

GENE THERAPY FOR HEARING DISORDERS

SEQUENCE LISTING

[001] The instant application contains a Sequence Listing which has been submitted electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on June 14, 2024, is named 67000-1321_WO_11223WO01_SL.xml and is 69,930 bytes in size.

CROSS REFERENCE TO RELATED APPLICATIONS

[002] The present application claims priority to US Provisional Application No. 63/521,118, filed 15-June-2023, which is hereby incorporated in its entirety.

FIELD

[003] The present disclosure relates to polynucleotides containing regions of the transmembrane serine protease 3 ("TMPRSS3") promoter, as well as vectors comprising the same, that can be used to promote expression of a transgene in a subset of cells within the inner ear.

BACKGROUND

[004] Hearing loss is the most common sensory deficit in humans. There are about 466 million people worldwide living with disabling hearing loss (432 million adults and 34 million children). The number of people with disabling hearing loss will grow to 630 million by 2030 and to over 900 million by 2050 (1 in 10 people). Over 90% of persons with disabling hearing loss (420 million) reside in the low-income regions of the world. There are currently no approved therapeutic agents for preventing or treating hearing loss or deafness.

SUMMARY

[005] The present disclosure provides polynucleotides that regulate transcription of a transgene in a subset of cells within the inner ear (e.g., cochlear hair cells, vestibular hair cells and sensory support cells). The polynucleotides described herein may be operably linked together and also to a therapeutic transgene, and may be administered to a patient to treat, prevent, or reduce the

likelihood of hearing loss (e.g., sensorineural hearing loss) and/or vestibular dysfunction (e.g., vertigo, dizziness, or imbalance).

[006] The present disclosure provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 1 or a functional portion, or derivative thereof, wherein said region is not identical to SEQ ID NO: 2. In one embodiment, the polynucleotide may comprise SEQ ID NO: 16 or a functional portion, or derivative thereof. In another embodiment, the polynucleotide may consist of SEQ NO: 1 or a functional portion, or derivative thereof.

[007] The present disclosure also provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 3 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 26 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 4. In one embodiment, the polynucleotide may consist of SEQ NO: 3 or a functional portion or derivative thereof, or SEQ ID NO: 26 or a functional portion or derivative thereof.

[008] The present disclosure further provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 5 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 6. In one embodiment, the polynucleotide may comprise SEQ ID NO: 17 or a functional portion, or derivative thereof. In another embodiment, the polynucleotide may consist of SEQ NO: 5 or a functional portion, or derivative thereof.

[009] The present disclosure also provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 8 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 27 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 9.

In one embodiment, the polynucleotide may consist of SEQ NO: 8 or a functional portion or derivative thereof, or SEQ ID NO: 27 or a functional portion or derivative thereof.

[0010] The present disclosure further provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 10 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 11. In one embodiment, the polynucleotide may consist of SEQ NO: 10 or a functional portion or derivative thereof.

[0011] The present disclosure also provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 12 or a functional portion or derivative thereof. In one embodiment, the polynucleotide may consist of SEQ NO: 12 or a functional portion or derivative thereof.

[0012] The present disclosure further provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 13 or a functional portion or derivative thereof. In one embodiment, the polynucleotide may consist of SEQ NO: 13 or a functional portion or derivative thereof.

[0013] The present disclosure also provides a polynucleotide comprising a region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 22 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 23. In one embodiment, the polynucleotide may consist of SEQ NO: 22 or a functional portion or derivative thereof.

[0014] The present disclosure also provides a polynucleotide comprising in a 5' to 3' direction: a) a first region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 1, or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 2 or a functional portion or derivative thereof; b) a second region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or

more, sequence identity) to SEQ ID NO: 3 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 26 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 4 or a functional portion, or derivative thereof; and c) a linker having a length from 0 to two thousand nucleotides between the first region and the second region, wherein the first region is operably linked to the second region and the polynucleotide is not identical to SEQ ID NO: 14. In one embodiment, the polynucleotide may comprise SEQ ID NO: 16 or a functional portion or derivative thereof. In another embodiment, the linker may have a length from 0 to 1000 nucleotides, 100 to 700 nucleotides, or 300 to 500 nucleotides. In another embodiment, the linker may have a length of 0, 50, 100, 250, 500, or 1000 nucleotides. In another embodiment, the polynucleotide comprises a third region in the 3' direction of the second region coding for a protein. In another embodiment, the protein may be TMPRSS3 or a reporter gene such as a green fluorescent protein. In yet another embodiment, the polynucleotide may have the sequence of SEQ ID NO: 18. In another embodiment, the protein is not TMPRSS3 or a part thereof.

[0015] The present disclosure further provides a polynucleotide comprising in a 5' to 3' direction: a) a first region having at least 85% sequence identity to SEQ ID NO: 1 or a functional portion or derivative thereof, at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 2 or a functional portion, or derivative thereof; b) a second region having at least 85% sequence identity to SEQ ID NO: 3 or a functional portion, or derivative thereof, at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 26 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 4 or a functional portion or derivative thereof; c) a third region coding for a protein; d) a fourth region having at least 85% sequence identity to SEQ ID NO: 5 or a functional portion or derivative thereof, at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 6 or a functional portion or derivative thereof; e) a first linker having a length from 0 to two thousand nucleotides between the first region and the second region; f) a second linker having a length from 0 to 100 nucleotides between the second region and the third region; g) a third linker

between having a length from 0 to 100 nucleotides between the third region and the fourth region, wherein the first region, the second region, third region and the fourth region are operably linked together and the polynucleotide is not identical to SEQ ID NO: 15. In one embodiment, the polynucleotide may comprise SEQ ID NO: 16 or a functional portion or derivative thereof. In another embodiment, the polynucleotide may comprise SEQ ID NO: 17 or a functional portion or derivative thereof. In another embodiment, the first linker may have a length from 0 to 1000 nucleotides, 100 to 700 nucleotides, or 300 to 500 nucleotides. In yet another embodiment, the first linker may have a length of 0, 50, 100, 250, 500, or 1000 nucleotides. In another embodiment, each of the second and third linkers may independently have a length of 0 to 75 nucleotides, 10 to 50 nucleotides, or 20 to 30 nucleotides. In another embodiment, each of the second and third linkers may independently have a length of 0, 10, 20, 30, 50, or 75 nucleotides. In yet another embodiment, each of the first linker, second linker and third linker may independently have a length of 0 nucleotides. In another embodiment, the protein may be TMPRSS3 or a reporter gene such as a green fluorescent protein. In yet another embodiment, the polynucleotide may have the sequence of SEQ ID NO: 19. In another embodiment, the protein is not TMPRSS3 or a part thereof.

[0016] The present disclosure also provides a polynucleotide comprising in a 5' to 3' direction: a) a first region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 1, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 2 or a functional portion, or derivative thereof; b) a second region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 7 or a functional portion, or derivative thereof; c) a third region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 8, at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 27 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 9 or a functional portion, or derivative thereof; d) a fourth region coding for a protein; e) a fifth region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 5, at least 85% sequence identity (e.g.,

85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 6 or a functional portion, or derivative thereof; f) a first linker having a length from 0 to two thousand nucleotides between the first region and the second region; g) a second linker having a length from 0 to 100 nucleotides between the second region and the third region, wherein the third linker is not identical to SEQ ID NO: 11; h) a third linker between having a length from 0 to 100 nucleotides between the third region and the fourth region; i) a fourth linker having a length from 0 to 100 nucleotides between the fourth region and the fifth region, wherein the first region, second region, third region, fourth region and fifth region are operably linked together. In one embodiment, the first linker may have a length from 0 to 1000 nucleotides, 100 to 700 nucleotides, or 300 to 500 nucleotides. In another embodiment, the first linker may have a length of 0, 50, 100, 250, 500, or 1000 nucleotides. In another embodiment, each of the second and third linkers may independently have a length of 0 to 75 nucleotides, 10 to 50 nucleotides, or 20 to 30 nucleotides. In another embodiment, each of the second and third linkers may independently have a length of 0, 10, 20, 30, 50, or 75 nucleotides. In yet another embodiment, each of the first linker, second linker, third linker and fourth linker may independently have a length of 0 nucleotides. In another embodiment, the protein may be TMPRSS3 or a reporter gene such as a green fluorescent protein. In yet another embodiment, the polynucleotide may have the sequence of SEQ ID NO: 21. In another embodiment, the protein is not TMPRSS3 or a part thereof.

[0017] The present disclosure further provides a polynucleotide comprising in a 5' to 3' direction: a) a first region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 1 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 2 or a functional portion or derivative thereof; b) a second region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 3 or a functional portion or derivative thereof, at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 26 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 4 or a functional portion or derivative thereof; c) a third region coding for a protein

or a part thereof; d) a fourth region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 22 or a functional portion or derivative thereof, or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 23 or a functional portion or derivative thereof; e) a first linker having a length from 0 to two thousand nucleotides between the first region and the second region; f) a second linker having a length from 0 to 100 nucleotides between the second region and the third region; g) a third linker between having a length from 0 to 100 nucleotides between the third region and the fourth region, wherein the first region, the second region, third region and the fourth region are operably linked together and the polynucleotide is not identical to SEQ ID NO: 14 or SEQ ID NO: 15. In one embodiment, the third region of the polynucleotide codes for a part of the protein and the polynucleotide further comprises a fifth region in the 3' direction of the fourth region coding for another part of the same protein coded by the third region. In another embodiment, the polynucleotide comprises a fourth linker having a length from 0 to 100 nucleotides between the fourth region and the fifth region. In one embodiment, the first linker may have a length from 0 to 1000 nucleotides, 100 to 700 nucleotides, or 300 to 500 nucleotides. In another embodiment, each of the second, third and fourth linkers may independently have a length of 0 to 75 nucleotides, 10 to 50 nucleotides, or 20 to 30 nucleotides. In another embodiment, each of the first, second, third, fourth may have a length of 0 nucleotides. In another embodiment, the protein may be TMPRSS3 or a reporter gene such as a green fluorescent protein. In another embodiment, the protein is not TMPRSS3 or a part thereof.

[0018] The present disclosure further provides a polynucleotide comprising in a 5' to 3' direction: a) at least one first region having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to a sequence, a functional portion or derivative thereof, wherein the region is selected from the group consisting of: SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 26, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 27, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 22 and SEQ ID NO: 23; and b) a second region coding for a protein or a part thereof, wherein the first region and the second region are operably linked together and the polynucleotide is not identical to SEQ ID NO: 14 or SEQ ID

NO: 15. In one embodiment, the second region codes for a part of a protein and the polynucleotide further comprises a third region in the 5' direction of the first region and which codes for another part of the same protein coded by the second region. In another embodiment, the first region the polynucleotide comprises a sequence having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 22 or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 23 and the second region codes for a part of a protein and the polynucleotide further comprises a third region in the 5' direction of the sequence having at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 22 or at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to SEQ ID NO: 23 and which codes for another part of the same protein coded by the second region. In another embodiment, the protein may be TMPRSS3 or a reporter gene such as a green fluorescent protein. In another embodiment, the protein is not TMPRSS3 or a part thereof.

[0019] In one embodiment, the polynucleotides of the present disclosure have sequences that are not naturally occurring.

[0020] In one embodiment, the polynucleotides of the present disclosure are operably linked to a regulatory nucleotide sequence (e.g., promoter or enhancer sequence). In another embodiment the promoter sequence is a constitutive promoter. Examples of constitutive promoters include, but are not limited to, cytomegalovirus immediate early promoter (CMV), simian virus (SV40) promoter, adenovirus major late (MLP) promoter, Rous sarcoma virus (RSV) promoter, elongation factor- α (EF-1 α) promoter, ubiquitin promoters, actin promoters, tubulin promoters, immunoglobulin promoters, functional fragments thereof, or combinations thereof.

[0021] The present disclosure also provides a nucleic acid vector comprising any one the polynucleotides of the present disclosure. In one embodiment, the polynucleotide is operably linked to a transgene coding for a protein. In another embodiment, the polynucleotide directs cell-specific expression of the protein from the nucleic acid sequence in a mammalian inner ear cell type. Inner ear cell types include: hair cells, border cells, inner phalangeal cells, inner pillar

cells, outer pillar cells, first row Deiter cells, second row Deiter cells, third row Deiter cells, Hensen's cells, Claudius cells, inner sulcus cells, outer sulcus cells, spiral prominence cells, root cells, interdental cells, basal cells of the stria vascularis, intermediate cells of the stria vascularis, marginal cells of the stria vascularis, spiral ganglion neurons, Schwann cells. For example, hair cell-specific expression refers to production of an RNA transcript or polypeptide primarily within hair cells (e.g., cochlear hair cells and/or vestibular hair cells) as compared to other cell types of the inner ear (e.g., spiral ganglion neurons, glia, or other inner ear cell types). Hair cell-specific expression of a transgene can be confirmed by comparing transgene expression (e.g., RNA or protein expression) between various cell types of the inner ear (e.g., hair cells vs. non-hair cells) using any standard technique (e.g., quantitative RT PCR, immunohistochemistry, western blot analysis, or measurement of the fluorescence of a reporter (e.g., GFP) operably linked to a promoter). A hair cell-specific promoter induces expression (e.g., RNA or protein expression) of a transgene to which it is operably linked that is at least 50% greater (e.g., 50%, 75%, 100%, 125%, 150%, 175%, 200% greater or more) in hair cells compared to at least 3 (e.g., 3, 4, 5, 6, 7, 8, 9, 10, or more) other inner ear cell types.

[0022] In yet another embodiment, the protein encoded by a transgene (protein/transgene) may be Actin, gamma 1/ACTG1, ADP-ribosyltransferase 1/ART1, Activating transcription factor 6 /ATF6, Atonal bHLH transcription factor 1/ATOH1, ATPase plasma membrane Ca²⁺ transporting 2/ATP2B2, Brain-derived neurotrophic factor/BDNF, Binding immunoglobulin protein/BIP, Calcium binding protein 2/CABP2, Calcium channel, voltage-dependent, alpha 2/delta subunit 4/CACNA2D4, Coiled-Coil Domain Containing 92/CCER2, Cadherin-related 23/CDH23, Cholinergic receptor nicotinic alpha 9 subunit/CHRNA9, Calcium and integrin binding family member 2/CIB2, Calcium and integrin binding family member 3/CIB3, Clarin 1/CLRN1, Clarin 2/CLRN2, Pejvakin/DFNB59, Dystonin/DYTN, ELFN1 adhesion G protein-coupled receptor/ELFN1, Epidermal growth factor receptor kinase substrate 8/EPS8, EPS8-like 2/EPS8L2, Espin/ESPN, Espin-like/ESPNL, Fc receptor-like B/FCRLB, Fascin actin-bundling protein 2/FSCN2, Growth factor independent 1 transcriptional repressor/GFI1, Glutaredoxin domain-containing cysteine-rich protein 1/GRXCR1, Glutaredoxin domain-containing cysteine-rich protein 2/GRXCR2, Heat shock protein 70/HSP70, Heat shock protein 90/HSP90, Inositol-requiring enzyme 1/IRE1, Potassium voltage-gated channel subfamily A member 10/KCNA10,

Kielin/chordin-like protein/KCP, Kinocilin/KNCN, Late cornified envelope protein 6A/LCE6A, LIM homeobox 3/LHX3, Lipoygenase homology domains 1/LOXHD1, Leucine-rich repeats and transmembrane domains 2/LRTM2, Leucine-rich transmembrane and O-methyltransferase domain containing 1/LRTOMT1, Leucine-rich transmembrane and O-methyltransferase domain containing 2/LRTOMT2, Makorin ring finger protein 20/MKRN20S, Myosin XV/MY015, Myosin IIIA/MY03A, Myosin IIIB/MY03B, Myosin VI/MY06, Myosin VIIA/MY07A, Nescient helix-loop-helix 1/NHLH1, Neurotrophin 3/NTF3, Otoferlin/OTOF, Protocadherin-related 15/PCDH15, Protein kinase R (PKR)-like endoplasmic reticulum kinase/PERK, POU class 4 homeobox 3/POU4F3, Peripherin 2/PRPH2, Protein tyrosine phosphatase receptor type Q/PTPRQ, Radixin/RDX, Serpin family E member 3/SERPINE3, SKI family transcriptional corepressor 1/SKOR1, Solute carrier family 17, member 8 (vesicular glutamate transporter 3)/SLC17A8, Solute carrier family 8, member 2 (sodium/calcium exchanger 2)/SLC8A2, Stereocilin/STRC, Tctex1 domain containing 1/TCTEX1D1, Transmembrane channel-like 1/TMC1, Transmembrane protease, serine 3/TMPRSS3, Taperin/TPRN, TRIO and F-actin binding protein/TRIOBP, Tetratricopeptide repeat domain 24/TTC24, Usher syndrome 1C (harmonin)/USH1 C, Solute carrier family 26, member 4 (pendrin)/SLC26A4, Xin actin-binding repeat containing 2/XIRP2, or Zinc finger CCHC-type containing 12/ZCCHC12.

[0023] In one embodiment, the nucleic acid vector of the present disclosure may be a plasmid, cosmid, artificial chromosome, or viral vector. In another embodiment, the nucleic acid vector may be a viral vector selected from the group consisting of an adeno-associated virus (AAV), an adenovirus, and a lentivirus. In a preferred embodiment, the viral vector may be an AAV vector including, for example, an AAV vector having a serotype of AAV1, AAV2, AAV2quad(Y-F), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, rh10, rh39, rh43, rh74, Anc80, Anc80L65, DJ/8, DJ/9, 7m8, PHP.B, PHP.eb, or PHP.S.

[0024] The present disclosure further provides a composition comprising a nucleic acid vector of the present disclosure. In one embodiment, the composition may further comprise a pharmaceutically acceptable excipient.

[0025] The present disclosure also provides a method of increasing expression of a protein in a mammalian inner ear cell, comprising contacting the mammalian inner ear cell with a nucleic acid vector or a composition according to the present disclosure. In one embodiment, the expression of the protein may be increased specifically in cochlear hair cells, inner hair cells, outer hair cells, vestibular hair cells, sensory support cells and interdental cells, or a combination thereof. In another embodiment, the expression of the protein may not be substantially increased in inner ear cells that are not hair cells.

[0026] The present disclosure further provides a method of treating a subject in need thereof having or at risk of developing a condition, the method comprising administering to the subject a therapeutically effective amount of the nucleic acid vector or the composition according to the present disclosure. In one embodiment, the condition may be hearing loss. In another embodiment, the hearing loss may be a genetic hearing loss such as for example autosomal dominant hearing loss, autosomal recessive hearing loss, or X-linked hearing loss. In another embodiment, the hearing loss may be acquired hearing loss such for example, noise-induced hearing loss, age- related hearing loss, disease or infection-related hearing loss, head trauma-related hearing loss, or ototoxic drug-induced hearing loss. In yet another embodiment, the condition may be vestibular dysfunction such as for example vertigo, dizziness, or imbalance. In another embodiment, the condition may be tinnitus.

