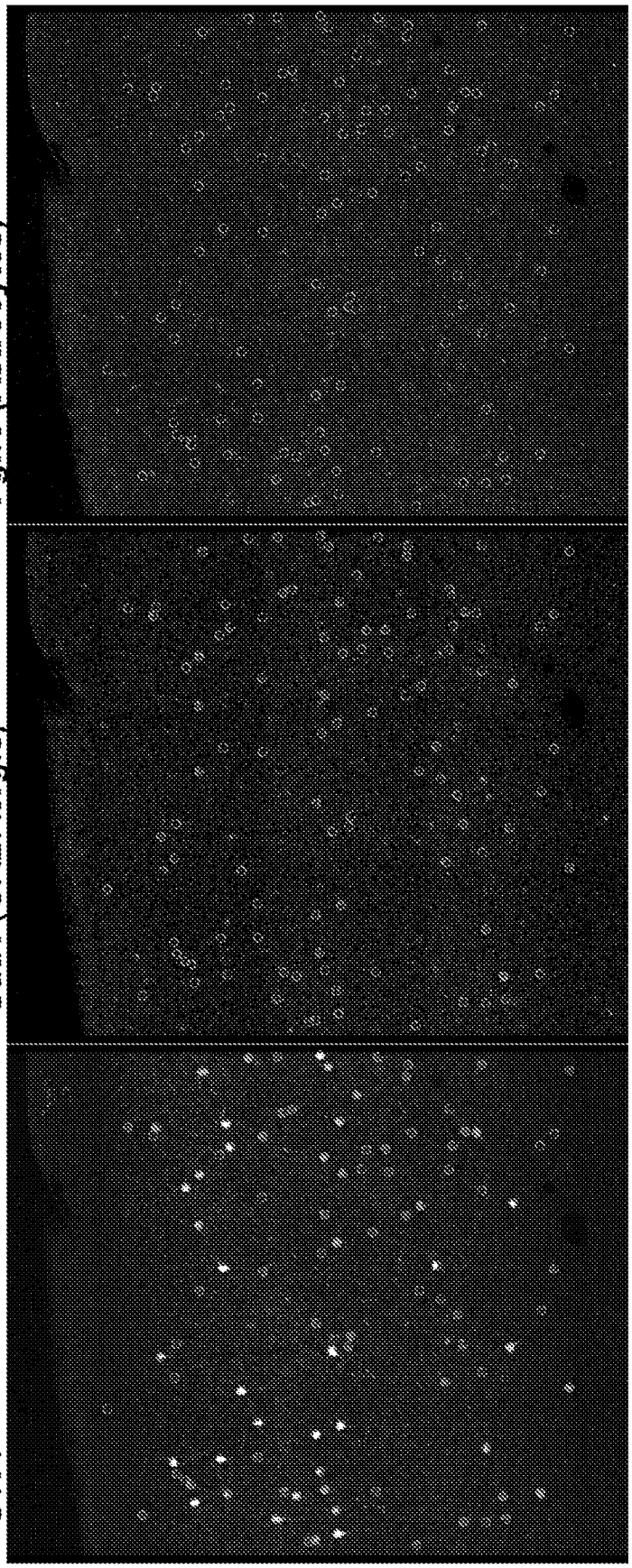

SYFP
Gad1 (GABAergic)
Fgfr3 (Astrocytes)
= SYFP only 7%
= SYFP+ Gad1+ 93%
= SYFP+ Fgfr3+ 0%

FIG. 7

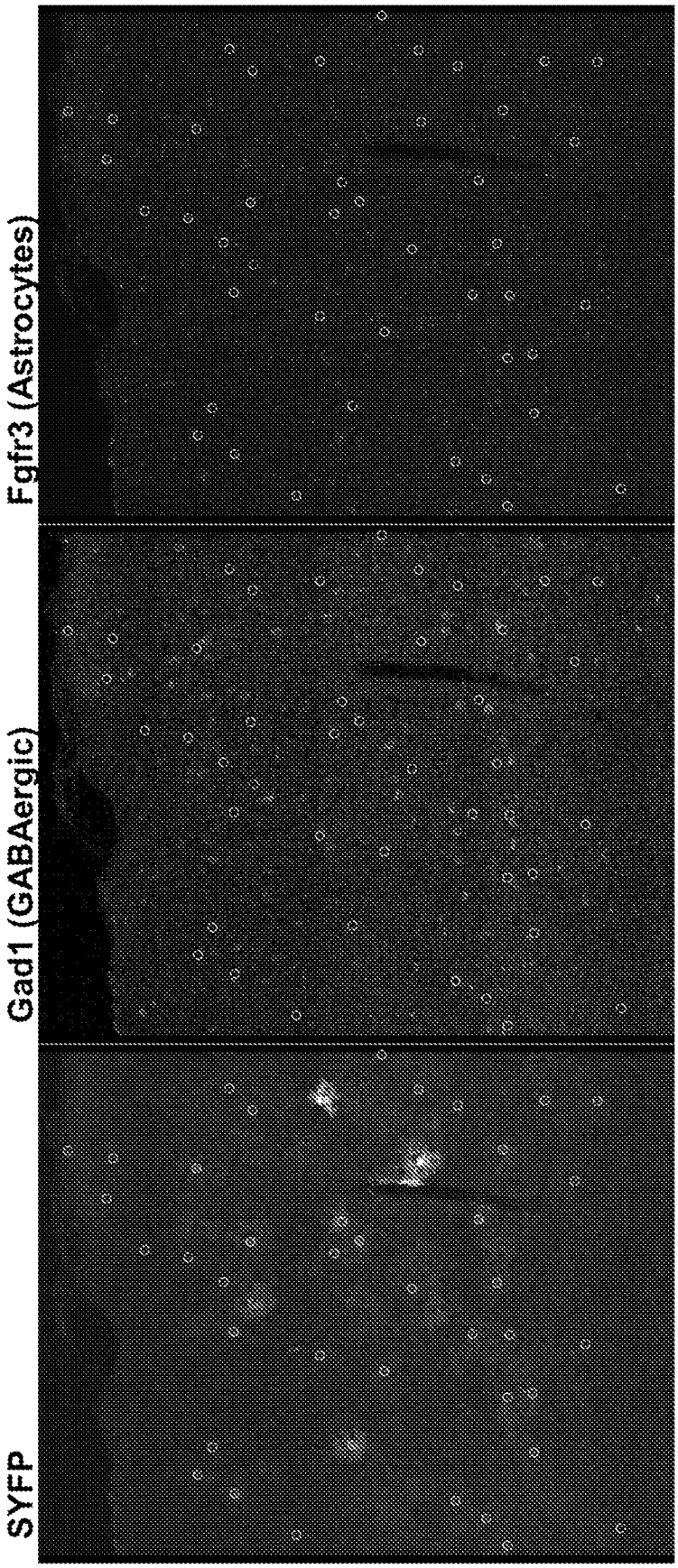

SYFP
Gad1 (GABAergic)
Fgfr3 (Astrocytes)
= SYFP only 2%
= SYFP+ Gad1+ 0%
= SYFP+ Fgfr3+ 98%

FIG. 8

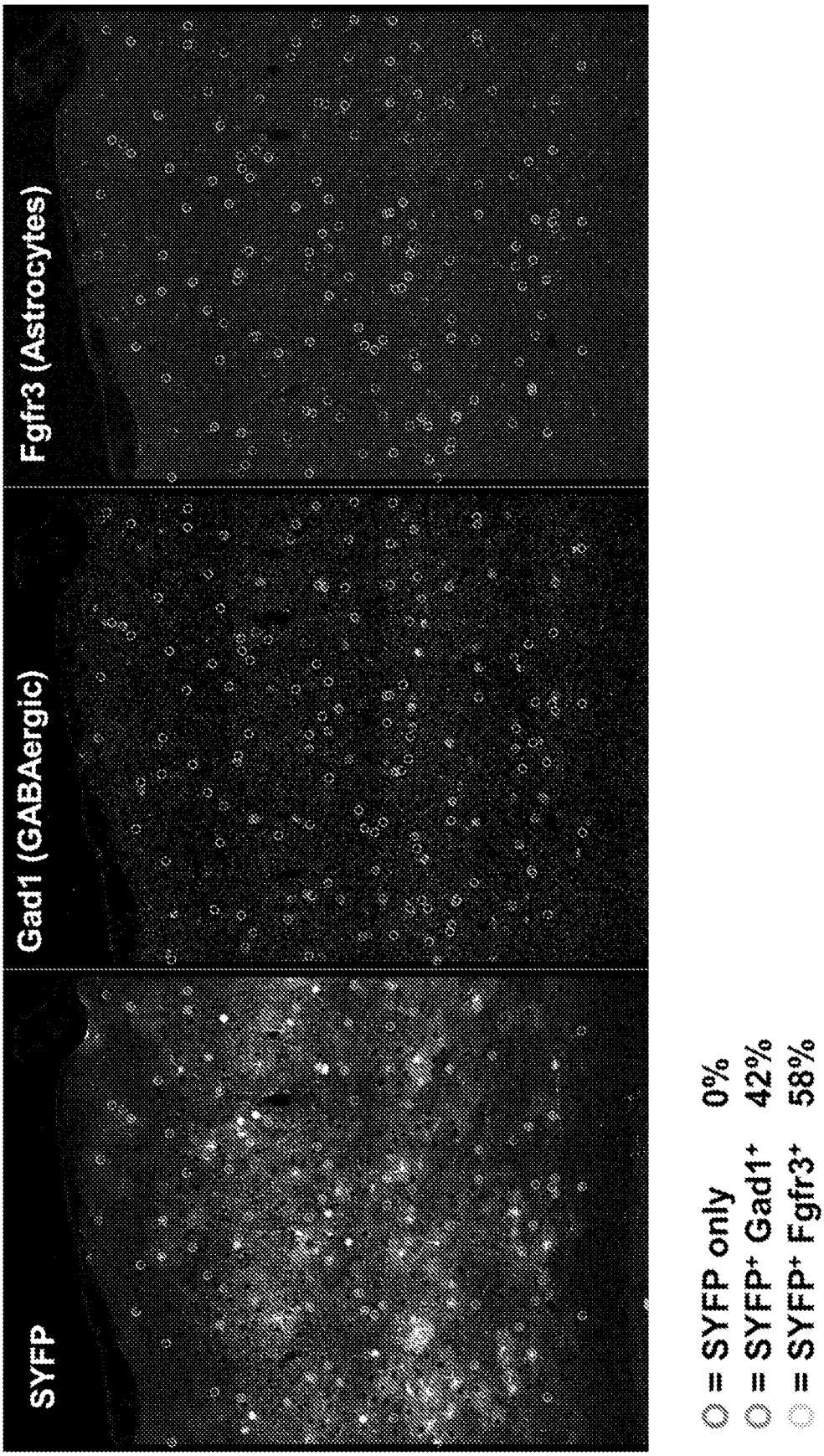

SYFP
Gad1 (GABAergic)
Fgfr3 (Astrocytes)
= SYFP only 0%
= SYFP+ Gad1+ 42%
= SYFP+ Fgfr3+ 58%

FIG. 9

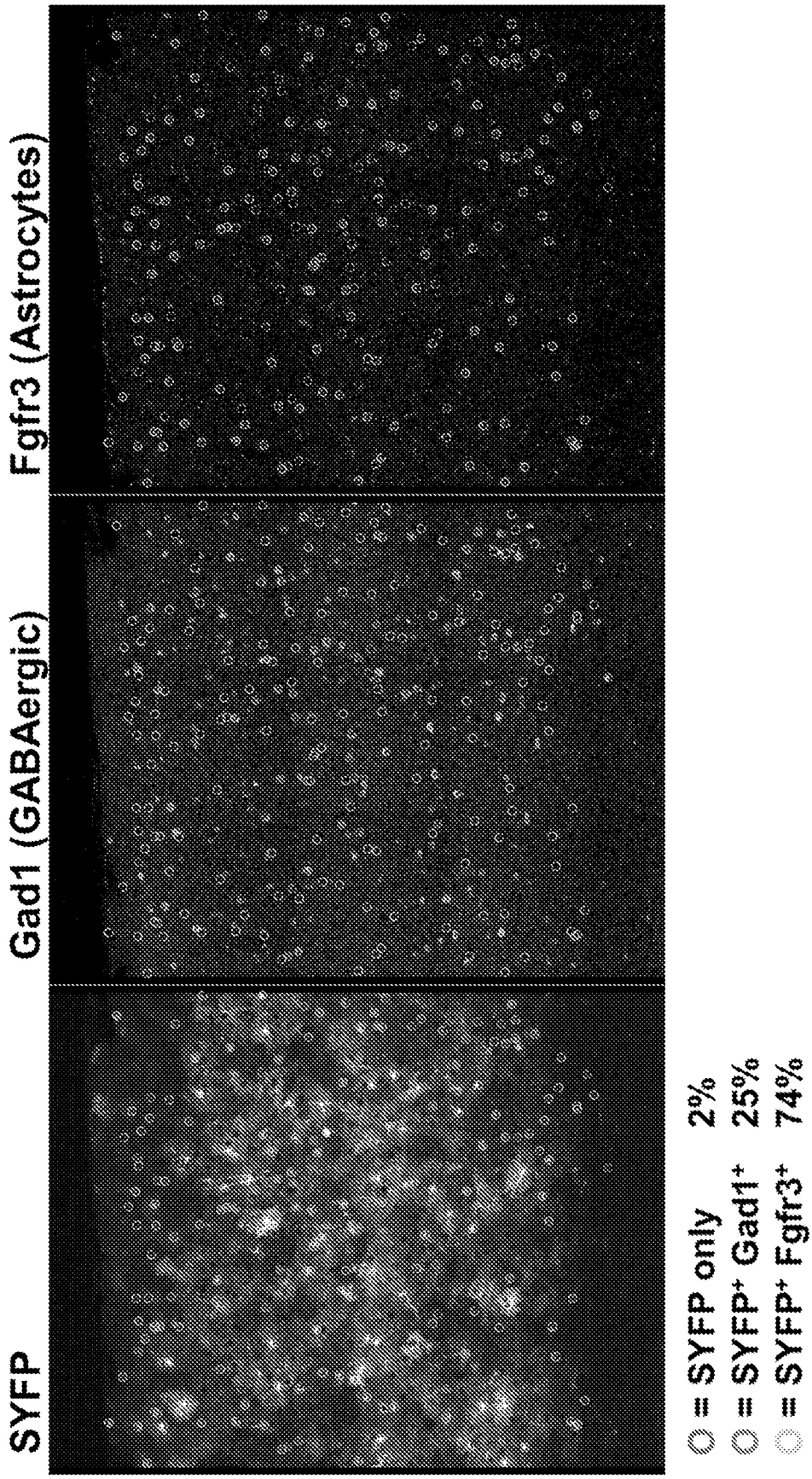

SYFP
Gad1 (GABAergic)
Fgfr3 (Astrocytes)
○ = SYFP only 2%
○ = SYFP+ Gad1+ 25%
○ = SYFP+ Fgfr3+ 74%

FIG. 10A

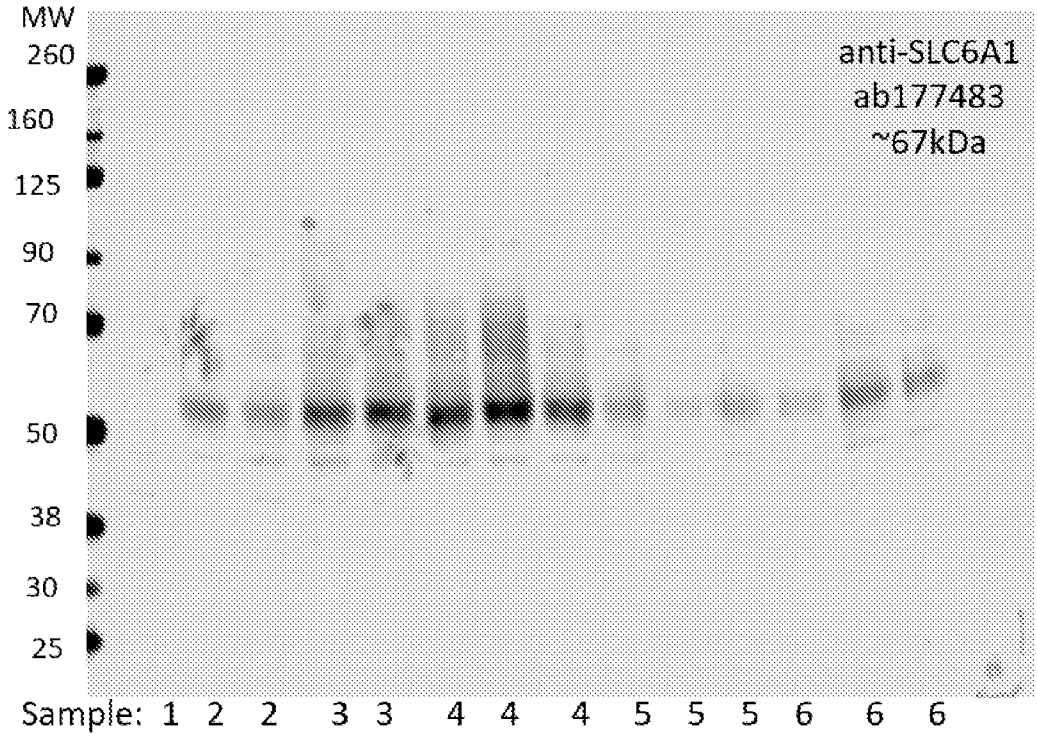


FIG. 10B

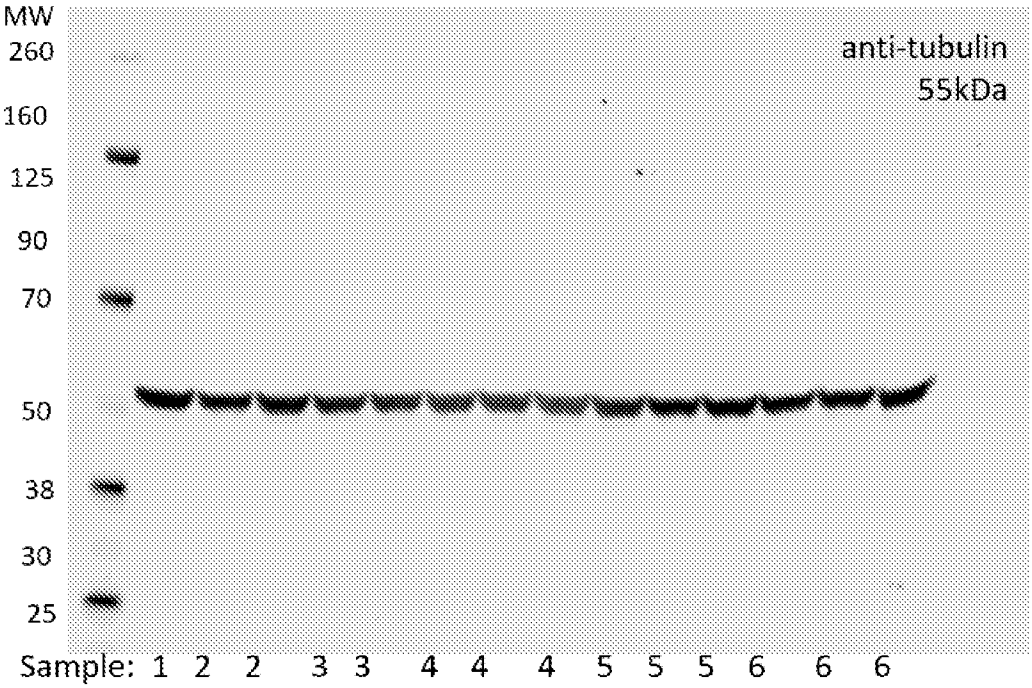
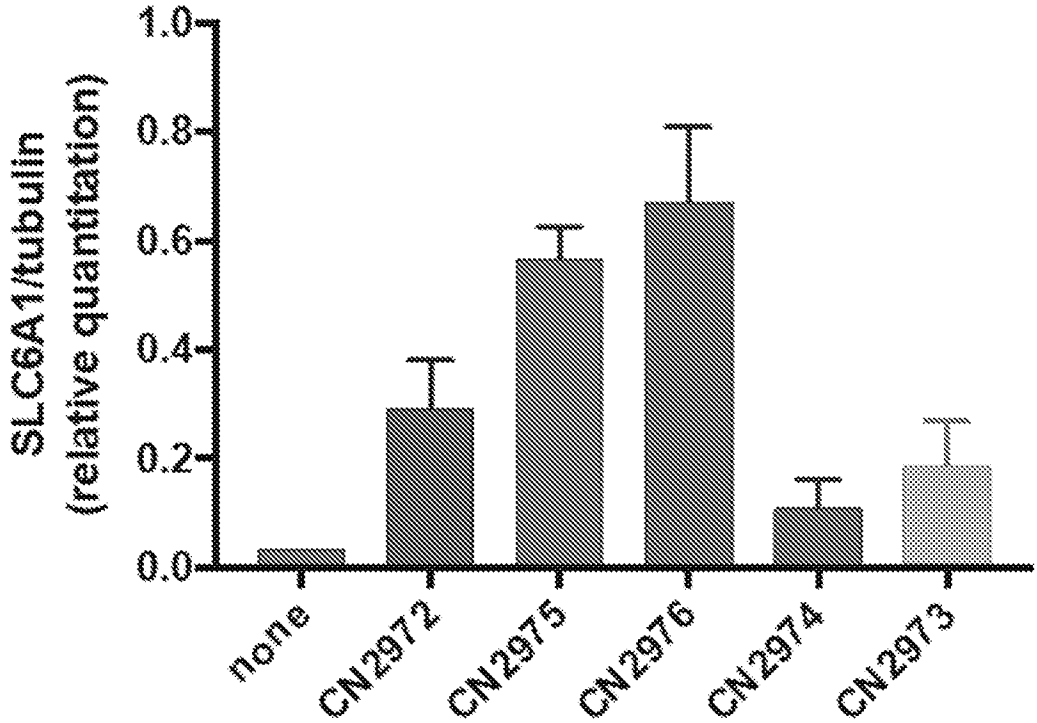


FIG. 10C


SLC6A1/tubulin
(relative quantitation)
1.0
0.8
0.6
0.4
0.2
0.0
none
CN2972
CN2975
CN2976
CN2974
CN2973