[0027] The present disclosure also provides a method of promoting hair cell regeneration in a subject in need thereof, comprising administering to the subject a therapeutically effective amount of a nucleic acid vector or a composition according to the present disclosure. In one embodiment, the hair cell may be a cochlear hair cell or a vestibular hair cell.

[0028] The present disclosure further provides a method of preventing, reducing the likelihood of developing, or reducing ototoxic drug-induced hair cell damage or death, comprising administering to the subject a therapeutically effective amount of a nucleic acid vector or a composition according to the present disclosure. In one embodiment, the ototoxic drug may be an aminoglycoside, an antineoplastic drug, an ethacrynic acid, a furosemide, a salicylate, or a quinine.

[0029] The present disclosure also provides a method of preventing, reducing the likelihood of developing, or reducing hair cell damage or death in a subject in need thereof, comprising administering to the subject a therapeutically effective amount of a nucleic acid vector or a composition according to the present disclosure.

[0030] The present disclosure further provides a method of increasing hair cell survival in a subject in need thereof, comprising administering to the subject a therapeutically effective amount of a nucleic acid vector or a composition according to the present disclosure. In one embodiment, the nucleic acid vector or composition may be locally administered.

[0031] The present disclosure also provides a kit comprising a nucleic acid vector or a composition according to the present disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

[0032] **Figure 1:** Figure 1 shows a diagram of mouse cochlea with gray nuclei as indications of the expression of TMPRSS3 mRNA in various cell types including in the spiral prominence, outer sulcus cells, supporting cells, outer hair cells, pillar cells, inner hair cells, inner sulcus cells, and interdental cells.

[0033] **Figure 2:** Figure 2 shows a map of polynucleotide REGN-huAB (SEQ ID NO: 18) showing the green fluorescent protein (GFP) reporter gene under the control of regulatory elements from the human TMPRSS3 gene Region A (SEQ ID NO: 1) and modified Region B (SEQ ID NO: 3).

[0034] **Figure 3:** Figure 3 shows a map of polynucleotide REGN-huABC (SEQ ID NO: 19) showing the green fluorescent protein (GFP) reporter gene under the control of regulatory elements from the human TMPRSS3 gene, modified Region A (SEQ ID NO: 1), modified Region B (SEQ ID NO: 3), and modified Region C (SEQ ID NO: 5).

[0035] Figure 4: Figure 4 shows a map of polynucleotide REGN-huABE showing the green fluorescent protein (GFP) reporter gene under the control of regulatory elements from the human TMPRSS3 gene, modified Region A (SEQ ID NO: 1), modified Region B (SEQ ID NO: 3), and modified Region E (SEQ ID NO: 22).

[0036] Figure 5: Figure 5 shows a map of polynucleotide REGN-huAB2B3 (SEQ ID NO: 20) showing the green fluorescent protein (GFP) reporter gene under the control of regulatory elements from the human TMPRSS3 gene, modified Region A (SEQ ID NO: 1), Region B2 (SEQ ID NO: 7), and modified Region B3 (SEQ ID NO: 8).

[0037] Figure 6: Figure 6 shows a map of polynucleotide REGN –huAB2B3C (SEQ ID NO: 21) showing the green fluorescent protein (GFP) reporter gene under the control of regulatory elements from the human TMPRSS3 gene, modified Region A (SEQ ID NO: 1), Region B2 (SEQ ID NO: 7), modified Region B3 (SEQ ID NO: 8), and modified Region C (SEQ ID NO: 5).

[0038] Figure 7: Figure 7 shows an ex-vivo qualitative summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0039] Figure 8: Figure 8 shows an ex-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0040] Figure 9: Figure 9 shows an in-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0041] Figure 10: Figure 10 shows an ex-vivo qualitative summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huABC and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0042] Figure 11: Figure 11 shows an ex-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huABC and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0043] Figure 12: Figure 12 shows an in-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huABC and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0044] Figure 13: Figure 13 shows an ex-vivo qualitative summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3

promoter elements described in REGN-huABE and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0045] Figure 14: Figure 14 shows an ex-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huABE and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0046] Figure 15: Figure 15 shows an in-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huABE and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0047] Figure 16: Figure 16 shows an ex-vivo qualitative summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB2B3 and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0048] Figure 17: Figure 17 shows an ex-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB2B3 and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0049] Figure 18: Figure 18 shows an in-vivo quantification summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB2B3 and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0050] Figure 19: Figure 19 shows an ex-vivo qualitative summary diagram of mouse cochlea with gray nuclei as indications of the expression of GFP in mouse cochlea transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB2B3C and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0051] Figure 20: Figure 20 shows the levels of ex-vivo GFP expression in hair cells transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB, REGN-huABC, REGN-huABE and REGN-huAB2B3 and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0052] Figure 21: Figure 21 shows the levels of ex-vivo GFP expression in sensory supporting cells transduced with an adeno-associated virus (AAV) vector expressing GFP under the control of the human TMPRSS3 promoter elements described in REGN-huAB, REGN-huABC, REGN-huABE and REGN-huAB2B3 and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

[0053] Figure 22: Figure 22 shows the levels of in-vivo GFP expression in various types of ear cells transduced with an adeno-associated virus (AAV) vector expressing GFP under the control

of the human TMPRSS3 promoter elements described in REGN-huAB, REGN-huABC, REGN-huABE and REGN-huAB2B3 and which are driving H2B-GFP. An AAV vector expressing GFP (CMV-H2B-GFP) under the control of the cytomegalovirus (CMV) promoter was used as a positive control.

DETAILED DESCRIPTION

[0054] Sensorineural hearing loss is a type of hearing loss caused by defects in the cells of the inner ear, such as cochlear hair cells, or the neural pathways that project from the inner ear to the brain. This type of hearing loss is often acquired, and has a variety of causes, including, for example, acoustic trauma, disease or infection, head trauma, ototoxic drugs, and aging. There are also genetic causes of sensorineural hearing loss, such as mutations in genes involved in the development and function of the inner ear. Mutations in over 90 such genes have been identified, including mutations inherited in an autosomal recessive, autosomal dominant, and X-linked pattern.

[0055] Patients carrying mutations that disrupt hair cell development, function, or survival can present with both hearing loss and vestibular dysfunction, or either disorder alone. There are currently no approved therapeutic agents for preventing or treating hearing loss or deafness and hence there is a need for new therapeutics to target hair cells for the treatment of sensorineural hearing loss or vestibular dysfunction.

[0056] Described herein are compositions and methods for inducing transgene expression in subpopulations of inner ear cells (e.g., cochlear hair cells, vestibular hair cells, interdental cells, and/or sensory support cells of the cochlea). The present disclosure provides polynucleotides containing regions of the TMPRSS3 promoter that are capable of expressing a transgene in subpopulations of inner ear cells. The present disclosure also provides nucleic acid vectors containing these promoters operably linked to polynucleotides encoding polypeptides. The compositions and methods described herein can be used to express polynucleotides encoding inner ear proteins in subpopulations of inner ear cells, and, therefore, the compositions described

herein can be administered to a subject (such as a mammalian subject, for instance, a human) to treat disorders caused by dysfunction of inner ear cells, such as hearing loss or vestibular dysfunction.

[0057] It is to be understood that the foregoing general description and the following detailed description are exemplary and explanatory only, and are not restrictive of any subject matter claimed.

[0058] Headings are used solely for organizational purposes, and are not intended to limit the disclosure in any way.

[0059] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the disclosures belong. All patents, patent applications, published applications and publications, websites and other published materials referred to throughout the entire disclosure herein, unless noted otherwise, are incorporated by reference in their entirety for any purpose. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, the preferred methods are described.

Definitions

[0060] The term “or” refers to any one member of a particular list and also includes any combination of members of that list.

[0061] The singular forms of the articles “a,” “an” and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a protein” or “at least one protein” can include a plurality of proteins, including mixtures thereof.

[0062] Unless otherwise indicated, statistically significant means $p < 0.05$.

[0063] As used herein, the term “about” is intended to also include the exact amount. Hence “about 5 percent” means “about 5 percent” and also “5 percent.” “About” means within typical

experimental error for the application or purpose intended. Unless otherwise apparent from the context, the term “about” encompasses values within a standard margin of error of measurement (e.g., SEM) of a stated value.

[0064] As used herein, the term “administration” and the like refers to and includes the administration of a composition to a subject or system (e.g., to a cell, organ, tissue, organism, or relevant component or set of components thereof). The skilled artisan will appreciate that route of administration may vary depending, for example, on the subject or system to which the composition is being administered, the nature of the composition, the purpose of the administration, etc. For example, in certain embodiments, administration to an animal subject (e.g., to a human or a rodent) may be intracochlear, bronchial (including by bronchial instillation), buccal, enteral, interdermal, intraarterial, intradermal, intragastric, intramedullary, intramuscular, intranasal, intraperitoneal, intrathecal, intravenous, intraventricular, mucosal, intranasal, oral, rectal, subcutaneous, sublingual, topical, tracheal (including by intratracheal instillation), transdermal, vaginal and/or vitreal. In some embodiments, administration may involve intermittent dosing. In some embodiments, administration may involve continuous dosing (e.g., perfusion) for at least a selected period of time. In one embodiment, the term “administration” may refer to providing or giving a subject a therapeutic agent (e.g., a nucleic acid vector containing a TMPRSS3 promoter operably linked to a transgene), by any effective route. Exemplary routes of administration are described herein above.

[0065] As used herein, the term “cell type” refers to a group of cells sharing a phenotype that is statistically separable based on gene expression data. For instance, cells of a common cell type may share similar structural and/or functional characteristics, such as similar gene activation patterns and antigen presentation profiles. Cells of a common cell type may include those that are isolated from a common tissue (e.g., epithelial tissue, neural tissue, connective tissue, or muscle tissue) and/or those that are isolated from a common organ, tissue system, blood vessel, or other structure and/or region in an organism.

[0066] As used herein, the term “protein” also includes an entire protein, a part of a protein including a functional or non-functional part thereof.

[0067] As used herein, the terms “conservative mutation,” “conservative substitution,” and “conservative amino acid substitution” refer to a substitution of one or more amino acids for one or more different amino acids that exhibit similar physicochemical properties, such as polarity, electrostatic charge, and steric volume. The abbreviations for each of the twenty naturally-occurring amino acids are listed in table 1 below.

Table 1. Abbreviations of naturally-occurring amino acids.

Amino Acid	Abbreviations
Alanine	Ala; A
Arginine	Arg; R
Asparagine	Asn; N
Aspartic acid	Asp; D
Cysteine	Cys; C
Glutamic acid	Glu; E
Glutamine	Gln; Q
Glycine	Gly; G
Histidine	His; H
Isoleucine	Ile; I
Leucine	Leu; L
Lysine	Lys; K
Methionine	Met; M
Phenylalanine	Phe; F
Proline	Pro; P
Serine	Ser; S
Threonine	Thr; T
Tyrosine	Tyr; Y
Tryptophan	Trp; W
Valine	Val; V

[0068] As used herein, the term “effective” applied to dose or amount refers to that quantity of a compound or pharmaceutical composition that is sufficient to result in a desired activity upon administration to a subject in need thereof. Note that when a combination of active ingredients is administered, the effective amount of the combination may or may not include amounts of each ingredient that would have been effective if administered individually. The exact amount required will vary from subject to subject, depending on the species, age, and general condition of the subject, the severity of the condition being treated, the particular drug or drugs employed, the mode of administration, and the like.

[0069] As used herein, the term “endogenous” describes a molecule (e.g., a polypeptide, nucleic acid, or cofactor) that is found naturally in a particular organism (e.g., a human) or in a particular location within an organism (e.g., an organ, a tissue, or a cell, such as a human cell, e.g., a human hair cell).

[0070] As used herein, the term “express” refers to one or more of the following events: (1) production of an RNA template from a DNA sequence (e.g., by transcription); (2) processing of an RNA transcript (e.g., by splicing, editing, 5' cap formation, and/or 3' end processing); (3) translation of an RNA into a polypeptide or protein; and (4) post-translational modification of a polypeptide or protein.

[0071] As used herein, the term “exogenous” describes a molecule (e.g., a polypeptide, nucleic acid, or cofactor) that is not found naturally in a particular organism (e.g., a human) or in a particular location within an organism (e.g., an organ, a tissue, or a cell, such as a human cell, e.g., a human hair cell). Exogenous materials include those that are provided from an external source to an organism or to cultured matter extracted there from.

[0072] As used herein, the terms “increasing” and “decreasing” refer to modulating resulting in, respectively, greater or lesser amounts, of function, expression, or activity of a metric relative to a reference. For example, subsequent to administration of a composition in a method described herein, the amount of a marker of a metric (e.g., transgene expression) as described herein may be increased or decreased in a subject by at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 98% or more relative to the amount of the marker prior to administration. Generally, the metric is measured subsequent to administration at a time that the administration has had the recited effect, e.g., at least one week, one month, 3 months, or 6 months, after a treatment regimen has begun.

[0073] As used herein, the term “intron” refers to a region within the coding region of a gene, the nucleotide sequence of which is not translated into the amino acid sequence of the corresponding protein. The term intron also refers to the corresponding region of the RNA transcribed from a

gene. Introns are transcribed into pre-mRNA, but are removed during processing, and are not included in the mature mRNA.

[0074] As used herein, the term “linker” refers to one or more nucleotides that connect two different regions of a polynucleotide. A linker does not disrupt the function of the two regions of the polynucleotide that it connects. A linker with zero nucleotides is an indication of the absence of a linker.

[0075] As used herein, “locally” or “local administration” means administration at a particular site of the body intended for a local effect and not a systemic effect. Examples of local administration are epicutaneous, inhalational, intra-articular, intrathecal, intravaginal, intravitreal, intrauterine, intra-lesional administration, lymph node administration, intratumoral administration, administration to the inner ear, and administration to a mucous membrane of the subject, wherein the administration is intended to have a local and not a systemic effect.

[0076] As used herein, the term “operably linked” refers to a first molecule that can be joined to a second molecule, wherein the molecules are so arranged that the first molecule affects the function of the second molecule. The term “operably linked” includes the juxtaposition of two or more components (e.g., a promoter and another sequence element) such that both components function normally and allow for the possibility that at least one of the components can mediate a function that is exerted upon at least one of the other components. The two molecules may or may not be part of a single contiguous molecule and may or may not be adjacent. For example, a promoter is operably linked to a transcribable polynucleotide molecule if the promoter modulates transcription of the transcribable polynucleotide molecule of interest in a cell. In additional embodiments, two portions of a transcription regulatory element are operably linked to one another if they are joined such that the transcription-activating functionality of one portion is not adversely affected by the presence of the other portion. Two transcription regulatory elements may be operably linked to one another by way of a linker nucleic acid (e.g., an intervening non-coding nucleic acid) or may be operably linked to one another with no intervening nucleotides present.

[0077] As used herein, the term “plasmid” refers to a to an extrachromosomal circular double stranded DNA molecule into which additional DNA segments may be ligated. A plasmid is a type of vector, a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. Certain plasmids are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial plasmids having a bacterial origin of replication and episomal mammalian plasmids). Other vectors (e.g., non-episomal mammalian vectors) can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Certain plasmids are capable of directing the expression of genes to which they are operably linked.

[0078] As used herein, the terms “nucleic acid” and “polynucleotide,” used interchangeably herein, refer to a polymeric form of nucleosides in any length. Typically, a polynucleotide is composed of nucleosides that are naturally found in DNA or RNA (e.g., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine) joined by phosphodiester bonds. The term encompasses molecules comprising nucleosides or nucleoside analogs containing chemically or biologically modified bases, modified backbones, etc., whether or not found in naturally occurring nucleic acids, and such molecules may be preferred for certain applications. Where this application refers to a polynucleotide it is understood that both DNA, RNA, and in each case both single- and double-stranded forms (and complements of each single-stranded molecule) are provided. “Polynucleotide sequence” as used herein can refer to the polynucleotide material itself and/or to the sequence information (i.e., the succession of letters used as abbreviations for bases) that biochemically characterizes a specific nucleic acid. A polynucleotide sequence presented herein is presented in a 5' to 3' direction unless otherwise indicated.

[0079] As used herein, the terms “complementarity” or “complementary” of nucleic acids means that a nucleotide sequence in one strand of nucleic acid, due to orientation of its nucleobase groups, forms hydrogen bonds with another sequence on an opposing nucleic acid strand. The complementary bases in DNA are typically A with T and C with G. In RNA, they are typically C with G and U with A. Complementarity can be perfect or substantial/sufficient. Perfect complementarity between two nucleic acids means that the two nucleic acids can form a duplex

in which every base in the duplex is bonded to a complementary base by Watson-Crick pairing. “Substantial” or “sufficient” complementary means that a sequence in one strand is not completely and/or perfectly complementary to a sequence in an opposing strand, but that sufficient bonding occurs between bases on the two strands to form a stable hybrid complex in set of hybridization conditions (e.g., salt concentration and temperature). Such conditions can be predicted by using the sequences and standard mathematical calculations to predict the T_m (melting temperature) of hybridized strands, or by empirical determination of T_m by using routine methods. T_m includes the temperature at which a population of hybridization complexes formed between two nucleic acid strands are 50% denatured (i.e., a population of double-stranded nucleic acid molecules becomes half dissociated into single strands). At a temperature below the T_m , formation of a hybridization complex is favored, whereas at a temperature above the T_m , melting or separation of the strands in the hybridization complex is favored. T_m may be estimated for a nucleic acid having a known G+C content in an aqueous 1 M NaCl solution by using, e.g., $T_m = 81.5 + 0.41 (\% \text{ G+C})$, although other known T_m computations take into account nucleic acid structural characteristics.

[0080] As used herein, the term “promoter” refers to a regulatory region of DNA usually comprising a TATA box capable of directing RNA polymerase II to initiate RNA synthesis at the appropriate transcription initiation site for a particular polynucleotide sequence. A promoter may additionally comprise other regions which influence the transcription initiation rate. The promoter sequences disclosed herein modulate transcription of an operably linked polynucleotide. A promoter can be active in one or more of the cell types disclosed herein (e.g., a eukaryotic cell, a non-human mammalian cell, a human cell, a rodent cell, a pluripotent cell, a one-cell stage embryo, a differentiated cell, or a combination thereof). A promoter can be, for example, a constitutively active promoter, a conditional promoter, an inducible promoter, a temporally restricted promoter (e.g., a developmentally regulated promoter), or a spatially restricted promoter (e.g., a cell-specific or tissue-specific promoter). Examples of promoters can be found, for example, in WO 2013/176772, herein incorporated by reference in its entirety for all purposes. A “promoter/regulatory sequence” means a nucleic acid sequence which is required for expression of a gene product operably linked to the promoter/regulatory sequence.