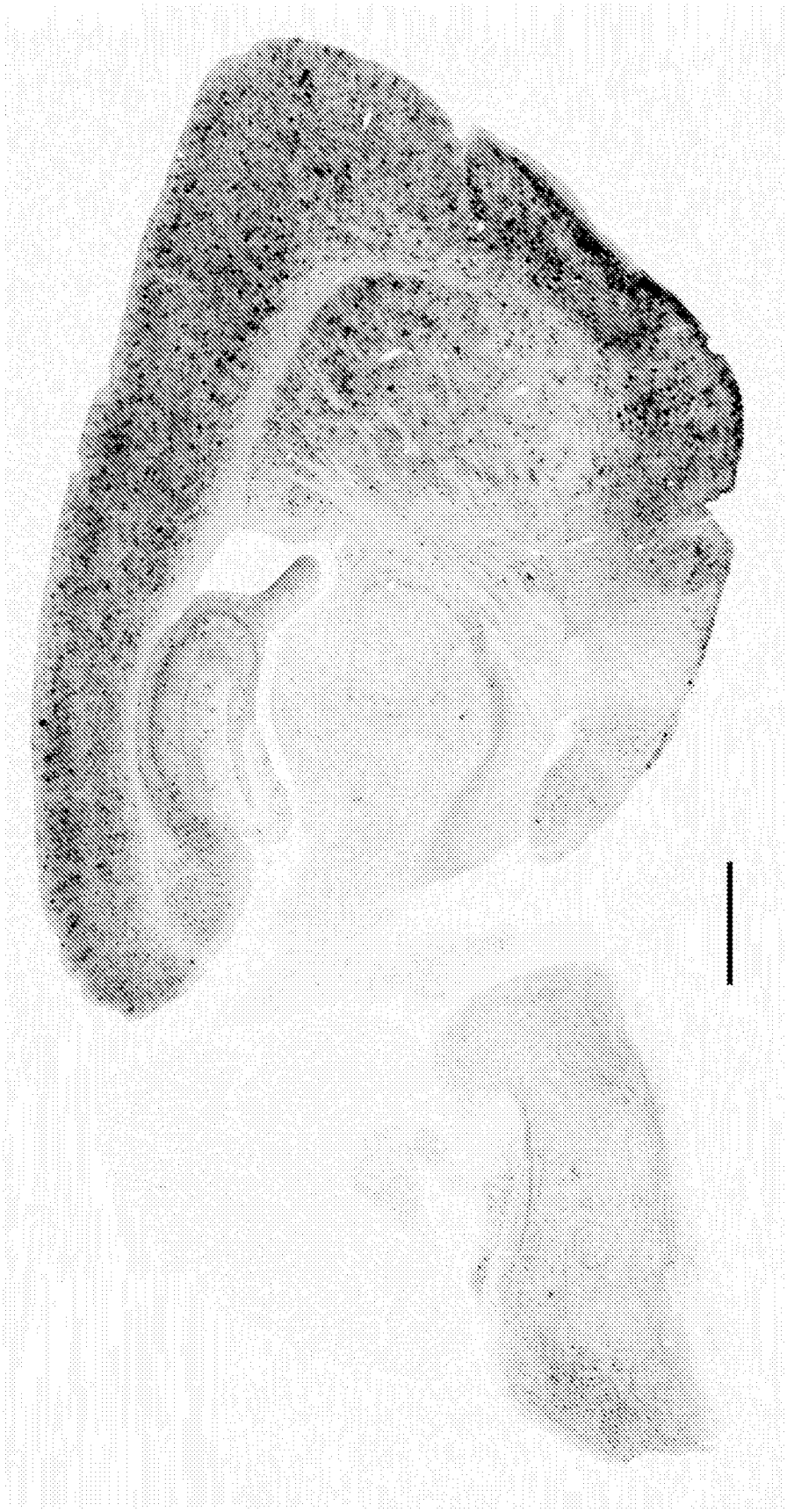


FIG. 11A

FIG. 11B

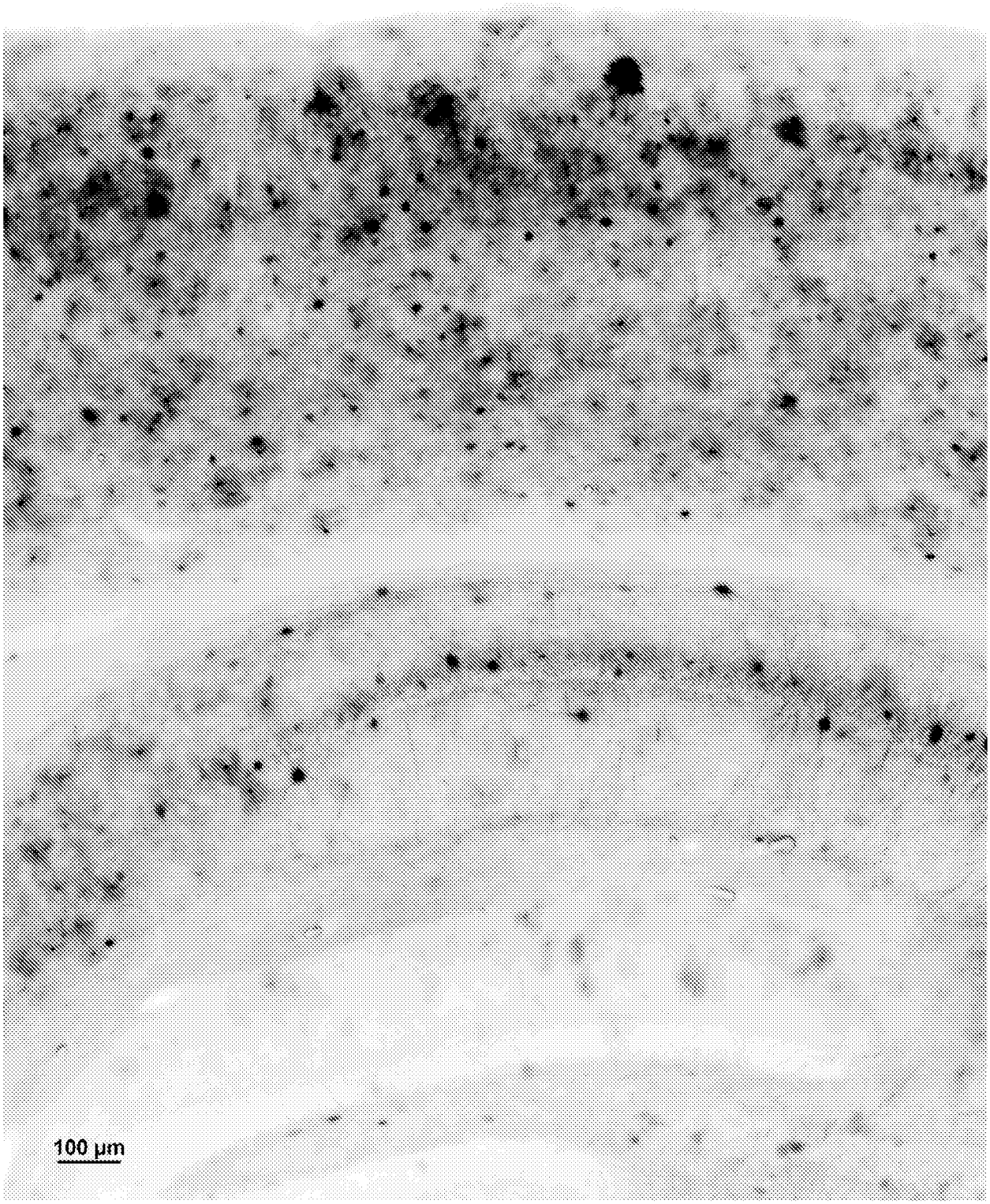


FIG. 12A

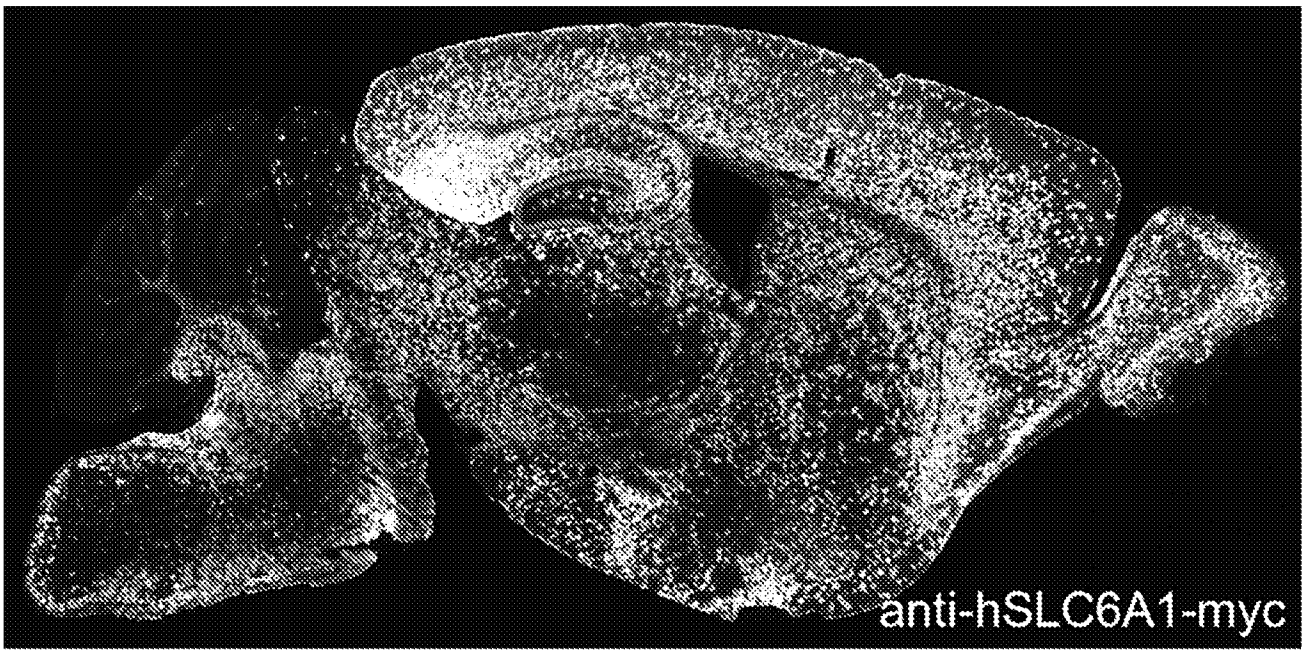


FIG. 12B

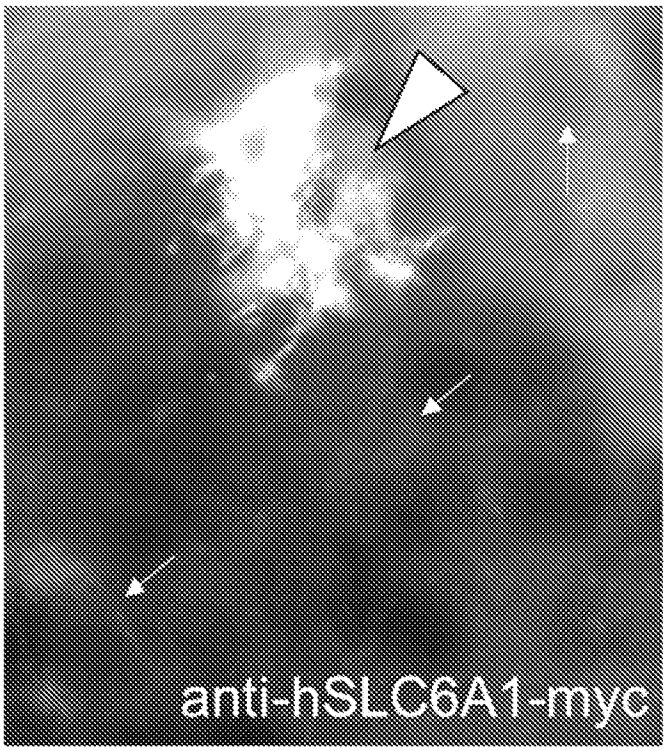


FIG. 12C

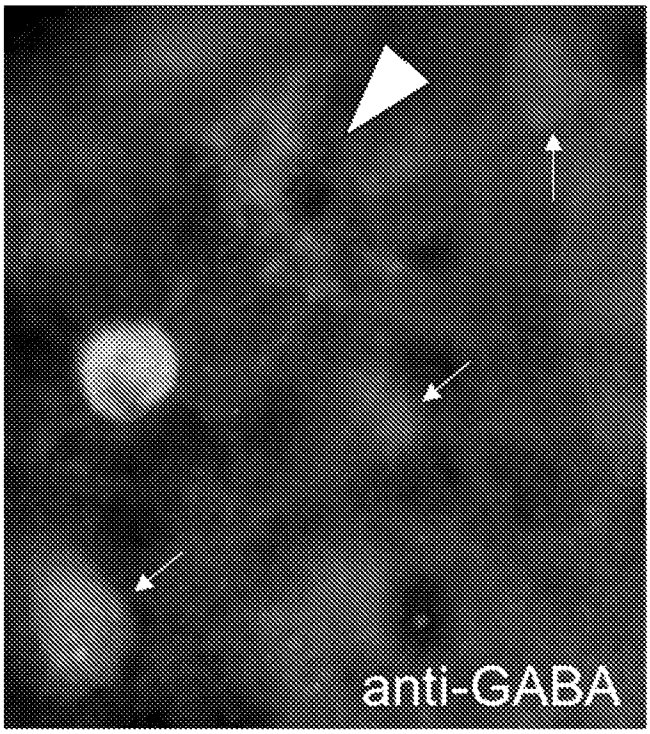


FIG. 13A

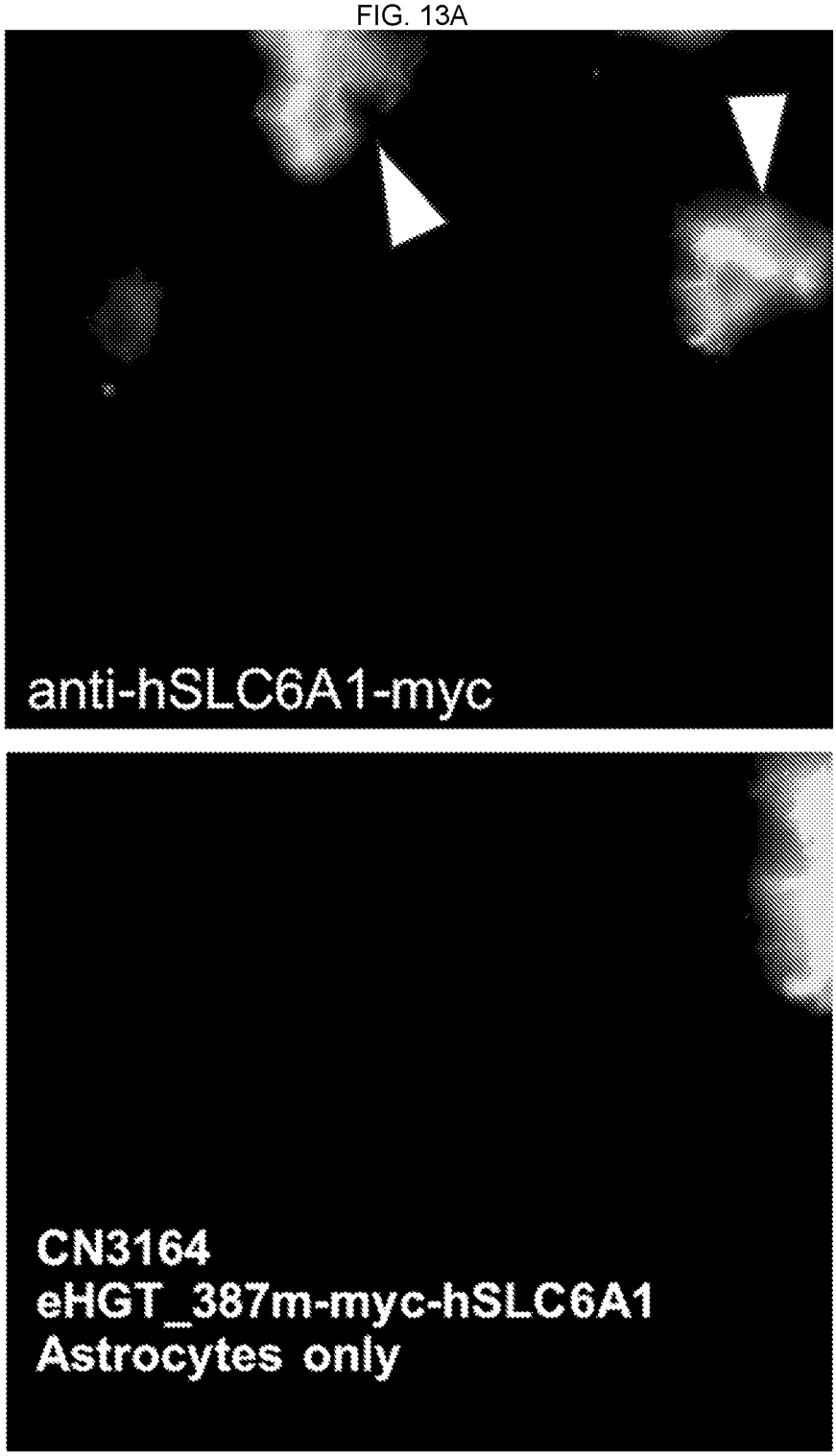

anti-hSLC6A1-myc
CN3164
eHGT_387m-myc-hSLC6A1
Astrocytes only

FIG. 13B

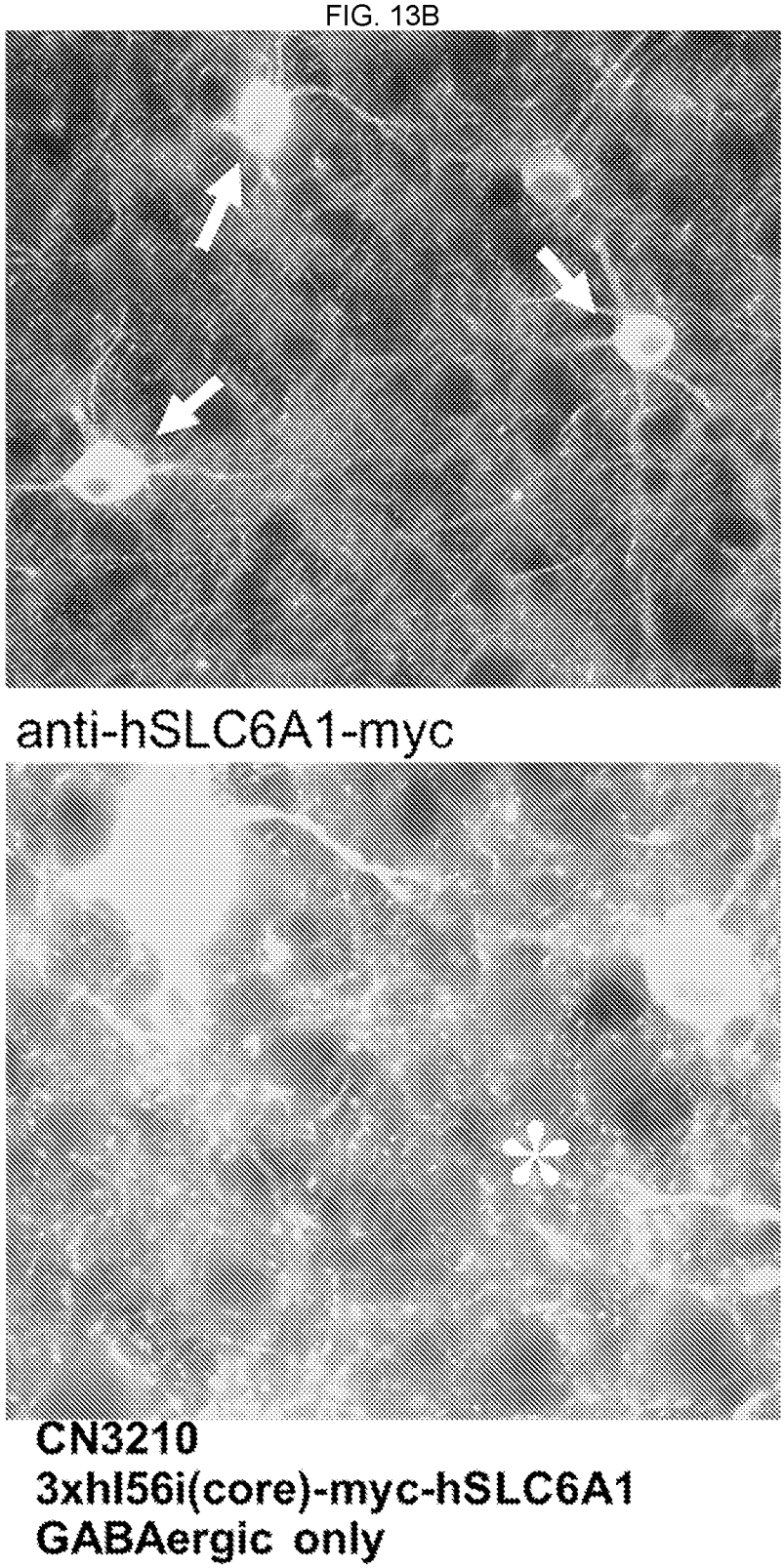

anti-hSLC6A1-myc
CN3210
3xhI56i(core)-myc-hSLC6A1
GABAergic only

FIG. 13C

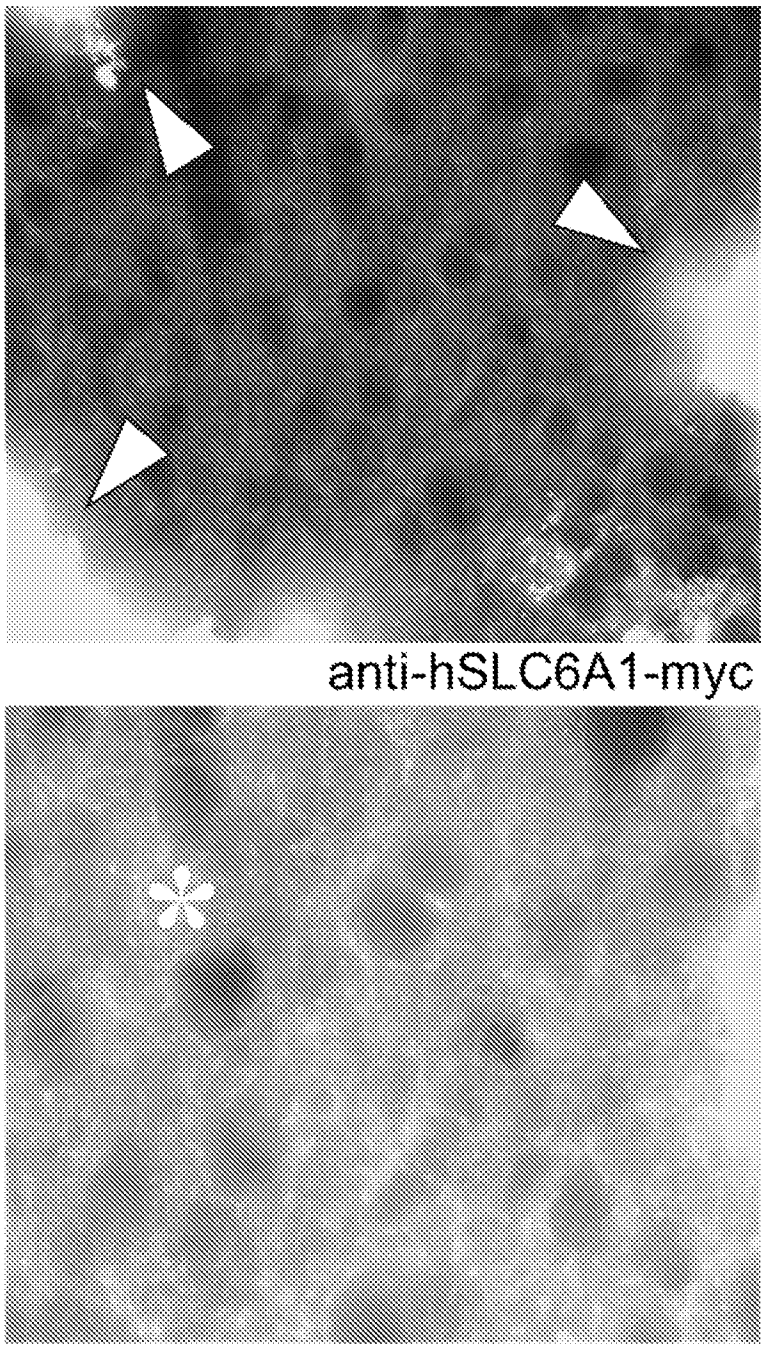

anti-hSLC6A1-myc
CN3213
3xhl56i(core)_eHGT_387m-
myc-hSLC6A1
Astro+GABAergic

FIG. 14

hl56i – full length human hl56i enhancer
TATGCACTCACAGTGGTTTGGCATGCATCTGGTGAATTTTTTTTAACGAAAAATTAGTGTTG
GTTTCGATGTATGGTAGCATTCTCCCTAACGTAATTTGAATAATTCAGCAAAGCCCCACTAC
CAGCTGTACTTCTGCAGCCTCTTCCATTCTTTTCAGCATTATAATTTTGGTTAATTTTCAATT
TTAGGTCCTACGTCTCTGCAATTTGTGTATGAATAACAGAATAATTTCCCTCTTTTGTTTCGC
CTTTCCTGTTCCTGAATCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACA
GGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAAT
TATGGCTGCATTTAAGAGAATGGAAAAAAACCTTCTTGTGGATAAAAACCTTAAATTGTCCC
CAATGTCTGCTTCAAATTGGATGGCACTGCAGCTGGAGGCTTTGTTCAGAATTGATCCTGG
GGAGCTACGAACCCAAAGTTTCACAGTAGG (SEQ ID NO: 1)

Murine l56i Enhancer (core is the same as human):
TATACACTCACAGTGGTTTGGCATATATTTGGTGAAATTTTTTAAGGAAAAATTAGTGTTGGT
TTCGATATATGGTAGCTTTTTCTCTAACATAATTTGAATAATTCAGCAAAGCCCTACTACCAG
CTGTACTTCTGCAGCCTCTTCCATTCTTTCCAGCATTATAATTTTGGTTAATTTTCAATTTTA
GGTCCTACGTCTCTGCAATTTGTGTATGAATAACAGAATAATTTCCCTCTTTTGTTTCGCCTT
TCCTGTTCCTGAATCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGG
TAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTA
TGGCTGCATTTAAGAGAATGGAAAAAAACCTTCTTGTGGATAAAAACCTTAAATTGTCCCCA
ATGTCTGCTTCAAATTGGATGGCACTGCAGCTGGAGGCTTTGTTCAGAATTGATCCTGGGG
AGCTACGAACCCAAAGTTTCACAGTAGG (SEQ ID NO: 2)

Zebrafish l46i Enhancer:
ACATTGTAATTTTAGATAATATCCCAAGCGTTCACTCTCCTCGGCAATTTGTACATGAATAAC
CGAATAATTTCATCTTTTGTTTCGTCTTTGCCACTTCAAATCCAAATAAAGATGCCTTTTAGT
ATTAAAAGTGGTAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGG
GCTACATCAAAAATTACCCTAATTATGTCTGCATTTATGAGAATGGAAAAAAACCCTCTCTT
GGATAAAACCCATAAATTGTCCCAAATATCT (SEQ ID NO: 3)

hl56i core - human hl56i enhancer core
CTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGA
GAATGG (SEQ ID NO: 4)

Core of the Zebrafish l46i Enhancer:
CCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGTCTGCATTTATGA
GAATGG (SEQ ID NO: 5)

3xhl56i(core)
CTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGA
GAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATC
TTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCA
TTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGT
AATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTAT
GGCTGCATTTAAGAGAATGG (SEQ ID NO: 6)

FIG. 14 cont'd

3x Concatamerized Core of the Zebrafish I46i Enhancer:
CCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGTCTGCATTTATGA
GAATGGCCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAATTATCT
TTGACGGTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGTCTGCAT
TTATGAGAATGGCCAAATAAAGATGCCTTTTAGTATTAAAAGTGGTAGAAAATTACAGGTAA
TTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATCAAAAATTACCCTAATTATGT
CTGCATTTATGAGAATGG (SEQ ID NO: 7)

eHGT_375h
GGCCACGTTTTGCCTAGGGAAGATGAAGTCTTTGATGATAGCCAACTTAGTGTCAATAAGT
GGCTTTCTTTGAGACATATTCAGATGGGAACGTCTTGCTTGCCAATTGCCATAGAAATCTTA
ACACACCATGAAGATTGTCCGAGCGCCAAGCCTTCCGTTCTGGACTAAATTACTTTGAAGT
GGCGCAGTACGAGCAGTGGTCAATTTTAACTCTATAGACTGGACAGAGAATGCTGGGAGT
GGGAGATTGTGCGTTTTAAAGCAGAAAATAAGAAGGGGAAACTTGTTTTATACACTCTATAC
AAGGTTCTGCTCATTGTCAGA (SEQ ID NO: 8)

eHGT_376h
GGGGCCATTGAGATTGCTGGATCCTGATTCTTTGAAGCATTTCATTCATTATGGTAAAGAAG
GGTTTAAGTGGCCACCAACAGAGAGACGTGAAGTACATGAAACAATTAGGTTCTCTGTATC
TCCAGCAGAATTGGCCCCAGAAGAGGGTCAGGCTTTGCAAAGACACAGAACATTTTTCCC
GCTGGGCTCCTTGGGAAAAGGTCTCAGCATTATGGAGGGTGTCTCTGGCTATTCACAGCTT
GCCAGTGGGAACAGCCAAGAAGGAGAAGAGGTCACATGGCCAGGCCTGCTGCCGGCACA
GAATGTTCTGTGCAGCTGATGGCTGGGCAGGGATGAGGTTTGCCCGATCCCCTGCTGAGG
CTTCCTATGAAAGGTAGGCCTGAGAGTTGCCAAAACATATTCTGCAACAGGGTCTATGAAG
GTCACCATAAAGCAGGATTCAGACTCCATATTAGTTGGTTATCTCTTCTATACAGAATGTTA
TGCCAAGCATTTCTATATATATTCCTTTTGTTTGTTTGTTTTTGTTTTTGTTTTGAGAGGGAGT
CTTGC (SEQ ID NO: 9)

eHGT_390h
CCCCAAGCAAGGTCAACAGCAGCTTCAATTACAAGACAAACTTAACAAGAGTTGCTATAAA
CCAAGTACTCTGTATTGACTTAAGAACAGGCTCCACTCCACATATTGCCACTAGCAGGATT
GTACAGGAATGCATATTGTAAAATAAAGGAAGGGGTGAGCTTTTTTCTTTGCCAGAATTTGC
GAGTGCACAGCGACTCCTCATTCACCTCTCTCCAACCAGCTAGCCGCTCAGCTCAATTCAC
CCCACACAAAGGCTGGAGCCTAGACCTCAATGGACCGAGTGAAACATGTTCAAAACTAGG
CTCTCTATTGTGACTGAATTTCTTAACATCTTTTCAAAAAGCGGAGAATGCCTTGAGGCTAA
AGGAAGAAACAGGCTAATGGTGAATTGGGAATTCTGAGCAAATTTCAGAGCCCTTTCCTCC
TAGCTTTTGAGGTTGAAAGCAAGCTCTTTCCTTTCAAGTTTCAAAGTCCTTTTTCCTCCCGC
AGTGTCACAGAAGGATTTGAAAAGAAGGTAATTGTGCTCGCAGTCTCCCTGATCAGAGCTT
ACGTCCTATTTCTGGTATTTCGGAATACTTCTTGCAATAATAGTGCATATAGCTCAATCCCTT
AACCGGCCTGCACTCTGCAATTGCTCATTAAATGAACAATTGCGGGTATAAAATGCCTTTTA
TGTTCAAGGTCTGGA (SEQ ID NO: 10)

eHGT_373m

FIG. 14 cont'd

ACTGATCAGCCATGCCAGCTGCCCAGGCATTTTCAATGATCCCTAGCTTTGGCTTTGGCCA
GGACACTCAACTCTTGTGGACACTGGCCCTGTGTCCTGCAGCAAAAGGGGAGCCTTGGCC
TCTTTTCCTTTTATGCAAAAGTTCTTTTAAGGCAGCGCTCTGGCCCAGTGGGCAGGCAGCT
GGAGTCCTCTCTGTTCCCGGCAGGGGCACTCTTCAAGGGTTGTGCCCCAGCTCCTTCAGC
AGCTGGAAGGCTGGTTGGCTTTTGTTTGCTTTTTCTTTCCTTTTTTTTTTTTTTTTTTCCACTT
CCATCCCTTACCCCAAACATTTCTCTGAAGGAATG (SEQ ID NO: 11)

eHGT_375m
AGATATCTGACAATGAGCAGAAGCTTGTATAGAGTGTATAAAACAAGTTTCCCCTTCTTATT
TTCTGCTTTAAAACGCACAATCTCCCACTCCCAGCCTCTTCTGTCCAGTCTATAGAGTTAAA
ATTGACCACTGCTCGTCCTGAGCCACTTCAAAGTAATTTAGTCCAGAACGGAAGGCTTGGC
GCTTGGACAATCTTCATGGTGTGTTAAGATTTCTATGGCAATTGGCAAGCAAGACGTTCCC
ATCTGAATATGTCTCAAAGAAAGCCACTTATTGACACTAAGTTGGCTATCATCAAAGACTTC
ATCTTCCCTCTGCA (SEQ ID NO: 12)

eHGT_386m
ACAGGGCTGGGGATCAAAGGGGGCCAGCACCATGTCTAGAGGGCAGAGATCCTGCCACT
GGCTGGTGGCCCCTGGTTGTCTTGGAACTCCACCTGCTGTTCATCAGCAGGGCTATGGGA
ATCTGGACACACCTCCTCTGGCTCAGACCAATCCCAGCTGCTGCTGGCCAGGGCGGGCCT
GGGCCATGGGATCTGGAGGAAAGGTCACTTGTGACTGGAAGGACAGTGGGCAGGAAGAA
AGAGCCAAGCTAGAGTAGGGGAGGGCACCCCTCCTTGTCCTCTAGACTTCCTGTTCCATC
TCCTGTCCCTGTAGGGCTCCAGCTAAAAGC (SEQ ID NO: 13)

eHGT_387m
TGGCTCTGGCAAAGGATGGAGATTTTTTTTTTAACCTATTCTTTTGAGAATAAGGAGGGTCT
TTGTTTCTCTTCCCTGGGAACCGAGCAGCCCCTTTCTCTGAGGGTAGAGTTGGGGCATGG
CTCAGCAGAACGACAGTCATCCTGGCTAGTAAGACTCAGAGGCTGGCCTTCAAAGGCTTG
AGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGC
GGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAA
TTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAG
GGGAAGGGGATAGGTTGCAATGGTCTGCTTAGCCTCCCGAAGGCTGCAGAAGCTGCTAAG
ACAGGGGTAACTGAAAGATAGACTCTGTGAGGTTGAACAGCAATCTTTATGACCTGACCCC
CTAGGCA (SEQ ID NO: 14)

eHGT_387m(core2)
GCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGCG
GGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAAT
TTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGG
GG (SEQ ID NO: 84)

eHGT_390m
TCCAGACCTTGAACATAAAAGGCATTTTATACCCGCAATTGTTCATTTAATGAGCAATTGCG
GAGTGCAGGCCGGTTAAGGGATTGAGCTATATGCACTATTATTGCAAGAAGTATTCCGAAA
TACCAGAAATAGGACGTAAGCTCTGATCAGGGAGACTGCGAGCACAATTACCTTCTTTTCA
AATCCTTCTGTGACACTGCGGGAGGAAAAAGGACTTTGAAACTTGAAAGGAAAGAGCTTGC

FIG. 14 cont'd

TTTCAACCTCAAAAGCTAGGAGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACC
ATTAGCCTGTTTCTTCCTTTAGCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAAT
TCAGTCACAATAGAGAGCCTAGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCT
CCAGCCTTTGTGTGGGGTGAATTGAGCTGAGCGGCTAGCTGGTTGGAGAGAGGTGAATGA
GAAGTCGCTGTGCAGTTGCAAATTCTGGCAAAGAAAAAAAAAAAGCTCACCCCTTCCTTTAT
TTTGTAATATGCATTCCTGTACAATCCTGCCAGTGGCAATCTGTGGAGTTCAGTGTGTCCCT
AAGTCAATATGGAGTACTTGGTTTATAGCAACTCTTGTTAAGTTTGTCTTGTAATTGAAGCT
GCTGTTGACCTTGCTTGGGG (SEQ ID NO: 15)

eHGT_390m(core2)
GAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGGCTCTGAAATTTGCTC
AGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGGCATTCTCCGCTTTTT
GAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAACATGTTTCACTCGGT
CCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTGAGCGGCTAGCTGGT
TGGAG (SEQ ID NO: 85)

Combination Concatenated Enhancer (eHGT_387m(core2)-hI56i(core))
GCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGCG
GGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAAT
TTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGG
GGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGA
CGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAA
GAGAATGGG (SEQ ID NO: 95)

Combination Concatenated Enhancer (eHGT_390m(core2)-hI56i(core))
GAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGGCTCTGAAATTTGCTC
AGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGGCATTCTCCGCTTTTT
GAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAACATGTTTCACTCGGT
CCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTGAGCGGCTAGCTGGT
TGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTT
TGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATT
TAAGAGAATGG (SEQ ID NO: 86)

3xhI56i(core)_eHGT_390m Enhancer
CTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACG
GTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGA
GAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATC
TTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCA
TTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGT
AATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTAT
GGCTGCATTTAAGAGAATGGAGCTCTCCAGACCTTGAACATAAAAGGCATTTTATACCCGC
AATTGTTCATTTAATGAGCAATTGCGGAGTGCAGGCCGGTTAAGGGATTGAGCTATATGCA
CTATTATTGCAAGAAGTATTCCGAAATACCAGAAATAGGACGTAAGCTCTGATCAGGGAGA
CTGCGAGCACAATTACCTTCTTTTCAAATCCTTCTGTGACACTGCGGGAGGAAAAAGGACT
TTGAAACTTGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGGCTCTGA
AATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGGCATTCT
CCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAACATGTT

FIG. 14 cont'd

TCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTGAGCGG
CTAGCTGGTTGGAGAGAGGTGAATGAGAAGTCGCTGTGCAGTTGCAAATTCTGGCAAAGA
AAAAAAAAAAGCTCACCCCTTCCTTTATTTTGTAATATGCATTCCTGTACAATCCTGCCAGT
GGCAATCTGTGGAGTTCAGTGTGTCCCTAAGTCAATATGGAGTACTTGGTTTATAGCAACT
CTTGTTAAGTTTGTCTTGTAATTGAAGCTGCTGTTGACCTTGCTTGGGG (SEQ ID NO: 87)

3X Combination Concatenated Enhancer (eHGT_387m(core2)-hl56i(core)-eHGT_387m(core2)-hl56i(core)- eHGT_387m(core2)-hl56i(core))
GCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGCG
GGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAAT
TTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGG
GGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGA
CGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAA
GAGAATGGGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTC
CCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAAC
TGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTT
CCTTAGAGGGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAA
TTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGG
CTGCATTTAAGAGAATGGGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCT
TGAATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTC
TTAAAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCA
GGAGGAGTTCCTTAGAGGGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGG (SEQ ID NO: 88)

3X Combination Concatenated Enhancer (eHGT_390m(core2)-hl56i(core)-eHGT_390m(core2)-hl56i(core)- eHGT_390m(core2)-hl56i(core))
GAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGGCTCTGAAATTTGCTC
AGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGGCATTCTCCGCTTTTT
GAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAACATGTTTCACTCGGT
CCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTGAGCGGCTAGCTGGT
TGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTT
TGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATT
TAAGAGAATGGGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGGCTC
TGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGGCAT
TCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAACAT
GTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTGAG
CGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTAC
AGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTAA
TTATGGCTGCATTTAAGAGAATGGGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGG
AGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTA
GCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCT
AGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGA
ATTGAGCTGAGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTG
GAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGA
AAAATTACTCTAATTATGGCTGCATTTAAGAGAATGG (SEQ ID NO: 89)

FIG. 14 cont'd

Beta-Globin Minimal Promoter (pBGmin/minBGlobin/minBGprom):
GGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTG (SEQ ID NO: 16)

minCMV Promoter:
GAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGA TCGCCTGG (SEQ ID NO: 17)

Mutated minCMV Promoter (SacI RE site removed):
GAGGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGA TCGCCTGG (SEQ ID NO: 18)

minRho Promoter:
GATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCCTGGGGGGGGAGTTG GAGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGAGCTCCGGCCTCAGAAGCATCC CC (SEQ ID NO: 19)

minRho* Promoter:
GATTCAGCCGGGAGCTTAGGGAGGGGAGGTCACTTCATAAGGGCTTGGGGGGGGAGTTG GAGCCACGAGTCGTCCAGCCGGAGCCCCGTGTGGCTGTGCTCCGGCCTCAGAAGCATCC CC (SEQ ID NO: 20)

Hsp68 minimal Promoter (proHsp68):
CAGGAACATCCAAACTGAGCAGCCGGGGTCCCCCCCACCCCCCACCCCGCCCCACGCGG CAACTTTGAGCCTGTGCTGGGACAGAGCCTCTAGTTCCTAAATTAGTCCATGAGGTCAGAG GCAGCACTGCCATTGTAACGCGATTGGAGAGGATCACGTCACCGGACACGCCCCCAGGC ATCTCCCTGGGTCTCCTAAACTTGGCGGGGAGAAGTTTTAGCCCTTAAGTTTTAGCCTTTAA CCCCCATATTCAGAACTGTGCGAGTTGGCGAAACCCCACAAATCACAACAAACTGTACACA ACACCGAGCTAGAGGTGATCTTTCTTGTCCATTCCACACAGGCCTTAGTAATGCGTCGCCA TAGCAACAGTGTCACTAGTAGCACCAGCACTTCCCCACACCCTCCCCCTCAGGAATCCGTA CTCTCCAGTGAACCCCAGAAACCTCTGGAGAGTTCTGGACAAGGGCGGAACCCACAACTC CGATTACTCAAGGGAGGCGGGGAAGCTCCACCAGACGCGAAACTGCTGGAAGATTCCTG GCCCCAAGGCCTCCTCCGGCTCGCTGATTGGCCCAGCGGAGAGTGGGCGGGGCCGGTG AAGACTCCTTAAAGGCGCAGGGCGGCGAGCAGGTCACCAGACGCTGACAGCTACTCAGA ACCAAATCTGGTTCCATCCAGAGACAAGCGAAGACAAGAGAAGCAGAGCGAGCGGCGCGT TCCCGATCCTCGGCCAGGACCAGCCTTCCCCAGAGCATCCCTGCCGCGGAGCGCAACCT TCCCAGGAGCATCCCTGCCGCGGAGCGCAACTTTCCCCGGAGCATCCACGCCGCGGAGC GCAGCCTTCCAGAAGCAGAGCGCGGCGCC (SEQ ID NO: 21)

SLC6A1 encoding sequence from CN2972
ATGGCTACGAATGGCTCAAAAGTCGCTGATGGTCAGATCAGCACGGAAGTGAGCGAGGCC CCTGTGGCCAACGACAAGCCTAAAACTCTGGTCGTAAAAGTACAGAAGAAAGCAGCGGAT CTGCCAGACCGTGATACATGGAAAGGACGTTTCGATTTCCTGATGAGCTGCGTAGGCTATG CAATCGGACTGGGAAATGTCTGGCGCTTCCCTTACCTGTGTGGTAAAAATGGTGGAGGCG CCTTTCTCATTCCTTATTTTCTCACCCTCATTTTTGCCGGGGTTCCCTTGTTTCTCCTGGAAT

FIG. 14 cont'd

GCAGCCTCGGGCAATATACCAGCATTGGCGGACTCGGTGTCTGGAAACTGGCGCCTATGT
TTAAGGGCGTTGGGCTGGCTGCAGCTGTGCTGTCCTTCTGGCTGAACATCTATTACATTGT
CATCATTAGCTGGGCGATCTACTATCTGTACAATAGCTTCACTACGACTCTGCCGTGGAAA
CAATGTGACAATCCTTGGAATACGGATCGCTGCTTTAGCAATTATAGCATGGTTAACACCAC
TAACATGACTTCCGCTGTAGTGGAGTTCTGGGAGAGGAACATGCATCAGATGACCGACGG
ACTTGACAAACCCGGACAGATTAGGTGGCCCCTCGCCATTACGCTCGCAATCGCCTGGAT
CCTTGTCTATTTTTGTATCTGGAAAGGTGTCGGCTGGACTGGAAAGGTGGTCTATTTCAGC
GCGACGTATCCATACATTATGCTCATTATCCTGTTTTTCCGGGGAGTAACGCTCCCGGGTG
CCAAAGAAGGGATCCTGTTCTACATTACGCCTAATTTTAGGAAGCTCTCTGATAGTGAGGTT
TGGCTGGATGCTGCCACGCAAATTTTTTTCTCCTACGGTCTGGGCCTCGGGTCCCTGATCG
CCCTGGGCAGTTATAATTCTTTTCATAACAATGTGTACAGAGATAGCATTATCGTGTGCTGT
ATCAATTCTTGCACATCCATGTTTGCCGGTTTCGTGATCTTCTCCATCGTCGGGTTTATGGC
CCACGTCACCAAACGTTCTATTGCGGATGTCGCTGCCTCAGGTCCCGGTCTGGCCTTTCTT
GCGTACCCTGAGGCTGTTACACAACTGCCAATCAGTCCTCTGTGGGCCATTCTTTTCTTTT
CCATGCTTCTGATGCTGGGCATTGACTCTCAGTTTTGTACTGTTGAGGGTTTCATCACCGC
CCTGGTTGACGAATATCCCCGTCTCTTGAGAAATCGTCGCGAACTCTTCATTGCCGCAGTG
TGTATTATCTCCTACCTCATTGGCCTGTCCAATATCACGCAGGGAGGCATTTACGTATTCAA
ACTCTTTGATTACTATTCCGCAAGTGGAATGAGCCTCTTGTTTCTCGTGTTTTTCGAGTGTG
TTAGCATCTCCTGGTTCTACGGAGTAAACCGTTTTTATGATAATATCCAGGAAATGGTGGG
GAGTCGTCCGTGTATCTGGTGGAAACTTTGTTGGTCATTCTTTACGCCCATCATTGTCGCT
GGTGTGTTCATTTTTTCTGCCGTGCAAATGACCCAACTGACGATGGGAAATTACGTATTCC
CCAAATGGGGCCAGGGCGTGGGGTGGCTCATGGCTCTGAGTTCCATGGTTTTGATCCCTG
GGTATATGGCCTACATGTTCCTGACACTGAAAGGGTCACTGAAACAGCGCATTCAAGTGAT
GGTTCAACCCAGCGAGGACATTGTCAGACCGGAAAACGGGCCTGAGCAGCCCCAAGCGG
GATCCTCTACAAGTAAAGAGGCCTATATCACTCGCACGCGCCCCCTG (SEQ ID NO: 22)

Myc-DDK tag sequence in CN2972
GAGCAAAAGCTCATCTCTGAGGAAGACCTGGCGGCTAACGATATTCTTGACTATAAAGACG
ATGACGATAAA (SEQ ID NO: 23)

CN2972
ATGGCTACGAATGGCTCAAAAGTCGCTGATGGTCAGATCAGCACGGAAGTGAGCGAGGCC
CCTGTGGCCAACGACAAGCCTAAAACTCTGGTCGTAAAAGTACAGAAGAAAGCAGCGGAT
CTGCCAGACCGTGATACATGGAAAGGACGTTTCGATTTCCTGATGAGCTGCGTAGGCTATG
CAATCGGACTGGGAAATGTCTGGCGCTTCCCTTACCTGTGTGGTAAAAATGGTGGAGGCG
CCTTTCTCATTCCTTATTTTCTCACCCTCATTTTTGCCGGGGTTCCCTTGTTTCTCCTGGAAT
GCAGCCTCGGGCAATATACCAGCATTGGCGGACTCGGTGTCTGGAAACTGGCGCCTATGT
TTAAGGGCGTTGGGCTGGCTGCAGCTGTGCTGTCCTTCTGGCTGAACATCTATTACATTGT
CATCATTAGCTGGGCGATCTACTATCTGTACAATAGCTTCACTACGACTCTGCCGTGGAAA
CAATGTGACAATCCTTGGAATACGGATCGCTGCTTTAGCAATTATAGCATGGTTAACACCAC
TAACATGACTTCCGCTGTAGTGGAGTTCTGGGAGAGGAACATGCATCAGATGACCGACGG
ACTTGACAAACCCGGACAGATTAGGTGGCCCCTCGCCATTACGCTCGCAATCGCCTGGAT
CCTTGTCTATTTTTGTATCTGGAAAGGTGTCGGCTGGACTGGAAAGGTGGTCTATTTCAGC
GCGACGTATCCATACATTATGCTCATTATCCTGTTTTTCCGGGGAGTAACGCTCCCGGGTG
CCAAAGAAGGGATCCTGTTCTACATTACGCCTAATTTTAGGAAGCTCTCTGATAGTGAGGTT

FIG. 14 cont'd

TGGCTGGATGCTGCCACGCAAATTTTTTTCTCCTACGGTCTGGGCCTCGGGTCCCTGATCG
CCCTGGGCAGTTATAATTCTTTTCATAACAATGTGTACAGAGATAGCATTATCGTGTGCTGT
ATCAATTCTTGCACATCCATGTTTGCCGGTTTCGTGATCTTCTCCATCGTCGGGTTTATGGC
CCACGTCACCAAACGTTCTATTGCGGATGTCGCTGCCTCAGGTCCCGGTCTGGCCTTTCTT
GCGTACCCTGAGGCTGTTACACAACTGCCAATCAGTCCTCTGTGGGCCATTCTTTTCTTTT
CCATGCTTCTGATGCTGGGCATTGACTCTCAGTTTTGTACTGTTGAGGGTTTCATCACCGC
CCTGGTTGACGAATATCCCCGTCTCTTGAGAAATCGTCGCGAACTCTTCATTGCCGCAGTG
TGTATTATCTCCTACCTCATTGGCCTGTCCAATATCACGCAGGGAGGCATTTACGTATTCAA
ACTCTTTGATTACTATTCCGCAAGTGGAATGAGCCTCTTGTTTCTCGTGTTTTTCGAGTGTG
TTAGCATCTCCTGGTTCTACGGAGTAAACCGTTTTTATGATAATATCCAGGAAATGGTGGG
GAGTCGTCCGTGTATCTGGTGGAAACTTTGTTGGTCATTCTTTACGCCCATCATTGTCGCT
GGTGTGTTCATTTTTTCTGCCGTGCAAATGACCCAACTGACGATGGGAAATTACGTATTCC
CCAAATGGGGCCAGGGCGTGGGGTGGCTCATGGCTCTGAGTTCCATGGTTTTGATCCCTG
GGTATATGGCCTACATGTTCCTGACACTGAAAGGGTCACTGAAACAGCGCATTCAAGTGAT
GGTTCAACCCAGCGAGGACATTGTCAGACCGGAAAACGGGCCTGAGCAGCCCCAAGCGG
GATCCTCTACAAGTAAAGAGGCCTATATCACTCGCACGCGCCCCCTGGAGCAAAAGCTCAT
CTCTGAGGAAGACCTGGCGGCTAACGATATTCTTGACTATAAAGACGATGACGATAAAGTG
TAA (SEQ ID NO: 24)

SLC6A1 encoding sequence from CN2973

ATGGCAACTAATGGGTCCAAGGTCGCTGACGGCCAGATTAGCACAGAAGTGAGTGAAGCC
CCGGTTGCGAATGATAAACCCAAAACACTTGTGGTAAAAGTGCAGAAAAAAGCAGCTGATC
TGCCGGACCGGGACACCTGGAAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCTGC
TTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAA
TCATGTTCATACCTCTTATCTTCCTCCCACAGGGCAGATTTGACTTTTTGATGTCCTGTGTG
GGCTACGCTATCGGCCTCGGCAACGTCTGGAGATTCCCGTACTTGTGTGGAAAGAACGGA
GGCGGAGCGTTTCTTATCCCCTATTTCCTGACCTTGATCTTTGCTGGAGTTCCGTTGTTTCT
CCTGGAATGCTCTTTGGGCCAATATACGTCAATAGGTGGACTTGGTGTGTGGAAGCTTGCT
CCCATGTTCAAGGGAGTGGGACTTGCTGCCGCCGTCTTGAGTTTTTGGTTGAACATCTATT
ACATTGTTATCATCTCATGGGCAATATACTACCTTTATAACTCCTTCACTACCACCCTCCCG
TGGAAGCAATGTGACAACCCATGGAACACTGATCGATGTTTCTCAAACTACAGCATGGTCA
ATACCACAAATATGACTTCCGCAGTTGTAGAGTTCTGGGAGCGCAACATGCATCAAATGAC
AGACGGACTCGATAAACCGGGACAAATCCGGTGGCCTCTTGCGATTACGCTTGCTATTGC
CTGGATCCTCGTGTACTTTTGTATCTGGAAAGGAGTCGGGTGGACAGGGAAAGTCGTCTAT
TTTAGTGCAACATACCCGTACATCATGCTGATCATCCTGTTCTTCCGAGGTGTCACCCTTCC
CGGCGCGAAAGAGGGGATACTTTTTTATATCACACCGAATTTCCGAAAACTGTCAGACAGT
GAAGTGTGGCTGGACGCCGCTACTCAGATATTCTTCTCTTATGGTCTGGGTCTGGGTAGCC
TGATTGCTCTTGGTTCCTACAACTCTTTCCATAATAATGTGTACAGAGACTCTATCATTGTCT
GCTGCATCAATAGCTGTACGAGCATGTTTGCGGGTTTCGTAATCTTTTCTATTGTTGGTTTC
ATGGCGCATGTAACTAAGCGGAGTATTGCTGACGTCGCGGCGAGCGGCCCTGGGTTGGC
ATTTTTGGCATACCCGGAGGCTGTGACACAACTGCCTATTAGTCCGTTGTGGGCGATTCTG
TTCTTCTCCATGCTTCTCATGCTTGGCATTGACTCACAATTCTGTACTGTGGAAGGTTTCAT
CACAGCATTGGTGGATGAATATCCCCGGTTGCTCAGAAATAGGCGGGAGCTGTTCATTGC
GGCTGTGTGTATAATTTCTTACCTCATCGGGTTGAGCAACATCACGCAAGGGGGAATCTAT
GTATTTAAACTGTTTGATTACTACAGCGCTAGTGGAATGTCTTTGCTCTTTCTTGTATTCTTC

FIG. 14 cont'd

GAGTGTGTGAGTATCAGTTGGTTCTATGGGGTCAACAGATTTTATGATAATATACAAGAAAT
GGTCGGGTCTAGGCCATGTATTTGGTGGAAACTCTGTTGGTCTTTTTTTACCCCAATAATAG
TTGCTGGTGTGTTCATCTTCTCCGCGGTCCAAATGACCCAGCTTACCATGGGGAATTACGT
GTTCCCGAAATGGGGCCAAGGGGTGGGCTGGCTCATGGCCCTTAGCTCAATGGTACTCAT
CCCTGGATATATGGCTTATATGTTTTTGACCCTCAAGGGGAGTCTCAAACAGCGGATCCAG
GTTATGGTACAACCAAGCGAGGACATAGTCCGACCGGAGAATGGGCCAGAGCAACCCCAA
GCCGGCAGCAGTACCAGCAAGGAGGCTTATATTACCCGAACCAGGCCCTTG (SEQ ID NO:
25)

Myc-DDK tag sequence in CN2973
GAACAGAAGCTCATTTCAGAGGAGGATCTGGCGGCCAATGACATTCTCGATTACAAAGACG
ATGATGATAAA (SEQ ID NO: 26)

CN2973
ATGGCAACTAATGGGTCCAAGGTCGCTGACGGCCAGATTAGCACAGAAGTGAGTGAAGCC
CCGGTTGCGAATGATAAACCCAAAACACTTGTGGTAAAAGTGCAGAAAAAAGCAGCTGATC
TGCCGGACCGGGACACCTGGAAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCTGC
TTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAA
TCATGTTCATACCTCTTATCTTCCTCCCACAGGGCAGATTTGACTTTTTGATGTCCTGTGTG
GGCTACGCTATCGGCCTCGGCAACGTCTGGAGATTCCCGTACTTGTGTGGAAAGAACGGA
GGCGGAGCGTTTCTTATCCCCTATTTCCTGACCTTGATCTTTGCTGGAGTTCCGTTGTTTCT
CCTGGAATGCTCTTTGGGCCAATATACGTCAATAGGTGGACTTGGTGTGTGGAAGCTTGCT
CCCATGTTCAAGGGAGTGGGACTTGCTGCCGCCGTCTTGAGTTTTTGGTTGAACATCTATT
ACATTGTTATCATCTCATGGGCAATATACTACCTTTATAACTCCTTCACTACCACCCTCCCG
TGGAAGCAATGTGACAACCCATGGAACACTGATCGATGTTTCTCAAACTACAGCATGGTCA
ATACCACAAATATGACTTCCGCAGTTGTAGAGTTCTGGGAGCGCAACATGCATCAAATGAC
AGACGGACTCGATAAACCGGGACAAATCCGGTGGCCTCTTGCGATTACGCTTGCTATTGC
CTGGATCCTCGTGTACTTTTGTATCTGGAAAGGAGTCGGGTGGACAGGGAAAGTCGTCTAT
TTTAGTGCAACATACCCGTACATCATGCTGATCATCCTGTTCTTCCGAGGTGTCACCCTTCC
CGGCGCGAAAGAGGGGATACTTTTTTATATCACACCGAATTTCCGAAAACTGTCAGACAGT
GAAGTGTGGCTGGACGCCGCTACTCAGATATTCTTCTCTTATGGTCTGGGTCTGGGTAGCC
TGATTGCTCTTGGTTCCTACAACTCTTTCCATAATAATGTGTACAGAGACTCTATCATTGTCT
GCTGCATCAATAGCTGTACGAGCATGTTTGCGGGTTTCGTAATCTTTTCTATTGTTGGTTTC
ATGGCGCATGTAACTAAGCGGAGTATTGCTGACGTCGCGGCGAGCGGCCCTGGGTTGGC
ATTTTTGGCATACCCGGAGGCTGTGACACAACTGCCTATTAGTCCGTTGTGGGCGATTCTG
TTCTTCTCCATGCTTCTCATGCTTGGCATTGACTCACAATTCTGTACTGTGGAAGGTTTCAT
CACAGCATTGGTGGATGAATATCCCCGGTTGCTCAGAAATAGGCGGGAGCTGTTCATTGC
GGCTGTGTGTATAATTTCTTACCTCATCGGGTTGAGCAACATCACGCAAGGGGGAATCTAT
GTATTTAAACTGTTTGATTACTACAGCGCTAGTGGAATGTCTTTGCTCTTTCTTGTATTCTTC
GAGTGTGTGAGTATCAGTTGGTTCTATGGGGTCAACAGATTTTATGATAATATACAAGAAAT
GGTCGGGTCTAGGCCATGTATTTGGTGGAAACTCTGTTGGTCTTTTTTTACCCCAATAATAG
TTGCTGGTGTGTTCATCTTCTCCGCGGTCCAAATGACCCAGCTTACCATGGGGAATTACGT
GTTCCCGAAATGGGGCCAAGGGGTGGGCTGGCTCATGGCCCTTAGCTCAATGGTACTCAT
CCCTGGATATATGGCTTATATGTTTTTGACCCTCAAGGGGAGTCTCAAACAGCGGATCCAG
GTTATGGTACAACCAAGCGAGGACATAGTCCGACCGGAGAATGGGCCAGAGCAACCCCAA

FIG. 14 cont'd

GCCGGCAGCAGTACCAGCAAGGAGGCTTATATTACCCGAACCAGGCCCTTGGAACAGAAG
CTCATTTCAGAGGAGGATCTGGCGGCCAATGACATTCTCGATTACAAAGACGATGATGATA
AAGTGTAA(SEQ ID NO: 27)

SLC6A1 encoding sequence from CN2974

ATGGCAACTAATGGGTCCAAGGTCGCTGACGGCCAGATTAGCACAGAAGTGAGTGAAGCC
CCGGTTGCGAATGATAAACCCAAAACACTTGTGGTAAAAGTGCAGAAAAAAGCAGCTGATC
TGCCGGACCGGGACACCTGGAAGGGCAGATTTGACTTTTTGATGTCCTGTGTGGGCTACG
CTATCGGCCTCGGCAACGTCTGGAGATTCCCGTACTTGTGTGGAAAGAACGGAGGCGGAG
CGTTTCTTATCCCCTATTTCCTGACCTTGATCTTTGCTGGAGTTCCGTTGTTTCTCCTGGAA
TGCTCTTTGGGCCAATATACGTCAATAGGTGGACTTGGTGTGTGGAAGCTTGCTCCCATGT
TCAAGGGAGTGGGACTTGCTGCCGCCGTCTTGAGTTTTTGGTTGAACATCTATTACATTGT
TATCATCTCATGGGCAATATACTACCTTTATAACTCCTTCACTACCACCCTCCCGTGGAAGC
AATGTGACAACCCATGGAACACTGATCGATGTTTCTCAAACTACAGCATGGTCAATACCAC
AAATATGACTTCCGCAGTTGTAGAGTTCTGGGAGCGCAACATGCATCAAATGACAGACGGA
CTCGATAAACCGGGACAAATCCGGTGGCCTCTTGCGATTACGCTTGCTATTGCCTGGATCC
TCGTGTACTTTTGTATCTGGAAAGGAGTCGGGTGGACAGGGAAAGTCGTCTATTTTAGTGC
AACATACCCGTACATCATGCTGATCATCCTGTTCTTCCGAGGTGTCACCCTTCCCGGCGCG
AAAGAGGGGATACTTTTTTATATCACACCGAATTTCCGAAAACTGTCAGACAGTGAAGTGTG
GCTGGACGCCGCTACTCAGATATTCTTCTCTTATGGTCTGGGTCTGGGTAGCCTGATTGCT
CTTGGTTCCTACAACTCTTTCCATAATAATGTGTACAGAGACTCTATCATTGTCTGCTGCAT
CAATAGCTGTACGAGCATGTTTGCGGGTTTCGTAATCTTTTCTATTGTTGGTTTCATGGCGC
ATGTAACTAAGCGGAGTATTGCTGACGTCGCGGCGAGCGGCCCTGGGTTGGCATTTTTGG
CATACCCGGAGGCTGTGACACAACTGCCTATTAGTCCGTTGTGGGCGATTCTGTTCTTCTC
CATGCTTCTCATGCTTGGCATTGACTCACAATTCTGTACTGTGGAAGGTTTCATCACAGCAT
TGGTGGATGAATATCCCCGGTTGCTCAGAAATAGGCGGGAGCTGTTCATTGCGGCTGTGT
GTATAATTTCTTACCTCATCGGGTTGAGCAACATCACGCAAGGGGGAATCTATGTATTTAAA
CTGTTTGATTACTACAGCGCTAGTGGAATGTCTTTGCTCTTTCTTGTATTCTTCGAGTGTGT
GAGTATCAGTTGGTTCTATGGGGTCAACAGATTTTATGATAATATACAAGAAATGGTCGGGT
CTAGGCCATGTATTTGGTGGAAACTCTGTTGGTCTTTTTTTACCCCAATAATAGTTGCTGGT
GTGTTCATCTTCTCCGCGGTCCAAATGACCCAGCTTACCATGGGGAATTACGTGTTCCCGA
AATGGGGCCAAGGGGTGGGCTGGCTCATGGCCCTTAGCTCAATGGTACTCATCCCTGGAT
ATATGGCTTATATGTTTTTGACCCTCAAGGGGAGTCTCAAACAGCGGATCCAGGTTATGGT
ACAACCAAGCGAGGACATAGTCCGACCGGAGAATGGGCCAGAGCAACCCCAAGCCGGCA
GCAGTACCAGCAAGGAGGCTTATATTACCCGAACCAGGCCCTTG (SEQ ID NO: 28)