This sequence may be the core promoter sequence, or it may also include an enhancer sequence and other regulatory elements which are required for expression of the gene product.

[0081] “Percent (%) sequence identity” with respect to a reference polynucleotide or polypeptide sequence is defined as the percentage of nucleic acids or amino acids in a candidate sequence that are identical to the nucleic acids or amino acids in the reference polynucleotide or polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid or amino acid sequence identity can be achieved in various ways that are within the capabilities of one of skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, or Megalign software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For example, percent sequence identity values may be generated using the sequence comparison computer program BLAST. As an illustration, the percent sequence identity of a given nucleic acid or amino acid sequence, A, to, with, or against a given nucleic acid or amino acid sequence, B, (which can alternatively be phrased as a given nucleic acid or amino acid sequence, A that has a certain percent sequence identity to, with, or against a given nucleic acid or amino acid sequence, B) is calculated as follows: 100 multiplied by (the fraction X/Y), where X is the number of nucleotides or amino acids scored as identical matches by a sequence alignment program (e.g., BLAST) in that program’s alignment of A and B, and where Y is the total number of nucleic acids in B. It will be appreciated that where the length of nucleic acid or amino acid sequence A is not equal to the length of nucleic acid or amino acid sequence B, the percent sequence identity of A to B will not equal the percent sequence identity of B to A.

[0082] The term “derivative” as used herein refers to a nucleic acid, peptide, or protein or a variant or analog thereof comprising one or more mutations and/or chemical modifications as compared to a corresponding full-length wild-type nucleic acid, peptide, or protein. Non-limiting examples of chemical modifications involving nucleic acids include, for example, modifications to the base moiety, sugar moiety, phosphate moiety, phosphate-sugar backbone, or a combination thereof. In one embodiment, a derivative nucleic acid, peptide, or protein or a variant molecule

thereof has 1% to 100%, 5% to 90%, 10% to 80%, 10% to 50%, or 20 to 30% of its nucleotides or amino acids mutated or chemically modified.

[0083] As used herein, the term “pharmaceutical composition” refers to a mixture containing a therapeutic agent, optionally in combination with one or more pharmaceutically acceptable excipients, diluents, and/or carriers, to be administered to a subject, such as a mammal, e.g., a human, in order to prevent, treat or control a particular disease or condition affecting or that may affect the subject.

[0084] As used herein, the term “pharmaceutically acceptable” refers to those compounds, materials, compositions and/or dosage forms, which are suitable for contact with the tissues of a subject, such as a mammal (e.g., a human) without excessive toxicity, irritation, allergic response and other problem complications commensurate with a reasonable benefit/risk ratio. Preferably, the term “pharmaceutically acceptable” means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in mammals, and more particularly in humans.

[0085] The term “reporter gene” refers to a nucleic acid having a sequence encoding a gene product (typically an enzyme) that is easily and quantifiably assayed when a construct comprising the reporter gene sequence operably linked to an endogenous or heterologous promoter and/or enhancer element is introduced into cells containing (or which can be made to contain) the factors necessary for the activation of the promoter and/or enhancer elements. Examples of reporter genes include, but are not limited, to genes encoding beta-galactosidase (lacZ), the bacterial chloramphenicol acetyltransferase (cat) genes, firefly luciferase genes, genes encoding beta-glucuronidase (GUS), and genes encoding fluorescent proteins. A “reporter protein” refers to a protein encoded by a reporter gene.

[0086] The term “fluorescent reporter protein” as used herein means a reporter protein that is detectable based on fluorescence wherein the fluorescence may be either from the reporter protein directly, activity of the reporter protein on a fluorogenic substrate, or a protein with affinity for binding to a fluorescent tagged compound. Examples of fluorescent proteins include

green fluorescent proteins (e.g., GFP, GFP-2, tagGFP, turboGFP, eGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, and ZsGreen1), yellow fluorescent proteins (e.g., YFP, eYFP, Citrine, Venus, YPet, PhiYFP, and ZsYellow1), blue fluorescent proteins (e.g., BFP, eBFP, eBFP2, Azurite, mKalamal, GFPuv, Sapphire, and T-sapphire), cyan fluorescent proteins (e.g., CFP, eCFP, Cerulean, CyPet, AmCyan1, and Midoriishi-Cyan), red fluorescent proteins (e.g., RFP, mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, and Jred), orange fluorescent proteins (e.g., mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, and tdTomato), and any other suitable fluorescent protein whose presence in cells can be detected by flow cytometry methods.

[0087] As used herein, the term “sample” refers to a specimen (e.g., blood, blood component (e.g., serum or plasma), urine, saliva, amniotic fluid, cerebrospinal fluid, tissue (e.g., placental or dermal), pancreatic fluid, chorionic villus sample, and cells) isolated from a subject.

[0088] As used herein, the term “transcription regulatory element” refers to a nucleic acid that controls, at least in part, the transcription of a gene of interest. Transcription regulatory elements may include promoters, enhancers, and other nucleic acids (e.g., polyadenylation signals) that control or help to control gene transcription. Examples of transcription regulatory elements are described, for example, in Lorence, *Recombinant Gene Expression: Reviews and Protocols* (Humana Press, New York, NY, 2012).

[0089] As used herein, the term “transfection” refers to any of a wide variety of techniques commonly used for the introduction of exogenous DNA into a prokaryotic or eukaryotic host cell, e.g., electroporation, lipofection, calcium phosphate precipitation, DEAE-dextran transfection, Nucleofection, squeeze-poration, sonoporation, optical transfection, magnetofection, impalefection and the like.

[0090] As used herein, the term “subject” means a living organism. Preferably, a subject is a mammal, such as a human, non-human primate, rodent, a dog, cat, cow, pig, etc. A subject to be treated according to the methods described herein may be one who has been diagnosed with

hearing loss (e.g., sensorineural hearing loss) or vestibular dysfunction (e.g., dizziness, vertigo, or imbalance) or one at risk of developing these conditions.

[0091] Diagnosis may be performed by any method or technique known in the art. One skilled in the art will understand that a subject to be treated according to the present disclosure may have been subjected to standard tests or may have been identified, without examination, as one at risk due to the presence of one or more risk factors associated with the disease or condition.

[0092] As used herein, the terms "transduction" and "transduce" refer to a method of introducing a vector construct or a part thereof into a cell. Wherein the vector construct is contained in a viral vector such as for example an AAV vector, transduction refers to viral infection of the cell and subsequent transfer and integration of the vector construct or part thereof into the cell genome.

[0093] As used herein, "treatment" and "treating" of a state, disorder or condition can include: (1) preventing, delaying, or reducing the incidence and/or likelihood of the appearance of at least one clinical or sub-clinical symptom of the state, disorder or condition developing in a subject that may be afflicted with or predisposed to the state, disorder or condition, but does not yet experience or display clinical or subclinical symptoms of the state, disorder or condition; or (2) inhibiting the state, disorder or condition, i.e., arresting, reducing or delaying the development of the disease or a relapse thereof or at least one clinical or sub-clinical symptom thereof; or (3) relieving the disease, i.e., causing regression of the state, disorder or condition or at least one of its clinical or sub-clinical symptoms. The benefit to a subject to be treated is either statistically significant or at least perceptible to the patient or to the physician.

[0094] As used herein, the term "vector" includes a nucleic acid vector, e.g., a DNA vector, such as a plasmid, cosmid, or artificial chromosome, an RNA vector, a virus, or any other suitable replicon (e.g., viral vector). A variety of vectors have been developed for the delivery of polynucleotides encoding exogenous proteins into a prokaryotic or eukaryotic cell. Examples of such expression vectors are described in, e.g., Gellissen, Production of Recombinant Proteins: Novel Microbial and Eukaryotic Expression Systems (John Wiley & Sons, Marblehead, MA, 2006). Expression vectors suitable for use with the compositions and methods described herein

contain a polynucleotide sequence as well as, e.g., additional sequence elements used for the expression of proteins and/or the integration of these polynucleotide sequences into the genome of a mammalian cell. Certain vectors that can be used for the expression of transgene as described herein include vectors that contain regulatory sequences, such as promoter and enhancer regions, which direct gene transcription. Other useful vectors for expression of a transgene contain polynucleotide sequences that enhance the rate of translation of the transgene or improve the stability or nuclear export of the mRNA that results from gene transcription. These sequence elements include, e.g., 5' and 3' untranslated regions and a polyadenylation signal site in order to direct efficient transcription of the gene carried on the expression vector. The expression vectors suitable for use with the compositions and methods described herein may also contain a polynucleotide encoding a marker for selection of cells that contain such a vector. Examples of a suitable marker include genes that encode resistance to antibiotics, such as ampicillin, chloramphenicol, kanamycin, or nourseothricin.

[0095] As used herein, the term “vestibular hair cell” refers to group of specialized cells in the inner ear that are involved in sensing movement and contribute to the sense of balance and spatial orientation. Vestibular hair cells are located in the semicircular canals and otoliths of the inner ear. Damage to vestibular hair cells and genetic mutations that disrupt vestibular hair cell function are implicated in vestibular dysfunction such as vertigo and imbalance disorders.

[0096] As used herein, the term “wild-type” refers to a genotype with the highest frequency for a particular gene in a given organism.

Inner Ear Structure

[0097] The human inner ear is a complex and intricately designed organ responsible for hearing and balance. It comprises two main parts: the cochlea, which is involved in hearing, and the vestibular system, which is critical for balance and directional movement.

A. Cochlea

[0098] The cochlea is a spiral-shaped, fluid-filled tube that is divided into three parallel chambers: the scala vestibuli, scala media, and scala tympani. It contains the organ of Corti, the sensory receptor organ for hearing. Cell types in the Cochlea include the following:

I. Sensory Cells: The organ of Corti houses sensory cells called hair cells, which are mechanosensitive cells that convert sound vibrations into electrical signals. These hair cells are divided into:

- a) **Inner Hair Cells:** Single row of about 3,500 cells that primarily transmit sound information to the brain via the auditory nerve.
- b) **Outer Hair Cells:** Three rows totaling about 12,000 cells that amplify low-level sounds and improve the cochlea's frequency selectivity.

II. Nonsensory Supporting Cells: These include Deiters' cells, Hensen's cells, Claudius' cells, and inner sulcus cells. These cells provide structural support, ionic balance, and potentially influence the electromotility of the outer hair cells.

B. Vestibular System

[0099] The vestibular system consists of the semicircular canals, which detect rotational movements, and the otolith organs (utricle and saccule), which detect linear accelerations. Cell types in the vestibular system include the following:

I. Sensory Cells: Hair cells similar to those in the cochlea, but these are involved in sensing balance and movement. They are also categorized into Type I and Type II hair cells, with differences in their shape and innervation patterns.

II. Nonsensory Supporting Cells: These cells play roles in structural support, maintaining the endolymph fluid that surrounds the hair cells, and possibly affecting hair cell function. Examples include supporting cells in the maculae of the saccule and utricle and those in the cristae of the semicircular canals.

Hair Cells

[00100] Hair cells are sensory cells of the auditory and vestibular systems that reside in the inner ear. Cochlear hair cells are the sensory cells of the auditory system and consist of one row of inner hair cells and three rows of outer hair cells. The inner hair cells are the actual sensory receptors, and 95% of the fibers of the auditory nerve that project to the brain

arise from this subpopulation. The terminations on the outer hair cells are almost all from efferent axons that arise from cells in the brain.. Vestibular hair cells are located in the semicircular canals and otolith organs of the inner ear and are involved in the sensation of movement that contributes to the sense of balance and spatial orientation. Hair cells are named for the stereocilia that protrude from the apical surface of the cell, forming a hair cell bundle. Deflection of the stereocilia (e.g., by sound waves in cochlear hair cells, or by rotation or linear acceleration in vestibular hair cells) leads to the opening of mechanically gated ion channels, which allows hair cells to release neurotransmitters to activate nerves, thereby converting mechanical sound or motion signals into electrical signals that can be transmitted to the brain. Cochlear hair cells are essential for normal hearing, and damage to cochlear hair cells and genetic mutations that disrupt cochlear hair cell function are implicated in hearing loss and deafness. Damage to vestibular hair cells and genetic mutations that disrupt vestibular hair cell function are implicated in vestibular dysfunction, such as loss of balance and vertigo (e.g., dizziness).

[00101] Sensorineural hearing loss (SNHL) is the most common neurodegenerative disease in humans. SNHL can be caused by genetic disorders as well as acquired through injuries such as sound trauma and ototoxicity. Genetic diagnostics have demonstrated that there are at least 100 genes causing nonsyndromic SNHL. Recent advances in genetics and gene therapy techniques have shown that rescue of a number of recessive types of deafness is possible through gene therapy (Askew et al., 2015, Science Translational Medicine, 7(295), 295ral08). An ideal disease target for translational research in this domain is a recessive genetic hearing loss that affects a defined group of cells within the inner ear and occurs postnatally after the development of speech. Prevalence of the mutation is an additional consideration.

Supporting Cells

[00102] Supporting cells play essential roles beyond structural support, including ionic homeostasis and maintenance of the extracellular environment critical for hair cell function. The development, function, and maintenance of inner ear sensory epithelia are heavily dependent upon the supporting cells, which are non-sensory cells that reside between hair cells. Unlike hair cells, which contact only the luminal surface of the epithelium, supporting cells span the entire

depth of the epithelium, from the basal lamina to the lumen. Supporting cells are linked to each other and to hair cells by tight and adherens junctions; and they communicate directly with other supporting cells by gap junctions. Supporting cells serve a diverse set of functions in the sensory epithelia. They have rigid cytoskeletons that maintain the structural integrity of the sensory organs during sound stimulation and head movements. Supporting cells also help to maintain an environment in the epithelium that enables hair cells to function. For instance, hair cells recycle K^+ ions, which help to maintain the driving force for generating the receptor potential. They generate components of the tectorial membrane in the organ of Corti, the otoconial membrane, and otoconial components in the macular organs, and the cupula in the cristae ampullaris. Following trauma or toxicity, supporting cells can eject injured hair cells from the epithelium, phagocytose hair cell debris, and in some cases, generate new hair cells.

Transmembrane Protease, Serine 3 (TMPRSS3)

[00103] Auditory and vestibular systems share numerous molecular constituents, as well as a common molecular evolution, although these systems differ in their structure, cellular components, and physiological properties. Defects in the conserved features may consequently impair both hearing and balance. One such family of molecules includes type II transmembrane serine proteases, which are classified by their N-terminal anchor to membranes and contain an active C-terminal serine protease domain. One such type II transmembrane serine protease, TMPRSS3, is required for proper mammalian hearing, and mutations in TMPRSS3 cause congenital and early-onset hearing loss.

[00104] The human transmembrane protease, serine 3 (TMPRSS3; also referred to as DFNB10, DFNB8, ECHOS1, TADG12; Acc: HGNC: 11877) was identified by its association with both congenital (present at birth) and childhood onset autosomal recessive deafness. Mutations in the TMPRSS3 gene are associated with Autosomal Recessive Nonsyndromic Hearing Impairment type DFNB8 and 10. TMPRSS3 is a 1646 base pair gene that codes for a serine protease and is associated with DFNA 8/10 and may make up to 1-5% of patients with hearing loss undergoing cochlear implantation. Loss of function of TMPRSS3 can result in a broad spectrum of hearing phenotypes depending on the site of the mutation. Both congenital and adult onset progressive hearing loss have been associated with the loss of this gene.

[00105] TMPRSS3 mutations are a fairly common cause of hearing loss that is severe enough to warrant cochlear implantation. However, patients with mutations in TMPRSS3 may not respond to cochlear implantation as well as patients with other mutations. The present disclosure provides methods for targeting TMPRSS3, or other genes such as LOXHD1, as a stand-alone therapeutic or in combination with other therapeutic agents and/or cochlear implantation to improve implant outcomes for this disorder. Due to the size of the TMPRSS3 gene, it may be built into existing vectors such as for example AAV vectors.

Human TMPRSS3 Promoter

[00106] The polynucleotides of the compositions and methods of the present disclosure include nucleic acid sequences from regions of the human TMPRSS3 locus that are capable of expressing a transgene in subpopulations of inner ear cells, or variants thereof, such as a nucleic acid sequences that have at least 85% sequence identity (e.g., 85%, 90%, 95%, 96%, 97%, 98%, 99%, or more, sequence identity) to regions of the human TMPRSS3 locus that are capable of expressing a transgene in subpopulations of inner ear cells. The regulatory regions of the human TMPRSS3 gene tested for driving the expression of a transgene in mouse cochlea cells are listed in Table 2 below. Certain regulatory regions in Table 2 below are termed as modified to indicate one or more changes in the wild-type sequence to remove certain enzyme restriction sites such as for example: 1) the Sba1 restriction site in Region A; 2) two Cla1 restriction sites and an Xba1 restriction site in Region B; and 3) an Xba1 restriction site, an XmaI restriction site, an Acl1 restriction site and an BamH1 restriction site in Region C.

Table 2 – A list of the human TMPRSS3 gene regulatory regions and the corresponding chromosome positions.

Name	Chromosome Position	Sequence ID NO
Modified Region A	chr21:42,396,804-42,397,117 (hg38_dna range=chr21:42396804-42397117)	SEQ ID NO: 1

Region A	chr21:42,396,804-42,397,117 (hg38_dna range=chr21:42396804-42397117)	SEQ ID NO: 2
Modified Region B	chr21:42,395,419-42,396,092 (hg38_dna range=chr21:42395419-42396092)	SEQ ID NO: 3
Second Modified Region B	chr21:42,395,419-42,396,092 (hg38_dna range=chr21:42395419-42396092)	SE ID NO: 26
Region B	chr21:42,395,419-42,396,092 (hg38_dna range=chr21:42395419-42396092)	SEQ ID NO: 4
Modified Region C	chr21:42,392,583-42,393,926 (hg38_dna range=chr21:42392583-42393926)	SEQ ID NO: 5
Region C	chr21:42,392,583-42,393,926 (hg38_dna range=chr21:42392583-42393926)	SEQ ID NO: 6
Region B2	chr21:42,395,933-42,396,092 (hg38_dna range=chr21:42395933-42396092)	SEQ ID NO: 7
Modified Region B3	chr21:42,395,419-42,395,665 (hg38_dna range=chr21:42395418-42395665)	SEQ ID NO: 8
Second Modified Region B3	chr21:42,395,419-42,395,665 (hg38_dna range=chr21:42395418-42395665)	SEQ ID NO: 27
Region B3	chr21:42,395,419-42,395,665 (hg38_dna range=chr21:42395418-42395665)	SEQ ID NO: 9
Modified Region E	chr21:42389046-42389926	SEQ ID NO: 22

	(hg38_dna range=chr21:42389046-42389926)	
Region E	chr21:42389046-42389926 (hg38_dna range=chr21:42389046-42389926)	SEQ ID NO: 23

Expression of exogenous nucleic acids in various types of cells

[00107] Mutations in a variety of genes, such as MYO7A, POU4F3, SLC17A8, and TMC1, have been linked to sensorineural hearing loss, and some of these mutations, such as mutations in MYO7A, are also associated with vestibular dysfunction. The compositions and methods of the present disclosure can be used to induce or increase the expression of proteins encoded by genes of interest (e.g., the wild-type form of genes implicated in hearing loss and/or vestibular dysfunction, or genes involved in hair cell development, function, cell fate specification, regeneration, survival, or maintenance) specifically in subpopulations of inner ear cells (e.g., cochlear and/or vestibular hair cells and/or sensory support cells) by administering a nucleic acid vector according to the present disclosure. A wide array of methods can be used for the delivery of proteins to mammalian cells and for the stable expression of genes encoding proteins in mammalian cells.