Myc-DDK tag sequence in CN2974

GAACAGAAGCTCATTTCAGAGGAGGATCTGGCGGCCAATGACATTCTCGATTACAAAGACG
ATGATGATAAA (SEQ ID NO: 26)

CN2974

ATGGCAACTAATGGGTCCAAGGTCGCTGACGGCCAGATTAGCACAGAAGTGAGTGAAGCC
CCGGTTGCGAATGATAAACCCAAAACACTTGTGGTAAAAGTGCAGAAAAAAGCAGCTGATC
TGCCGGACCGGGACACCTGGAAGGGCAGATTTGACTTTTTGATGTCCTGTGTGGGCTACG
CTATCGGCCTCGGCAACGTCTGGAGATTCCCGTACTTGTGTGGAAAGAACGGAGGCGGAG

FIG. 14 cont'd

CGTTTCTTATCCCCTATTTCCTGACCTTGATCTTTGCTGGAGTTCCGTTGTTTCTCCTGGAA
TGCTCTTTGGGCCAATATACGTCAATAGGTGGACTTGGTGTGTGGAAGCTTGCTCCCATGT
TCAAGGGAGTGGGACTTGCTGCCGCCGTCTTGAGTTTTTGGTTGAACATCTATTACATTGT
TATCATCTCATGGGCAATATACTACCTTTATAACTCCTTCACTACCACCCTCCCGTGGAAGC
AATGTGACAACCCATGGAACACTGATCGATGTTTCTCAAACTACAGCATGGTCAATACCAC
AAATATGACTTCCGCAGTTGTAGAGTTCTGGGAGCGCAACATGCATCAAATGACAGACGGA
CTCGATAAACCGGGACAAATCCGGTGGCCTCTTGCGATTACGCTTGCTATTGCCTGGATCC
TCGTGTACTTTTGTATCTGGAAAGGAGTCGGGTGGACAGGGAAAGTCGTCTATTTTAGTGC
AACATACCCGTACATCATGCTGATCATCCTGTTCTTCCGAGGTGTCACCCTTCCCGGCGCG
AAAGAGGGGATACTTTTTTATATCACACCGAATTTCCGAAAACTGTCAGACAGTGAAGTGTG
GCTGGACGCCGCTACTCAGATATTCTTCTCTTATGGTCTGGGTCTGGGTAGCCTGATTGCT
CTTGGTTCCTACAACTCTTTCCATAATAATGTGTACAGAGACTCTATCATTGTCTGCTGCAT
CAATAGCTGTACGAGCATGTTTGCGGGTTTCGTAATCTTTTCTATTGTTGGTTTCATGGCGC
ATGTAACTAAGCGGAGTATTGCTGACGTCGCGGCGAGCGGCCCTGGGTTGGCATTTTTGG
CATACCCGGAGGCTGTGACACAACTGCCTATTAGTCCGTTGTGGGCGATTCTGTTCTTCTC
CATGCTTCTCATGCTTGGCATTGACTCACAATTCTGTACTGTGGAAGGTTTCATCACAGCAT
TGGTGGATGAATATCCCCGGTTGCTCAGAAATAGGCGGGAGCTGTTCATTGCGGCTGTGT
GTATAATTTCTTACCTCATCGGGTTGAGCAACATCACGCAAGGGGGAATCTATGTATTTAAA
CTGTTTGATTACTACAGCGCTAGTGGAATGTCTTTGCTCTTTCTTGTATTCTTCGAGTGTGT
GAGTATCAGTTGGTTCTATGGGGTCAACAGATTTTATGATAATATACAAGAAATGGTCGGGT
CTAGGCCATGTATTTGGTGGAAACTCTGTTGGTCTTTTTTTACCCCAATAATAGTTGCTGGT
GTGTTCATCTTCTCCGCGGTCCAAATGACCCAGCTTACCATGGGGAATTACGTGTTCCCGA
AATGGGGCCAAGGGGTGGGCTGGCTCATGGCCCTTAGCTCAATGGTACTCATCCCTGGAT
ATATGGCTTATATGTTTTTGACCCTCAAGGGGAGTCTCAAACAGCGGATCCAGGTTATGGT
ACAACCAAGCGAGGACATAGTCCGACCGGAGAATGGGCCAGAGCAACCCCAAGCCGGCA
GCAGTACCAGCAAGGAGGCTTATATTACCCGAACCAGGCCCTTGGAACAGAAGCTCATTTC
AGAGGAGGATCTGGCGGCCAATGACATTCTCGATTACAAAGACGATGATGATAAAGTGTAA
(SEQ ID NO: 30)

SLC6A1 encoding sequence from CN2975

ATGGCGACCAACGGCAGCAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGAGGC
CCCTGTGGCCAATGACAAGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGGCAGA
CCTCCCCGACCGGGACACGTGGAAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCT
GCTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCT
AATCATGTTCATACCTCTTATCTTCCTCCCACAGGGCCGCTTCGACTTCCTCATGTCCTGTG
TGGGCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATG
GTGGGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTT
CCTGCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGC
TGGCTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACA
TCTACTACATCGTCATCATCTCCTGGGCCATTTACTACCTGTACAACTCCTTCACCACGACA
CTGCCGTGGAAACAGTGCGACAACCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGC
ATGGTCAACACTACCAACATGACCAGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCAT
CAGATGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCT
GGCCATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAA
GGTGGTCTACTTTTCAGCCACATACCCCTACATCATGCTGATCATCCTGTTCTTCCGTGGA

FIG. 14 cont'd

GTGACGCTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACATCACACCCAACTTCCGCAAG
CTGTCTGACTCCGAGGTGTGGCTGGATGCGGCAACCCAGATCTTCTTCTCATACGGGCTG
GGCCTGGGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTTCCACAACAATGTCTACAGG
GACTCCATCATCGTCTGCTGCATCAATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCT
TCTCCATCGTGGGCTTCATGGCCCATGTCACGAAGAGGTCCATTGCTGATGTGGCGGCCT
CAGGCCCCGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTGACCCAGCTGCCTATCTCC
CCACTCTGGGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTTCT
GCACTGTGGAGGGCTTCATCACAGCCCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACC
GCAGAGAGCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCTGATCGGTCTCTCTAACAT
CACTCAGGGGGGTATTTATGTCTTCAAACTCTTTGACTACTACTCTGCCAGTGGCATGAGC
CTGCTGTTCCTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTTTACGGTGTCAACCGATT
CTATGACAATATCCAAGAGATGGTTGGATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGG
TCTTTCTTCACACCAATCATTGTGGCGGGCGTGTTCATTTTCAGTGCTGTGCAGATGACGC
AACTCACCATGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGG
CTCTGTCTTCCATGGTCCTCATCCCCGGGTACATGGCCTACATGTTCCTCACCTTAAAGGG
CTCCCTGAAGCAGCGCATCCAAGTCATGGTCCAGCCCAGCGAAGACATCGTTCGCCCAGA
GAATGGTCCTGAGCAGCCCCAGGCGGGCAGCTCCACCAGCAAGGAGGCCTACATCACGC
GTACGCGGCCGCTC (SEQ ID NO: 31)

Myc-DDK tag sequence in CN2975
GAGCAGAAACTCATCTCAGAAGAGGATCTGGCAGCAAATGATATCCTGGATTACAAGGATG
ACGACGATAAG (SEQ ID NO: 32)

CN2975
ATGGCGACCAACGGCAGCAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGAGGC
CCCTGTGGCCAATGACAAGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGGCAGA
CCTCCCCGACCGGGACACGTGGAAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCT
GCTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCT
AATCATGTTCATACCTCTTATCTTCCTCCCACAGGGCCGCTTCGACTTCCTCATGTCCTGTG
TGGGCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATG
GTGGGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTT
CCTGCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGC
TGGCTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACA
TCTACTACATCGTCATCATCTCCTGGGCCATTTACTACCTGTACAACTCCTTCACCACGACA
CTGCCGTGGAAACAGTGCGACAACCCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGC
ATGGTCAACACTACCAACATGACCAGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCAT
CAGATGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCT
GGCCATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAA
GGTGGTCTACTTTTCAGCCACATACCCCTACATCATGCTGATCATCCTGTTCTTCCGTGGA
GTGACGCTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACATCACACCCAACTTCCGCAAG
CTGTCTGACTCCGAGGTGTGGCTGGATGCGGCAACCCAGATCTTCTTCTCATACGGGCTG
GGCCTGGGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTTCCACAACAATGTCTACAGG
GACTCCATCATCGTCTGCTGCATCAATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCT
TCTCCATCGTGGGCTTCATGGCCCATGTCACGAAGAGGTCCATTGCTGATGTGGCGGCCT
CAGGCCCCGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTGACCCAGCTGCCTATCTCC

FIG. 14 cont'd

CCACTCTGGGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTTCT
GCACTGTGGAGGGCTTCATCACAGCCCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACC
GCAGAGAGCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCTGATCGGTCTCTCTAACAT
CACTCAGGGGGGTATTTATGTCTTCAAACTCTTTGACTACTACTCTGCCAGTGGCATGAGC
CTGCTGTTCCTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTTTACGGTGTCAACCGATT
CTATGACAATATCCAAGAGATGGTTGGATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGG
TCTTTCTTCACACCAATCATTGTGGCGGGCGTGTTCATTTTCAGTGCTGTGCAGATGACGC
AACTCACCATGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGG
CTCTGTCTTCCATGGTCCTCATCCCCGGGTACATGGCCTACATGTTCCTCACCTTAAAGGG
CTCCCTGAAGCAGCGCATCCAAGTCATGGTCCAGCCCAGCGAAGACATCGTTCGCCCAGA
GAATGGTCCTGAGCAGCCCCAGGCGGGCAGCTCCACCAGCAAGGAGGCCTACATCACGC
GTACGCGGCCGCTCGAGCAGAAACTCATCTCAGAAGAGGATCTGGCAGCAAATGATATCC
TGGATTACAAGGATGACGACGATAAGGTTTAA (SEQ ID NO: 33)

SLC6A1 encoding sequence from CN2976

ATGGCTACGAATGGCTCAAAAGTCGCTGATGGTCAGATCAGCACGGAAGTGAGCGAGGCC
CCTGTGGCCAACGACAAGCCTAAAACTCTGGTCGTAAAAGTACAGAAGAAAGCAGCGGAT
CTGCCAGACCGTGATACATGGAAAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCTG
CTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTA
ATCATGTTCATACCTCTTATCTTCCTCCCACAGGACGTTTCGATTTCCTGATGAGCTGCGTA
GGCTATGCAATCGGACTGGGAAATGTCTGGCGCTTCCCTTACCTGTGTGGTAAAAATGGTG
GAGGCGCCTTTCTCATTCCTTATTTTCTCACCCTCATTTTTGCCGGGGTTCCCTTGTTTCTC
CTGGAATGCAGCCTCGGGCAATATACCAGCATTGGCGGACTCGGTGTCTGGAAACTGGCG
CCTATGTTTAAGGGCGTTGGGCTGGCTGCAGCTGTGCTGTCCTTCTGGCTGAACATCTATT
ACATTGTCATCATTAGCTGGGCGATCTACTATCTGTACAATAGCTTCACTACGACTCTGCCG
TGGAAACAATGTGACAATCCTTGGAATACGGATCGCTGCTTTAGCAATTATAGCATGGTTAA
CACCACTAACATGACTTCCGCTGTAGTGGAGTTCTGGGAGAGGAACATGCATCAGATGAC
CGACGGACTTGACAAACCCGGACAGATTAGGTGGCCCCTCGCCATTACGCTCGCAATCGC
CTGGATCCTTGTCTATTTTTGTATCTGGAAAGGTGTCGGCTGGACTGGAAAGGTGGTCTAT
TTCAGCGCGACGTATCCATACATTATGCTCATTATCCTGTTTTTCCGGGGAGTAACGCTCC
CGGGTGCCAAAGAAGGGATCCTGTTCTACATTACGCCTAATTTTAGGAAGCTCTCTGATAG
TGAGGTTTGGCTGGATGCTGCCACGCAAATTTTTTTCTCCTACGGTCTGGGCCTCGGGTCC
CTGATCGCCCTGGGCAGTTATAATTCTTTTCATAACAATGTGTACAGAGATAGCATTATCGT
GTGCTGTATCAATTCTTGCACATCCATGTTTGCCGGTTTCGTGATCTTCTCCATCGTCGGGT
TTATGGCCCACGTCACCAAACGTTCTATTGCGGATGTCGCTGCCTCAGGTCCCGGTCTGG
CCTTTCTTGCGTACCCTGAGGCTGTTACACAACTGCCAATCAGTCCTCTGTGGGCCATTCT
TTTCTTTTCCATGCTTCTGATGCTGGGCATTGACTCTCAGTTTTGTACTGTTGAGGGTTTCA
TCACCGCCCTGGTTGACGAATATCCCCGTCTCTTGAGAAATCGTCGCGAACTCTTCATTGC
CGCAGTGTGTATTATCTCCTACCTCATTGGCCTGTCCAATATCACGCAGGGAGGCATTTAC
GTATTCAAACTCTTTGATTACTATTCCGCAAGTGGAATGAGCCTCTTGTTTCTCGTGTTTTTC
GAGTGTGTTAGCATCTCCTGGTTCTACGGAGTAAACCGTTTTTATGATAATATCCAGGAAAT
GGTGGGGAGTCGTCCGTGTATCTGGTGGAAACTTTGTTGGTCATTCTTTACGCCCATCATT
GTCGCTGGTGTGTTCATTTTTTCTGCCGTGCAAATGACCCAACTGACGATGGGAAATTACG
TATTCCCCAAATGGGGCCAGGGCGTGGGGTGGCTCATGGCTCTGAGTTCCATGGTTTTGA
TCCCTGGGTATATGGCCTACATGTTCCTGACACTGAAAGGGTCACTGAAACAGCGCATTCA

FIG. 14 cont'd

AGTGATGGTTCAACCCAGCGAGGACATTGTCAGACCGGAAAACGGGCCTGAGCAGCCCCA
AGCGGGATCCTCTACAAGTAAAGAGGCCTATATCACTCGCACGCGCCCCCTG (SEQ ID
NO: 34)

Myc-DDK tag sequence in CN2976
GAGCAAAAGCTCATCTCTGAGGAAGACCTGGCGGCTAACGATATTCTTGACTATAAAGACG
ATGACGATAAA (SEQ ID NO: 23)

CN2976
ATGGCTACGAATGGCTCAAAAGTCGCTGATGGTCAGATCAGCACGGAAGTGAGCGAGGCC
CCTGTGGCCAACGACAAGCCTAAAACTCTGGTCGTAAAAGTACAGAAGAAAGCAGCGGAT
CTGCCAGACCGTGATACATGGAAAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCTG
CTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTA
ATCATGTTCATACCTCTTATCTTCCTCCCACAGGACGTTTCGATTTCCTGATGAGCTGCGTA
GGCTATGCAATCGGACTGGGAAATGTCTGGCGCTTCCCTTACCTGTGTGGTAAAAATGGTG
GAGGCGCCTTTCTCATTCCTTATTTTCTCACCCTCATTTTTGCCGGGGTTCCCTTGTTTCTC
CTGGAATGCAGCCTCGGGCAATATACCAGCATTGGCGGACTCGGTGTCTGGAAACTGGCG
CCTATGTTTAAGGGCGTTGGGCTGGCTGCAGCTGTGCTGTCCTTCTGGCTGAACATCTATT
ACATTGTCATCATTAGCTGGGCGATCTACTATCTGTACAATAGCTTCACTACGACTCTGCCG
TGGAAACAATGTGACAATCCTTGGAATACGGATCGCTGCTTTAGCAATTATAGCATGGTTAA
CACCACTAACATGACTTCCGCTGTAGTGGAGTTCTGGGAGAGGAACATGCATCAGATGAC
CGACGGACTTGACAAACCCGGACAGATTAGGTGGCCCCTCGCCATTACGCTCGCAATCGC
CTGGATCCTTGTCTATTTTTGTATCTGGAAAGGTGTCGGCTGGACTGGAAAGGTGGTCTAT
TTCAGCGCGACGTATCCATACATTATGCTCATTATCCTGTTTTTCCGGGGAGTAACGCTCC
CGGGTGCCAAAGAAGGGATCCTGTTCTACATTACGCCTAATTTTAGGAAGCTCTCTGATAG
TGAGGTTTGGCTGGATGCTGCCACGCAAATTTTTTTCTCCTACGGTCTGGGCCTCGGGTCC
CTGATCGCCCTGGGCAGTTATAATTCTTTTCATAACAATGTGTACAGAGATAGCATTATCGT
GTGCTGTATCAATTCTTGCACATCCATGTTTGCCGGTTTCGTGATCTTCTCCATCGTCGGGT
TTATGGCCCACGTCACCAAACGTTCTATTGCGGATGTCGCTGCCTCAGGTCCCGGTCTGG
CCTTTCTTGCGTACCCTGAGGCTGTTACACAACTGCCAATCAGTCCTCTGTGGGCCATTCT
TTTCTTTTCCATGCTTCTGATGCTGGGCATTGACTCTCAGTTTTGTACTGTTGAGGGTTTCA
TCACCGCCCTGGTTGACGAATATCCCCGTCTCTTGAGAAATCGTCGCGAACTCTTCATTGC
CGCAGTGTGTATTATCTCCTACCTCATTGGCCTGTCCAATATCACGCAGGGAGGCATTTAC
GTATTCAAACTCTTTGATTACTATTCCGCAAGTGGAATGAGCCTCTTGTTTCTCGTGTTTTTC
GAGTGTGTTAGCATCTCCTGGTTCTACGGAGTAAACCGTTTTTATGATAATATCCAGGAAAT
GGTGGGGAGTCGTCCGTGTATCTGGTGGAAACTTTGTTGGTCATTCTTTACGCCCATCATT
GTCGCTGGTGTGTTCATTTTTTCTGCCGTGCAAATGACCCAACTGACGATGGGAAATTACG
TATTCCCCAAATGGGGCCAGGGCGTGGGGTGGCTCATGGCTCTGAGTTCCATGGTTTTGA
TCCCTGGGTATATGGCCTACATGTTCCTGACACTGAAAGGGTCACTGAAACAGCGCATTCA
AGTGATGGTTCAACCCAGCGAGGACATTGTCAGACCGGAAAACGGGCCTGAGCAGCCCCA
AGCGGGATCCTCTACAAGTAAAGAGGCCTATATCACTCGCACGCGCCCCCTGGAGCAAAA
GCTCATCTCTGAGGAAGACCTGGCGGCTAACGATATTCTTGACTATAAAGACGATGACGAT
AAAGTGTAA (SEQ ID NO: 36)

CN2478

FIG. 14 cont'd

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATG
CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCG
CGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCT
TAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGA
GTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCG
CCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGAC
GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA
GTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA
CCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGG
GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT
ACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTG
GCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACGGG
CCCTCTAGACTCGAGGCAGTTCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGC
CGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTT
TGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCT
TTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTG
GAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCA
CCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCG
GCACAACCCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCT
CAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTG
ATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAACGTCTAGG
CCCCCCGAACCACGGGGACGTGGTTTTCCTTTGAAAAACACGATGATAATATGGCCACAAC
CCGCGGCCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATC
CTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGA
GGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGC
CCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGC
TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTC
CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAA
GTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG
ACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCA
CCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGG
ACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCC
GTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAAC
GAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGG
CATGGACGAGCTGTACAAGTAAGAATTCCACCACACTGGACTAGTGGATCCGAGCTCGGT
ACCAAGCTTAAGTTTAAACCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCT
GTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTT
CCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGG
TGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGG
GATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTA
TCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCT

FIG. 14 cont'd

CGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCG
ATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGT
GGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATA
GTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTA
TAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAA
CGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCA
GCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCC
CCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATA
GTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCG
CCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGC
TATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGG
AGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATT
GAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTAT
GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCA
GGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGA
CGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCG
ACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGAT
CTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGC
GGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCG
AGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGC
ATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGC
GAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGC
CGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATA
GCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTC
GTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACG
AGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCC
ATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTC
CGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCA
CCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCAC
AAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTA
TCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTT
CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGT
GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC
CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGG
GAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC
AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTAT
CTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAG
CCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGAC
TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTA

FIG. 14 cont'd

TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
ACAAACCACCGCTGGTAGCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT
CACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAAT
TAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAA
TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTG
ACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGC
AATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGC
CGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAA
TTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCC
ATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT
CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCA
GCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA
CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGT
TCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCA
CTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAA
ACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTC
ATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC
ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGT (SEQ ID NO: 37)

Solute carrier family 6 member 1 (SLC6A1), transcript variant 1 (NM_003042.4) [Homo sapiens]

GCCCTCGGAAGACCGAGACAGCGGAGAGGTTGCGGGTGAGCTGCGCTGAGCCCAGGAG
CCGAGGAGTCGGGAGCGCAGTAGCGCTGAGCCCGAGCCCGAGCGGCCCCGCGTCCCGA
GCGCATCGGAGCGGCCGAGCCGCCCGGATGCAGCGCCTGTCCCGGGCAGCGCAGCCCC
GGCCGCAGGATCTCACCCAGGGTGGCAGAAGGAGGCCTTCTGGAGCTGACCCACCCCCG
ACGACCATCAGGGTGAGGCAACTCCAAGGTCCTACTCTCTTTCTGTGCCTGTTACCCACCC
CGTCCTCCTAGGGTGCCCTTGAGCCGCAAAACTGCTGTCCACGTGGACCGGGGGTGACAT
CGCACGTCCATCTGCCAGGACCCCTGCGTCCAAATTCCGAGACATGGCGACCAACGGCAG
CAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGAGGCCCCTGTGGCCAATGACA
AGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGGCAGACCTCCCCGACCGGGAC
ACGTGGAAGGGCCGCTTCGACTTCCTCATGTCCTGTGTGGGCTATGCCATCGGCCTGGGC
AACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATGGTGGGGGAGCCTTCCTGATCCCC
TATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTTCCTGCTGGAGTGCTCCCTGGGCC
AGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGCTGGCTCCTATGTTCAAGGGCGTG
GGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACATCTACTACATCGTCATCATCTCCT
GGGCCATTTACTACCTGTACAACTCCTTCACCACGACACTGCCGTGGAAACAGTGCGACAA
CCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGCATGGTCAACACTACCAACATGACC
AGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCATCAGATGACGGACGGGCTGGATAA
GCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCTGGCCATCGCCTGGATCCTTGTGTA
TTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAAGGTGGTCTACTTTTCAGCCACATAC
CCCTACATCATGCTGATCATCCTGTTCTTCCGTGGAGTGACGCTGCCCGGGGCCAAGGAG
GGCATCCTCTTCTACATCACACCCAACTTCCGCAAGCTGTCTGACTCCGAGGTGTGGCTG

FIG. 14 cont'd

GATGCGGCAACCCAGATCTTCTTCTCATACGGGCTGGGCCTGGGGTCCCTGATCGCTCTC
GGGAGCTACAACTCTTTCCACAACAATGTCTACAGGGACTCCATCATCGTCTGCTGCATCA
ATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCTTCTCCATCGTGGGCTTCATGGCCCA
TGTCACCAAGAGGTCCATTGCTGATGTGGCGGCCTCAGGCCCCGGGCTGGCGTTCCTGG
CATACCCAGAGGCGGTGACCCAGCTGCCTATCTCCCCACTCTGGGCCATCCTCTTCTTCTC
CATGCTGTTGATGCTGGGCATTGACAGCCAGTTCTGCACTGTGGAGGGCTTCATCACAGC
CCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACCGCAGAGAGCTCTTCATTGCTGCTGT
CTGCATCATCTCCTACCTGATCGGTCTCTCTAACATCACTCAGGGGGGTATTTATGTCTTCA
AACTCTTTGACTACTACTCTGCCAGTGGCATGAGCCTGCTGTTCCTCGTGTTCTTTGAATGT
GTCTCTATTTCCTGGTTTTACGGTGTCAACCGATTCTATGACAATATCCAAGAGATGGTTGG
ATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGGTCTTTCTTCACACCAATCATTGTGGCG
GGCGTGTTCATTTTCAGTGCTGTGCAGATGACGCCACTCACCATGGGAAACTATGTTTTCC
CCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGGCTCTGTCTTCCATGGTCCTCATCCCCG
GGTACATGGCCTACATGTTCCTCACCTTAAAGGGCTCCCTGAAGCAGCGCATCCAAGTCAT
GGTCCAGCCCAGCGAAGACATCGTTCGCCCAGAGAATGGTCCTGAGCAGCCCCAGGCGG
GCAGCTCCACCAGCAAGGAGGCCTACATCTAGGGTGGGGGCCACTCACCGACCCGACAC
TCTCACCCCCCGACCTGGCTGAGTGCGACCACCACTTGATGTCTGAGGATACCTTCCATCT
CAACCTACCTCGAGTGGCGAGTCCAGACACCATCACCACGCAGAGAGGGGAGGTGGGAG
GACAGTTAGACCCCTGGGTGGGCCCTGCCGTGGGCAAGGATACCCGGTGGCTTCTGGCA
CCTGGCGGGCTGGTGACCTTTTTAATCCAGGCCCCATCAGCATCCCACGATCGGCCTTGG
TAACCGCCGCGGTAGATCATTTTTATCCCGCCAGGGAGTGTGATGCAGGAAGACCACATG
CGCTCCTGGCTTTTAAACCTGTTCCTGACTGTTCTCTTACTGCCGAAACCCTTGACTGTTAT
CTCGGACTTTGCAGGAGTTCCTTTCCCTCCGAACGCTGCTCCATGCACAGGAAAAGGGCA
TTTTGTACAATGGGGACTTCCCGGGAACGCTTGCTCTTAAGTACCAGAAGCCGGCGGAGC
TCTGGCTTTCGTGTTTTTGGTTTTCTCCTTCCCAAGGCAGCTGGATTGAAAAAACAAAACAA
AACAAAAAAACCCAGGGGCGTCAGTCGATATTCCCAGGGCCGCTTCTCCTGCAGTCTGTG
GAGCGTCCTTGTCCCCGCCGCCGGAATGAATGAGCATTCTGCAGCCCGATGTCCCTGTCC
CCTCCTCGCCGGGCCATTCTGATTGGACCTGGCCCAGTGCAATCTGTCCAGACAAGCCCT
GCTTGCTGGAAAACTGCCACAAGCACAATTGATCTCTTTTTATCGCCATTCCAGGGGCCTC
AGGTCCTACTGGGGAAACTTCCTATACCGGAGCTCCAGTTTCTCTTAAGCTGCCCAATTTC
ACAGAGTACAAAATAGTTGTAGGGGAAATCAAGGTGAAGGATCTGTCCGACAGTCAAGAC
GGATCCACAGTAATCTTTCGGTCTCCTTAAACTACCACCCTCGCTGCCACCCACCCCAAGC
TGCTGCCGCCTCACCTTCCTTGAAATTTCTCAGCGGGAGTCTCCTCACTGCCACTAAAATC
CACCCAGCCCACTAACTGAGGAGCTAGTGTTAATCCAGAGAACCCCCCGCAATGTGCTTC
CGAGATTCAGACTGCTTCATTGGGAAGTATGATTTGTTCCTTTCTGGAATTGGGCTCCGTG
GTGGCGGCGGCACTTCAAGCAAAGACAGTTTCTTGCAAGCTCCAGTAGCTCCGCGTGTCT
CATTTGCCAGGAAGATGGGTTCCCACGTAGCAAATCGTACATTGTGCCCTGTAGCTCCTTA
GCTAGTTAGCTCACAAGCCGTGTTTTATGACTAATCCTTAATAACTATGGTAAATAACTGTG
ACTGTGGGGTTTTTAATCTCTTGTCATTCTCATCCAAAAGTGACCAGCATACCAGTTCTTGC
AATAAGATATTACCCTCAGAATATTAAGCACATTATTGTAGAGAAAAAAAAATATGTGTACAC
ATATGAACGCACAACATGCACATTCATCCTCACATGTGGCACGTAAGGTCTCATTTGATATT
GTGTAGGAAATCTGAAGCCTTTTCCTGAGGTCATCTGTAAAATAGTCTCATTGCCAAGGCAT
CCCCAGTGCCAGCTGGTGAATCCATGATCAAAATGCATACGTATTGTTAAATGATAAGGTTT
AGAATGACAGGAACCCATCACTGTGTCTCATGGTCCCACTTCCCCATCTGTGTGTGAATTC
CTTTAGACTAAGGGCAGGAAGACTTCCAGCTTTCTCTTTGTTCTTCAATGTGAAACTGAGAC

FIG. 14 cont'd

CAAGTCTCTCTAAGACAAATGCAGTGTATTTAATGTTTGTAAGCAATTCTAAGTGAGATGTTT
GGCAAGAAATCCCCTAACTGATTTCCATCCAAACCTACCTTATAGAGCACAATATTAAGTGT
TGTACAATTACTGTGAGAACTGTGAATATGTGTAACTTTTTTTTAGTATTTGCCCGGGGGGA
AAAAGATATTGTATTATCATATATGCTTTTTTGCAATAAGGATTTATTCTCAGAACACCAAGT
AAATCTATCTCTATATAAAAAATATATGTAATATATACATATTCAAAGTATATACAGAGCCTGT
TTTAAAAAATACAGTATTATTTAGTAAAATTATCTGTTCTATGGACCAAATGTAAAATATTTAT
AAATGAAGATGCATTTTAAATGTCTATAAATGGTGTCATAACTAGAGCACGGGCGTTATGTA
AGTTTCTAAGAATTTAGAGGATAAATAATAAAGGTTCTATGATATACAA (SEQ ID NO: 38)

Solute carrier family 6 member 1 (SLC6A1), transcript variant 2 (NM_001348250.2) [Homo sapiens]

GCCCTCGGAAGACCGAGACAGCGGAGAGGTTGCGGGTGAGCTGCGCTGAGCCCAGGAG
CCGAGGAGTCGGGAGCGCAGTAGCGCTGAGCCCGAGCCCGAGCGGCCCCGCGTCCCGA
GCGCATCGGAGCGGCCGAGCCGCCCGGATGCAGCGCCTGTCCCGGGCAGCGCAGCCCC
GGCCGCAGGATCTCACCCAGGGTGGCAGAAGGAGGCCTTCTGGAGCTGACCCACCCCCG
ACGACCATCAGGGTGCCCTTGAGCCGCAAAACTGCTGTCCACGTGGACCGGGGGTGACA
TCGCACGTCCATCTGCCAGGACCCCTGCGTCCAAATTCCGAGACATGGCGACCAACGGCA
GCAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGAGGCCCCTGTGGCCAATGAC
AAGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGGCAGACCTCCCCGACCGGGA
CACGTGGAAGGGCCGCTTCGACTTCCTCATGTCCTGTGTGGGCTATGCCATCGGCCTGGG
CAACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATGGTGGGGGAGCCTTCCTGATCCC
CTATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTTCCTGCTGGAGTGCTCCCTGGGC
CAGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGCTGGCTCCTATGTTCAAGGGCGTG
GGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACATCTACTACATCGTCATCATCTCCT
GGGCCATTTACTACCTGTACAACTCCTTCACCACGACACTGCCGTGGAAACAGTGCGACAA
CCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGCATGGTCAACACTACCAACATGACC
AGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCATCAGATGACGGACGGGCTGGATAA
GCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCTGGCCATCGCCTGGATCCTTGTGTA
TTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAAGGTGGTCTACTTTTCAGCCACATAC
CCCTACATCATGCTGATCATCCTGTTCTTCCGTGGAGTGACGCTGCCCGGGGCCAAGGAG
GGCATCCTCTTCTACATCACACCCAACTTCCGCAAGCTGTCTGACTCCGAGGTGTGGCTG
GATGCGGCAACCCAGATCTTCTTCTCATACGGGCTGGGCCTGGGGTCCCTGATCGCTCTC
GGGAGCTACAACTCTTTCCACAACAATGTCTACAGGGACTCCATCATCGTCTGCTGCATCA
ATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCTTCTCCATCGTGGGCTTCATGGCCCA
TGTCACCAAGAGGTCCATTGCTGATGTGGCGGCCTCAGGCCCCGGGCTGGCGTTCCTGG
CATACCCAGAGGCGGTGACCCAGCTGCCTATCTCCCCACTCTGGGCCATCCTCTTCTTCTC
CATGCTGTTGATGCTGGGCATTGACAGCCAGTTCTGCACTGTGGAGGGCTTCATCACAGC
CCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACCGCAGAGAGCTCTTCATTGCTGCTGT
CTGCATCATCTCCTACCTGATCGGTCTCTCTAACATCACTCAGGGGGGTATTTATGTCTTCA
AACTCTTTGACTACTACTCTGCCAGTGGCATGAGCCTGCTGTTCCTCGTGTTCTTTGAATGT
GTCTCTATTTCCTGGTTTTACGGTGTCAACCGATTCTATGACAATATCCAAGAGATGGTTGG
ATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGGTCTTTCTTCACACCAATCATTGTGGCG
GGCGTGTTCATTTTCAGTGCTGTGCAGATGACGCCACTCACCATGGGAAACTATGTTTTCC
CCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGGCTCTGTCTTCCATGGTCCTCATCCCCG
GGTACATGGCCTACATGTTCCTCACCTTAAAGGGCTCCCTGAAGCAGCGCATCCAAGTCAT

FIG. 14 cont'd

GGTCCAGCCCAGCGAAGACATCGTTCGCCCAGAGAATGGTCCTGAGCAGCCCCAGGCGG
GCAGCTCCACCAGCAAGGAGGCCTACATCTAGGGTGGGGGCCACTCACCGACCCGACAC
TCTCACCCCCCGACCTGGCTGAGTGCGACCACCACTTGATGTCTGAGGATACCTTCCATCT
CAACCTACCTCGAGTGGCGAGTCCAGACACCATCACCACGCAGAGAGGGGAGGTGGGAG
GACAGTTAGACCCCTGGGTGGGCCCTGCCGTGGGCAAGGATACCCGGTGGCTTCTGGCA
CCTGGCGGGCTGGTGACCTTTTTAATCCAGGCCCCATCAGCATCCCACGATCGGCCTTGG
TAACCGCCGCGGTAGATCATTTTTATCCCGCCAGGGAGTGTGATGCAGGAAGACCACATG
CGCTCCTGGCTTTTAAACCTGTTCCTGACTGTTCTCTTACTGCCGAAACCCTTGACTGTTAT
CTCGGACTTTGCAGGAGTTCCTTTCCCTCCGAACGCTGCTCCATGCACAGGAAAAGGGCA
TTTTGTACAATGGGGACTTCCCGGGAACGCTTGCTCTTAAGTACCAGAAGCCGGCGGAGC
TCTGGCTTTCGTGTTTTTGGTTTTCTCCTTCCCAAGGCAGCTGGATTGAAAAAACAAAACAA
AACAAAAAAACCCAGGGGCGTCAGTCGATATTCCCAGGGCCGCTTCTCCTGCAGTCTGTG
GAGCGTCCTTGTCCCCGCCGCCGGAATGAATGAGCATTCTGCAGCCCGATGTCCCTGTCC
CCTCCTCGCCGGGCCATTCTGATTGGACCTGGCCCAGTGCAATCTGTCCAGACAAGCCCT
GCTTGCTGGAAAACTGCCACAAGCACAATTGATCTCTTTTTATCGCCATTCCAGGGGCCTC
AGGTCCTACTGGGGAAACTTCCTATACCGGAGCTCCAGTTTCTCTTAAGCTGCCCAATTTC
ACAGAGTACAAAATAGTTGTAGGGGAAATCAAGGTGAAGGATCTGTCCGACAGTCAAGAC
GGATCCACAGTAATCTTTCGGTCTCCTTAAACTACCACCCTCGCTGCCACCCACCCCAAGC
TGCTGCCGCCTCACCTTCCTTGAAATTTCTCAGCGGGAGTCTCCTCACTGCCACTAAAATC
CACCCAGCCCACTAACTGAGGAGCTAGTGTTAATCCAGAGAACCCCCCGCAATGTGCTTC
CGAGATTCAGACTGCTTCATTGGGAAGTATGATTTGTTCCTTTCTGGAATTGGGCTCCGTG
GTGGCGGCGGCACTTCAAGCAAAGACAGTTTCTTGCAAGCTCCAGTAGCTCCGCGTGTCT
CATTTGCCAGGAAGATGGGTTCCCACGTAGCAAATCGTACATTGTGCCCTGTAGCTCCTTA
GCTAGTTAGCTCACAAGCCGTGTTTTATGACTAATCCTTAATAACTATGGTAAATAACTGTG
ACTGTGGGGTTTTTAATCTCTTGTCATTCTCATCCAAAAGTGACCAGCATACCAGTTCTTGC
AATAAGATATTACCCTCAGAATATTAAGCACATTATTGTAGAGAAAAAAAAATATGTGTACAC
ATATGAACGCACAACATGCACATTCATCCTCACATGTGGCACGTAAGGTCTCATTTGATATT
GTGTAGGAAATCTGAAGCCTTTTCCTGAGGTCATCTGTAAAATAGTCTCATTGCCAAGGCAT
CCCCAGTGCCAGCTGGTGAATCCATGATCAAAATGCATACGTATTGTTAAATGATAAGGTTT
AGAATGACAGGAACCCATCACTGTGTCTCATGGTCCCACTTCCCCATCTGTGTGTGAATTC
CTTTAGACTAAGGGCAGGAAGACTTCCAGCTTTCTCTTTGTTCTTCAATGTGAAACTGAGAC
CAAGTCTCTCTAAGACAAATGCAGTGTATTTAATGTTTGTAAGCAATTCTAAGTGAGATGTTT
GGCAAGAAATCCCCTAACTGATTTCCATCCAAACCTACCTTATAGAGCACAATATTAAGTGT
TGTACAATTACTGTGAGAACTGTGAATATGTGTAACTTTTTTTTTAGTATTTGCCCGGGGGGA
AAAAGATATTGTATTATCATATATGCTTTTTTGCAATAAGGATTTATTCTCAGAACACCAAGT
AAATCTATCTCTATATAAAAAATATATGTAATATATACATATTCAAAGTATATACAGAGCCTGT
TTTAAAAAATACAGTATTATTTAGTAAAATTATCTGTTCTATGGACCAAATGTAAAATATTTAT
AAATGAAGATGCATTTTAAATGTCTATAAATGGTGTCATAACTAGAGCACGGGCGTTATGTA
AGTTTCTAAGAATTTAGAGGATAAATAATAAAGGTTCTATGATATACAA (SEQ ID NO: 39)

Solute carrier family 6 member 1 (SLC6A1), transcript variant X5 (XM_011534027.3) [Homo sapiens]

TAATCGTGGTATAGTATGCACCTGAAGGAATGTGGAGCCGCATGCTGGGTCTGCAGAATG
TTGTCACCTGATGAAGAAGAACTGGAAGACGCTGGAAGAAGGCTTTGGAAAAATGCCCAG
ATCATGAGAAGGTGAGAGGTGTTTCTTCCGCCTGCTCCACCAGCAGGTAAAGGAGGCTGA

FIG. 14 cont'd

TCACAGGCTGGCACTCAGGGCGAGGTGGAAGAAATGCGTGGTCTTGCCACTACAGACCC
GCAGAGCTGTCCTCCAACATGATCCAAGAGAGCCACAAAACACAGGGGTGTGTCTTGGGA
AAGGAAAAGTAACAACTTGCAAAAAGGTTGGTTCTGCTTCAGAGAAATGTTCTAGAAAGAC
AGCCAAGTTACCCTGCCGTGAACTCACACACACCTGTGCTTTCTGCATGTTATGTTTCATAG
GAGAGGGAAGTGCCGGGAATTAAAATTGCGGATGCACTGTTCCCCAGGTGTGTGGGTGAG
GCACGCTTCCTGGGACCACTCCATTTAAGCATTGGGTCCCAGGCATGTTCACTACTAGCAG
GAGGCTCTTTATCAATTTCCTCCCGTGGACCAGATGATTGGGATTTTTTTTTCTACCAGACT
TTATTACAGAATAACTATTTTTGTATTAACTAAGCAAACAAATGTACCGCTCCCAGCTTCTGA
CTGGCCCAAGGTGCTGAGCTAGACCTCTTTGCTTTGGTACAAAGAGAGCCCAAAGCCAAG
AGACTTGGCATTTCTGTTAGCTGGTATTCCCCAGGGTGACTAAAAGGATGATTTCCATTGG
ATTTTTTAAAATGCTGAAGACACTGAAAGACCTGAATAGTAAGCCCGTAATAAACGTTATTC
GGTATTATTATTGTTGTTGTTGTAGCCTTCTGGGAAGTTGCGAATGCCAAATTAAGAACAAT
GGACCCATACCCAGATGGTGGCGATTGCAGTGATAACAATGGCTGAGATTTCCTAAGCCCT
AACTATATGCCAGGCATAGTTGTACATGCTTTACACACATGATCTCATTTAATCATCACCATA
ACCCTCGGAAGGAGTTATTATTATCTTCCCATTTTATAGATGAGGGTCAGAGAGGTGAAGT
AACTCACCCAGCATTAGTAGGTGTTGGAGGGAGCCGAGGTCTGTCTGACTCAGCACCAGG
GTGGCTTTGCTGCTGTGTAAGTCAAAGAGCCTTCTCCGTTCTAAATCCTCAAAGAAGTATG
AGTTATATCAATTCATGGCTTTACTTTGTTTATCTTCATATGTGCCACATCTTCTAAAAGGAA
AACTCCAAATATAGTCATGTGGCTCCTCCTAGGGAATTATAGGCATTTGGCTGAATCTGAAT
ACATATCCCCGTATTCCAGAAATTGGAAATCACATGTATGTCTGCTTTATGGTGAGGTCTTC
AGAGCAGCCACCGTCTCTGATCGTTCAGGGCTTGGCATGAGGAGGTGCTCAATATAGTCA
TGGAGAAGAACTCATGGACTGAACTTCCTGGTGAACCAGAGGGATGGTGCAGAGATTTAA
GATTGCAACAGGAACCCCCAAGCCAATGTAAGCAAAACATCTATTTCATGACTATGGGAAT
GTTTTGTAAAATCCAAGCACCAGCCTGAAGAAAGCCTGGAACTGCAAACTAGGAGCTTTGA
GACACTCTATCTTTTGATGTCTAGTTTTCTCCTTCCTCAGCAGACCAAATCTCACTCTGAGT
ACAAGATGGGCAGAGGATGGCTGCCCCAGAGTCCCCAAGTCACTAGCAAAACTGCCCTTC
TTCCTGGGTCCCATGCCAAATTCCAGAGAAGGAATTTGATTGGCTAGCCTGGGTCAGGTG
GTCAGTCCTGATTCAAGGTGATTGGATCATTTTGAACAAATATGGCAGCTGGAGCCCACAA
GGGTAGATCATGGGACCGTTAAGGGCATCATTGTTACCAGAGTTAATGGCTGAGGCTGCT
CTCTCCTGCGTATCCCAAGAGTCCCTGGTAGACCACAGAATCGGCCGCCTCCTACTATCCT
GCCATTCTCCAGAAGGGGAGGCAGGCACCCTACTGAGTCCAGAGCTCTCTCTAACATTGC
CATGCCCAGGCCAGGATCTAGATGGGCATCCAAGACCAGCCTGTGTGCCGTGAAGGAGTT
CAAAGTGAGCAGGCCATCCCTGAGAACGAGGAGCAGGGGTGTTAGGAAAGTTTGGATTAT
CTGAAGATGGCCTGGTTCTCCGTGAACTCTGGGAGGAGATGATGGTGTTTGGATCTCACC
CAGGGTGGCAGAAGGAGGCCTTCTGGAGCTGACCCACCCCCGACGACCATCAGGGTGCC
CTTGAGCCGCAAAACTGCTGTCCACGTGGACCGGGGGTGACATCGCACGTCCATCTGCCA
GGACCCCTGCGTCCAAATTCCGAGACATGGCGACCAACGGCAGCAAGGTGGCCGACGGG
CAGATCTCCACCGAGGTCAGCGAGGCCCCTGTGGCCAATGACAAGCCCAAAACCTTGGTG
GTCAAGGTGCAGAAGAAGGCGGCAGACCTCCCCGACCGGGACACGTGGAAGGGCCGCTT
CGACTTCCTCATGTCCTGTGTGGGCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCC
CTATCTCTGCGGGAAAAATGGTGGGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATC
TTTGCGGGGGTCCCACTCTTCCTGCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGG
GGGCTAGGGGTATGGAAGCTGGCTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGT
GCTATCATTCTGGCTGAACATCTACTACATCGTCATCATCTCCTGGGCCATTTACTACCTGT
ACAACTCCTTCACCACGACACTGCCGTGGAAACAGTGCGACAACCCCTGGAACACAGACC

FIG. 14 cont'd

GCTGCTTCTCCAACTACAGCATGGTCAACACTACCAACATGACCAGCGCTGTGGTGGAGTT
CTGGGAGCGCAACATGCATCAGATGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCT
GGCCACTGGCCATCACGCTGGCCATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGG
GTGTTGGCTGGACTGGAAAGGTGGTCTACTTTTCAGCCACATACCCCTACATCATGCTGAT
CATCCTGTTCTTCCGTGGAGTGACGCTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACAT
CACACCCAACTTCCGCAAGCTGTCTGACTCCGAGGTGTGGCTGGATGCGGCAACCCAGAT
CTTCTTCTCATACGGGCTGGGCCTGGGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTT
CCACAACAATGTCTACAGGGACTCCATCATCGTCTGCTGCATCAATTCGTGCACCAGCATG
TTCGCAGGATTCGTCATCTTCTCCATCGTGGGCTTCATGGCCCATGTCACCAAGAGGTCCA
TTGCTGATGTGGCGGCCTCAGGCCCCGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTG
ACCCAGCTGCCTATCTCCCCACTCTGGGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGG
GCATTGACAGCCAGTTCTGCACTGTGGAGGGCTTCATCACAGCCCTGGTGGATGAGTACC
CCAGGCTCCTCCGCAACCGCAGAGAGCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCT
GATCGGTCTCTCTAACATCACTCAGGGGGGTATTTATGTCTTCAAACTCTTTGACTACTACT
CTGCCAGTGGCATGAGCCTGCTGTTCCTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTT
TACGGTGTCAACCGATTCTATGACAATATCCAAGAGATGGTTGGATCCAGGCCCTGCATCT
GGTGGAAACTCTGCTGGTCTTTCTTCACACCAATCATTGTGGCGGGCGTGTTCATTTTCAG
TGCTGTGCAGATGACGCCACTCACCATGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGG
TGTGGGCTGGCTGATGGCTCTGTCTTCCATGGTCCTCATCCCCGGGTACATGGCCTACAT
GTTCCTCACCTTAAAGGGCTCCCTGAAGCAGCGCATCCAAGTCATGGTCCAGCCCAGCGA
AGACATCGTTCGCCCAGAGAATGGTCCTGAGCAGCCCCAGGCGGGCAGCTCCACCAGCA
AGGAGGCCTACATCTAGGGTGGGGGCCACTCACCGACCCGACACTCTCACCCCCCGACC
TGGCTGAGTGCGACCACCACTTGATGTCTGAGGATACCTTCCATCTCAACCTACCTCGAGT
GGCGAGTCCAGACACCATCACCACGCAGAGAGGGGAGGTGGGAGGACAGTTAGACCCCT
GGGTGGGCCCTGCCGTGGGCAAGGATACCCGGTGGCTTCTGGCACCTGGCGGGCTGGT
GACCTTTTTAATCCAGGCCCCATCAGCATCCCACGATCGGCCTTGGTAACCGCCGCGGTA
GATCATTTTTATCCCGCCAGGGAGTGTGATGCAGGAAGACCACATGCGCTCCTGGCTTTTA
AACCTGTTCCTGACTGTTCTCTTACTGCCGAAACCCTTGACTGTTATCTCGGACTTTGCAGG
AGTTCCTTTCCCTCCGAACGCTGCTCCATGCACAGGAAAAGGGCATTTTGTACAATGGGGA
CTTCCCGGGAACGCTTGCTCTTAAGTACCAGAAGCCGGCGGAGCTCTGGCTTTCGTGTTTT
TGGTTTTCTCCTTCCCAAGGCAGCTGGATTGAAAAAACAAAACAAAACAAAAAAACCCAGG
GGCGTCAGTCGATATTCCCAGGGCCGCTTCTCCTGCAGTCTGTGGAGCGTCCTTGTCCCC
GCCGCCGGAATGAATGAGCATTCTGCAGCCCGATGTCCCTGTCCCCTCCTCGCCGGGCCA
TTCTGATTGGACCTGGCCCAGTGCAATCTGTCCAGACAAGCCCTGCTTGCTGGAAAACTGC
CACAAGCACAATTGATCTCTTTTTATCGCCATTCCAGGGGCCTCAGGTCCTACTGGGGAAA
CTTCCTATACCGGAGCTCCAGTTTCTCTTAAGCTGCCCAATTTCACAGAGTACAAAATAGTT
GTAGGGGAAATCAAGGTGAAGGATCTGTCCGACAGTCAAGACGGATCCACAGTAATCTTTC
GGTCTCCTTAAACTACCACCCTCGCTGCCACCCACCCCAAGCTGCTGCCGCCTCACCTTC
CTTGAAATTTCTCAGCGGGAGTCTCCTCACTGCCACTAAAATCCACCCAGCCCACTAACTG
AGGAGCTAGTGTTAATCCAGAGAACCCCCCGCAATGTGCTTCCGAGATTCAGACTGCTTCA
TTGGGAAGTATGATTTGTTCCTTTCTGGAATTGGGCTCCGTGGTGGCGGCGGCACTTCAAG
CAAAGACAGTTTCTTGCAAGCTCCAGTAGCTCCGCGTGTCTCATTTGCCAGGAAGATGGGT
TCCCACGTAGCAAATCGTACATTGTGCCCTGTAGCTCCTTAGCTAGTTAGCTCACAAGCCG
TGTTTTATGACTAATCCTTAATAACTATGGTAAATAACTGTGACTGTGGGGTTTTTAATCTCT
TGTCATTCTCATCCAAAAGTGACCAGCATACCAGTTCTTGCAATAAGATATTACCCTCAGAA

FIG. 14 cont'd

TATTAAGCACATTATTGTAGAGAAAAAAAAATATGTGTACACATATGAACGCACAACATGCA
CATTCATCCTCACATGTGGCACGTAAGGTCTCATTTGATATTGTGTAGGAAATCTGAAGCCT
TTTCCTGAGGTCATCTGTAAAATAGTCTCATTGCCAAGGCATCCCCAGTGCCAGCTGGTGA
ATCCATGATCAAAATGCATACGTATTGTTAAATGATAAGGTTTAGAATGACAGGAACCCATC
ACTGTGTCTCATGGTCCCACTTCCCCATCTGTGTGTGAATTCCTTTAGACTAAGGGCAGGA
AGACTTCCAGCTTTCTCTTTGTTCTTCAATGTGAAACTGAGACCAAGTCTCTCTAAGACAAA
TGCAGTGTATTTAATGTTTGTAAGCAATTCTAAGTGAGATGTTTGGCAAGAAATCCCCTAAC
TGATTTCCATCCAAACCTACCTTATAGAGCACAATATTAAGTGTTGTACAATTACTGTGAGA
ACTGTGAATATGTGTAACTTTTTTTTTAGTATTTGCCCGGGGGGGAAAAAGATATTGTATTATCA
TATATGCTTTTTTGCAATAAGGATTTATTCTCAGAACACCAAGTAAATCTATCTCTATATAAA
AAATATATGTAATATATACATATTCAAAGTATATACAGAGCCTGTTTTAAAAAATACAGTATTA
TTTAGTAAAATTATCTGTTCTATGGACCAAATGTAAAATATTTATAAATGAAGATGCATTTTAA
ATGTCTATAAATGGTGTCATAACTAGAGCACGGGCGTTATGTAAGTTTCTAAGAATTTAGAG
GATAAATAATAAAGGTTCTATGA (SEQ ID NO: 40)

Solute carrier family 6 member 1 (SLC6A1), transcript variant X1 (XM_011534025.3) [Homo sapiens]

TAATCGTGGTATAGTATGCACCTGAAGGAATGTGGAGCCGCATGCTGGGTCTGCAGAATG
TTGTCACCTGATGAAGAAGAACTGGAAGACGCTGGAAGAAGGCTTTGGAAAAATGCCCAG
ATCATGAGAAGGTGAGAGGTGTTTCTTCCGCCTGCTCCACCAGCAGGTAAAGGAGGCTGA
TCACAGGCTGGCACTCAGGGCGAGGTGGAAGAAATGCGTGGTCTTGCCACTACAGACCC
GCAGAGCTGTCCTCCAACATGATCCAAGAGAGCCACAAAACACAGGGGTGTGTCTTGGGA
AAGGAAAAGTAACAACTTGCAAAAAGGTTGGTTCTGCTTCAGAGAAATGTTCTAGAAAGAC
AGCCAAGTTACCCTGCCGTGAACTCACACACACCTGTGCTTTCTGCATGTTATGTTTCATAG
GAGAGGGAAGTGCCGGGAATTAAAATTGCGGATGCACTGTTCCCCAGGTGTGTGGGTGAG
GCACGCTTCCTGGGACCACTCCATTTAAGCATTGGGTCCCAGGCATGTTCACTACTAGCAG
GAGGCTCTTTATCAATTTCCTCCCGTGGACCAGATGATTGGGATTTTTTTTTCTACCAGACT
TTATTACAGAATAACTATTTTTGTATTAACTAAGCAAACAAATGTACCGCTCCCAGCTTCTGA
CTGGCCCAAGGTGCTGAGCTAGACCTCTTTGCTTTGGTACAAAGAGAGCCCAAAGCCAAG
AGACTTGGCATTTCTGTTAGCTGGTATTCCCCAGGGTGACTAAAAGGATGATTTCCATTGG
ATTTTTTAAAATGCTGAAGACACTGAAAGACCTGAATAGTAAGCCCGTAATAAACGTTATTC
GGTATTATTATTGTTGTTGTTGTAGCCTTCTGGGAAGTTGCGAATGCCAAATTAAGAACAAT
GGACCCATACCCAGATGGTGGCGATTGCAGTGATAACAATGGCTGAGATTTCCTAAGCCCT
AACTATATGCCAGGCATAGTTGTACATGCTTTACACACATGATCTCATTTAATCATCACCATA
ACCCTCGGAAGGAGTTATTATTATCTTCCCATTTTATAGATGAGGGTCAGAGAGGTGAAGT
AACTCACCCAGCATTAGTAGGTGTTGGAGGGAGCCGAGGTCTGTCTGACTCAGCACCAGG
GTGGCTTTGCTGCTGTGTAAGTCAAAGAGCCTTCTCCGTTCTAAATCCTCAAAGAAGTATG
AGTTATATCAATTCATGGCTTTACTTTGTTTATCTTCATATGTGCCACATCTTCTAAAAGGAA
AACTCCAAATATAGTCATGTGGCTCCTCCTAGGGAATTATAGGCATTTGGCTGAATCTGAAT
ACATATCCCCGTATTCCAGAAATTGGAAATCACATGTATGTCTGCTTTATGGTGAGGTCTTC
AGAGCAGCCACCGTCTCTGATCGTTCAGGGCTTGGCATGAGGAGGTGCTCAATATAGTCA
TGGAGAAGAACTCATGGACTGAACTTCCTGGTGAACCAGAGGGATGGTGCAGAGATTTAA
GATTGCAACAGGAACCCCCAAGCCAATGTAAGCAAAACATCTATTTCATGACTATGGGAAT
GTTTTGTAAAATCCAAGCACCAGCCTGAAGAAAGCCTGGAACTGCAAACTAGGAGCTTTGA
GACACTCTATCTTTTGATGTCTAGTTTTCTCCTTCCTCAGCAGACCAAATCTCACTCTGAGT