[00108] Proteins that may be expressed in connection with the compositions of the present disclosure include proteins that are expressed in healthy hair cells (e.g., cochlear and/or vestibular hair cells, e.g., proteins that play a role in hair cell development, function, regeneration, cell fate specification, survival, or maintenance, or proteins that are deficient in subjects with sensorineural hearing loss or vestibular dysfunction) or other therapeutic proteins of interest. Proteins that can be expressed in hair cells using the compositions and methods described herein include Actin, gamma 1; ADP-ribosyltransferase 1; Activating transcription factor 6; Atonal bHLH transcription factor 1; ATPase plasma membrane Ca²⁺ transporting 2; Brain-derived neurotrophic factor; Binding immunoglobulin protein; Calcium binding protein 2; Calcium channel, voltage-dependent, alpha 2/delta subunit 4; Coiled-Coil Domain Containing

92; Cadherin-related 23; Cholinergic receptor nicotinic alpha 9 subunit; Calcium and integrin binding family member 2; Calcium and integrin binding family member 3; Clarin 1; Clarin 2; Pejvakin; Dystonin; ELFN1 adhesion G protein-coupled receptor; Epidermal growth factor receptor kinase substrate 8; EPS8-like 2; Espin; Espin-like; Fc receptor-like B; Fascin actin-bundling protein 2; Growth factor independent 1 transcriptional repressor; Glutaredoxin domain-containing cysteine-rich protein 1; Glutaredoxin domain-containing cysteine-rich protein 2; Heat shock protein 70; Heat shock protein 90; Inositol-requiring enzyme 1; Potassium voltage-gated channel subfamily A member 10; Kielin/chordin-like protein; Kinocilin; Late cornified envelope protein 6A; LIM homeobox 3; Lipoxigenase homology domains 1; Leucine-rich repeats and transmembrane domains 2; Leucine-rich transmembrane and O-methyltransferase domain containing 1; Leucine-rich transmembrane and O-methyltransferase domain containing 2; Makorin ring finger protein 20; Myosin XV; Myosin IIIA; Myosin IIIB; Myosin VI; Myosin VIIA; Nescient helix-loop-helix 1; Neurotrophin 3; Otoferlin; Protocadherin-related 15; Protein kinase R (PKR)-like endoplasmic reticulum kinase; POU class 4 homeobox 3; Peripherin 2; Protein tyrosine phosphatase receptor type Q; Radixin; Serpin family E member 3; SKI family transcriptional corepressor 1; Solute carrier family 17, member 8 (vesicular glutamate transporter 3); Solute carrier family 8, member 2; Stereocilin; Tctex1 domain containing 1; Transmembrane channel-like 1; Transmembrane protease, serine 3; Taperin; TRIO and F-actin binding protein; Tetratricopeptide repeat domain 24; Usher syndrome 1C; Solute carrier family 26, member 4; Xin actin-binding repeat containing 2; and Zinc finger CCHC-type containing 12.

[00109] One platform that can be used to achieve therapeutically effective intracellular concentrations of proteins of interest in mammalian cells is via the stable expression of the gene encoding the protein of interest (e.g., by integration into the nuclear or mitochondrial genome of a mammalian cell, or by episomal concatemer formation in the nucleus of a mammalian cell). The gene is a polynucleotide that encodes the primary amino acid sequence of the corresponding protein. In order to introduce exogenous genes into a mammalian cell, genes can be incorporated into a vector. Vectors can be introduced into a cell by a variety of methods, including transformation, transfection, transduction, direct uptake, projectile bombardment, and by encapsulation of the vector in a liposome. Examples of suitable methods of transfecting or transforming cells include calcium phosphate precipitation, electroporation, microinjection,

infection, lipofection and direct uptake. Such methods are described in more detail, for example, in Green, et al., *Molecular Cloning: A Laboratory Manual*, Fourth Edition (Cold Spring Harbor University Press, New York 2014); and Ausubel, et al., *Current Protocols in Molecular Biology* (John Wiley & Sons, New York 2015), the disclosures of each of which are incorporated herein by reference.

[00110] Recognition and binding of the polynucleotide encoding a protein of interest by mammalian RNA polymerase is important for gene expression. As such, sequence elements may be included within the polynucleotide that exhibit a high affinity for transcription factors that recruit RNA polymerase and promote the assembly of the transcription complex at the transcription initiation site. Such sequence elements include, e.g., a mammalian promoter, the sequence of which can be recognized and bound by specific transcription initiation factors and ultimately RNA polymerase. Examples of mammalian promoters include, for example, modified Pol III promoter (*TetO-PRR1*), RPR1 promoter, *TEF1* promoter, *MFa1* promoter, or *MFa1* promoter.

[00111] Once a polynucleotide encoding a protein of interest has been incorporated into a mammalian cell, the transcription of this polynucleotide can be induced by methods known in the art. For example, expression can be induced by exposing the mammalian cell to an external chemical reagent, such as an agent that modulates the binding of a transcription factor and/or RNA polymerase to the mammalian promoter and thus regulates gene expression. The chemical reagent can serve to facilitate the binding of RNA polymerase and/or transcription factors to the mammalian promoter, e.g., by removing a repressor protein that has bound the promoter. Alternatively, the chemical reagent can serve to enhance the affinity of the mammalian promoter for RNA polymerase and/or transcription factors such that the rate of transcription of the gene located downstream of the promoter is increased in the presence of the chemical reagent. Examples of chemical reagents that potentiate polynucleotide transcription by the above mechanisms include tetracycline and doxycycline. These reagents are commercially available (Life Technologies, Carlsbad, CA) and can be administered to a mammalian cell to promote gene expression according to established protocols.

[00112] Other DNA sequence elements that may be included in polynucleotides for use in the compositions and methods of the present disclosure include enhancer sequences. Enhancers represent another class of regulatory elements that induce a conformational change in the polynucleotide comprising the gene of interest such that the DNA adopts a three-dimensional orientation that is favorable for binding of transcription factors and RNA polymerase at the transcription initiation site. Thus, polynucleotides for use in the compositions and methods described herein include those that encode a protein of interest and additionally include a mammalian enhancer sequence. Many enhancer sequences are now known from mammalian genes, and examples include enhancers from the genes that encode mammalian globin, elastase, albumin, α -fetoprotein, and insulin. Enhancers for use in the compositions and methods described herein also include those that are derived from the genetic material of a virus capable of infecting a eukaryotic cell. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. Additional enhancer sequences that induce activation of eukaryotic gene transcription include the CMV enhancer and RSV enhancer. An enhancer may be spliced into a vector containing a polynucleotide encoding a protein of interest, for example, at a position 5' or 3' relative to the polynucleotide encoding a protein of interest. In a preferred orientation, the enhancer is positioned at the 5' side of the promoter, which in turn is located 5' relative to the polynucleotide encoding a protein of interest.

[00113] In some embodiments, the nucleic acid vectors containing a TMPRSS3 promoter described herein include a reporter sequence, which can be useful in verifying the expression of a gene operably linked to a TMPRSS3 promoter, for example, in cells and tissues (e.g., in hair cells, such as cochlear and/or vestibular hair cells). Reporter sequences that may be provided in a transgene disclosed herein include, for example, DNA sequences encoding β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, and others well known in the art. When associated with regulatory elements that drive their expression, such as a TMPRSS3 promoter, the reporter sequences provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating

cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for b-galactosidase activity. Where the transgene is green fluorescent protein or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

Delivery of exogenous nucleic acids to target cells

[00114] Various techniques can be used to introduce a transgene, such as a transgene operably linked to a polynucleotide according to the present disclosure, into a target cell are well known in the art. Such techniques include, electroporation which can be used to permeabilize mammalian cells (e.g., human target cells) by the application of an electrostatic potential to the cell of interest. Additional techniques useful for the transfection of target cells include the squeeze-poration methodology. This technique induces the rapid mechanical deformation of cells in order to stimulate the uptake of exogenous DNA through membranous pores that form in response to the applied stress. Lipofection represents another technique useful for transfection of target cells. This method involves the loading of nucleic acids into a liposome, which often presents cationic functional groups, such as quaternary or protonated amines, towards the liposome exterior. This promotes electrostatic interactions between the liposome and a cell due to the anionic nature of the cell membrane, which ultimately leads to uptake of the exogenous nucleic acids, for instance, by direct fusion of the liposome with the cell membrane or by endocytosis of the complex. Another tool for inducing the uptake of exogenous nucleic acids by target cells is laserfection, also called optical transfection, a technique that involves exposing a cell to electromagnetic radiation of a particular wavelength to gently permeabilize the cells and allow polynucleotides to penetrate the cell membrane.

[00115] Various vectors can also be used for delivery of exogenous nucleic acids to target cells. Stable expression of an exogenous gene in a mammalian cell can be achieved by integration of the polynucleotide comprising a gene according to the present disclosure into the nuclear genome of the mammalian cell. A variety of vectors for the delivery and integration of polynucleotides encoding exogenous proteins into the nuclear DNA of a mammalian cell have been developed. Examples of expression vectors are described in, e.g., Gellissen, Production of

Recombinant Proteins: Novel Microbial and Eukaryotic Expression Systems (John Wiley & Sons, Marblehead, MA, 2006). Vectors that can contain a TMPRSS3 promoter operably linked to a transgene encoding a protein of interest include plasmids (e.g., circular DNA molecules that can autonomously replicate inside a cell), cosmids, artificial chromosomes (e.g., a human artificial chromosome (HAC), a yeast artificial chromosome (YAC), a bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC), and viral vectors. Certain vectors that can be used for the expression of a protein of interest include plasmids that contain regulatory sequences, such as enhancer regions, which direct gene transcription.

Viral vectors for nucleic acid delivery

[00116] Viral genomes provide a rich source of vectors that can be used for the efficient delivery of a gene of interest into the genome of a target cell (e.g., a mammalian cell, such as a human cell). Viral genomes are particularly useful vectors for gene delivery because the polynucleotides contained within such genomes are typically incorporated into the nuclear genome of a mammalian cell by generalized or specialized transduction. These processes occur as part of the natural viral replication cycle, and do not require added proteins or reagents to induce gene integration. Examples of viral vectors include a retrovirus (e.g., Retroviridae family viral vector), adenovirus (e.g., Ad5, Ad26, Ad34, Ad35, and Ad48), parvovirus (e.g., adeno-associated viruses), coronavirus, negative strand RNA viruses such as orthomyxovirus (e.g., influenza virus), rhabdovirus (e.g., rabies and vesicular stomatitis virus), paramyxovirus (e.g. measles and Sendai), positive strand RNA viruses, such as picornavirus and alphavirus, and double stranded DNA viruses including adenovirus, herpesvirus (e.g., Herpes Simplex virus types 1 and 2, Epstein-Barr virus, cytomegalovirus), and poxvirus (e.g., vaccinia, modified vaccinia Ankara (MVA), fowlpox and canarypox). Other viruses include Norwalk virus, togavirus, flavivirus, reoviruses, papovavirus, hepadnavirus, human papilloma virus, human foamy virus, and hepatitis virus, for example. Examples of retroviruses include: avian leukosis-sarcoma, avian C-type viruses, mammalian C-type, B-type viruses, D-type viruses, oncoretroviruses, HTLV-BLV group, lentivirus, alpharetrovirus, gammaretrovirus, spumavirus (Coffin, J. M., Retroviridae: The viruses and their replication, Virology, Third Edition (Lippincott-Raven, Philadelphia, 1996)). Other examples include murine leukemia viruses, murine sarcoma viruses, mouse mammary tumor virus, bovine leukemia virus, feline leukemia virus, feline sarcoma virus, avian leukemia virus, human T-cell leukemia virus, baboon endogenous virus, Gibbon ape leukemia virus,

Mason Pfizer monkey virus, simian immunodeficiency virus, simian sarcoma virus, Rous sarcoma virus and lentiviruses. Other examples of vectors include, for example, a vector comprising a first and a second recombining sites such that the vector is internalized by a cell, and then provides the cell with a site-specific recombinase that effects recombination between the first and second recombining sites of the vector.

AAV vectors for nucleic acid delivery

[00117] Adeno-associated virus (AAV) is a small, replication-deficient parvovirus. AAV is about 20-24 nm long, with a density of about 1.40-1.41 g/cc. AAV contains a single-stranded linear genomic DNA molecule approximately 4.7 kb in length. The single-stranded AAV genomic DNA can be either a plus strand, or a minus strand. AAV contains two open reading frames. Rep and Cap, flanked by two 145 base inverted terminal repeats (ITRs). AAVs contain a single iniron. Cis-acting sequences directing viral DNA replication (Rep), encapsidation/packaging and host cell chromosome integration are contained within the ITRs. Three AAV promoters, p5, p19, and p40 (named for their relative map locations) drive the expression of the two AAV internal open reading frames encoding rep and cap genes. The p5 and p 19 are the rep promoters. When coupled with the differential splicing of the single AAV intron, the two rep promoters result in the production of four rep proteins (rep 78, rep 68, rep 52, and rep 40) from the rep gene. The rep proteins have multiple enzymatic properties that are responsible for replicating the viral genome. The cap gene is expressed from the p40 promoter, and encodes the three capsid proteins VP1, VP2, and VP3. Alternative splicing and non-consensus translational start sites are responsible for the production of the three related capsid proteins. A single polyadenylationn site is located at map position 95 of the AAV genome. Muzyczka reviews the life cycle and genetics of AAV (Muzyczka, Current Topics in Microbiology and Immunology, 158:97-129 (1992)).

[00118] The single stranded AAV genomic DNA can be either a plus strand, or a minus strand. In certain embodiments, the term “AAV” or “AAV vector” refers to an AAV that has been modified so that a therapeutic, such as for example, a CRISPR complex, replaces the Rep and Cap open reading frames between the inverted terminal repeats (ITRs) of the AAV genome.

[00119] AAV infection is non-cytopathic in cultured cells. Natural infection of humans and other animals is silent and asymptomatic (does not cause disease). Because AAV infects many mammalian cells, there is the possibility of targeting many different tissues in-vivo. In addition to dividing cells, AAV transduces slowly dividing and non-dividing cells, and can persist essentially for the lifetime of those cells as a transcriptionally active nuclear episome (i.e. extrachromosomal element). The AAV proviral genome is infective as cloned DNA in plasmids, which makes construction of recombinant genomes possible. Moreover, because the signals directing AAV replication, genome encapsidation, and integration are all contained within the ITRs of the AAV genome, some or all of the approximately 4.3 kb of the genome, encoding replication and structural capsid proteins (rep-cap) are contained within the ITRs of the AAV genome, and can be replaced with heterologous DNA, such as a gene cassette containing a promoter, a DNA of interest, and a polyadenylation signal. The rep and cap proteins may be provided in trans. AAV is a very stable and robust virus, and easily withstands conditions used to inactivate adenovirus (56°C to 65°C for several hours), therefore cold preservation of AAV is less critical. And, AAV-infected cells are not resistant to super-infection. These unique properties of AAV a suitable vector for delivering foreign DNA to cells or subjects, for example, in gene therapy.

[00120] In some embodiments, the polynucleotides of the present disclosure are incorporated into rAAV vectors and/or virions to facilitate their introduction into a cell. Viral sequences may include those sequences of AAV that are required in cis for replication and packaging (e.g., functional ITRs) of the DNA into a virion. A transgene as disclosed herein can encode a therapeutic protein that can promote hair cell development, hair cell function, hair cell regeneration, hair cell fate specification, hair cell survival, or hair cell maintenance. A transgene as disclosed herein can also encode a wild-type form of a hair cell protein that is mutated in subjects suffering from hereditary hearing loss or vestibular dysfunction. Such rAAV vectors may also contain marker or reporter genes. Useful rAAV vectors have one or more of the AAV WT genes deleted in whole or in part, but retain functional flanking ITR sequences. The AAV ITRs may be of any serotype suitable for a particular application. For use in the methods and compositions described herein, the ITRs can be AAV2 ITRs. Methods for using rAAV vectors are described, for example, in Tal et al., J. Biomed. Sci. 7:279 (2000), and Monahan and

Samulski, Gene Delivery 7:24 (2000), the disclosures of each of which are incorporated herein by reference as they pertain to AAV vectors for gene delivery.

[00121] Viral vectors of the present disclosure may comprise AAV particles which can include any AAV particles or derivatives thereof. An AAV particle may comprise or be derived from any natural or recombinant AAV serotype. According to the present disclosure, the AAV particles may utilize or be based on a serotype selected from any of the following AAV1, AAV2, AAV2G9, AAV3, AAV3a, AAV3b, AAV3-3, AAV4, AAV4-4, AAV5, AAV6, AAV6.1, AAV6.2, AAV6.1.2, AAV7, AAV7.2, AAV8, AAV9, AAV9.11, AAV9.13, AAV9.16, AAV9.24, AAV9.45, AAV9.47, AAV9.61, AAV9.68, AAV9.84, AAV9.9, AAV10, AAV11, AAV12, AAV16.3, AAV24.1, AAV27.3, AAV42.12, AAV42-1b, AAV42-2, AAV42-3a, AAV42-3b, AAV42-4, AAV42-5a, AAV42-5b, AAV42-6b, AAV42-8, AAV42-10, AAV42-11, AAV42-12, AAV42-13, AAV42-15, AAV42-aa, AAV43-1, AAV43-12, AAV43-20, AAV43-21, AAV43-23, AAV43-25, AAV43-5, AAV44.1, AAV44.2, AAV44.5, AAV223.1, AAV223.2, AAV223.4, AAV223.5, AAV223.6, AAV223.7, AAV1-7/rh.48, AAV1-81rh.49, AAV2-15/rh.62, AAV2-3/rh.61, AAV2-4/rh.50, AAV2-5/rh.51, AAV3.1/hu.6, AAV3.1/hu.9, AAV3-9/rh.52, AAV3-11/rh.53, AAV4-8/rh.64, AAV4-9/rh.54, AAV4-19/rh.55, AAV5-3/rh.57, AAV5-22/rh.58, AAV7.3/hu.7, AAV16.8/hu.10, AAV16.12/hu.11, AAV29.3/bb.1, AAV29.5/bb.2, AAV106.1/hu.37, AAV114.3/hu.40, AAV127.2/hu.41, AAV127.5/hu.42, AAV128.3/hu.44, AAV130.4/hu.48, AAV145.1/hu.53, AAV145.5/hu.54, AAV145.6/hu.55, AAV161.10/hu.60, AAV61.6/hu.61, AAV33.12/hu.17, AAV33.4/hu.15, AAV33.8/hu.16, AAV52/hu.19, AAV52.1/hu.20, AAV58.2/hu.25, AAV3.3, AAV3.4, AAV3.5, AAV3.7, AAVC1, AAVC2, AAVC5, AAV-DJ, AAV-DJ8, AAVF3, AAVF5, AAVH2, AAVrh.72, AAVhu.8, AAVrh.68, AAVrh.70, AAVpi.1, AAVpi.3, AAVpi.2, AAVrh.60, AAVrh.44, AAVrh.65, AAVrh.55, AAVrh.47, AAVrh.69, AAVrh.45, AAVrh.59, AAVhu.12, AAVH6, AAVLK03, AAVH-1/hu.1, AAVH-5/hu.3, AAVLG-10/rh.40, AAVLG-4/rh.38, AAVLG-9/hu.39, AAVN721-8/rh.43, AAVCh.5, AAVCh.5R1, AAVcy.2, AAVcy.3, AAVcy.4, AAVcy.5, AAVCy.5R1, AAVCy.5R2, AAVCy.5R3, AAVCy.5R4, AAVcy.6, AAVhu.1, AAVhu.2, AAVhu.3, AAVhu.4, AAVhu.5, AAVhu.6, AAVhu.7, AAVhu.9, AAVhu.10, AAVhu.11, AAVhu.13, AAVhu.15, AAVhu.16, AAVhu.17, AAVhu.18, AAVhu.20, AAVhu.21, AAVhu.22, AAVhu.23.2, AAVhu.24, AAVhu.25, AAVhu.27, AAVhu.28, AAVhu.29,

AAVhu.29R, AAVhu.31, AAVhu.32, AAVhu.34, AAVhu.35, AAVhu.37, AAVhu.39.
AAVhu.40, AAVhu.41, AAVhu.42, AAVhu.43, AAVhu.44, AAVhu.44R1, AAVhu.44R2,
AAVhu.44R3, AAVhu.45, AAVhu.46, AAVhu.47, AAVhu.48, AAVhu.48R1, AAVhu.48R2.
AAVhu.48R3, AAVhu.49, AAVhu.51, AAVhu.52, AAVhu.54, AAVhu.55, AAVhu.56,
AAVhu.57, AAVhu.58, AAVhu.60, AAVhu.61, AAVhu.63, AAVhu.64, AAVhu.66,
AAVhu.67, AAVhu.14/9, AAVhu.t 19, AAVrh.2, AAVrh.2R, AAVrh.8, AAVrh.8R, AAVrh.10,
AAVrh.12, AAVrh.13, AAVrh.13R, AAVrh.14, AAVrh.17, AAVrh.18, AAVrh.19, AAVrh.20.
AAVrh.21, AAVrh.22, AAVrh.23, AAVrh.24, AAVrh.25, AAVrh.31, AAVrh.32, AAVrh.33.
AAVrh.34, AAVrh.35, AAVrh.36, AAVrh.37, AAVrh.37R2, AAVrh.38, AAVrh.39, AAVrh.40,
AAVrh.46, AAVrh.48, AAVrh.48.1, AAVrh.48.1.2, AAVrh.48.2, AAVrh.49, AAVrh.51,
AAVrh.52, AAVrh.53, AAVrh.54, AAVrh.56, AAVrh.57, AAVrh.58, AAVrh.61, AAVrh.64,
AAVrh.64R1, AAVrh.64R2, AAVrh.67, AAVrh.73, AAVrh.74, AAVrh8R, AAVrh8R A586R
mutant, AAVrh8R R533A mutant, AAV, BAAV, caprine AAV, bovine AAV. AAVhE1.1,
AAVhEr1.5, AAVhEr1.14, AAVhEr1.8, AAVhEr1.16, AAVhEr1.18, AAVhEr1.35,
AAVhEr1.7, AAVhEr1.36, AAVhEr2.29, AAVhEr2.4, AAVhEr2.16, AAVhEr2.30,
AAVhEr2.31, AAVhEr2.36, AAVhEr1.23, AAVhEr3.1, AAV2.5T. AAV-PAEC, AAV-LK01.
AAV-LK02, AAV-LK03, AAV-LK04, AAV-LK05, AAV-LK06, AAV-LK07, AAV-LK08,
AAV-LK09, AAV-LK10, AAV-LK11, AAV-LK12, AAV-LK13, AAV-LK14, AAV-LK15,
AAV-LK16, AAV-LK17, AAV-LK18, AAV-LK19, AAV-PAEC2, AAV-PAEC4, AAV-
PAEC6, AAV-PAEC7, AAV-PAEC8, AAV-PAEC11, AAV-PAEC12, AAV-2-pre-miRNA-101,
AAV-8h, AAV-8b. AAV-h, AAV-b, AAV SM 10-2, AAV Shuffle 100-1, AAV Shuffle 100-3,
AAV Shuffle 100-7, AAV Shuffle 10-2, AAV Shuffle 10-6, AAV Shuffle 10-8, AAV Shuffle
100-2, AAV SM 10-1, AAV SM 10-8, AAV SM 100-3, AAV SM 100-10, BNP61 AAV. BNP62
AAV, BNP63 AAV, AAVrh.50, AAVrh.43, AAVrh.62, AAVrh.48, AAVhu.19, AAVhu.11,
AAVhu.53, AAV4-8/rh.64, AAVLG-9/hu.39, AAV54.5/hu.23, AAV54.2/hu.22,
AAV54.7/hu.24, AAV54.1/hu.21, AAV54.4R/hu.27, AAV46.2/hu.28, AAV46.6/hu.29,
AAV128.1/hu.43, true type AAV (tAAV), UPENN AAV10, Japanese AAV10 serotypes, AAV
CBr-7.1, AAV CBr-7.10, AAV CBr-7.2, AAV CBr-7.3, AAV CBr-7.4, AAV CBr-7.5, AAV
CBr-7.7, AAV CBr-7.8, AAV CBr-B7.3, AAV CBr-B7.4, AAV CBr-E1, AAV CBr-E2, AAV
CBr-E3, AAV CBr-E4, AAV CBr-E5, AAV CBr-e5, AAV CBr-E6, AAV CBr-E7, AAV CBr-
E8, AAV CHt-1, AAV CHt-2, AAV CH-3, AAV CHt-6.1, AAV CHt-6.10, AAV CHt-6.5. AAV

CHt-6.6, AAV CHt-6.7, AAV CHt-6.8, AAV CHt-P1, AAV CHt-P2, AAV CHt-P5, AAV CHt-P6, AAV CHt-P8, AAV CHt-P9, AAV CKd-1, AAV CKd-10, AAV CKd-2, AAV CKd-3, AAV CKd-4, AAV CKd-6, AAV CKd-7, AAV CKd-8, AAV CKd-B1, AAV CKd-B2, AAV CKd-B3, AAV CKd-B4, AAV CKd-B5, AAV CKd-B6, AAV CKd-B7, AAV CKd-B8, AAV CKd-H, AAV CKd-H2, AAV CKd-H3, AAV CKd-H4, AAV CKd-H5, AAV CKd-H6, AAV CKd-N3, AAV CKd-N4, AAV CKd-N9, AAV CLg-F1, AAV CLg-F2, AAV CLg-F3, AAV CLg-F4, AAV CLg-F5, AAV CLg-F6, AAV CLg-F7, AAV CLg-F8, AAV CLv-1, AAV CLv1-1, AAV CLv1-10, AAV CLv1-2, AAV CLv-12, AAV CLv1-3, AAV CLv-13, AAV CLv1-4, AAV CLv1-7, AAV CLv1-8, AAV CLv1-9, AAV CLV-2, AAV CLv-3, AAV CLV-4, AAV CLv-6, AAV CLv-8, AAV CLv-D1, AAV CLv-D2, AAV CLv-D3, AAV CLv-D4, AAV CL %D5, AAV CLv-D6, AAV CLv-D7, AAV CLv-D8, AAV CLv-E1, AAV CLv-K1, AAV CLv-K3, AAV CLv-K6, AAV CLV-L4, AAV CLV-L5, AAV CLv-L6, AAV CLv-M1, AAV CLv-M11, AAV CLv-M2, AAV CLv-M5, AAV CLv-M6, AAV CLv-M7, AAV CLv-M8, AAV CLv-M9, AAV CLv-R1, AAV CLv-R2, AAV CLv-R3, AAV CLv-R4, AAV CLv-R5, AAV CLv-R6, AAV CLv-R7, AAV CLv-R8, AAV CLv-R9, AAV CSp-1, AAV CSp-10, AAV CSp-11, AAV CSp-2, AAV CSp-3, AAV CSp-4, AAV CSp-6, AAV CSp-7, AAV CSp-8, AAV CSp-8.10, AAV CSp-8.2, AAV CSp-8.4, AAV CSp-8.5, AAV CSp-8.6, AAV CSp-8.7, AAV CSp-8.8, AAV CSp-8.9, AAV CSp-9, AAV.hu.48R3, AAV.VR-355, AAV3B, AAV4, AAV5, AAVF1/HSC1, AAVF11/HSC11, AAVF12/HSC:12, AAVF13/HSC13, AAVF14/HSC14, AAVF15/HSC15, AAVF16/HSC16, AAVF17/HSC17, AAVF2/HSC2, AAVF3/HSC3, AAVF4/HSC4, AAVF5/HSC5, AAVF6/HSC6, AAVF7/HSC7, AAVF8/HSC8, AAVF9/HSC9, PHP.B, PHP.A, G2B-26, G2B-13, TH1.1-32, Anc80, Anc80L65 and/or TH1.1-35 and variants thereof.

[00122] Also useful in conjunction with the compositions and methods described in the present disclosure are pseudotyped rAAV vectors. Pseudotyped vectors include AAV vectors of a given serotype (e.g., AAV9) pseudotyped with a capsid gene derived from a serotype other than the given serotype (e.g., AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, etc.). Techniques involving the construction and use of pseudotyped rAAV virions are known in the art and are described, for instance, in Duan et al., *J. Virol.* 75:7662 (2001); Halbert et al., *J. Virol.* 74:1524 (2000); Zolotukhin et al., *Methods*, 28:158 (2002); and Auricchio et al., *Hum. Molec. Genet.* 10:3075 (2001).

[00123] In a preferred embodiment, the AAV vector comprises the polynucleotides of the present disclosure. In another embodiment, more than one AAV vector such as a dual AAV vector may be used with such vectors comprising one or more polynucleotides of the present disclosure.

[00124] In certain embodiments, the AAV vectors may be delivered without being enclosed in any particle or lipid vessels. In other embodiments, the AAV vectors may be enclosed in a lipid nanoparticle, liposome, non-lipid nanoparticle, or viral capsid for delivery.

Pharmaceutical Compositions

[00125] The present disclosure provides a pharmaceutical composition comprising the nucleic acid vector of the present disclosure, and a pharmaceutically acceptable excipient.

Pharmaceutically acceptable excipients include, but are not limited to, a water-insoluble diluent, a water-soluble diluent, a disintegrant, a binder, a wetting agent, a solubilizer, a glidant, a lubricant, and a granulating solvent.

[00126] Pharmaceutical compositions of the present disclosure may be prepared in water suitably mixed with one or more excipients, carriers, or diluents. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations may contain a preservative to prevent the growth of microorganisms. The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (described in US 5,466,468, the disclosure of which is incorporated herein by reference). In any case, the formulation may be sterile and may be fluid to the extent that easy syringability exists. Formulations may be stable under the conditions of manufacture and storage and may be preserved 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 by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Lipid Nanoparticles

[00127] In preferred embodiments the pharmaceutical composition of the present disclosure is a lipid nanoparticle (LNP) pharmaceutical composition. LNP formulations are described in WO 2017/173054. Lipid nanoparticles (“LNPs”) are examples of vectors according to the present disclosure. LNPs are particles comprising a plurality of lipid molecules physically associated with each other by intermolecular forces. These include microspheres (including unilamellar and multilamellar vesicles, e.g., liposomes), a dispersed phase in an emulsion, micelles, or an internal phase in a suspension. Such lipid nanoparticles can be used to encapsulate one or more nucleic acids or proteins for delivery. Formulations which contain cationic lipids are useful for delivering polyanions such as nucleic acids. Other lipids that can be included are neutral lipids (i.e., uncharged or zwitterionic lipids), anionic lipids, helper lipids that enhance transfection, and stealth lipids that increase the length of time for which nanoparticles can exist in-vivo. Examples of suitable cationic lipids, neutral lipids, anionic lipids, helper lipids, and stealth lipids can be found in WO 2016/010840 A1 and WO 2017/173054 A1, each of which is herein incorporated by reference in its entirety for all purposes. An exemplary lipid nanoparticle can comprise a cationic lipid and one or more other components. In one example, the other component can comprise a helper lipid such as cholesterol. In another example, the other components can comprise a helper lipid such as cholesterol and a neutral lipid such as DSPC. In another example, the other components can comprise a helper lipid such as cholesterol, an optional neutral lipid such as DSPC, and a stealth lipid such as S010, S024, S027, S031, or S033.

[00128] The LNP may contain one or more or all of the following: (i) a lipid for encapsulation and for endosomal escape; (ii) a neutral lipid for stabilization; (iii) a helper lipid for stabilization;

and (iv) a stealth lipid. See, e.g., Finn et al. (2018) Cell Rep. 22(9):2227-2235 and WO 2017/173054 A1, each of which is herein incorporated by reference in its entirety for all purposes. In a preferred embodiment, the LNP includes a nucleic acid molecule coding for the chimeric protein of the present disclosure.

[00129] The lipid for encapsulation and endosomal escape can be a cationic lipid. The lipid can also be a biodegradable lipid, such as a biodegradable ionizable lipid. One example of a suitable lipid is Lipid A or LP01, which is (9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)-)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl-)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate. See, e.g., Finn et al. (2018) Cell Rep. 22(9):2227-2235 and WO 2017/173054 A1, each of which is herein incorporated by reference in its entirety for all purposes. Another example of a suitable lipid is Lipid B, which is ((5-(((dimethylamino)methyl)-1,3-phenylene)bis(oxy))bis(octane-8,1-diyl)bis(decanoate), also called ((5-(((dimethylamino)methyl)-1,3-phenylene)bis(oxy))bis(octane-8,1-diyl)bi- s(decanoate). Another example of a suitable lipid is Lipid C, which is 2-((4-(((3-(dimethylamino)propoxy)carbonyl)oxy)hexadecanoyl)oxy)propane-1- ,3-diyl(9Z,9'Z,12Z,12'Z)-bis(octadeca-9,12-dienoate). Another example of a suitable lipid is Lipid D, which is 3-(((3-(dimethylamino)propoxy)carbonyl)oxy)-13-(octanoyloxy)tridecyl 3-octylundecanoate. Other suitable lipids include heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (also known as [(6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl] 4-(dimethylamino)butanoate or Dlin-MC3-DMA (MC3))).

[00130] Some such lipids suitable for use in the LNPs described herein are biodegradable in-vivo. For example, LNPs comprising such a lipid include those where at least 75% of the lipid is cleared from the plasma within 8, 10, 12, 24, or 48 hours, or 3, 4, 5, 6, 7, or 10 days. As another example, at least 50% of the LNP is cleared from the plasma within 8, 10, 12, 24, or 48 hours, or 3, 4, 5, 6, 7, or 10 days.

[00131] Such lipids may be ionizable depending upon the pH of the medium they are in. For example, in a slightly acidic medium, the lipids may be protonated and thus bear a positive

charge. Conversely, in a slightly basic medium, such as, for example, blood where pH is approximately 7.35, the lipids may not be protonated and thus bear no charge. In some embodiments, the lipids may be protonated at a pH of at least about 9, 9.5, or 10. The ability of such a lipid to bear a charge is related to its intrinsic pKa. For example, the lipid may, independently, have a pKa in the range of from about 5.8 to about 6.2.

[00132] Neutral lipids function to stabilize and improve processing of the LNPs. Examples of suitable neutral lipids include a variety of neutral, uncharged or zwitterionic lipids. Examples of neutral phospholipids suitable for use in the present disclosure include, but are not limited to, 5-heptadecylbenzene-1,3-diol (resorcinol), dipalmitoylphosphatidylcholine (DPPC), distearoylphosphatidylcholine or 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), phosphocholine (DOPC), dimyristoylphosphatidylcholine (DMPC), phosphatidylcholine (PLPC), 1,2-diarachidonoyl-sn-glycero-3-phosphocholine (DAPC), phosphatidylethanolamine (PE), egg phosphatidylcholine (EPC), dilauryloylphosphatidylcholine (DLPC), dimyristoylphosphatidylcholine (DMPC), 1-myristoyl-2-palmitoyl phosphatidylcholine (MPPC), 1-palmitoyl-2-myristoyl phosphatidylcholine (PMPC), 1-palmitoyl-2-stearoyl phosphatidylcholine (PSPC), 1,2-diarachidoyl-sn-glycero-3-phosphocholine (DBPC), 1-stearoyl-2-palmitoyl phosphatidylcholine (SPPC), 1,2-dieicosenoyl-sn-glycero-3-phosphocholine (DEPC), palmitoyloleoyl phosphatidylcholine (POPC), lysophosphatidyl choline, dioleoyl phosphatidylethanolamine (DOPE), dilinoleoylphosphatidylcholine distearoylphosphatidylethanolamine (DSPE), dimyristoyl phosphatidylethanolamine (DMPE), dipalmitoyl phosphatidylethanolamine (DPPE), palmitoyloleoyl phosphatidylethanolamine (POPE), lysophosphatidylethanolamine, 1-stearoyl-2-oleoyl-sn-glycero-3-phosphocholine (SOPC), and combinations thereof. For example, the neutral phospholipid may be selected from the group consisting of distearoylphosphatidylcholine (DSPC) and dimyristoyl phosphatidyl ethanolamine (DMPE).

[00133] Helper lipids include lipids that enhance transfection. The mechanism by which the helper lipid enhances transfection can include enhancing particle stability. In certain cases, the helper lipid can enhance membrane fusogenicity. Helper lipids include steroids, sterols, and alkyl resorcinols. Examples of suitable helper lipids suitable include cholesterol, 5-

heptadecylresorcinol, and cholesterol hemisuccinate. In one example, the helper lipid may be cholesterol or cholesterol hemisuccinate.