FIG. 14 cont'd

ACAAGATGGGCAGAGGATGGCTGCCCCAGAGTCCCCAAGTCACTAGCAAAACTGCCCTTC
TTCCTGGGTCCCATGCCAAATTCCAGAGAAGGAATTTGATTGGCTAGCCTGGGTCAGGTG
GTCAGTCCTGATTCAAGGTGATTGGATCATTTTGAACAAATATGGCAGCTGGAGCCCACAA
GGGTAGATCATGGGACCGTTAAGGGCATCATTGTTACCAGAGTTAATGGCTGAGGCTGCT
CTCTCCTGCGTATCCCAAGAGTCCCTGGTAGACCACAGAATCGGCCGCCTCCTACTATCCT
GCCATTCTCCAGAAGGGGAGGCAGGCACCCTACTGAGTCCAGAGCTCTCTCTAACATTGC
CATGCCCAGGCCAGGATCTAGATGGGCATCCAAGACCAGCCTGTGTGCCGTGAAGGAGTT
CAAAGTGAGCAGGCCATCCCTGAGAACGAGGAGCAGGGGTGTTAGGAAAGTTTGGATTAT
CTGAAGATGGCCTGGTTCTCCGTGAACTCTGGGAGGAGATGATGGTGTTTGGATCTCACC
CAGGGTGGCAGAAGGAGGCCTTCTGGAGCTGACCCACCCCCGACGACCATCAGGGTGAG
GCAACTCCAAGGTCCTACTCTCTTTCTGTGCCTGTTACCCACCCCGTCCTCCTAGGGTGCC
CTTGAGCCGCAAAACTGCTGTCCACGTGGACCGGGGGGTGACATCGCACGTCCATCTGCCA
GGACCCCTGCGTCCAAATTCCGAGACATGGCGACCAACGGCAGCAAGGTGGCCGACGGG
CAGATCTCCACCGAGGTCAGCGAGGCCCCTGTGGCCAATGACAAGCCCAAAACCTTGGTG
GTCAAGGTGCAGAAGAAGGCGGCAGACCTCCCCGACCGGGACACGTGGAAGGGCCGCTT
CGACTTCCTCATGTCCTGTGTGGGCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCC
CTATCTCTGCGGGAAAAATGGTGGGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATC
TTTGCGGGGGTCCCACTCTTCCTGCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGG
GGGCTAGGGGTATGGAAGCTGGCTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGT
GCTATCATTCTGGCTGAACATCTACTACATCGTCATCATCTCCTGGGCCATTTACTACCTGT
ACAACTCCTTCACCACGACACTGCCGTGGAAACAGTGCGACAACCCCTGGAACACAGACC
GCTGCTTCTCCAACTACAGCATGGTCAACACTACCAACATGACCAGCGCTGTGGTGGAGTT
CTGGGAGCGCAACATGCATCAGATGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCT
GGCCACTGGCCATCACGCTGGCCATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGG
GTGTTGGCTGGACTGGAAAGGTGGTCTACTTTTCAGCCACATACCCCTACATCATGCTGAT
CATCCTGTTCTTCCGTGGAGTGACGCTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACAT
CACACCCAACTTCCGCAAGCTGTCTGACTCCGAGGTGTGGCTGGATGCGGCAACCCAGAT
CTTCTTCTCATACGGGCTGGGCCTGGGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTT
CCACAACAATGTCTACAGGGACTCCATCATCGTCTGCTGCATCAATTCGTGCACCAGCATG
TTCGCAGGATTCGTCATCTTCTCCATCGTGGGCTTCATGGCCCATGTCACCAAGAGGTCCA
TTGCTGATGTGGCGGCCTCAGGCCCCGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTG
ACCCAGCTGCCTATCTCCCCACTCTGGGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGG
GCATTGACAGCCAGTTCTGCACTGTGGAGGGCTTCATCACAGCCCTGGTGGATGAGTACC
CCAGGCTCCTCCGCAACCGCAGAGAGCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCT
GATCGGTCTCTCTAACATCACTCAGGGGGGTATTTATGTCTTCAAACTCTTTGACTACTACT
CTGCCAGTGGCATGAGCCTGCTGTTCCTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTT
TACGGTGTCAACCGATTCTATGACAATATCCAAGAGATGGTTGGATCCAGGCCCTGCATCT
GGTGGAAACTCTGCTGGTCTTTCTTCACACCAATCATTGTGGCGGGCGTGTTCATTTTCAG
TGCTGTGCAGATGACGCCACTCACCATGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGG
TGTGGGCTGGCTGATGGCTCTGTCTTCCATGGTCCTCATCCCCGGGTACATGGCCTACAT
GTTCCTCACCTTAAAGGGCTCCCTGAAGCAGCGCATCCAAGTCATGGTCCAGCCCAGCGA
AGACATCGTTCGCCCAGAGAATGGTCCTGAGCAGCCCCAGGCGGGCAGCTCCACCAGCA
AGGAGGCCTACATCTAGGGTGGGGGCCACTCACCGACCCGACACTCTCACCCCCCGACC
TGGCTGAGTGCGACCACCACTTGATGTCTGAGGATACCTTCCATCTCAACCTACCTCGAGT
GGCGAGTCCAGACACCATCACCACGCAGAGAGGGGAGGTGGGAGGACAGTTAGACCCCT

FIG. 14 cont'd

GGGTGGGCCCTGCCGTGGGCAAGGATACCCGGTGGCTTCTGGCACCTGGCGGGCTGGT
GACCTTTTTAATCCAGGCCCCATCAGCATCCCACGATCGGCCTTGGTAACCGCCGCGGTA
GATCATTTTTATCCCGCCAGGGAGTGTGATGCAGGAAGACCACATGCGCTCCTGGCTTTTA
AACCTGTTCCTGACTGTTCTCTTACTGCCGAAACCCTTGACTGTTATCTCGGACTTTGCAGG
AGTTCCTTTCCCTCCGAACGCTGCTCCATGCACAGGAAAAGGGCATTTTGTACAATGGGGA
CTTCCCGGGAACGCTTGCTCTTAAGTACCAGAAGCCGGCGGAGCTCTGGCTTTCGTGTTTT
TGGTTTTCTCCTTCCCAAGGCAGCTGGATTGAAAAAACAAAACAAAACAAAAAAACCCAGG
GGCGTCAGTCGATATTCCCAGGGCCGCTTCTCCTGCAGTCTGTGGAGCGTCCTTGTCCCC
GCCGCCGGAATGAATGAGCATTCTGCAGCCCGATGTCCCTGTCCCCTCCTCGCCGGGCCA
TTCTGATTGGACCTGGCCCAGTGCAATCTGTCCAGACAAGCCCTGCTTGCTGGAAAACTGC
CACAAGCACAATTGATCTCTTTTTATCGCCATTCCAGGGGCCTCAGGTCCTACTGGGGAAA
CTTCCTATACCGGAGCTCCAGTTTCTCTTAAGCTGCCCAATTTCACAGAGTACAAAATAGTT
GTAGGGGAAATCAAGGTGAAGGATCTGTCCGACAGTCAAGACGGATCCACAGTAATCTTTC
GGTCTCCTTAAACTACCACCCTCGCTGCCACCCACCCCAAGCTGCTGCCGCCTCACCTTC
CTTGAAATTTCTCAGCGGGAGTCTCCTCACTGCCACTAAAATCCACCCAGCCCACTAACTG
AGGAGCTAGTGTTAATCCAGAGAACCCCCCGCAATGTGCTTCCGAGATTCAGACTGCTTCA
TTGGGAAGTATGATTTGTTCCTTTCTGGAATTGGGCTCCGTGGTGGCGGCGGCACTTCAAG
CAAAGACAGTTTCTTGCAAGCTCCAGTAGCTCCGCGTGTCTCATTTGCCAGGAAGATGGGT
TCCCACGTAGCAAATCGTACATTGTGCCCTGTAGCTCCTTAGCTAGTTAGCTCACAAGCCG
TGTTTTATGACTAATCCTTAATAACTATGGTAAATAACTGTGACTGTGGGGTTTTTAATCTCT
TGTCATTCTCATCCAAAAGTGACCAGCATACCAGTTCTTGCAATAAGATATTACCCTCAGAA
TATTAAGCACATTATTGTAGAGAAAAAAAAATATGTGTACACATATGAACGCACAACATGCA
CATTCATCCTCACATGTGGCACGTAAGGTCTCATTTGATATTGTGTAGGAAATCTGAAGCCT
TTTCCTGAGGTCATCTGTAAAATAGTCTCATTGCCAAGGCATCCCCAGTGCCAGCTGGTGA
ATCCATGATCAAAATGCATACGTATTGTTAAATGATAAGGTTTAGAATGACAGGAACCCATC
ACTGTGTCTCATGGTCCCACTTCCCCATCTGTGTGTGAATTCCTTTAGACTAAGGGCAGGA
AGACTTCCAGCTTTCTCTTTGTTCTTCAATGTGAAACTGAGACCAAGTCTCTCTAAGACAAA
TGCAGTGTATTTAATGTTTGTAAGCAATTCTAAGTGAGATGTTTGGCAAGAAATCCCCTAAC
TGATTTCCATCCAAACCTACCTTATAGAGCACAATATTAAGTGTTGTACAATTACTGTGAGA
ACTGTGAATATGTGTAACTTTTTTTTAGTATTTGCCCGGGGGGAAAAAGATATTGTATTATCA
TATATGCTTTTTTGCAATAAGGATTTATTCTCAGAACACCAAGTAAATCTATCTCTATATAAA
AAATATATGTAATATATACATATTCAAAGTATATACAGAGCCTGTTTTAAAAAATACAGTATTA
TTTAGTAAAATTATCTGTTCTATGGACCAAATGTAAAATATTTATAAATGAAGATGCATTTTAA
ATGTCTATAAATGGTGTCATAACTAGAGCACGGGCGTTATGTAAGTTTCTAAGAATTTAGAG
GATAAATAATAAAGGTTCTATGA (SEQ ID NO: 41)

Solute carrier family 6 member 1 (SLC6A1), transcript variant X2 (XM_005265410.5) [Homo sapiens]

TAATCGTGGTATAGTATGCACCTGAAGGAATGTGGAGCCGCATGCTGGGTCTGCAGAATG
TTGTCACCTGATGAAGAAGAACTGGAAGACGCTGGAAGAAGGCTTTGGAAAAATGCCCAG
ATCATGAGAAGGTGAGAGGTGTTTCTTCCGCCTGCTCCACCAGCAGGTAAAGGAGGCTGA
TCACAGGCTGGCACTCAGGGCGAGGTGGAAGAAATGCGTGGTCTTGCCACTACAGACCC
GCAGAGCTGTCCTCCAACATGATCCAAGAGAGCCACAAAACACAGGGGTGTGTCTTGGGA
AAGGAAAAGTAACAACTTGCAAAAAGGTTGGTTCTGCTTCAGAGAAATGTTCTAGAAAGAC
AGCCAAGTTACCCTGCCGTGAACTCACACACACCTGTGCTTTCTGCATGTTATGTTTCATAG

FIG. 14 cont'd

GAGAGGGAAGTGCCGGGAATTAAAATTGCGGATGCACTGTTCCCCAGGTGTGTGGGTGAG
GCACGCTTCCTGGGACCACTCCATTTAAGCATTGGGTCCCAGGCATGTTCACTACTAGCAG
GAGGCTCTTTATCAATTTCCTCCCGTGGACCAGATGATTGGGATTTTTTTTTCTACCAGACT
TTATTACAGAATAACTATTTTTGTATTAACTAAGCAAACAAATGTACCGCTCCCAGCTTCTGA
CTGGCCCAAGGTGCTGAGCTAGACCTCTTTGCTTTGGTACAAAGAGAGCCCAAAGCCAAG
AGACTTGGCATTTCTGTTAGCTGGTATTCCCCAGGGTGACTAAAAGGATGATTTCCATTGG
ATTTTTTAAAATGCTGAAGACACTGAAAGACCTGAATAGTAAGCCCGTAATAAACGTTATTC
GGTATTATTATTGTTGTTGTTGTAGCCTTCTGGGAAGTTGCGAATGCCAAATTAAGAACAAT
GGACCCATACCCAGATGGTGGCGATTGCAGTGATAACAATGGCTGAGATTTCCTAAGCCCT
AACTATATGCCAGGCATAGTTGTACATGCTTTACACACATGATCTCATTTAATCATCACCATA
ACCCTCGGAAGGAGTTATTATTATCTTCCCATTTTATAGATGAGGGTCAGAGAGGTGAAGT
AACTCACCCAGCATTAGTAGGTGTTGGAGGGAGCCGAGGTCTGTCTGACTCAGCACCAGG
GTGGCTTTGCTGCTGTGTAAGTCAAAGAGCCTTCTCCGTTCTAAATCCTCAAAGAAGTATG
AGTTATATCAATTCATGGCTTTACTTTGTTTATCTTCATATGTGCCACATCTTCTAAAAGGAA
AACTCCAAATATAGTCATGTGGCTCCTCCTAGGGAATTATAGGCATTTGGCTGAATCTGAAT
ACATATCCCCGTATTCCAGAAATTGGAAATCACATGTATGTCTGCTTTATGGTGAGGTCTTC
AGAGCAGCCACCGTCTCTGATCGTTCAGGGCTTGGCATGAGGAGGTGCTCAATATAGTCA
TGGAGAAGAACTCATGGACTGAACTTCCTGGTGAACCAGAGGGATGGTGCAGAGATTTAA
GATTGCAACAGGAACCCCCAAGCCAATGATCTCACCCAGGGTGGCAGAAGGAGGCCTTCT
GGAGCTGACCCACCCCCGACGACCATCAGGGTGAGGCAACTCCAAGGTCCTACTCTCTTT
CTGTGCCTGTTACCCACCCCGTCCTCCTAGGGTGCCCTTGAGCCGCAAAACTGCTGTCCA
CGTGGACCGGGGGTGACATCGCACGTCCATCTGCCAGGACCCCTGCGTCCAAATTCCGA
GACATGGCGACCAACGGCAGCAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGA
GGCCCCTGTGGCCAATGACAAGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGG
CAGACCTCCCCGACCGGGACACGTGGAAGGGCCGCTTCGACTTCCTCATGTCCTGTGTGG
GCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATGGTG
GGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTTCCT
GCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGCTGG
CTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACATCTA
CTACATCGTCATCATCTCCTGGGCCATTTACTACCTGTACAACTCCTTCACCACGACACTGC
CGTGGAAACAGTGCGACAACCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGCATGG
TCAACACTACCAACATGACCAGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCATCAGA
TGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCTGGCC
ATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAAGGTGG
TCTACTTTTCAGCCACATACCCCTACATCATGCTGATCATCCTGTTCTTCCGTGGAGTGACG
CTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACATCACACCCAACTTCCGCAAGCTGTCT
GACTCCGAGGTGTGGCTGGATGCGGCAACCCAGATCTTCTTCTCATACGGGCTGGGCCTG
GGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTTCCACAACAATGTCTACAGGGACTCCA
TCATCGTCTGCTGCATCAATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCTTCTCCAT
CGTGGGCTTCATGGCCCATGTCACCAAGAGGTCCATTGCTGATGTGGCGGCCTCAGGCCC
CGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTGACCCAGCTGCCTATCTCCCCACTCTG
GGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTTCTGCACTGTG
GAGGGCTTCATCACAGCCCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACCGCAGAGA
GCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCTGATCGGTCTCTCTAACATCACTCAGG
GGGGTATTTATGTCTTCAAACTCTTTGACTACTACTCTGCCAGTGGCATGAGCCTGCTGTTC

FIG. 14 cont'd

CTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTTTACGGTGTCAACCGATTCTATGACAA
TATCCAAGAGATGGTTGGATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGGTCTTTCTTC
ACACCAATCATTGTGGCGGGCGTGTTCATTTTCAGTGCTGTGCAGATGACGCCACTCACCA
TGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGGCTCTGTCTT
CCATGGTCCTCATCCCCGGGTACATGGCCTACATGTTCCTCACCTTAAAGGGCTCCCTGAA
GCAGCGCATCCAAGTCATGGTCCAGCCCAGCGAAGACATCGTTCGCCCAGAGAATGGTCC
TGAGCAGCCCCAGGCGGGCAGCTCCACCAGCAAGGAGGCCTACATCTAGGGTGGGGGCC
ACTCACCGACCCGACACTCTCACCCCCCGACCTGGCTGAGTGCGACCACCACTTGATGTC
TGAGGATACCTTCCATCTCAACCTACCTCGAGTGGCGAGTCCAGACACCATCACCACGCA
GAGAGGGGAGGTGGGAGGACAGTTAGACCCCTGGGTGGGCCCTGCCGTGGGCAAGGAT
ACCCGGTGGCTTCTGGCACCTGGCGGGCTGGTGACCTTTTTAATCCAGGCCCCATCAGCA
TCCCACGATCGGCCTTGGTAACCGCCGCGGTAGATCATTTTTATCCCGCCAGGGAGTGTG
ATGCAGGAAGACCACATGCGCTCCTGGCTTTTAAACCTGTTCCTGACTGTTCTCTTACTGC
CGAAACCCTTGACTGTTATCTCGGACTTTGCAGGAGTTCCTTTCCCTCCGAACGCTGCTCC
ATGCACAGGAAAAGGGCATTTTGTACAATGGGGACTTCCCGGGAACGCTTGCTCTTAAGTA
CCAGAAGCCGGCGGAGCTCTGGCTTTCGTGTTTTTGGTTTTCTCCTTCCCAAGGCAGCTG
GATTGAAAAAACAAAACAAAACAAAAAAACCCAGGGGCGTCAGTCGATATTCCCAGGGCCG
CTTCTCCTGCAGTCTGTGGAGCGTCCTTGTCCCCGCCGCCGGAATGAATGAGCATTCTGC
AGCCCGATGTCCCTGTCCCCTCCTCGCCGGGCCATTCTGATTGGACCTGGCCCAGTGCAA
TCTGTCCAGACAAGCCCTGCTTGCTGGAAAACTGCCACAAGCACAATTGATCTCTTTTTATC
GCCATTCCAGGGGCCTCAGGTCCTACTGGGGAAACTTCCTATACCGGAGCTCCAGTTTCT
CTTAAGCTGCCCAATTTCACAGAGTACAAAATAGTTGTAGGGGAAATCAAGGTGAAGGATC
TGTCCGACAGTCAAGACGGATCCACAGTAATCTTTCGGTCTCCTTAAACTACCACCCTCGC
TGCCACCCACCCCAAGCTGCTGCCGCCTCACCTTCCTTGAAATTTCTCAGCGGGAGTCTC
CTCACTGCCACTAAAATCCACCCAGCCCACTAACTGAGGAGCTAGTGTTAATCCAGAGAAC
CCCCCGCAATGTGCTTCCGAGATTCAGACTGCTTCATTGGGAAGTATGATTTGTTCCTTTCT
GGAATTGGGCTCCGTGGTGGCGGCGGCACTTCAAGCAAAGACAGTTTCTTGCAAGCTCCA
GTAGCTCCGCGTGTCTCATTTGCCAGGAAGATGGGTTCCCACGTAGCAAATCGTACATTGT
GCCCTGTAGCTCCTTAGCTAGTTAGCTCACAAGCCGTGTTTTATGACTAATCCTTAATAACT
ATGGTAAATAACTGTGACTGTGGGGTTTTTAATCTCTTGTCATTCTCATCCAAAAGTGACCA
GCATACCAGTTCTTGCAATAAGATATTACCCTCAGAATATTAAGCACATTATTGTAGAGAAA
AAAAAATATGTGTACACATATGAACGCACAACATGCACATTCATCCTCACATGTGGCACGTA
AGGTCTCATTTGATATTGTGTAGGAAATCTGAAGCCTTTTCCTGAGGTCATCTGTAAAATAG
TCTCATTGCCAAGGCATCCCCAGTGCCAGCTGGTGAATCCATGATCAAAATGCATACGTAT
TGTTAAATGATAAGGTTTAGAATGACAGGAACCCATCACTGTGTCTCATGGTCCCACTTCCC
CATCTGTGTGTGAATTCCTTTAGACTAAGGGCAGGAAGACTTCCAGCTTTCTCTTTGTTCTT
CAATGTGAAACTGAGACCAAGTCTCTCTAAGACAAATGCAGTGTATTTAATGTTTGTAAGCA
ATTCTAAGTGAGATGTTTGGCAAGAAATCCCCTAACTGATTTCCATCCAAACCTACCTTATA
GAGCACAATATTAAGTGTTGTACAATTACTGTGAGAACTGTGAATATGTGTAACTTTTTTTTA
GTATTTGCCCGGGGGGAAAAAGATATTGTATTATCATATATGCTTTTTTGCAATAAGGATTT
ATTCTCAGAACACCAAGTAAATCTATCTCTATATAAAAAATATATGTAATATATACATATTCAA
AGTATATACAGAGCCTGTTTTAAAAAATACAGTATTATTTAGTAAAATTATCTGTTCTATGGA
CCAAATGTAAAATATTTATAAATGAAGATGCATTTTAAATGTCTATAAATGGTGTCATAACTA
GAGCACGGGCGTTATGTAAGTTTCTAAGAATTTAGAGGATAAATAATAAAGGTTCTATGA
(SEQ ID NO: 42)

FIG. 14 cont'd

Solute carrier family 6 member 1 (SLC6A1), transcript variant X6 (XM_005265411.5) [Homo sapiens]

TAATCGTGGTATAGTATGCACCTGAAGGAATGTGGAGCCGCATGCTGGGTCTGCAGAATG
TTGTCACCTGATGAAGAAGAACTGGAAGACGCTGGAAGAAGGCTTTGGAAAAATGCCCAG
ATCATGAGAAGGTGAGAGGTGTTTCTTCCGCCTGCTCCACCAGCAGGTAAAGGAGGCTGA
TCACAGGCTGGCACTCAGGGCGAGGTGGAAGAAATGCGTGGTCTTGCCACTACAGACCC
GCAGAGCTGTCCTCCAACATGATCCAAGAGAGCCACAAAACACAGGGGTGTGTCTTGGGA
AAGGAAAAGTAACAACTTGCAAAAAGGTTGGTTCTGCTTCAGAGAAATGTTCTAGAAAGAC
AGCCAAGTTACCCTGCCGTGAACTCACACACACCTGTGCTTTCTGCATGTTATGTTTCATAG
GAGAGGGAAGTGCCGGGAATTAAAATTGCGGATGCACTGTTCCCCAGGTGTGTGGGTGAG
GCACGCTTCCTGGGACCACTCCATTTAAGCATTGGGTCCCAGGCATGTTCACTACTAGCAG
GAGGCTCTTTATCAATTTCCTCCCGTGGACCAGATGATTGGGATTTTTTTTTCTACCAGACT
TTATTACAGAATAACTATTTTTGTATTAACTAAGCAAACAAATGTACCGCTCCCAGCTTCTGA
CTGGCCCAAGGTGCTGAGCTAGACCTCTTTGCTTTGGTACAAAGAGAGCCCAAAGCCAAG
AGACTTGGCATTTCTGTTAGCTGGTATTCCCCAGGGTGACTAAAAGGATGATTTCCATTGG
ATTTTTTAAAATGCTGAAGACACTGAAAGACCTGAATAGTAAGCCCGTAATAAACGTTATTC
GGTATTATTATTGTTGTTGTTGTAGCCTTCTGGGAAGTTGCGAATGCCAAATTAAGAACAAT
GGACCCATACCCAGATGGTGGCGATTGCAGTGATAACAATGGCTGAGATTTCCTAAGCCCT
AACTATATGCCAGGCATAGTTGTACATGCTTTACACACATGATCTCATTTAATCATCACCATA
ACCCTCGGAAGGAGTTATTATTATCTTCCCATTTTATAGATGAGGGTCAGAGAGGTGAAGT
AACTCACCCAGCATTAGTAGGTGTTGGAGGGAGCCGAGGTCTGTCTGACTCAGCACCAGG
GTGGCTTTGCTGCTGTGTAAGTCAAAGAGCCTTCTCCGTTCTAAATCCTCAAAGAAGTATG
AGTTATATCAATTCATGGCTTTACTTTGTTTATCTTCATATGTGCCACATCTTCTAAAAGGAA
AACTCCAAATATAGTCATGTGGCTCCTCCTAGGGAATTATAGGCATTTGGCTGAATCTGAAT
ACATATCCCCGTATTCCAGAAATTGGAAATCACATGTATGTCTGCTTTATGGTGAGGTCTTC
AGAGCAGCCACCGTCTCTGATCGTTCAGGGCTTGGCATGAGGAGGTGCTCAATATAGTCA
TGGAGAAGAACTCATGGACTGAACTTCCTGGTGAACCAGAGGGATGGTGCAGAGATTTAA
GATTGCAACAGGAACCCCCAAGCCAATGATCTCACCCAGGGTGGCAGAAGGAGGCCTTCT
GGAGCTGACCCACCCCCGACGACCATCAGGGTGCCCTTGAGCCGCAAAACTGCTGTCCA
CGTGGACCGGGGGTGACATCGCACGTCCATCTGCCAGGACCCCTGCGTCCAAATTCCGA
GACATGGCGACCAACGGCAGCAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGA
GGCCCCTGTGGCCAATGACAAGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGG
CAGACCTCCCCGACCGGGACACGTGGAAGGGCCGCTTCGACTTCCTCATGTCCTGTGTGG
GCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATGGTG
GGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTTCCT
GCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGCTGG
CTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACATCTA
CTACATCGTCATCATCTCCTGGGCCATTTACTACCTGTACAACTCCTTCACCACGACACTGC
CGTGGAAACAGTGCGACAACCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGCATGG
TCAACACTACCAACATGACCAGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCATCAGA
TGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCTGGCC
ATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAAGGTGG
TCTACTTTTCAGCCACATACCCCTACATCATGCTGATCATCCTGTTCTTCCGTGGAGTGACG
CTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACATCACACCCAACTTCCGCAAGCTGTCT

FIG. 14 cont'd

GACTCCGAGGTGTGGCTGGATGCGGCAACCCAGATCTTCTTCTCATACGGGCTGGGCCTG
GGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTTCCACAACAATGTCTACAGGGACTCCA
TCATCGTCTGCTGCATCAATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCTTCTCCAT
CGTGGGCTTCATGGCCCATGTCACCAAGAGGTCCATTGCTGATGTGGCGGCCTCAGGCCC
CGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTGACCCAGCTGCCTATCTCCCCACTCTG
GGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTTCTGCACTGTG
GAGGGCTTCATCACAGCCCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACCGCAGAGA
GCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCTGATCGGTCTCTCTAACATCACTCAGG
GGGGTATTTATGTCTTCAAACTCTTTGACTACTACTCTGCCAGTGGCATGAGCCTGCTGTTC
CTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTTTACGGTGTCAACCGATTCTATGACAA
TATCCAAGAGATGGTTGGATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGGTCTTTCTTC
ACACCAATCATTGTGGCGGGCGTGTTCATTTTCAGTGCTGTGCAGATGACGCCACTCACCA
TGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGGCTCTGTCTT
CCATGGTCCTCATCCCCGGGTACATGGCCTACATGTTCCTCACCTTAAAGGGCTCCCTGAA
GCAGCGCATCCAAGTCATGGTCCAGCCCAGCGAAGACATCGTTCGCCCAGAGAATGGTCC
TGAGCAGCCCCAGGCGGGCAGCTCCACCAGCAAGGAGGCCTACATCTAGGGTGGGGGCC
ACTCACCGACCCGACACTCTCACCCCCCGACCTGGCTGAGTGCGACCACCACTTGATGTC
TGAGGATACCTTCCATCTCAACCTACCTCGAGTGGCGAGTCCAGACACCATCACCACGCA
GAGAGGGGAGGTGGGAGGACAGTTAGACCCCTGGGTGGGCCCTGCCGTGGGCAAGGAT
ACCCGGTGGCTTCTGGCACCTGGCGGGCTGGTGACCTTTTTAATCCAGGCCCCATCAGCA
TCCCACGATCGGCCTTGGTAACCGCCGCGGTAGATCATTTTTATCCCGCCAGGGAGTGTG
ATGCAGGAAGACCACATGCGCTCCTGGCTTTTAAACCTGTTCCTGACTGTTCTCTTACTGC
CGAAACCCTTGACTGTTATCTCGGACTTTGCAGGAGTTCCTTTCCCTCCGAACGCTGCTCC
ATGCACAGGAAAAGGGCATTTTGTACAATGGGGACTTCCCGGGAACGCTTGCTCTTAAGTA
CCAGAAGCCGGCGGAGCTCTGGCTTTCGTGTTTTTGGTTTTCTCCTTCCCAAGGCAGCTG
GATTGAAAAAACAAAACAAAACAAAAAAACCCAGGGGCGTCAGTCGATATTCCCAGGGCCG
CTTCTCCTGCAGTCTGTGGAGCGTCCTTGTCCCCGCCGCCGGAATGAATGAGCATTCTGC
AGCCCGATGTCCCTGTCCCCTCCTCGCCGGGCCATTCTGATTGGACCTGGCCCAGTGCAA
TCTGTCCAGACAAGCCCTGCTTGCTGGAAAACTGCCACAAGCACAATTGATCTCTTTTTATC
GCCATTCCAGGGGCCTCAGGTCCTACTGGGGAAACTTCCTATACCGGAGCTCCAGTTTCT
CTTAAGCTGCCCAATTTCACAGAGTACAAAATAGTTGTAGGGGAAATCAAGGTGAAGGATC
TGTCCGACAGTCAAGACGGATCCACAGTAATCTTTCGGTCTCCTTAAACTACCACCCTCGC
TGCCACCCACCCCAAGCTGCTGCCGCCTCACCTTCCTTGAAATTTCTCAGCGGGAGTCTC
CTCACTGCCACTAAAATCCACCCAGCCCACTAACTGAGGAGCTAGTGTTAATCCAGAGAAC
CCCCCGCAATGTGCTTCCGAGATTCAGACTGCTTCATTGGGAAGTATGATTTGTTCCTTTCT
GGAATTGGGCTCCGTGGTGGCGGCGGCACTTCAAGCAAAGACAGTTTCTTGCAAGCTCCA
GTAGCTCCGCGTGTCTCATTTGCCAGGAAGATGGGTTCCCACGTAGCAAATCGTACATTGT
GCCCTGTAGCTCCTTAGCTAGTTAGCTCACAAGCCGTGTTTTATGACTAATCCTTAATAACT
ATGGTAAATAACTGTGACTGTGGGGTTTTTAATCTCTTGTCATTCTCATCCAAAAGTGACCA
GCATACCAGTTCTTGCAATAAGATATTACCCTCAGAATATTAAGCACATTATTGTAGAGAAA
AAAAAATATGTGTACACATATGAACGCACAACATGCACATTCATCCTCACATGTGGCACGTA
AGGTCTCATTTGATATTGTGTAGGAAATCTGAAGCCTTTTCCTGAGGTCATCTGTAAAATAG
TCTCATTGCCAAGGCATCCCCAGTGCCAGCTGGTGAATCCATGATCAAAATGCATACGTAT
TGTTAAATGATAAGGTTTAGAATGACAGGAACCCATCACTGTGTCTCATGGTCCCACTTCCC
CATCTGTGTGTGAATTCCTTTAGACTAAGGGCAGGAAGACTTCCAGCTTTCTCTTTGTTCTT