[00134] Stealth lipids include lipids that alter the length of time the nanoparticles can exist in-vivo. Stealth lipids may assist in the formulation process by, for example, reducing particle aggregation and controlling particle size. Stealth lipids may modulate pharmacokinetic properties of the LNP. Suitable stealth lipids include lipids having a hydrophilic head group linked to a lipid moiety.

[00135] The hydrophilic head group of stealth lipid can comprise, for example, a polymer moiety selected from polymers based on PEG (sometimes referred to as poly(ethylene oxide)), poly(oxazoline), poly(vinyl alcohol), poly(glycerol), poly(N-vinylpyrrolidone), polyaminoacids, and poly N-(2-hydroxypropyl)methacrylamide. The term PEG means any polyethylene glycol or other polyalkylene ether polymer. In certain LNP formulations, the PEG, is a PEG-2K, also termed PEG 2000, which has an average molecular weight of about 2,000 daltons. See, e.g., WO 2017/173054 A1, herein incorporated by reference in its entirety for all purposes.

[00136] The lipid moiety of the stealth lipid may be derived, for example, from diacylglycerol or diacylglycamide, including those comprising a dialkylglycerol or dialkylglycamide group having alkyl chain length independently comprising from about C4 to about C40 saturated or unsaturated carbon atoms, wherein the chain may comprise one or more functional groups such as, for example, an amide or ester. The dialkylglycerol or dialkylglycamide group can further comprise one or more substituted alkyl groups.

[00137] As one example, the stealth lipid may be selected from PEG-dilauroylglycerol, PEG-dimyristoylglycerol (PEG-DMG), PEG-dipalmitoylglycerol, PEG-distearoylglycerol (PEG-DSPE), PEG-dilaurylglycamide, PEG-dimyristylglycamide, PEG-dipalmitoylglycamide, and PEG-distearoylglycamide, PEG-cholesterol (1-[8'-(Cholest-5-en-3[beta]-oxy)carboxamido-3',6'-dioxaoctanyl]carbamoyl- -[omega]-methyl-poly(ethylene glycol), PEG-DMB (3,4-ditetradecoxylbenzyl-[omega]-methyl-poly(ethylene glycol)ether), 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (PEG2k-DMG), 1,2-distearoyl-

sn-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-2000] (PEG2k-DSPE), 1,2-distearoyl-sn-glycerol, methoxypoly ethylene glycol (PEG2k-DSG), poly(ethylene glycol)-2000-dimethacrylate (PEG2k-DMA), and 1,2-distearoyloxypropyl-3-amine-N-[methoxy(polyethylene glycol)-2000] (PEG2k-DSA). In one particular example, the stealth lipid may be PEG2k-DMG.

[00138] The LNPs can comprise different respective molar ratios of the component lipids in the formulation. The mol-% of the CCD lipid may be, for example, from about 30 mol-% to about 60 mol-%, from about 35 mol-% to about 55 mol-%, from about 40 mol-% to about 50 mol-%, from about 42 mol-% to about 47 mol-%, or about 45%. The mol-% of the helper lipid may be, for example, from about 30 mol-% to about 60 mol-%, from about 35 mol-% to about 55 mol-%, from about 40 mol-% to about 50 mol-%, from about 41 mol-% to about 46 mol-%, or about 44 mol-%. The mol-% of the neutral lipid may be, for example, from about 1 mol-% to about 20 mol-%, from about 5 mol-% to about 15 mol-%, from about 7 mol-% to about 12 mol-%, or about 9 mol-%. The mol-% of the stealth lipid may be, for example, from about 1 mol-% to about 10 mol-%, from about 1 mol-% to about 5 mol-%, from about 1 mol-% to about 3 mol-%, about 2 mol-%, or about 1 mol-%.

[00139] The LNPs can have different ratios between the positively charged amine groups of the biodegradable lipid (N) and the negatively charged phosphate groups (P) of the nucleic acid to be encapsulated. This may be mathematically represented by the equation N/P. For example, the N/P ratio may be from about 0.5 to about 100, from about 1 to about 50, from about 1 to about 25, from about 1 to about 10, from about 1 to about 7, from about 3 to about 5, from about 4 to about 5, about 4, about 4.5, or about 5. The N/P ratio can also be from about 4 to about 7 or from about 4.5 to about 6. In specific examples, the N/P ratio can be 4.5 or can be 6.

[00140] A specific example of a suitable LNP has a nitrogen-to-phosphate (N/P) ratio of 4.5 and contains biodegradable cationic lipid, cholesterol, DSPC, and PEG2k-DMG in a 45:44:9:2 molar ratio. The biodegradable cationic lipid can be (9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)-)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl-)oxy)methyl)propyl

(9Z,12Z)-octadeca-9,12-dienoate. See, e.g., Finn et al. (2018) Cell Rep. 22(9):2227-2235, herein incorporated by reference in its entirety for all purposes. Another specific example of a suitable LNP contains Dlin-MC3-DMA (MC3), cholesterol, DSPC, and PEG-DMG in a 50:38.5:10:1.5 molar ratio.

[00141] Another specific example of a suitable LNP has a nitrogen-to-phosphate (N/P) ratio of 6 and contains biodegradable cationic lipid, cholesterol, DSPC, and PEG2k-DMG in a 50:38:9:3 molar ratio. The biodegradable cationic lipid can be (9Z,12Z)-3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)-)carbonyl)oxy)methyl)propyl octadeca-9,12-dienoate, also called 3-((4,4-bis(octyloxy)butanoyl)oxy)-2-(((3-(diethylamino)propoxy)carbonyl-)oxy)methyl)propyl (9Z,12Z)-octadeca-9,12-dienoate.

[00142] Another specific example of a suitable LNP has a nitrogen-to-phosphate (N/P) ratio of 3 and contains a cationic lipid, a structural lipid, cholesterol (e.g., cholesterol (ovine) (Avanti 700000)), and PEG2k-DMG (e.g., PEG-DMG 2000 (NOF America-STJNBRIGHT.RTM. GM-020(DMG-PEG)) in a 50:10:38.5:1.5 ratio or a 47:10:42:1 ratio. The structural lipid can be, for example, DSPC (e.g., DSPC (Avanti 850365)), SOPC, DOPC, or DOPE. The cationic/ionizable lipid can be, for example, Dlin-MC3-DMA (e.g., Dlin-MC3-DMA (Biofine International)).

[00143] Another specific example of a suitable LNP contains Dlin-MC3-DMA, DSPC, cholesterol, and a PEG lipid in a 45:9:44:2 ratio. Another specific example of a suitable LNP contains Dlin-MC3-DMA, DOPE, cholesterol, and PEG lipid or PEG DMG in a 50:10:39:1 ratio. Another specific example of a suitable LNP has Dlin-MC3-DMA, DSPC, cholesterol, and PEG2k-DMG at a 55:10:32.5:2.5 ratio. Another specific example of a suitable LNP has Dlin-MC3-DMA, DSPC, cholesterol, and PEG-DMG in a 50:10:38.5:1.5 ratio. Another specific example of a suitable LNP has Dlin-MC3-DMA, DSPC, cholesterol, and PEG-DMG in a 50:10:38.5:1.5 ratio.

Methods of treatment

[00144] The compositions of the present disclosure may be administered to a subject with sensorineural hearing loss and/or vestibular dysfunction by a variety of routes, such as local

administration to the inner ear (e.g., administration into the perilymph or endolymph, e.g., through the oval window, round window, or a semicircular canal (e.g., the horizontal canal), e.g., administration to a cochlear or vestibular hair cell), intravenous, parenteral, intradermal, transdermal, intramuscular, intranasal, subcutaneous, percutaneous, intratracheal, intraperitoneal, intraarterial, intravascular, inhalation, perfusion, lavage, and oral administration. The most suitable route for administration in any given case will depend on the particular composition administered, the patient, pharmaceutical formulation methods, administration methods (e.g., administration time and administration route), the patient's age, body weight, sex, severity of the disease being treated, the patient's diet, and the patient's excretion rate. Compositions may be administered once, or more than once (e.g., once annually, twice annually, three times annually, bi-monthly, or monthly).

[00145] Subjects that may be treated as described herein are subjects having or at risk of developing sensorineural hearing loss and/or vestibular dysfunction (e.g., subjects having or at risk of developing hearing loss, vestibular dysfunction, or both). The compositions and methods described herein can be used to treat subjects having or at risk of developing damage to cochlear hair cells (e.g., damage related to acoustic trauma, disease or infection, head trauma, ototoxic drugs, or aging), subjects having or at risk of developing damage to vestibular hair cells (e.g., damage related to disease or infection, head trauma, ototoxic drugs, or aging), subjects having or at risk of developing sensorineural hearing loss, deafness, or auditory neuropathy, subjects having or at risk of developing vestibular dysfunction (e.g., dizziness, vertigo, or imbalance), subjects having tinnitus (e.g., tinnitus alone, or tinnitus that is associated with sensorineural hearing loss or vestibular dysfunction), subjects having a genetic mutation associated with hearing loss and/or vestibular dysfunction, or subjects with a family history of hereditary hearing loss, deafness, auditory neuropathy, tinnitus, or vestibular dysfunction. In some embodiments, the subject has hearing loss and/or vestibular dysfunction that is associated with or results from loss of hair cells (e.g., cochlear or vestibular hair cells).

[00146] The compositions and methods described herein may also be administered as a preventative treatment to patients at risk of developing hearing loss and/or vestibular dysfunction, e.g., patients who have a family history of hearing loss or vestibular dysfunction (e.g., inherited hearing loss or vestibular dysfunction), patients carrying a genetic mutation

associated with hearing loss or vestibular dysfunction who do not yet exhibit hearing impairment or vestibular dysfunction or patients exposed to risk factors for acquired hearing loss (e.g., disease or infection, head trauma, ototoxic drugs, or aging) or vestibular dysfunction (e.g., acoustic trauma, disease or infection, head trauma, ototoxic drugs, or aging).

[00147] The compositions and methods described herein can be used to promote or induce hair cell regeneration in a subject (e.g., cochlear and/or vestibular hair cell regeneration). Subjects that may benefit from compositions that promote or induce hair cell regeneration include subjects suffering from hearing loss or vestibular dysfunction as a result of loss of hair cells (e.g., loss of hair cells related to trauma (e.g., acoustic trauma or head trauma), disease or infection, ototoxic drugs, or aging), and subjects with abnormal hair cells (e.g., hair cells that do not function properly when compared to normal hair cells), damaged hair cells (e.g., hair cell damage related to trauma (e.g., acoustic trauma or head trauma), disease or infection, ototoxic drugs, or aging), or reduced hair cell numbers due to genetic mutations or congenital abnormalities. The compositions and methods described herein can also be used to promote or increase hair cell survival (e.g., increase survival of damaged hair cells, promote repair of damaged hair cells, or preserve hair cells in a subject at risk of loss of hair cells (e.g., loss of hair cells due to age, exposure to loud noise, disease or infection, head trauma or ototoxic drugs)).

[00148] The compositions and methods described herein can also be used to prevent or reduce ototoxic drug-induced hair cell damage or death (e.g., cochlear and/or vestibular hair cell damage or death) in subjects who have been treated with ototoxic drugs, or who are currently undergoing or soon to begin treatment with ototoxic drugs. Ototoxic drugs are toxic to the cells of the inner ear, and can cause sensorineural hearing loss, vestibular dysfunction (e.g., vertigo, dizziness, or imbalance), tinnitus, or a combination of these symptoms. Drugs that have been found to be ototoxic include aminoglycoside antibiotics (e.g., gentamycin, neomycin, streptomycin, tobramycin, kanamycin, vancomycin, and amikacin), viomycin, antineoplastic drugs (e.g., platinum-containing chemotherapeutic agents, such as cisplatin, carboplatin, and oxaliplatin), loop diuretics (e.g., ethacrynic acid and furosemide), salicylates (e.g., aspirin, particularly at high doses), and quinine. In some embodiments, the methods described herein prevent or reduce hair cell damage or death related to acoustic trauma, disease or infection, head trauma, or aging.

[00149] In some embodiments, the proteins may be expressed by a transgene operably linked to a TMPRSS3 promoter of the present disclosure for treatment of a subject as described herein. The transgene includes ACTG1, ART1, ATF6, ATOH1, ATP2B2, BDNF, BIP, CABP2, CACNA2D4, CCER2, CDH23, CHRNA9, CIB2, CIB3, CLRN1, CLRN2, DFNB59, DYT1N, ELFN1, EPS8, EPS8L2, ESPN, ESPNL, FCRLB, FSCN2, GFI1, GRXCR1, GRXCR2, HSP70, HSP90, IRE1, KCNA10, KCP, KNCN, LCE6A, LHX3, LOXHD1, LRTM2, LRTOMT1, LRTOMT2, MKRN20S, MY015, MY03A, MY03B, MY06, MY07A, NHLH1, NTF3, OTOF, PCDH15, PERK, POU4F3, PRPH2, PTPRQ, RDX, SERPINE3, SKOR1, SLC17A8, SLC8A2, STRC, TCTEX1 D1, TMC1, TMPRSS3, TPRN, TRIOBP, TTC24, USH1 C, SLC26A4, XIRP2, and ZCCHC12. In certain embodiments, the transgene is packaged into one or more AAV vectors such as for example dual AAV vectors.

[00150] Dual vectors for use in the methods and compositions described herein are designed such that a portion of the transgene is contained within each vector (e.g., each vector contains a polynucleotide that encodes a portion of the transgene). The determination of how to split the polynucleotide sequence between the two nucleic acid vectors can be made based on the size of the promoter and the locations of the portions of the polynucleotide that encode the transgene.

[00151] Treatment may include administration of a composition containing the nucleic acid vectors (e.g., AAV viral vectors) containing a TMPRSS3 promoter described herein in various unit doses. Each unit dose will ordinarily contain a predetermined-quantity of the therapeutic composition. The quantity to be administered, and the particular route of administration and formulation, are within the skill of those in the clinical arts. A unit dose need not be administered as a single injection but may comprise continuous infusion over a set period of time. Dosing may be performed using a syringe pump to control infusion rate in order to minimize damage to the inner ear (e.g., the cochlea). The viral vectors may be administered to the patient at a dose of, for example, from about 1×10^{10} vector genomes (VG) to 1×10^{15} VG.

CRISPR/Cas Systems

[00152] The methods and compositions disclosed herein can utilize Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/CRISPR-associated (Cas) systems or components of such systems to modify a genome within a cell. CRISPR/Cas systems include

transcripts and other elements involved in the expression of, or directing the activity of, Cas genes. A CRISPR/Cas system can be, for example, a type I, a type II, a type III system, or a type V system (e.g., subtype V-A or subtype V-B). The methods and compositions disclosed herein can employ CRISPR/Cas systems by utilizing CRISPR complexes (comprising a guide RNA (gRNA) complexed with a Cas protein) for site-directed binding or cleavage of nucleic acids. A CRISPR/Cas system targeting a target locus comprises a Cas protein (or a nucleic acid encoding the Cas protein) and one or more guide RNAs (or DNAs encoding the one or more guide RNAs), with each of the one or more guide RNAs targeting a different guide RNA target sequence in the target locus. Optionally, such CRISPR/Cas systems targeting a target locus can further comprise one or more exogenous donor sequences (e.g., targeting vectors) that target the target locus.

[00153] CRISPR/Cas systems used in the compositions and methods disclosed herein can be non-naturally occurring. A non-naturally occurring system includes anything indicating the involvement of the hand of man, such as one or more components of the system being altered or mutated from their naturally occurring state, being at least substantially free from at least one other component with which they are naturally associated in nature, or being associated with at least one other component with which they are not naturally associated. For example, some CRISPR/Cas systems employ non-naturally occurring CRISPR complexes comprising a gRNA and a Cas protein that do not naturally occur together, employ a Cas protein that does not occur naturally, or employ a gRNA that does not occur naturally.

A. Cas Proteins

[00154] Cas proteins generally comprise at least one RNA recognition or binding domain that can interact with guide RNAs. Cas proteins can also comprise nuclease domains (e.g., DNase domains or RNase domains), DNA-binding domains, helicase domains, protein-protein interaction domains, dimerization domains, and other domains. Some such domains (e.g., DNase domains) can be from a native Cas protein. Other such domains can be added to make a modified Cas protein. A nuclease domain possesses catalytic activity for nucleic acid cleavage, which includes the breakage of the covalent bonds of a nucleic acid molecule. Cleavage can produce blunt ends or staggered ends, and it can be single-stranded or double-stranded. For example, a wild type Cas9 protein will typically create a blunt cleavage product. Alternatively, a wild type

Cpf1 protein (e.g., FnCpf1) can result in a cleavage product with a 5-nucleotide 5' overhang, with the cleavage occurring after the 18th base pair from the PAM sequence on the non-targeted strand and after the 23rd base on the targeted strand. A Cas protein can have full cleavage activity to create a double-strand break at a target genomic locus (e.g., a double-strand break with blunt ends), or it can be a nickase that creates a single-strand break at a target genomic locus.

B. Guide RNAs

[00155] A “guide RNA” or “gRNA” is an RNA molecule that binds to a Cas protein (e.g., Cas9 protein) and targets the Cas protein to a specific location within a target DNA. Guide RNAs can comprise two segments: a “DNA-targeting segment” (also called a “guide sequence”) and a “protein-binding segment.” “Segment” includes a section or region of a molecule, such as a contiguous stretch of nucleotides in an RNA. Some gRNAs, such as those for Cas9, can comprise two separate RNA molecules: an “activator-RNA” (e.g., tracrRNA) and a “targeter-RNA” (e.g., CRISPR RNA or crRNA). Other gRNAs are a single RNA molecule (single RNA polynucleotide), which can also be called a “single-molecule gRNA,” a “single-guide RNA,” or an “sgRNA.” *See, e.g.,* WO 2013/176772, WO 2014/065596, WO 2014/089290, WO 2014/093622, WO 2014/099750, WO 2013/142578, and WO 2014/131833, each of which is herein incorporated by reference in its entirety for all purposes. A guide RNA can refer to either a CRISPR RNA (crRNA), or the combination of a crRNA and a trans-activating CRISPR RNA (tracrRNA). The crRNA and tracrRNA can be associated as a single RNA molecule (single guide RNA or sgRNA) or in two separate RNA molecules (dual guide RNA or dgRNA). For Cas9, for example, a single-guide RNA can comprise a crRNA fused to a tracrRNA (e.g., via a linker). For Cpf1 and CasΦ, for example, only a crRNA is needed to achieve binding to a target sequence. The terms “guide RNA” and “gRNA” include both double-molecule (i.e., modular) gRNAs and single-molecule gRNAs. In some of the methods and compositions disclosed herein, a guide RNA is a *S. pyogenes* Cas9 gRNA or an equivalent thereof. In some of the methods and compositions disclosed herein, a guide RNA is a *S. aureus* Cas9 gRNA or an equivalent thereof.

Kits

[00156] The compositions described herein can be provided in a kit for use in treating sensorineural hearing loss or vestibular dysfunction. Compositions may include a polynucleotide of the present disclosure, nucleic acid vectors containing such polynucleotides, and nucleic acid vectors containing a polynucleotide described herein operably linked to a transgene encoding a protein of interest (e.g., a protein that can be expressed in hair cells to treat hearing loss and/or vestibular dysfunction). The nucleic acid vectors may be packaged in an AAV virus capsid (e.g., AAV1, AAV2, AAV2quad(Y-F), AAV6, AAV9, Anc80, Anc80L65, DJ/9, 7m8, or PHP.B). The kit can further include a package insert that instructs a user of the kit, such as a physician, to perform the methods described herein. The kit may optionally include a syringe or other device for administering the composition.

EXAMPLE 1 – TESTING OF VARIOUS TMPRSS3 REGULATORY REGIONS

[00157] As shown in Figure 1, several cell types in the cochlea including supporting cell types, hair cells, and possibly interdental cells express TMPRSS3. However, overexpression with a ubiquitous promoter may result in ototoxicity due to expression of a transgene in non-native cells. Accordingly, it is important to design regulatory elements from the TMPRSS3 regulatory regions that preferentially targets certain cell types in the inner ear. AAV vectors were designed to contain various polynucleotides according to the present disclosure in order to for the ability of different regulatory regions of the human TMPRSS3 region to drive expression of a transgene such as the GFP reporter gene in mouse cochlea. Table 3 below provides a list of some of the polynucleotides tested including an indication of the regulatory regions included in each polynucleotide.

Table 3 – List of polynucleotides tested include which human TMPRSS3 gene regulatory regions are included therein.

Polynucleotide Name	Regulatory Regions	Sequence ID NO
REGN-huAB (Figure 2)	Modified Region A Modified Region B	SEQ ID NO: 18
REGN-huABC	Modified Region A	SEQ ID NO: 19

(Figure 3)	Modified Region B Modified Region C	
REGN-huABE (Figure 4)	Modified Region A Modified Region B Modified Region E	SEQ ID NO: 25
REGN-huAB2B3 (Figure 5)	Modified Region A Region B2 Modified Region B3	SEQ ID NO: 20
REGN-huAB2B3C (Figure 6)	Modified Region A Region B2 Modified Region B3 Modified Region C	SEQ ID NO: 21

Materials and Methods:

Viral vector production:

[00158] REGN-huAB promoter was derived from proximal predicted regulatory regions adjacent and within the human *TMPRSS3* locus. The potential regulatory element was cloned into viral vector in place of the CMV promoter. REGN-huABC is REGN-huAB with an additional regulatory element (Region C) placed 3' of the transgene PolyA sequence. This regulatory element was cloned from a predicted histone acetylation peak within the human *TMPRSS3* locus. REGN-huAB₂B₃ and REGN-huAB₂B₃C were derived from REGN-huAB and REGN-huABC respectively and contain a shortened 5' promoter, removing non-conserved nucleotides from the promoter sequence within Region B while maintaining splice donor and acceptor sites. Region E, within REGN-huABE, was derived from an intron with higher-than-expected conservation within the human *TMPRSS3* locus, retaining splice donor and acceptor sites. This regulatory element was cloned into reporter and transgene viral vectors within the reporter and transgene coding sequence (Between H2B and CMV), thereby serving as an intron to be spliced out to create a full-length mRNA of the reporter or transgene.

Ex-Vivo Culture:

[00159] One day prior to tissue collection, 35mm glass-bottom dishes were prepared with a collagen bubble (Gibco A10483-01) according to manufacturer's protocol and stored in 1x PBS at 4°C overnight. The following day, cochlear explants were prepared by dissecting P1-5 neonatal mouse cochlea in Lebovitz/L-15 media (Thermo Fisher 21083027). Neonatal mice were euthanized, and cochlea were removed from the skull. Sensory epithelium was microdissected from neonatal cochlea by lifting off the modiolus. After removing the stria, organ of Corti explants were placed on a collagen bubble, grown in 35 mm glass bottom dishes in complete culture media (DMEM/F-12 with 7% FBS, 2 mM L-glutamine, 100 U/mL penicillin G) at 37°C and 5% CO₂. On day zero in culture, explants were treated with 5e10 vg (MOI) of AAVDJ virus containing either one of an experimental promoter sequence driving H2B-GFP or a positive control CMV-H2B-GFP. A no virus control was used as a negative control. Three independent explants were treated with the same virus for each experiment. Treated explants were grown at 37°C and 5% CO₂ for 72 hours before being fixed in 4% PFA and analyzed using immunofluorescent staining.

Ex-Vivo Analysis:

[00160] Whole mount explants were analyzed using LSM confocal 780 imaging. For Qualitative analysis, antibodies against exogenous GFP and endogenous Myo7a (to label hair cells) were used. Localization of expression was determined based on established organ of Corti architecture and expression relative to Myo7a-positive inner and outer hair cells.

[00161] For Quantitative assessment, antibodies against exogenous GFP, Pou4f3 (hair cells) and Sox2 (supporting cells) were used. Relative GFP intensity per cell type was then calculated using Imaris image quantification software. Hair cell nuclei were detected in Imaris as nuclei labeled with Pou4f3. Supporting cell nuclei were detected in Imaris as nuclei labeled with Sox2. For each cell type, surface masks were created and intensity of GFP within the volume of the mask was extracted as GFP Intensity Sum. Background Fluorescence in GFP channel was determined by extracting the GFP Intensity Sum in the negative control explant cultures which were also counterstained with anti-GFP. Data plotted as individual cells with ~600-800 hair cells analyzed per treatment from three replicate cultures and ~950-1150 sensory support cells analyzed per treatment from three replicate cultures. Statistics performed in GraphPad Prism software using a

one-way ANOVA, Kruskal-Wallis nonparametric test followed by multiple comparison. P values of equal or less than 0.05 were considered as significant.

In-Vivo AAV Delivery and Analysis:

[00162] Postnatal day 3 or 4 mouse pups (B6.CAST-Cdh23Ahl+/Kjn) were subjected to bilateral posterior semicircular canal (PSCC) surgery to deliver viral transgenes packaged in AAVDJ to the inner ear in accordance with Regeneron's IACUC protocol. Pups were anesthetized using cryoanesthesia and a uniform dose of 3×10^9 viral genomes per ear was delivered in 1 microliter, diluted in 1xPBS with 0.05% fastgreen dye, to each ear. Each pup received the same virus in both ears. Cochlea were collected at postnatal day 20 or 21. A CMV ubiquitous promoter was used as a positive control while uninjected animals were used as negative controls. Cochlea were fixed in either 10% neutral buffered formalin for formalin-fixed, paraffin embedding (FFPE), or 4% paraformaldehyde for whole mount dissection and analysis. Cochlea for FFPE were fixed overnight while cochlea for whole mount were fixed for ~2 hours at room temperature. FFPE cochlea were then decalcified using Immunocal formic acid decalcifier for ~20 hours at room temperature prior to tissue processing and embedding using a Leica TP 1020 automatic tissue processor. Processed cochlea were sectioned and stained for GFP and cell-type specific expression was quantified. Mid-modiolar sections from each ear were selected for analysis. In each turn of the cochlea (apical, mid-, and basal), each cell type was quantified to obtain the total number of each cell type and the total number of GFP+ cells. Data from each turn was averaged to obtain the overall percentage of GFP+ cells for each cell type analyzed. Finally, percentages were converted to scores for heatmap data visualization (Scores: 0 = 0%; 1=0.1% to 33.3%; 2=33.4% to 66.6%; 3=66.7% to 100%).

Results:

[00163] In cochlear explants, REGN-huAB showed expression in sensory support cells of the cochlea (Deiters', pillar, inner and outer phalangeal cells – Figures 7 and 22). Expression in inner and outer hair cells is also detected but with variability. REGN-huABC similarly expressed GFP in sensory support cells of the cochlea (Deiters', pillar, inner and outer phalangeal cells) and had a higher and more consistent inner and outer hair cell expression pattern (Figures 10 and 22). Both REGN-huAB and REGN-huABC also showed high GFP expression in nuclei of the explant

proliferating front, which is a cell type of unknown origin and considered an artifact of the explant culture. Both REGN-huAB₂B₃ and REGN-huAB₂B₃C showed strong expression of GFP in outer sulcus cells with detectable but less expression in inner sulcus cells of the cochlear sensory epithelium (Figures 16 and 19). REGN-huABE also showed strong expression in sensory support cells, particularly with increased expression in the third row of Deiters' cells or Hensen's cells (Figure 13). Decreased hair cell expression was also observed for REGN-huABE (Figure 13).

[00164] To quantify observations from explant cultures, REGN-huAB, REGN-huABC, REGN-huAB₂B₃, and REGN-huABE were further characterized ex-vivo and in-vivo. Quantifying GFP intensity from Pou4f3-positive cells (hair cells) in cochlear explant cultures, it was determined that inclusion of the internal regulatory element Region E significantly decreased hair cell expression in explants treated with REGN-huABE (Figures 14, 15 and 20). Further, shortening of the 5' promoter in REGN-huAB₂B₃ also significantly decreased hair cell expression (Figures 17, 18 and 20). Inclusion of the 3' regulatory element Region C in REGN-huABC did not significantly affect reporter expression over huAB promoter alone (Figures 8, 9, 11, 12 and 20).

[00165] Sensory support cells of cochlear explants were labeled using SOX2 and intensity of GFP reporter was extracted from cells in proximity to sensory hair cells. Inclusion of the 3' regulatory element Region C in huABC significantly increased reporter expression over huAB promoter alone (Figure 21). Inclusion of the internal regulatory element Region E in huABE further increased sensory support cell reporter expression significantly (Figure 21). Expression of reporter driven by REGN-huAB₂B₃ was significantly lower than all other tested promoter and regulatory element combinations (Figure 21). Overall, huABE had the highest reporter expression in sensory support cell while shortening the 5' promoter in REGN-huAB₂B₃ led not only to a decrease of reporter expression in hair cells, but also in support cells (Figures 20 and 21).

[00166] To further characterize these elements, in-vivo expression patterns were quantified after direct injection via the posterior semicircular canal. REGN-huAB showed significant enrichment in sensory support cells but not hair cells (Figure 9). Expression was also observed in interdental

cells. REGN-huABC similarly expressed GFP in sensory support cells and interdental cells but also showed minor expression in inner hair cells (Figure 12). REGN-huAB₂B₃ had higher expression in both inner and outer hair cells compared to REGN-huAB and REGN-huABC (Figures 9, 12 and 18). REGN-huAB₂B₃ also had strong expression in Deiters, pillar cells, and interdental cells, but lacked expression in phalangeal cells (Figure 18). Inclusion of regulatory element E (REGN-huABE) drastically increased expression of the GFP reporter in most cell types of the cochlea (Figure 15). GFP expression was cytoplasmic in addition to nuclear, indicating an alternative transcription start site within Region E that is sufficient to drive GFP expression alone. Continued presence of nuclear GFP also supports splicing of Region E to generate the expected H2B-GFP reporter transcript.

[00167] Experiments conducted using the same polynucleotides as listed in Table 3 except that the second modified Region B (SEQ ID NO: 26) was used instead of modified region B (SEQ ID NO: 3) and the second modified Region B3 (SEQ ID NO: 27) was used instead of modified Region B3 (SEQ ID NO: 8), showed similar results from in-vivo and ex-vivo experiments.

[00168] Taken together, the data indicates regulatory elements derived from the human TMPRSS3 locus drive expression in sensory support cells of the cochlea. Expression of GFP in hair cells under the full-length human-derived promoter also indicate a functional role for TMPRSS3 in hair cells (Table 4).

Table 4 - Data summary of the ex-vivo GFP expression driven by various plasmids containing human TMPRSS3 regulatory elements

Plasmid Name	Description	Expression pattern description
REGN-huAB	Long Human-sequence derived promoter	Sensory support cells of the organ of Corti (Deiters' cells, Pillar cells, Phalangeal cells) and proliferating front of organ explant. Low inner and outer hair cell expression.

REGN-huABC	Long Human-sequence derived promoter with human regulatory element 3' of reporter transgene	Sensory support cells of the organ of Corti (Deiters' cells, Pillar cells, Phalangeal cells) and proliferating front of organ explant. Increased inner and outer hair cell expression compared to REGN-huAB.
REGN-huABE	Long Human-sequence derived promoter with a human regulatory intron within the coding sequence of reporter transgene	Sensory support cells of the organ Corti (Deiters' cells and Hensen's cells). Hair cell expression.
REGN-huAB2B3	Short Human-sequence derived promoter	Non sensory support cell expression (potentially inner and outer sulcus cells) and proliferating front. Low/no inner and outer hair cell expression.
REGN-huAB2B3C	Short Human-sequence derived promoter with human regulatory element 3' of reporter transgene	Non sensory support cell expression (potentially inner and outer sulcus cells) and proliferating front. Low/no inner and outer hair cell expression.

[00169] The present disclosure has been described in detail, including the preferred embodiments thereof. However, it will be appreciated that those skilled in the art, upon consideration of the present disclosure, may make modifications and/or improvements on this disclosure that fall within the scope and spirit of the disclosure.

SEQUENCE LISTING**SEQ ID NO: 1 – Modified Region A**

ATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTAAAGGGAAGCCACA
TTTTTCCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTTGATAGGCCACG
GCAGCTTGGGTACACAGATGCCCCGCGATAATGCCATTTGCTGCGATGATTGAGTCC
ACAAGGCTGAGTGACTTGGGGAAAAATATCCCCGATTGGACTAACCTGTTGGCGGCC
TGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGCTTTTAATGA
TGGCAGCTCCAGAAATGGTGAGTAAAT

SEQ ID NO: 2 – Unmodified Region A

ATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTAAAGGGAAGCCACA
TTTTTCCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTTGATAGGCCACG
GCAGCTTGGGTACACAGATGCCCCGCGATAATGCCATTTGCTGCGATGATTGAGTCC
ACAAGGCTGAGTGACTTGGGGAAAAATATCCCCGATTGGACTAACCTGTTGGCGGCC
TGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGCTTTTAATGA
TGGCAGCTCCAGAAATGGTGAGTAAAT

SEQ ID NO: 3 – Modified Region B

GACCGGGCACCGGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGA
GCCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTC
CGTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACCGAACTACTTA
TGCTTATGGTAGGTGTACCACACGCACAGGGAAAGAATTGCGTTTATGTGGGACAG
GTGAAAACAATCGCAAAAAAGCAATGGAAAGGGCTTTGAGAGTAATTTATCTTCTG
AACAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCTAAAGCAATG
GATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGCAGAATTAAGT
TCATACTTTTTTTTTTTTTTTTTTTTGCTAATTCAGAAGCTTTTCTGTTATCTGCTGT
AGCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTTGGTTAATGGATACTGAA
TGCGCTTTGTGTGGACTGTCGAATTTCAAAGATTTACCGTATGACCAAGATGCACCT
GATGCTACAAGTATAAATAGGGGAACAAATGCTTTCTGTTCTTCCTCGGCTAAGGAG
GTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC

SEQ ID NO: 4 – Unmodified Region B

GACCGGGCACCGGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGA
GCCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTC
CGTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACCGAACTACTTA
TGCTTATGGTAGGTGTACCACACGCACAGGGAAAGAATTGCGTTTATGTGGGACAG

GTGAAAACAATCGCAAAAAAGCAATGGAAAGGGCTTTGAGAGTAATTTATCTTCTG
AACAAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCTAAAGCAATC
GATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGCAGAATTAAGT
TCATACTTTTTTTTTTTTTTTTTTTTTTGGCTAATTCTAGAAGCTTTTCTGTTATCTGCTGT
AGCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTTGGTTAATCGATACTGAA
TGCCTTTGTGTGGACTGTGAATTTCAAAGATTTACCGTATGACCAAGATGCACCT
GATGCTACAAGTATAAATAGGGGAACAATGCTTTCTGTTCTTCTCGGCTAAGGAG
GTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCTGAAATAGTCAC

SEQ ID NO: 5 – Modified Region C

AATATCACCTTCACCTTTCTATGGCATTAATAATTAATAATTTTTCTAAATGTTGACA
AACTATGGCTGATGGGTCAAATCTGGCCCACCCCTTCTTTTGTAATAAAGTTTTAT
TGGAACACAGACATGCTCATTGTTCACATACTGTCTGTGGCTGCTTTCACGCCATA
ACAGCAGAGTGGGGTCGTTAGGACAGACCGCATGGCCCATAAAGTTTAAGTGCTGT
ATTTACTGTCTGGTCCCTTACAGAAAATCTCTCCAACCGTTTTCTATATCATCTTGTCT
CTCCTATAGGTATTCCATTATATGTGTACCTCATTGTTATTTACCCGTTCTTTTTTG
TGGGTAATTTGGTCCGTTTCCTGTGCATCAATATTATAAGTGCCATTCTTCTAAGTGT
CCTTGGGGGTGCGATATTTGCGTATATCTATGGTGACATTTTAGGATAAATTCCTAGA
AGTGGAAGTCTAGGTCAAACTTTCATCTGTTTTTGAGCTTTCCGGTGTCTAAGTTT
ACACAGCTGGGAAGTTTACACAGTTTCTACCCCTGTGACTGTAGGTGGGGGACCATT
TCCTTTCACTTCATGTGAATCATTGTGAAGTTGGCATGTAGTTATGTCTCATTGTGCAT
TGTTTTCACTTCTCAAGAAGTTGAAAGCTTTTTTTTTTCTTTTTCTTGAGACGAGGTC
TCGCTCTGTGCGCCAGGCTGGAGTGCAATGGCGTGATCTCGGCTCACTGCAACCTCC
AACCGCCAGGGTTCAAGCGATTCTCTTGCTCGGCCTCCTGAGTTGCTGGAATTACA
GGCACATGCCCCACACCCAGCTAATTTTTGTATTTTAGTAGAGGCGGGTTTCACC
GTGTTGGGCAGGCTGGTCAAGAAGTTGAAAGCTTTTTCATACTTGTGCGCCATTGTC
ATTCCTTCTTTGGAGAAATGCTGTACATTTTATTGTTTATTCCAGGCAATGTCTCAA
GGACCAAGCACAGAGGAGGCGTGGGGCACCTGCAGTGCCCTCAGGGCAGCCTGACA
GAGAAGCAGGGCAAACGGGAAGTACACAGCCTTTCGATTAAATGTCAATCTTTGA
TTTGACAATTTGGGGTCAATTTCTAAATCCAACCTTAAATACAGAGCAGCTGGAATGT
GAAGATGTCATTCTGGGGTTCTGTCAGATTTCTAAGTATAAGACGAAGGGTGGGCA
TCCTTATAGAGGAGAATTGTTAGTGTGCGGGGGATCGAGGTGTGAGGGTGGAGAGA
GAATCCTTTCTGCCAGCTCACCACCGCAGATGGGCTGTCCTGGAAGGACTTGGTCTC
AGACCCTGAGTTAAAACGGCCTTCAT