FIG. 14 cont'd

CAATGTGAAACTGAGACCAAGTCTCTCTAAGACAAATGCAGTGTATTTAATGTTTGTAAGCA
ATTCTAAGTGAGATGTTTGGCAAGAAATCCCCTAACTGATTTCCATCCAAACCTACCTTATA
GAGCACAATATTAAGTGTTGTACAATTACTGTGAGAACTGTGAATATGTGTAACTTTTTTTTA
GTATTTGCCCGGGGGGAAAAAGATATTGTATTATCATATATGCTTTTTTGCAATAAGGATTT
ATTCTCAGAACACCAAGTAAATCTATCTCTATATAAAAAATATATGTAATATATACATATTCAA
AGTATATACAGAGCCTGTTTTAAAAAATACAGTATTATTTAGTAAAATTATCTGTTCTATGGA
CCAAATGTAAAATATTTATAAATGAAGATGCATTTTAAATGTCTATAAATGGTGTCATAACTA
GAGCACGGGCGTTATGTAAGTTTCTAAGAATTTAGAGGATAAATAATAAAGGTTCTATGA
(SEQ ID NO: 43)

Solute carrier family 6 member 1 (SLC6A1), transcript variant X3 (XM_017007071.2) [Homo sapiens]

TGCCCTTTGAACGTGGCTTTGGTGGAATTAAGTTAATCGTGGTATAGTATGCACCTGAAGG
AATGTGGAGCCGCATGCTGGGTCTGCAGAATGTTGTCACCTGATGAAGAAGAACTGGAAG
ACGCTGGAAGAAGGCTTTGGAAAAATGCCCAGATCATGAGAAGGTCTTCAGAGCAGCCAC
CGTCTCTGATCGTTCAGGGCTTGGCATGAGGAGGTGCTCAATATAGTCATGGAGAAGAACT
CATGGACTGAACTTCCTGGTGAACCAGAGGGATGGTGCAGAGATTTAAGATTGCAACAGG
AACCCCCAAGCCAATGTAAGCAAAACATCTATTTCATGACTATGGGAATGTTTTGTAAAATC
CAAGCACCAGCCTGAAGAAAGCCTGGAACTGCAAACTAGGAGCTTTGAGACACTCTATCTT
TTGATGTCTAGTTTTCTCCTTCCTCAGCAGACCAAATCTCACTCTGAGTACAAGATGGGCAG
AGGATGGCTGCCCCAGAGTCCCCAAGTCACTAGCAAAACTGCCCTTCTTCCTGGGTCCCA
TGCCAAATTCCAGAGAAGGAATTTGATTGGCTAGCCTGGGTCAGGTGGTCAGTCCTGATTC
AAGGTGATTGGATCATTTTGAACAAATATGGCAGCTGGAGCCCACAAGGGTAGATCATGGG
ACCGTTAAGGGCATCATTGTTACCAGAGTTAATGGCTGAGGCTGCTCTCTCCTGCGTATCC
CAAGAGTCCCTGGTAGACCACAGAATCGGCCGCCTCCTACTATCCTGCCATTCTCCAGAA
GGGGAGGCAGGCACCCTACTGAGTCCAGAGCTCTCTCTAACATTGCCATGCCCAGGCCAG
GATCTAGATGGGCATCCAAGACCAGCCTGTGTGCCGTGAAGGAGTTCAAAGTGAGCAGGC
CATCCCTGAGAACGAGGAGCAGGGGTGTTAGGAAAGTTTGGATTATCTGAAGATGGCCTG
GTTCTCCGTGAACTCTGGGAGGAGATGATGGTGTTTGGATCTCACCCAGGGTGGCAGAAG
GAGGCCTTCTGGAGCTGACCCACCCCCGACGACCATCAGGGTGAGGCAACTCCAAGGTC
CTACTCTCTTTCTGTGCCTGTTACCCACCCCGTCCTCCTAGGGTGCCCTTGAGCCGCAAAA
CTGCTGTCCACGTGGACCGGGGGTGACATCGCACGTCCATCTGCCAGGACCCCTGCGTC
CAAATTCCGAGACATGGCGACCAACGGCAGCAAGGTGGCCGACGGGCAGATCTCCACCG
AGGTCAGCGAGGCCCCTGTGGCCAATGACAAGCCCAAAACCTTGGTGGTCAAGGTGCAGA
AGAAGGCGGCAGACCTCCCCGACCGGGACACGTGGAAGGGCCGCTTCGACTTCCTCATG
TCCTGTGTGGGCTATGCCATCGGCCTGGGCAACGTCTGGAGGTTCCCCTATCTCTGCGGG
AAAAATGGTGGGGGAGCCTTCCTGATCCCCTATTTCCTGACACTCATCTTTGCGGGGGTCC
CACTCTTCCTGCTGGAGTGCTCCCTGGGCCAGTACACCTCCATCGGGGGGCTAGGGGTAT
GGAAGCTGGCTCCTATGTTCAAGGGCGTGGGCCTTGCGGCTGCTGTGCTATCATTCTGGC
TGAACATCTACTACATCGTCATCATCTCCTGGGCCATTTACTACCTGTACAACTCCTTCACC
ACGACACTGCCGTGGAAACAGTGCGACAACCCCTGGAACACAGACCGCTGCTTCTCCAAC
TACAGCATGGTCAACACTACCAACATGACCAGCGCTGTGGTGGAGTTCTGGGAGCGCAAC
ATGCATCAGATGACGGACGGGCTGGATAAGCCAGGTCAGATCCGCTGGCCACTGGCCATC
ACGCTGGCCATCGCCTGGATCCTTGTGTATTTCTGTATCTGGAAGGGTGTTGGCTGGACTG
GAAAGGTGGTCTACTTTTCAGCCACATACCCCTACATCATGCTGATCATCCTGTTCTTCCGT

FIG. 14 cont'd

GGAGTGACGCTGCCCGGGGCCAAGGAGGGCATCCTCTTCTACATCACACCCAACTTCCGC
AAGCTGTCTGACTCCGAGGTGTGGCTGGATGCGGCAACCCAGATCTTCTTCTCATACGGG
CTGGGCCTGGGGTCCCTGATCGCTCTCGGGAGCTACAACTCTTTCCACAACAATGTCTACA
GGGACTCCATCATCGTCTGCTGCATCAATTCGTGCACCAGCATGTTCGCAGGATTCGTCAT
CTTCTCCATCGTGGGCTTCATGGCCCATGTCACCAAGAGGTCCATTGCTGATGTGGCGGC
CTCAGGCCCCGGGCTGGCGTTCCTGGCATACCCAGAGGCGGTGACCCAGCTGCCTATCT
CCCCACTCTGGGCCATCCTCTTCTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTT
CTGCACTGTGGAGGGCTTCATCACAGCCCTGGTGGATGAGTACCCCAGGCTCCTCCGCAA
CCGCAGAGAGCTCTTCATTGCTGCTGTCTGCATCATCTCCTACCTGATCGGTCTCTCTAAC
ATCACTCAGGGGGGTATTTATGTCTTCAAACTCTTTGACTACTACTCTGCCAGTGGCATGA
GCCTGCTGTTCCTCGTGTTCTTTGAATGTGTCTCTATTTCCTGGTTTTACGGTGTCAACCGA
TTCTATGACAATATCCAAGAGATGGTTGGATCCAGGCCCTGCATCTGGTGGAAACTCTGCT
GGTCTTTCTTCACACCAATCATTGTGGCGGGCGTGTTCATTTTCAGTGCTGTGCAGATGAC
GCCACTCACCATGGGAAACTATGTTTTCCCCAAGTGGGGCCAGGGTGTGGGCTGGCTGAT
GGCTCTGTCTTCCATGGTCCTCATCCCCGGGTACATGGCCTACATGTTCCTCACCTTAAAG
GGCTCCCTGAAGCAGCGCATCCAAGTCATGGTCCAGCCCAGCGAAGACATCGTTCGCCCA
GAGAATGGTCCTGAGCAGCCCCAGGCGGGCAGCTCCACCAGCAAGGAGGCCTACATCTA
GGGTGGGGGCCACTCACCGACCCGACACTCTCACCCCCCGACCTGGCTGAGTGCGACCA
CCACTTGATGTCTGAGGATACCTTCCATCTCAACCTACCTCGAGTGGCGAGTCCAGACACC
ATCACCACGCAGAGAGGGGAGGTGGGAGGACAGTTAGACCCCTGGGTGGGCCCTGCCGT
GGGCAAGGATACCCGGTGGCTTCTGGCACCTGGCGGGCTGGTGACCTTTTTAATCCAGGC
CCCATCAGCATCCCACGATCGGCCTTGGTAACCGCCGCGGTAGATCATTTTTATCCCGCCA
GGGAGTGTGATGCAGGAAGACCACATGCGCTCCTGGCTTTTAAACCTGTTCCTGACTGTTC
TCTTACTGCCGAAACCCTTGACTGTTATCTCGGACTTTGCAGGAGTTCCTTTCCCTCCGAA
CGCTGCTCCATGCACAGGAAAAGGGCATTTTGTACAATGGGGACTTCCCGGGAACGCTTG
CTCTTAAGTACCAGAAGCCGGCGGAGCTCTGGCTTTCGTGTTTTTGGTTTTCTCCTTCCCA
AGGCAGCTGGATTGAAAAAACAAAACAAAACAAAAAAACCCAGGGGCGTCAGTCGATATTC
CCAGGGCCGCTTCTCCTGCAGTCTGTGGAGCGTCCTTGTCCCCGCCGCCGGAATGAATGA
GCATTCTGCAGCCCGATGTCCCTGTCCCCTCCTCGCCGGGCCATTCTGATTGGACCTGGC
CCAGTGCAATCTGTCCAGACAAGCCCTGCTTGCTGGAAAACTGCCACAAGCACAATTGATC
TCTTTTTATCGCCATTCCAGGGGCCTCAGGTCCTACTGGGGAAACTTCCTATACCGGAGCT
CCAGTTTCTCTTAAGCTGCCCAATTTCACAGAGTACAAAATAGTTGTAGGGGAAATCAAGGT
GAAGGATCTGTCCGACAGTCAAGACGGATCCACAGTAATCTTTCGGTCTCCTTAAACTACC
ACCCTCGCTGCCACCCACCCCAAGCTGCTGCCGCCTCACCTTCCTTGAAATTTCTCAGCG
GGAGTCTCCTCACTGCCACTAAAATCCACCCAGCCCACTAACTGAGGAGCTAGTGTTAATC
CAGAGAACCCCCCCGCAATGTGCTTCCGAGATTCAGACTGCTTCATTGGGAAGTATGATTTG
TTCCTTTCTGGAATTGGGCTCCGTGGTGGCGGCGGCACTTCAAGCAAAGACAGTTTCTTGC
AAGCTCCAGTAGCTCCGCGTGTCTCATTTGCCAGGAAGATGGGTTCCCACGTAGCAAATC
GTACATTGTGCCCTGTAGCTCCTTAGCTAGTTAGCTCACAAGCCGTGTTTTATGACTAATCC
TTAATAACTATGGTAAATAACTGTGACTGTGGGGTTTTTAATCTCTTGTCATTCTCATCCAAA
AGTGACCAGCATACCAGTTCTTGCAATAAGATATTACCCTCAGAATATTAAGCACATTATTG
TAGAGAAAAAAAAAATATGTGTACACATATGAACGCACAACATGCACATTCATCCTCACATGT
GGCACGTAAGGTCTCATTTGATATTGTGTAGGAAATCTGAAGCCTTTTCCTGAGGTCATCT
GTAAAATAGTCTCATTGCCAAGGCATCCCCAGTGCCAGCTGGTGAATCCATGATCAAAATG
CATACGTATTGTTAAATGATAAGGTTTAGAATGACAGGAACCCATCACTGTGTCTCATGGTC

FIG. 14 cont'd

CCACTTCCCCATCTGTGTGTGAATTCCTTTAGACTAAGGGCAGGAAGACTTCCAGCTTTCT
CTTTGTTCTTCAATGTGAAACTGAGACCAAGTCTCTCTAAGACAAATGCAGTGTATTTAATG
TTTGTAAGCAATTCTAAGTGAGATGTTTGGCAAGAAATCCCCTAACTGATTTCCATCCAAAC
CTACCTTATAGAGCACAATATTAAGTGTTGTACAATTACTGTGAGAACTGTGAATATGTGTA
ACTTTTTTTTTAGTATTTGCCCGGGGGGGAAAAAGATATTGTATTATCATATATGCTTTTTTGCA
ATAAGGATTTATTCTCAGAACACCAAGTAAATCTATCTCTATATAAAAAATATATGTAATATAT
ACATATTCAAAGTATATACAGAGCCTGTTTTAAAAAATACAGTATTATTTAGTAAAATTATCT
GTTCTATGGACCAAATGTAAAATATTTATAAATGAAGATGCATTTTAAATGTCTATAAATGGT
GTCATAACTAGAGCACGGGCGTTATGTAAGTTTCTAAGAATTTAGAGGATAAATAATAAAGG
TTCTATGA (SEQ ID NO: 44)

Solute carrier family 6 member 1 (SLC6A1), transcript variant X4 (XM_017007072.2) [Homo sapiens]

TGCCCTTTGAACGTGGCTTTGGTGGAATTAAGTTAATCGTGGTATAGTATGCACCTGAAGG
AATGTGGAGCCGCATGCTGGGTCTGCAGAATGTTGTCACCTGATGAAGAAGAACTGGAAG
ACGCTGGAAGAAGGCTTTGGAAAAATGCCCAGATCATGAGAAGGTCTTCAGAGCAGCCAC
CGTCTCTGATCGTTCAGGGCTTGGCATGAGGAGGTGCTCAATATAGTCATGGAGAAGAACT
CATGGACTGAACTTCCTGGTGAACCAGAGGGATGGTGCAGAGATTTAAGATTGCAACAGG
AACCCCCAAGCCAATGATCTCACCCAGGGTGGCAGAAGGAGGCCTTCTGGAGCTGACCCA
CCCCCGACGACCATCAGGGTGAGGCAACTCCAAGGTCCTACTCTCTTTCTGTGCCTGTTAC
CCACCCCGTCCTCCTAGGGTGCCCTTGAGCCGCAAAACTGCTGTCCACGTGGACCGGGG
GTGACATCGCACGTCCATCTGCCAGGACCCCTGCGTCCAAATTCCGAGACATGGCGACCA
ACGGCAGCAAGGTGGCCGACGGGCAGATCTCCACCGAGGTCAGCGAGGCCCCTGTGGC
CAATGACAAGCCCAAAACCTTGGTGGTCAAGGTGCAGAAGAAGGCGGCAGACCTCCCCGA
CCGGGACACGTGGAAGGGCCGCTTCGACTTCCTCATGTCCTGTGTGGGCTATGCCATCGG
CCTGGGCAACGTCTGGAGGTTCCCCTATCTCTGCGGGAAAAATGGTGGGGGAGCCTTCCT
GATCCCCTATTTCCTGACACTCATCTTTGCGGGGGTCCCACTCTTCCTGCTGGAGTGCTCC
CTGGGCCAGTACACCTCCATCGGGGGGCTAGGGGTATGGAAGCTGGCTCCTATGTTCAAG
GGCGTGGGCCTTGCGGCTGCTGTGCTATCATTCTGGCTGAACATCTACTACATCGTCATCA
TCTCCTGGGCCATTTACTACCTGTACAACTCCTTCACCACGACACTGCCGTGGAAACAGTG
CGACAACCCCTGGAACACAGACCGCTGCTTCTCCAACTACAGCATGGTCAACACTACCAAC
ATGACCAGCGCTGTGGTGGAGTTCTGGGAGCGCAACATGCATCAGATGACGGACGGGCT
GGATAAGCCAGGTCAGATCCGCTGGCCACTGGCCATCACGCTGGCCATCGCCTGGATCCT
TGTGTATTTCTGTATCTGGAAGGGTGTTGGCTGGACTGGAAAGGTGGTCTACTTTTCAGCC
ACATACCCCTACATCATGCTGATCATCCTGTTCTTCCGTGGAGTGACGCTGCCCGGGGCC
AAGGAGGGCATCCTCTTCTACATCACACCCAACTTCCGCAAGCTGTCTGACTCCGAGGTGT
GGCTGGATGCGGCAACCCAGATCTTCTTCTCATACGGGCTGGGCCTGGGGTCCCTGATCG
CTCTCGGGAGCTACAACTCTTTCCACAACAATGTCTACAGGGACTCCATCATCGTCTGCTG
CATCAATTCGTGCACCAGCATGTTCGCAGGATTCGTCATCTTCTCCATCGTGGGCTTCATG
GCCCATGTCACCAAGAGGTCCATTGCTGATGTGGCGGCCTCAGGCCCCGGGCTGGCGTT
CCTGGCATACCCAGAGGCGGTGACCCAGCTGCCTATCTCCCCACTCTGGGCCATCCTCTT
CTTCTCCATGCTGTTGATGCTGGGCATTGACAGCCAGTTCTGCACTGTGGAGGGCTTCATC
ACAGCCCTGGTGGATGAGTACCCCAGGCTCCTCCGCAACCGCAGAGAGCTCTTCATTGCT
GCTGTCTGCATCATCTCCTACCTGATCGGTCTCTCTAACATCACTCAGGGGGGTATTTATG
TCTTCAAACTCTTTGACTACTACTCTGCCAGTGGCATGAGCCTGCTGTTCCTCGTGTTCTTT

FIG. 14 cont'd

GAATGTGTCTCTATTTCCTGGTTTTACGGTGTCAACCGATTCTATGACAATATCCAAGAGAT
GGTTGGATCCAGGCCCTGCATCTGGTGGAAACTCTGCTGGTCTTTCTTCACACCAATCATT
GTGGCGGGCGTGTTCATTTTCAGTGCTGTGCAGATGACGCCACTCACCATGGGAAACTAT
GTTTTCCCCAAGTGGGGCCAGGGTGTGGGCTGGCTGATGGCTCTGTCTTCCATGGTCCTC
ATCCCCGGGTACATGGCCTACATGTTCCTCACCTTAAAGGGCTCCCTGAAGCAGCGCATC
CAAGTCATGGTCCAGCCCAGCGAAGACATCGTTCGCCCAGAGAATGGTCCTGAGCAGCCC
CAGGCGGGCAGCTCCACCAGCAAGGAGGCCTACATCTAGGGTGGGGGCCACTCACCGAC
CCGACACTCTCACCCCCCGACCTGGCTGAGTGCGACCACCACTTGATGTCTGAGGATACC
TTCCATCTCAACCTACCTCGAGTGGCGAGTCCAGACACCATCACCACGCAGAGAGGGGAG
GTGGGAGGACAGTTAGACCCCTGGGTGGGCCCTGCCGTGGGCAAGGATACCCGGTGGCT
TCTGGCACCTGGCGGGCTGGTGACCTTTTTAATCCAGGCCCCATCAGCATCCCACGATCG
GCCTTGGTAACCGCCGCGGTAGATCATTTTTATCCCGCCAGGGAGTGTGATGCAGGAAGA
CCACATGCGCTCCTGGCTTTTAAACCTGTTCCTGACTGTTCTCTTACTGCCGAAACCCTTGA
CTGTTATCTCGGACTTTGCAGGAGTTCCTTTCCCTCCGAACGCTGCTCCATGCACAGGAAA
AGGGCATTTTGTACAATGGGGACTTCCCGGGAACGCTTGCTCTTAAGTACCAGAAGCCGG
CGGAGCTCTGGCTTTCGTGTTTTTGGTTTTCTCCTTCCCAAGGCAGCTGGATTGAAAAAAC
AAAACAAAACAAAAAAACCCAGGGGCGTCAGTCGATATTCCCAGGGCCGCTTCTCCTGCA
GTCTGTGGAGCGTCCTTGTCCCCGCCGCCGGAATGAATGAGCATTCTGCAGCCCGATGTC
CCTGTCCCCTCCTCGCCGGGCCATTCTGATTGGACCTGGCCCAGTGCAATCTGTCCAGAC
AAGCCCTGCTTGCTGGAAAACTGCCACAAGCACAATTGATCTCTTTTTATCGCCATTCCAG
GGGCCTCAGGTCCTACTGGGGAAACTTCCTATACCGGAGCTCCAGTTTCTCTTAAGCTGCC
CAATTTCACAGAGTACAAAATAGTTGTAGGGGAAATCAAGGTGAAGGATCTGTCCGACAGT
CAAGACGGATCCACAGTAATCTTTCGGTCTCCTTAAACTACCACCCTCGCTGCCACCCACC
CCAAGCTGCTGCCGCCTCACCTTCCTTGAAATTTCTCAGCGGGAGTCTCCTCACTGCCACT
AAAATCCACCCAGCCCACTAACTGAGGAGCTAGTGTTAATCCAGAGAACCCCCCGCAATGT
GCTTCCGAGATTCAGACTGCTTCATTGGGAAGTATGATTTGTTCCTTTCTGGAATTGGGCTC
CGTGGTGGCGGCGGCACTTCAAGCAAAGACAGTTTCTTGCAAGCTCCAGTAGCTCCGCGT
GTCTCATTTGCCAGGAAGATGGGTTCCCACGTAGCAAATCGTACATTGTGCCCTGTAGCTC
CTTAGCTAGTTAGCTCACAAGCCGTGTTTTATGACTAATCCTTAATAACTATGGTAAATAACT
GTGACTGTGGGGTTTTTAATCTCTTGTCATTCTCATCCAAAAGTGACCAGCATACCAGTTCT
TGCAATAAGATATTACCCTCAGAATATTAAGCACATTATTGTAGAGAAAAAAAAATATGTGTA
CACATATGAACGCACAACATGCACATTCATCCTCACATGTGGCACGTAAGGTCTCATTTGAT
ATTGTGTAGGAAATCTGAAGCCTTTTCCTGAGGTCATCTGTAAAATAGTCTCATTGCCAAGG
CATCCCCAGTGCCAGCTGGTGAATCCATGATCAAAATGCATACGTATTGTTAAATGATAAG
GTTTAGAATGACAGGAACCCATCACTGTGTCTCATGGTCCCACTTCCCCATCTGTGTGTGA
ATTCCTTTAGACTAAGGGCAGGAAGACTTCCAGCTTTCTCTTTGTTCTTCAATGTGAAACTG
AGACCAAGTCTCTCTAAGACAAATGCAGTGTATTTAATGTTTGTAAGCAATTCTAAGTGAGA
TGTTTGGCAAGAAATCCCCTAACTGATTTCCATCCAAACCTACCTTATAGAGCACAATATTA
AGTGTTGTACAATTACTGTGAGAACTGTGAATATGTGTAACTTTTTTTTTAGTATTTGCCCGG
GGGGAAAAAGATATTGTATTATCATATATGCTTTTTTGCAATAAGGATTTATTCTCAGAACAC
CAAGTAAATCTATCTCTATATAAAAAATATATGTAATATATACATATTCAAAGTATATACAGAG
CCTGTTTTAAAAAATACAGTATTATTTAGTAAAATTATCTGTTCTATGGACCAAATGTAAAATA
TTTATAAATGAAGATGCATTTTAAATGTCTATAAATGGTGTCATAACTAGAGCACGGGCGTT
ATGTAAGTTTCTAAGAATTTAGAGGATAAATAATAAAGGTTCTATGA (SEQ ID NO: 45)

FIG. 14 cont'd

Sodium- and chloride-dependent GABA transporter 1 isoform a [Homo sapiens] (NP_003033.3)
MATNGSKVADGQISTEVSEAPVANDKPKTLVVKVQKKAADLPDRDTWKGRFDFLMSCVGYAI
GLGNVWRFPYLCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLGQYTSIGGLGVWKLAPMFKGVG
LAAAVLSFWLNIYYIVIISWAIYYLYNSFTTTLPWKQCDNPWNTDRCFSNYSMVNTTNMTSAVVE
FWERNMHQMTDGLDKPGQIRWPLAITLAIAWILVYFCIWKGVGWTGKVVYFSATYPYIMLIILFF
RGVTLPGAKEGILFYITPNFRKLSDSEVWLDAATQIFFSYGLGLGSLIALGSYNSFHNNVYRDSII
VCCINSCTSMFAGFVIFSIVGFMAHVTKRSIADVAASGPGLAFLAYPEAVTQLPISPLWAILFFSM
LLMLGIDSQFCTVEGFITALVDEYPRLLRNRRELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYSAS
GMSLLFLVFFECVSISWFYGVNRFYDNIQEMVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQMT
PLTMGNYVFPKWGQGVGWLMALSSMVLIPGYMAYMFLTLKGSLKQRIQVMVQPSEDIVRPEN
GPEQPQAGSSTSKEAYI (SEQ ID NO: 46)

Myc-DDK tag
EQKLISEEDLAANDILDYKDDDDK (SEQ ID NO: 47)

SYFP2:
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCC
ACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAG
AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAA (SEQ ID NO: 48)

EGFP:
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCA
CCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGA
AGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT
TCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACA
ACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACA
TGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAA (SEQ ID NO: 49)

FIG. 14 cont'd

Optimized Flp recombinase (FlpO):

ATGGCTCCTAAGAAGAAGAGGAAGGTGATGAGCCAGTTCGACATCCTGTGCAAGACCCCC
CCCAAGGTGCTGGTGCGGCAGTTCGTGGAGAGATTCGAGAGGCCCAGCGGCGAGAAGAT
CGCCAGCTGTGCCGCCGAGCTGACCTACCTGTGCTGGATGATCACCCACAACGGCACCG
CCATCAAGAGGGCCACCTTCATGAGCTACAACACCATCATCAGCAACAGCCTGAGCTTCGA
CATCGTGAACAAGAGCCTGCAGTTCAAGTACAAGACCCAGAAGGCCACCATCCTGGAGGC
CAGCCTGAAGAAGCTGATCCCCGCCTGGGAGTTCACCATCATCCCTTACAACGGCCAGAA
GCACCAGAGCGACATCACCGACATCGTGTCCAGCCTGCAGCTGCAGTTCGAGAGCAGCG
AGGAGGCCGACAAGGGCAACAGCCACAGCAAGAAGATGCTGAAGGCCCTGCTGTCCGAG
GGCGAGAGCATCTGGGAGATCACCGAGAAGATCCTGAACAGCTTCGAGTACACCAGCAGG
TTCACCAAGACCAAGACCCTGTACCAGTTCCTGTTCCTGGCCACATTCATCAACTGCGGCA
GGTTCAGCGACATCAAGAACGTGGACCCCAAGAGCTTCAAGCTGGTGCAGAACAAGTACC
TGGGCGTGATCATTCAGTGCCTGGTGACCGAGACCAAGACAAGCGTGTCCAGGCACATCT
ACTTTTTCAGCGCCAGAGGCAGGATCGACCCCCTGGTGTACCTGGACGAGTTCCTGAGGA
ACAGCGAGCCCGTGCTGAAGAGAGTGAACAGGACCGGCAACAGCAGCAGCAACAAGCAG
GAGTACCAGCTGCTGAAGGACAACCTGGTGCGCAGCTACAACAAGGCCCTGAAGAAGAAC
GCCCCCTACCCCATCTTCGCTATCAAGAACGGCCCTAAGAGCCACATCGGCAGGCACCTG
ATGACCAGCTTTCTGAGCATGAAGGGCCTGACCGAGCTGACAAACGTGGTGGGCAACTGG
AGCGACAAGAGGGCCTCCGCCGTGGCCAGGACCACCTACACCCACCAGATCACCGCCAT
CCCCGACCACTACTTCGCCCTGGTGTCCAGGTACTACGCCTACGACCCCATCAGCAAGGA
GATGATCGCCCTGAAGGACGAGACCAACCCCATCGAGGAGTGGCAGCACATCGAGCAGC
TGAAGGGCAGCGCCGAGGGCAGCATCAGATACCCCGCCTGGAACGGCATCATCAGCCAG
GAGGTGCTGGACTACCTGAGCAGCTACATCAACAGGCGGATCTGA (SEQ ID NO: 50)

Improved Cre recombinase (iCre):

ATGGTGCCCAAGAAGAAGAGGAAAGTCTCCAACCTGCTGACTGTGCACCAAAACCTGCCT
GCCCTCCCTGTGGATGCCACCTCTGATGAAGTCAGGAAGAACCTGATGGACATGTTCAGG
GACAGGCAGGCCTTCTCTGAACACACCTGGAAGATGCTCCTGTCTGTGTGCAGATCCTGG
GCTGCCTGGTGCAAGCTGAACAACAGGAAATGGTTCCCTGCTGAACCTGAGGATGTGAGG
GACTACCTCCTGTACCTGCAAGCCAGAGGCCTGGCTGTGAAGACCATCCAACAGCACCTG
GGCCAGCTCAACATGCTGCACAGGAGATCTGGCCTGCCTCGCCCTTCTGACTCCAATGCT
GTGTCCCTGGTGATGAGGAGAATCAGAAAGGAGAATGTGGATGCTGGGGAGAGAGCCAA
GCAGGCCCTGGCCTTTGAACGCACTGACTTTGACCAAGTCAGATCCCTGATGGAGAACTC
TGACAGATGCCAGGACATCAGGAACCTGGCCTTCCTGGGCATTGCCTACAACACCCTGCT
GCGCATTGCCGAAATTGCCAGAATCAGAGTGAAGGACATCTCCCGCACCGATGGTGGGAG
AATGCTGATCCACATTGGCAGGACCAAGACCCTGGTGTCCACAGCTGGTGTGGAGAAGGC
CCTGTCCCTGGGGGTTACCAAGCTGGTGGAGAGATGGATCTCTGTGTCTGGTGTGGCTGA
TGACCCCAACAACTACCTGTTCTGCCGGGTCAGAAAGAATGGTGTGGCTGCCCCTTCTGC
CACCTCCCAACTGTCCACCCGGGCCCTGGAAGGGATCTTTGAGGCCACCCACCGCCTGAT
CTATGGTGCCAAGGATGACTCTGGGCAGAGATACCTGGCCTGGTCTGGCCACTCTGCCAG
AGTGGGTGCTGCCAGGGACATGGCCAGGGCTGGTGTGTCCATCCCTGAAATCATGCAGG
CTGGTGGCTGGACCAATGTGAACATTGTGATGAACTACATCAGAAACCTGGACTCTGAGAC
TGGGGCCATGGTGAGGCTGCTCGAGGATGGGGACTAA (SEQ ID NO: 51)

FIG. 14 cont'd

tet-Transactivator version 2 (tTA2):

ATGTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGAAGTCG
GTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCC
TGTACTGGCACGTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGGCAATCGAGATGCTGG
ACAGGCATCATACCCACTTCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTGCGGA
ACAACGCCAAGTCATTCCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCT
CGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTG
TCAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTAC
ACTGGGCTGCGTATTGGAGGATCAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACC
TACCACCGATTCTATGCCCCCACTTCTGAGACAAGCAATTGAGCTGTTCGACCATCAGGGA
GCCGAACCTGCCTTCCTTTTCGGCCTGGAACTAATCATATGTGGCCTGGAGAAACAGCTAA
AGTGCGAAAGCGGCGGGCCGGCCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAG
CCGATGCCCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGA
CCTTGACATGCTCCCCGGGTAA (SEQ ID NO: 52)

GCaMP6m:

ATGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCAAGATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGCTTTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCCACAGAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAGGGA
GCTACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 53)

GCaMP6s:

ATGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCCACATCCGCC

FIG. 14 cont'd

ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGCTTTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCCACAGAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAATGA
AATACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 54)

GCaMP6f:
ATGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCTCGCCACCATGGTCGACTCATCACGTC
GTAAGTGGAATAAGACAGGTCACGCAGTCAGAGCTATAGGTCGGCTGAGCTCACTCGAGA
ACGTCTATATCAAGGCCGACAAGCAGAAGAACGGCATCAAGGCGAACTTCAAGATCCGCC
ACAACATCGAGGACGGCGGCGTGCAGCTCGCCTACCACTACCAGCAGAACACCCCCATC
GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCGTGCAGTCCAAACTTTCG
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGTACCGGAGGGAGCATGGTGAGCA
AGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGTGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC
CACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCGAAGGCTACATCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGG
AGTACAACCTGCCGGACCAACTGACTGAAGAGCAGATCGCAGAATTTAAAGAGGAATTCTC
CCTATTTGACAAGGACGGGGATGGGACAATAACAACCAAGGAGCTGGGGACGGTGATGC
GGTCTCTGGGGCAGAACCCCACAGAAGCAGAGCTGCAGGACATGATCAATGAAGTAGATG
CCGACGGTGACGGCACAATCGACTTCCCTGAGTTCCTGACAATGATGGCAAGAAAAATGA
AATACAGGGACACGGAAGAAGAAATTAGAGAAGCGTTCGGTGTGTTTGATAAGGATGGCA
ATGGCTACATCAGTGCAGCAGAGCTTCGCCACGTGATGACAAACCTTGGAGAGAAGTTAA
CAGATGAAGAGGTTGATGAAATGATCAGGGAAGCAGACATCGATGGGGATGGTCAGGTAA
ACTACGAAGAGTTTGTACAAATGATGACAGCGAAGTGA (SEQ ID NO: 55)

FIG. 14 cont'd

SP10 insulator (SP10ins):
GAAGCTACCCCTAACACACTATTCTACACACAGAAAATGCTCTTCACTAG (SEQ ID NO: 56)

3xSP10ins:
GAAGCTACCCCTAACACACTATTCTACACACAGAAAATGCTCTTCACTAGGAAGCTACCCC
TAACACACTATTCTACACACAGAAAATGCTCTTCACTAGGAAGCTACCCCTAACACACTATT
CTACACACAGAAAATGCTCTTCACTAG (SEQ ID NO: 57)

WPRE3:
ATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTC
CTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATG
GCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGC
CGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG
(SEQ ID NO: 58)

WPRE:
GCTTATCGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTA
TGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTT
CCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAG
TTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCC
ACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCC
CTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGG
CTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGC
TCGCCTATGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCC
TCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTC
TTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCATCGATACC
G (SEQ ID NO: 59)

BGHpA:
CGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGAC
CCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATG (SEQ ID NO: 60)

HGHpA:
ACGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTC
CAGTGCCCACCAGCCTTGTCCTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCC
TTCTATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACA
ACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGG
CTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGT
TGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTTGGTAGAGACGGGG
TTTCACCATATTGGCCAGGCTGGTCTCCAACTCCTAATCTCAGGTGATCTACCCACCTTGG
CCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCCTT (SEQ ID
NO: 61)

FIG. 14 cont'd

P2A:
GGCAGCGGCGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAGAACCC
CGGCCCCGGAGCTAGCGGA (SEQ ID NO: 62)

T2A:
(GSG)EGRGSLLTCGDVEENPGP (SEQ ID NO: 63)

E2A:
(GSG)QCTNYALLKLAGDVESNPGPP (SEQ ID NO: 64)

F2A:
(GSG)VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 65)

Exemplary Plasmid Backbone 1 – Left ITR:
CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTT
TGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCA
CTAGGGGTTCCT (SEQ ID NO: 66)

Exemplary Plasmid Backbone 1 – Right ITR:
CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCG
TCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG
GCCAACTCCATCACTAGGGGTTCCT (SEQ ID NO: 67)

Exemplary Plasmid Backbone 2 – Left ITR:
CATGTCCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCG
GGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGG
AGTGGCCAACTCCATCACTAGGGGTTCCT (SEQ ID NO: 68)

Exemplary Plasmid Backbone 2 – Right ITR:
AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGG
CCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGA
GCGAGCGCGCAGCTGCCTGCAGGGGCGCCTG (SEQ ID NO: 69)

PHP.eB capsid:
MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDK
GEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKK
RLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDP
QPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRT
WALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWG
FRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADV
FMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLD
RLMNPLIDQYLYYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQN
NNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADK
VMITNEEEIKTTNPVATESYGQVATNHQSDGTLAVPFKAQAQTGWVQNQGILPGMVWQDRDV
YLQGPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYST

FIG. 14 cont'd

GQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL (SEQ ID NO: 70)

AAV9 VP1 capsid protein:
MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPGNGLDK GEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKK RLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDP QPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRT WALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWG FRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADV FMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLD RLMNPLIDQYLYYLSKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQN NNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADK VMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQGPIW AKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIE WELQKENSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL (SEQ ID NO: 71)

L-ITR to R-ITR for single and dual specificity vectors

CN1390
GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTGGGCATAAA
AGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTA
GCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC
CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGG
GCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAG
CTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGC
CCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTA
CGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGT
GAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG
AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATA
TCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCG
AGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGC
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCC
AACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTC
GGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAATTCGATATCA
TAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCC
TTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGG
CTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCC

FIG. 14 cont'd

GCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGC
TCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGT
GCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATT
GCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGC
AAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGTGCGGACCGAGCGGC
CGC (SEQ ID NO: 72)

CN2102

GCGGCCGCACGCGTTGGCTCTGGCAAAGGATGGAGATTTTTTTTTTTAACCTATTCTTTTGA
GAATAAGGAGGGTCTTTGTTTCTCTTCCCTGGGAACCGAGCAGCCCCTTTCTCTGAGGGTA
GAGTTGGGGCATGGCTCAGCAGAACGACAGTCATCCTGGCTAGTAAGACTCAGAGGCTGG
CCTTCAAAGGCTTGAGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGA
ATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTA
AAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGG
AGGAGTTCCTTAGAGGGGAAGGGGATAGGTTGCAATGGTCTGCTTAGCCTCCCGAAGGCT
GCAGAAGCTGCTAAGACAGGGGTAACTGAAAGATAGACTCTGTGAGGTTGAACAGCAATC
TTTATGACCTGACCCCCTAGGCAGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCAT
CTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCAC
CATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG
ACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC
TACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCC
CACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAG
AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCA
GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGATT
ACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGAT
ACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCT
TGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTG
CTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTG
TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGA
AGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTA
GGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAA
GACAATAGCAGGCATGAGATCTCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 73)

CN2720

GCGGCCGCACGCGTTGGCTCTGGCAAAGGATGGAGATTTTTTTTTTTAACCTATTCTTTTGA
GAATAAGGAGGGTCTTTGTTTCTCTTCCCTGGGAACCGAGCAGCCCCTTTCTCTGAGGGTA
GAGTTGGGGCATGGCTCAGCAGAACGACAGTCATCCTGGCTAGTAAGACTCAGAGGCTGG
CCTTCAAAGGCTTGAGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGA

FIG. 14 cont'd

ATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTA
AAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGG
AGGAGTTCCTTAGAGGGGAAGGGGATAGGTTGCAATGGTCTGCTTAGCCTCCCGAAGGCT
GCAGAAGCTGCTAAGACAGGGGTAACTGAAAGATAGACTCTGTGAGGTTGAACAGCAATC
TTTATGACCTGACCCCCTAGGCAGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTG
GAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGA
AAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATT
AAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGC
TACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTT
TAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCA
GCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTG
GGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTT
CGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGG
GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCA
CCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAG
TGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA
CTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAA
CGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCA
CAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCG
GCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGC
AAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAA
TTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACT
ATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTT
CCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAA
CTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAA
TTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC
CTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAAT
GAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGG
CAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGTGCGGA
CCGAGCGGCCGC (SEQ ID NO: 74)

CN2721

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCTGGCTCTGGCAAAG
GATGGAGATTTTTTTTTTAACCTATTCTTTTGAGAATAAGGAGGGTCTTTGTTTCTCTTCCCT
GGGAACCGAGCAGCCCCTTTCTCTGAGGGTAGAGTTGGGGCATGGCTCAGCAGAACGAC

FIG. 14 cont'd

AGTCATCCTGGCTAGTAAGACTCAGAGGCTGGCCTTCAAAGGCTTGAGCTCCTGGGTGAC
TGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGCGGGCTTTATGCAGT
CTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAATTTTGCACTGCATA
GCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGGGGAAGGGGATAG
GTTGCAATGGTCTGCTTAGCCTCCCGAAGGCTGCAGAAGCTGCTAAGACAGGGGTAACTG
AAAGATAGACTCTGTGAGGTTGAACAGCAATCTTTATGACCTGACCCCCTAGGCATCTAGA
AGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGA
TCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCT
GTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT
TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTG
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTA
CGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTC
CGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGA
AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACA
ACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCA
AGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAAC
ACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTC
CAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGAC
CGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGC
GGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGT
ATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCAT
GCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCC
ACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGG
CACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGT
TGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC
TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG
GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCA
CGTGCGGACCGAGCGGCCGC (SEQ ID NO: 75)

CN2722

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTGGGCATAAA
AGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTA
GCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC
CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGG
GCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAG
CTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGC
CCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTA
CGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGT

FIG. 14 cont'd

GAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGG
AGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATA
TCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCG
AGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGC
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCC
AACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTC
GGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAATTCGATATCA
TAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCC
TTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGG
CTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCC
GCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGC
TCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGT
GCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATT
GCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGC
AAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGTGCGGACCGTGGCTC
TGGCAAAGGATGGAGATTTTTTTTTTAACCTATTCTTTTGAGAATAAGGAGGGTCTTTGTTTC
TCTTCCCTGGGAACCGAGCAGCCCCTTTCTCTGAGGGTAGAGTTGGGGCATGGCTCAGCA
GAACGACAGTCATCCTGGCTAGTAAGACTCAGAGGCTGGCCTTCAAAGGCTTGAGCTCCT
GGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGCGGGCTTT
ATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAATTTTGCA
CTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGGGGAAG
GGGATAGGTTGCAATGGTCTGCTTAGCCTCCCGAAGGCTGCAGAAGCTGCTAAGACAGGG
GTAACTGAAAGATAGACTCTGTGAGGTTGAACAGCAATCTTTATGACCTGACCCCCTAGGC
ATCTAGAAGCGGCCGC (SEQ ID NO: 76)

CN2732

GCGGCCGCACGCGTTGGCTCTGGCAAAGGATGGAGATTTTTTTTTTAACCTATTCTTTTGA
GAATAAGGAGGGTCTTTGTTTCTCTTCCCTGGGAACCGAGCAGCCCCTTTCTCTGAGGGTA
GAGTTGGGGCATGGCTCAGCAGAACGACAGTCATCCTGGCTAGTAAGACTCAGAGGCTGG
CCTTCAAAGGCTTGAGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGA
ATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTA
AAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGG
AGGAGTTCCTTAGAGGGGAAGGGGATAGGTTGCAATGGTCTGCTTAGCCTCCCGAAGGCT
GCAGAAGCTGCTAAGACAGGGGTAACTGAAAGATAGACTCTGTGAGGTTGAACAGCAATC
TTTATGACCTGACCCCCTAGGCAGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCAT
CTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCAC
CATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG
ACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC
TACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCC
CACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACAT
GAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACA
CCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAG
AAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCA

FIG. 14 cont'd

GCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
ACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGT
ACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGATT
ACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGAT
ACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCT
TGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTG
CTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTG
TGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGA
AGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTA
GGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAA
GACAATAGCAGGCATGAGATCTCACGTGCGGACCGCTAAATAAAGATGGCTTTTTAGTATT
AAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGC
TACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTT
TAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCA
GCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGAT
GGCTTTTTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCT
GTAATCAGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGTCTAGA
GACCGAGCGGCCGC (SEQ ID NO: 77)

CN3213

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCTGGCTCTGGCAAAG
GATGGAGATTTTTTTTTTAACCTATTCTTTTGAGAATAAGGAGGGTCTTTGTTTCTCTTCCCT
GGGAACCGAGCAGCCCCTTTCTCTGAGGGTAGAGTTGGGGCATGGCTCAGCAGAACGAC
AGTCATCCTGGCTAGTAAGACTCAGAGGCTGGCCTTCAAAGGCTTGAGCTCCTGGGTGAC
TGAGCAGTGGAGAAACAGAATCCTGCCCTTGAATTGCTCCCCAGGCGGGCTTTATGCAGT
CTGGGGAAGCAAGGGATGCCCTGTGATTCTTAAAGAAACTGGTATAATTTTGCACTGCATA
GCAGACTCCCAAGACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGGGGAAGGGGATAG
GTTGCAATGGTCTGCTTAGCCTCCCGAAGGCTGCAGAAGCTGCTAAGACAGGGGTAACTG
AAAGATAGACTCTGTGAGGTTGAACAGCAATCTTTATGACCTGACCCCCTAGGCATCTAGA
AGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGA
TCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCACCATGGCTACGAATGGCTCAAAAGT
CGCTGATGGTCAGATCAGCACGGAAGTGAGCGAGGCCCCTGTGGCCAACGACAAGCCTA
AAACTCTGGTCGTAAAAGTACAGAAGAAAGCAGCGGATCTGCCAGACCGTGATACATGGA
AAGGTAAGTACTAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAA
GGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATCTT
CCTCCCACAGGACGTTTCGATTTCCTGATGAGCTGCGTAGGCTATGCAATCGGACTGGGA
AATGTCTGGCGCTTCCCTTACCTGTGTGGTAAAAATGGTGGAGGCGCCTTTCTCATTCCTT
ATTTTCTCACCCTCATTTTTGCCGGGGTTCCCTTGTTTCTCCTGGAATGCAGCCTCGGGCA

FIG. 14 cont'd

ATATACCAGCATTGGCGGACTCGGTGTCTGGAAACTGGCGCCTATGTTTAAGGGCGTTGG
GCTGGCTGCAGCTGTGCTGTCCTTCTGGCTGAACATCTATTACATTGTCATCATTAGCTGG
GCGATCTACTATCTGTACAATAGCTTCACTACGACTCTGCCGTGGAAACAATGTGACAATC
CTTGGAATACGGATCGCTGCTTTAGCAATTATAGCATGGTTAACACCACTAACATGACTTCC
GCTGTAGTGGAGTTCTGGGAGAGGAACATGCATCAGATGACCGACGGACTTGACAAACCC
GGACAGATTAGGTGGCCCCTCGCCATTACGCTCGCAATCGCCTGGATCCTTGTCTATTTTT
GTATCTGGAAAGGTGTCGGCTGGACTGGAAAGGTGGTCTATTTCAGCGCGACGTATCCAT
ACATTATGCTCATTATCCTGTTTTTCCGGGGAGTAACGCTCCCGGGTGCCAAAGAAGGGAT
CCTGTTCTACATTACGCCTAATTTTAGGAAGCTCTCTGATAGTGAGGTTTGGCTGGATGCT
GCCACGCAAATTTTTTTCTCCTACGGTCTGGGCCTCGGGTCCCTGATCGCCCTGGGCAGT
TATAATTCTTTTCATAACAATGTGTACAGAGATAGCATTATCGTGTGCTGTATCAATTCTTGC
ACATCCATGTTTGCCGGTTTCGTGATCTTCTCCATCGTCGGGTTTATGGCCCACGTCACCA
AACGTTCTATTGCGGATGTCGCTGCCTCAGGTCCCGGTCTGGCCTTTCTTGCGTACCCTGA
GGCTGTTACACAACTGCCAATCAGTCCTCTGTGGGCCATTCTTTTCTTTTCCATGCTTCTGA
TGCTGGGCATTGACTCTCAGTTTTGTACTGTTGAGGGTTTCATCACCGCCCTGGTTGACGA
ATATCCCCGTCTCTTGAGAAATCGTCGCGAACTCTTCATTGCCGCAGTGTGTATTATCTCCT
ACCTCATTGGCCTGTCCAATATCACGCAGGGAGGCATTTACGTATTCAAACTCTTTGATTAC
TATTCCGCAAGTGGAATGAGCCTCTTGTTTCTCGTGTTTTTCGAGTGTGTTAGCATCTCCTG
GTTCTACGGAGTAAACCGTTTTTATGATAATATCCAGGAAATGGTGGGGAGTCGTCCGTGT
ATCTGGTGGAAACTTTGTTGGTCATTCTTTACGCCCATCATTGTCGCTGGTGTGTTCATTTT
TTCTGCCGTGCAAATGACCCAACTGACGATGGGAAATTACGTATTCCCCAAATGGGGCCAG
GGCGTGGGGTGGCTCATGGCTCTGAGTTCCATGGTTTTGATCCCTGGGTATATGGCCTAC
ATGTTCCTGACACTGAAAGGGTCACTGAAACAGCGCATTCAAGTGATGGTTCAACCCAGCG
AGGACATTGTCAGACCGGAAAACGGGCCTGAGCAGCCCCAAGCGGGATCCTCTACAAGTA
AAGAGGCCTATATCACTCGCACGCGCCCCCTGGAGCAAAAGCTCATCTCTGAGGAAGACC
TGGCGGCTAACGATATTCTTGACTATAAAGACGATGACGATAAAGTGTAAGGCGCGCCGC
GGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGT
ATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCAT
GCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCC
ACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGG
CACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGT
TGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCC
TAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTG
GGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCA
CGTGCGGACCGAGCGGCCGC (SEQ ID NO: 90)

CN3322

GCGGCCGCACGCGTGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGG
CTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGG
CATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAA
CATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTG
AGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATT
ACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCT
AATTATGGCTGCATTTAAGAGAATGGGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTA
GGAGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTT

FIG. 14 cont'd

TAGCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGC
CTAGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGT
GAATTGAGCTGAGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAG
TGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACAT
GAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGGAAAGGAAAGAGCTTGCTTTCAA
CCTCAAAAGCTAGGAGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGC
CTGTTTCTTCCTTTAGCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGT
CACAATAGAGAGCCTAGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGC
CTTTGTGTGGGGTGAATTGAGCTGAGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTT
TTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATC
AGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGGAGCTCGGGC
TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCT
TTCGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCG
GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCAC
CACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGC
AGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCC
GCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATC
GACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCAC
AACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGC
CACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCAT
CGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGA
GCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCC
GGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCTTA
AAGAGACCGGTTCACTGTGACAGTAAAAGAGACCGGTTCACTGTGAGAATGAAAGAGACC
GGTTCACTGTGATCGGAAAAGAGACCGGTTCACTGTGAGCGGCCTTGAAACCCAGCAGAC
AATGTAGCTCAGTAGAAACCCAGCAGACAATGTAGCTGAATGGAAACCCAGCAGACAATGT
AGCTTCGGAGAAACCCAGCAGACAATGTAGCTGTCGACGAATTCGATATCATAATCAACCT
CTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTA
TGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTT
CTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTT
GCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGAT
CTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTT
GACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCAT
TGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA
GGATTGGGAAGACAATAGCAGGCATGCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 91)

CN3323

GCGGCCGCACGCGTGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTAGGAGGAAAGGG
CTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTTTAGCCTCAAGG
CATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGCCTAGTTTTGAA
CATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGTGAATTGAGCTG
AGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATT
ACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCT

FIG. 14 cont'd

AATTATGGCTGCATTTAAGAGAATGGGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAGCTA
GGAGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTCCTT
TAGCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAGAGC
CTAGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGGGGT
GAATTGAGCTGAGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTTTTAGTATTAAAAG
TGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACAT
GAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGGAAAGGAAAGAGCTTGCTTTCAA
CCTCAAAAGCTAGGAGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGC
CTGTTTCTTCCTTTAGCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGT
CACAATAGAGAGCCTAGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGC
CTTTGTGTGGGGTGAATTGAGCTGAGCGGCTAGCTGGTTGGAGCTAAATAAAGATGGCTTT
TTAGTATTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATC
AGCGGGCTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGGAGCTCGGGC
TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCT
TTCGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCG
GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCAC
CACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGC
AGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCC
GCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATC
GACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCAC
AACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGC
CACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCAT
CGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGA
GCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCC
GGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCGCG
AATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAA
CTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTG
CTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCG
GAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGA
CAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTG
CCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAA
AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTG
GGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGTGC
GGACCGAGCGGCCGC (SEQ ID NO: 92)

CN3887

GCGGCCGCACGCGTGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATCCTGCCCTTGA
ATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCTGTGATTCTTA
AAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGCCTTTTCCAGG
AGGAGTTCCTTAGAGGGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAATTA
CAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACTCTA
ATTATGGCTGCATTTAAGAGAATGGGCTCCTGGGTGACTGAGCAGTGGAGAAACAGAATC
CTGCCCTTGAATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAGGGATGCCCT

FIG. 14 cont'd

GTGATTCTTAAAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAGACACACAGC
CTTTTCCAGGAGGAGTTCCTTAGAGGGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGGCTCCTGGGTGACTGAGCAGTGGAGA
AACAGAATCCTGCCCTTGAATTGCTCCCCAGGCGGGCTTTATGCAGTCTGGGGAAGCAAG
GGATGCCCTGTGATTCTTAAAGAAACTGGTATAATTTTGCACTGCATAGCAGACTCCCAAG
ACACACAGCCTTTTCCAGGAGGAGTTCCTTAGAGGGGCTAAATAAAGATGGCTTTTTAGTA
TTAAAAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGG
CTACATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCGGGCTGGGCAT
AAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGGGATCCAGATCTTTCGAAG
CTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT
GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCG
AGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGCTGATCTGCACCACCGGC
AAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGGGCTACGGCGTGCAGTGCTT
CGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG
CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGA
GGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA
GGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTA
TATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACAT
CGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACG
GCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCAAGCTGAGCAAAGACC
CCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACT
CTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGCGCGCCGCGGCCGCGAATTCGATA
TCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGC
TCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTA
TGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATC
GCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGT
GGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCC
CGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAA
ATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGAC
AGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGAGATCTCACGTGCGGACCGAGC
GGCCGC (SEQ ID NO: 93)

CN3888

GCGGCCGCACGCGTGGTACCCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGGAAGAAAA
TTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAAAATTACT
CTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAAAAGTGG
AAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTACATGAAA
AATTACTCTAATTATGGCTGCATTTAAGAGAATGGCTAAATAAAGATGGCTTTTTAGTATTAA
AAGTGGAAGAAAATTACAGGTAATTATCTTTGACGGTAAAAACGCTGTAATCAGCGGGCTA
CATGAAAAATTACTCTAATTATGGCTGCATTTAAGAGAATGGAGCTCTCCAGACCTTGAACA
TAAAAGGCATTTTATACCCGCAATTGTTCATTTAATGAGCAATTGCGGAGTGCAGGCCGGT
TAAGGGATTGAGCTATATGCACTATTATTGCAAGAAGTATTCCGAAATACCAGAAATAGGAC
GTAAGCTCTGATCAGGGAGACTGCGAGCACAATTACCTTCTTTTCAAATCCTTCTGTGACA
CTGCGGGAGGAAAAAGGACTTTGAAACTTGAAAGGAAAGAGCTTGCTTTCAACCTCAAAAG

FIG. 14 cont'd

CTAGGAGGAAAGGGCTCTGAAATTTGCTCAGAATTCCCAATTCACCATTAGCCTGTTTCTTC
CTTTAGCCTCAAGGCATTCTCCGCTTTTTGAAAAGATGTTAAGAAATTCAGTCACAATAGAG
AGCCTAGTTTTGAACATGTTTCACTCGGTCCATTGAGGTCTGGGCTCCAGCCTTTGTGTGG
GGTGAATTGAGCTGAGCGGCTAGCTGGTTGGAGAGAGGTGAATGAGAAGTCGCTGTGCA
GTTGCAAATTCTGGCAAAGAAAAAAAAAAAGCTCACCCCTTCCTTTATTTTGTAATATGCATT
CCTGTACAATCCTGCCAGTGGCAATCTGTGGAGTTCAGTGTGTCCCTAAGTCAATATGGAG
TACTTGGTTTATAGCAACTCTTGTTAAGTTTGTCTTGTAATTGAAGCTGCTGTTGACCTTGCT
TGGGGGAGCTCGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCT
TCTGGGATCCAGATCTTTCGAAGCTAGCGCTACCGGTCGCCACCATGGTGAGCAAGGGCG
AGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC
CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
GAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCC
TGGGCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCT
TCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACG
GCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATC
GAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGC
CAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCT
ACCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGT
TCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGTCGACGGC
GCGCCGCGGCCGCGAATTCGATATCATAATCAACCTCTGGATTACAAAATTTGTGAAAGAT
TGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCT
TTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTA
GTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCG
GCTGTTGGGCACTGACAATTCCGTGGCTCGAGAGATCTTCGACTGTGCCTTCTAGTTGCCA
GCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACT
GTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCT
GGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCAT
GAGATCTCACGTGCGGACCGAGCGGCCGC (SEQ ID NO: 94)