SEQ ID NO: 6 – Unmodified Region C

AATATCACCTTCACCTTTCTATGGCATTAATAATTAATAATTTTTCTAGATGTTGACA
AACTATGGCTGATGGGTCAAATCTGGCCCACCCCTTCTTTTGTAATAAAGTTTTAT
TGGAACACAGACATGCTCATTGTTCACATACTGTCTGTGGCTGCTTTCACGCCATA
ACAGCAGAGTGGGGTCGTTAGGACAGACCGCATGGCCCATAAAGTTTAAGTGCTGT

ATTTACTGTCTGGTCCTTTACAGAAAATCTCTCCAACCGTTTTCTATATCATCTTGTC
CTCCTATAGGTATTCCATTATATGTGTACCTCATTGTTATTTACCCGTTCTTTTTTG
TGGGTAATTTGGTCCGTTTCCTGTGCATCAATATTATAAGTGCCATTCTTCTAAGTGT
CCTTGGGGGTCGATATTTGCGTATATCTATGGTGACATTTTAGGATAAAATTCCTAGA
AGTGGAAGTCTAGGTCAAACTTTTCATCTGTTTTTGAGCTTTCGGGTGTCTAAGTTT
ACACAGCTGGGAAGTTTACACAGTTTCTACCCCTGTGACTGTAGGTGGGGGACCATT
TCCTTTCACTTCATGTGAATCATTGTGAAGTTGGCATGTAGTTATGTCTCATTTCAT
TGTTTTCATTTCTCAAGAAGTTGAAAGCTTTTTTTTTTTCTTTTTCTTGAGACGAGGTC
TCGCTCTGTGCGCCAGGCTGGAGTGCAATGGCGTGATCTCGGCTCACTGCAACCTCC
AACCGCCCCGGGTTCAAGCGATTCTCTTGCCCTCGGCCCTCTGAGTTGCTGGAATTACA
GGCACATGCCCCACACCCAGCTAATTTTTGTATTTTAGTAGAGGCGGGTTTCACC
GTGTTGGGCAGGCTGGTCAAGAAGTTGAAAGCTTTTTCATACTTGTGCGCCATTTC
ATTCCTTCTTTGGAGAAATGCTGTACATTTTCATTGTTTATTCCAGGCAATGTCCTCAA
GGACCAAGCACAGAGGAGGCGTGGGGCCCCCTGCAGTGCCCTCAGGGCAGCCTGACA
GAGAAGCAGGGCAAACGGGAAGTACACAGCCTTTCGATTAAATGTCAATCTTTGA
TTTGACAATTTGGGGTCAATTTCTAAATCCAACCTAAATACAGAGCAGCTGGAATGT
GAAGATGTCATTCTGGGGTTCTGTCAGATTTCTAAGTATAAGACGAAGGGTGGGCA
TCCTTATAGAGGAGAATTGTTAGTGTGCGGGGGATCGAGGTGTGAGGGTGGAGAGA
GGATCCTTTCTGCCAGCTCACCACCGCAGATGGGCTGTCCTGGAAGGACTTGGTCTC
AGACCCTGAGTTAAAACGGCCTTCAT

SEQ ID NO: 7 – Unmodified region B2

GACCGGGCACCGGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCCGTGTGA
GCCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTC
CGTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACC

SEQ ID NO: 8 – Modified Region B3

ATCAAGAAGCTTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCAG
ATGACTGTTTTTGGTTAATGGATACTGAATGCGCTTTGTGTGGACTGTCGAATTTCAA
AGATTTACCGTATGACCAAGATGCACCTGATGCTACAAGTATAAAATAGGGGAACAA
ATGCTTTCTGTTCTTCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCA
GAGGTCCTGAAATAGTCAC

SEQ ID NO: 9 – Unmodified Region B3

ATTCTAGAAGCTTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCAG
ATGACTGTTTTTGGTTAATCGATACTGAATGCGCTTTGTGTGGACTGTCGAATTTCAA
AGATTTACCGTATGACCAAGATGCACCTGATGCTACAAGTATAAAATAGGGGAACAA

ATGCTTTCTGTTCTTCCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCA
GAGGTCCTGAAATAGTCAC

SEQ ID NO: 10 – Modified Region B4 (Region B- [Region B2 + Region B3])

GAAACTACTTATGCTTATGGTAGGTGTACCACACGCACAGGGAAAGAATTGCGTTTA
TGTGGGACAGGTGAAAACAATCGCAAAAAAGCAATGGAAAGGGCTTTGAGAGTAAT
TTATCTTCTGAACAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCT
AAAGCAATGGATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGC
AGAATTAAGTTCATACTTTTTTTTTTTTTTTTTTTTGCTA

SEQ ID NO: 11 - Unmodified Region B4 (Region B- [Region B2 + Region B3])

GAAACTACTTATGCTTATGGTAGGTGTACCACACGCACAGGGAAAGAATTGCGTTTA
TGTGGGACAGGTGAAAACAATCGCAAAAAAGCAATGGAAAGGGCTTTGAGAGTAAT
TTATCTTCTGAACAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCT
AAAGCAATCGATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGC
AGAATTAAGTTCATACTTTTTTTTTTTTTTTTTTTTGCTA

SEQ ID NO: 12- Unmodified Region B2 + modified Region B3

GACCGGGCACC GGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGA
GCCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTC
CGTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACCATTCAAGAAGC
TTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTT
GGTTAATGGATACTGAATGCGCTTTGTGTGGACTGTCGAATTTCAAAGATTTACCGT
ATGACCAAGATGCACCTGATGCTACAAGTATAAATAGGGGAACAAATGCTTTCTGTT
CTTCCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAA
ATAGTCAC

SEQ ID NO: 13- Unmodified Region B2 + unmodified Region B3

GACCGGGCACC GGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGA
GCCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTC
CGTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACCATTCTAGAAGC
TTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTT
GGTTAATCGATACTGAATGCGCTTTGTGTGGACTGTCGAATTTCAAAGATTTACCGT
ATGACCAAGATGCACCTGATGCTACAAGTATAAATAGGGGAACAAATGCTTTCTGTT
CTTCCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAA
ATAGTCAC

SEQ ID NO: 14 - hg38_dna range=chr21: 42,395,419-42397117

ATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTAAAGGGAAGCCACA
TTTTTCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTTGATAGGCCACG
GCAGCTTGGGTACACAGATGCCCCGCGATAATGCCATTTGCTGCGATGATTGAGTCG
ACAAGGCTGAGTGACTTGGGGAAAAATATCCCCGATTGGACTAACCTGTTGGCGGCC
TGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGCTTTTAATGA
TGGCAGCTCCAGAAATGGTGAGTAAATGGGTGTGAGGAGAGAAGATGCTTCTCAA
CACTTCTCAGTGTTGTACAGCTTGCCAACCACAGCCTGTTTGGGAGGGGATGTTCTC
ACCCAGGTTCCATCTTTCAGAAGCCATTTGCCCACTTATTCTTTTTATATGGATAAAC
GATGTAATTTGCTACATGCTGATAATAAACAGAGGTATCAGCAAGAATGAAAGCCA
GAAAAATGGCCCAACATGAACTTCTTGCTTTTCTAAGTTACTCTTTCAAGCTCTTCTTG
CGATAGAAATGAAATGAACATTTTGGCACAGAACTCTTGTTTAGATTTTACTATGA
TCTTGGGAAACACGGGTTGCTTCAAATGGCTTACTAGATCCCTTTGGAATTTCAAAC
GTTACAACAGGTAGTCACAAAGGGAATGAAAGCGGATGATCTGAGTGTCTGTGCTG
ATGTGCTGCTGGAGCCCCTGTGGATTTTAGCCCTTGAGGAAGCAGGTTCCCAGAG
CAGCCATGCCTGATATCTGGGAGAGGGAAGTAAATTAATTGGGAATCTTCTGCTA
TTTCCCTGCTTCCTCCGCCCTCTCAGAGTTACAGGGAATGACTTAGAAAGGTCTCTG
GTCTCCTTGGCATTGTGATAGGTGGAATTCTTTCGTAGGAAACACAGCCCCGGGGTT
ATTTTGATGCAAGCCGTTGGTCCAAAGCCACTCCCCTTTGCTATGTGGGGGAAGG
ACCGGGCACCGGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGAG
CCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCC
GTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACCGAACTACTTAT
GCTTATGGTAGGTGTACCACACGCACAGGGAAGAATTGCGTTTATGTGGGACAGG
TGAAAACAATCGCAAAAAAGCAATGGAAGGGCTTTGAGAGTAATTTATCTTCTGA
ACAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCTAAAGCAATCG
ATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGCAGAATTAAGTT
CATACTTTTTTTTTTTTTTTTTTTTGCTAATTCTAGAAGCTTTTCTGTTATCTGCTGTA
GCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTTGGTTAATCGATACTGAAT
GCGCTTTGTGTGGACTGTCGAATTTCAAAGATTTACCGTATGACCAAGATGCACCTG
ATGCTACAAGTATAAATAGGGGAACAAATGCTTTCTGTTCTTCCCTCGGCTAAGGAGG
TAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC

SEQ ID NO: 15 - hg38_dna range=chr21: 42,392,583 -42,397,117

ATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTAAAGGGAAGCCACA
TTTTTCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTTGATAGGCCACG
GCAGCTTGGGTACACAGATGCCCCGCGATAATGCCATTTGCTGCGATGATTGAGTCG
ACAAGGCTGAGTGACTTGGGGAAAAATATCCCCGATTGGACTAACCTGTTGGCGGCC
TGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGCTTTTAATGA
TGGCAGCTCCAGAAATGGTGAGTAAATGGGTGTGAGGAGAGAAGATGCTTCTCAA

CACTTCTCAGTGTTGTACAGCTTGCCAACCACAGCCTGTTTGGGAGGGGATGTTCTC
ACCCAGGTTCCATCTTTCAGAAGCCATTTGCCCACTTATTCTTTTTATATGGATAAAC
GATGTAATTTGCTACATGCTGATAATAAACAGAGGTATCAGCAAGAATGAAAGCCA
GAAAATGGCCCAACATGAACTTCTTGCTTTTCTAAGTTACTCTTCAAGCTCTTCTTG
CGATAGAAATGAAATGAACATTTTTGGCACAGAACTCTTGTTTAGATTTTACTATGA
TCTTGGGAAACACGGGTTGCTTCAAATGGCTTACTAGATCCCTTTGGAATTTCAAAC
GTTACAACAGGTAGTCACAAAGGGAATGAAAGCGGATGATCTGAGTGTCTGTGCTG
ATGTGCCTGCTGGAGCCCCGTGGATTTTAGCCCTTGGAGGAAGCAGGTTCCCAGAG
CAGCCATGCCTGATATCTGGGAGAGGGAAGTAAATTTAATTGGGAATCTTCTGCTA
TTCCCTGCTTCCCTCCGCCCTCTCAGAGTTACAGGGAATGACTTAGAAAGGTCTCTG
GTCTCCTTGGCATTGTGATAGGTGGAATTCTTTCGTAGGAAACACAGCCCCGGGGTT
ATTTTGATGCAAGCCGTTGGTCCAAAGCCACTCCCCACTTTGCTATGTGGGGGAAGG
ACCGGGCACCGGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGAG
CCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCC
GTGGTGCCATCTACATTTTTGGGACTCGGGAATTATGTGAGTACCGAACTACTTAT
GCTTATGGTAGGTGTACCACACGCACAGGGAAGAATTGCGTTTATGTGGGACAGG
TGAAAAAATCGCAAAAAAGCAATGGAAAGGGCTTTGAGAGTAATTTATCTTCTGA
ACAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCTAAAGCAATCG
ATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGCAGAATTAAGTT
CATACTTTTTTTTTTTTTTTTTTTTGCTAATTCTAGAAGCTTTTCTGTTATCTGCTGTA
GCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTTGGTTAATCGATACTGAAT
GCGCTTTGTGTGGACTGTCGAATTTCAAAGATTTACCGTATGACCAAGATGCACCTG
ATGCTACAAGTATAAATAGGGGAACAAATGCTTTCTGTTCTTCCCTCGGCTAAGGAGG
TAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCACCATGGGGGAA
AATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCCTTGATG
ATTTGAAAATAAGTCCTGTTGCACCAGGTAAGTGGACCCAGGTGAGGACTCTGTGAT
TTCTGCCCATAACCTCATGTAGGTGACCAATGTGACTGACTGTCCCTGTGGGGGAAA
TATCTCCCCAGCCATTCTGACACCCACAGGCTGGACACCTGCACTTCCCTAGATCTG
CAGAATCTCAGGGAGAAGGGGCATTGGAGAGGGGATCGTTTCGTAAAGCCCTTGC
TCTCTCCCTGGAGACCGGTGTTTTCTTCTTGTGAGGTTTTAGAGACTGGGGCT
CCACAATTGTCTGTCAATCCTGAAGGAGGTCAGATCCTGGCCAGGAAATCTCTGAG
TCCTCCAGGAAGTCCTGAGAAGCAGTGGCCACCAGTCCCTCCGAAAACCTCAGGATA
AACTGGATGTCGCCAAAGATATTAGAGGCCTAGATTAAAATCCAGCAGCAGATGAC
GGAGTGGCTCCTAACAACCTGAAGGAAAAAAGTTGGTTTCGTAAATAAGTCACAAAT
GCCTGATAAAACAGTGGACTCTCTAGCAACCTATACGGCTAGGATCTTTTTGACCAG
AGATAGTCAGGACTGGTCAAAGGGAGCTGTGGGGTTGCCGGTTCTGGAAGTCAGGG
CTTGGACGCAGGCAGATTCTGACACATTGGTGCTAGGGAGGAACACAGGAGCTC
TGCCTGAGGTTCCAGAAAGTGGGAGTGGGGGTACCCCTAATGATTTCTTTTAAACCA
ACTGGAAAATTGGCTTGGAGGAGGTTTGGAAAGGATGGGAGGTGGGAGAAGAGGTT
TCAAGTAACTGAAAAGACACACATTTAAAAATATGGCTCAAATTAGCTAAGTGTC
AATGAAAGAAAAAAAATCTTGCAATCACCTGACTCCCCAAACTTACCTGGTTTTTA
TGTCTGCAAAGTCTCTTCTGCTTTTGGCCGTATGTGTGGTTTCATGCACAGTACAGA
TATAGGCAGGTGGGTTTTTCCATAAAAGGTTTTTATAAATTTTTTCTTATGGAATCA

CAGTATTCCATCAAGTAGATATGCAATTTACACAGCCATTTGATCATAGTTTGACCT
ATAGTTGTTCCCTAATATTTTATAATATAAATAATATGAATATTAATAGCACTATAATA
TTCCTTTCTCTTATTTTTTGGATTATTTTGCTATGACTAAGTAGTGGGGAGGAACAAAA
ACTTTTCCTTGTAAGTAGAGTAATGTATTGTTTCATTAGAAGAACAATTTAGAAAATA
CAGTTCATTGAAAAGATTTTTTAAAAATTACCCACGACGGCATCATTACTCAGAGATA
ATCCTATTAGCAATTTATTATTGAATTTCTTGCCTTTTTCTATCTATATATCTATATTT
CCTTTGGAAGTAACATGTACTATTTGTATTATTTTGAAGCTTTTTAAAAAATTTTAACA
ATATCACCTTCACCTTTCTATGGCATTAAATAAATTAATATTTTCTAGATGTTGACAA
ACTATGGCTGATGGGTCAAATCTGGCCACCCCTTCTTTTGTAATAAAGTTTTATT
GGAACACAGACATGCTCATTGTTACATACTGTCTGTGGCTGCTTTCACGCCATAA
CAGCAGAGTGGGGTCGTTAGGACAGACCGCATGGCCATAAAGTTTAAGTGCTGTA
TTTACTGTCTGGTCCTTTACAGAAAATCTCTCAACCGTTTTCTATATCATCTTGTCTT
CCTATAGGTATTCCATTATATGTGTACCTCATTGTTATTTACCCGTTCTTTTTTGTG
GGTAATTTGGTCCGTTTCCTGTGCATCAATATTATAAGTGCCATTCTTCTAAGTGTCC
TTGGGGGTCGATATTTGCGTATATCTATGGTGACATTTTAGGATAAATTCCTAGAAG
TGGAAGTGCTAGGTCAAACTTTCATCTGTTTTTGAGCTTTCGGGTGTCTAAGTTTAC
ACAGCTGGGAAGTTTACACAGTTTCTACCCCTGTGACTGTAGGTGGGGGACCATTTC
CTTTCATTCATGTGAATCATTGTGAAGTTGGCATGTAGTTATGTCTCATTGTCATTG
TTTTCATTTCTCAAGAAGTTGAAAGCTTTTTTTTTTCTTTTCTTGAGACGAGGTCTC
GCTCTGTGCGCCAGGCTGGAGTGCAATGGCGTGATCTCGGCTCACTGCAACCTCCAA
CCGCCCCGGTTCAAGCGATTCTCTTGCCTCGGCCTCCTGAGTTGCTGGAATTACAGG
CACATGCCCCCACACCCAGCTAATTTTTGTATTTTAGTAGAGGCGGGTTTACCCGTG
TTGGGCAGGCTGGTCAAGAAGTTGAAAGCTTTTTCATACTTGTGCGCCATTTGCATT
CCTTCTTTGGAGAAATGCTGTACATTTTCATTGTTTATTCCAGGCAATGTCCTCAAGGA
CCAAGCACAGAGGAGGCGTGGGGCCCCCTGCAGTGCCCTCAGGGCAGCCTGACAGAG
AAGCAGGGCAAACGGGAAGTACACAGCCTTTCGATTAAATGTCAATCTTTGATTTG
ACAATTTGGGGTCATTTTCTAAATCCAACCTTAAATACAGAGCAGCTGGAATGTGAAG
ATGTCATTCCTGGGGTTCTGTCAGATTTCTAAGTATAAGACGAAGGGTGGGCATCCT
TATAGAGGAGAATTGTTAGTGTGCGGGGGATCGAGGTGTGAGGGTGGAGAGAGGAT
CCTTTCTGCCAGCTCACCACCGCAGATGGGCTGTCTGGAAGGACTTGGTCTCAGAC
CCTGAGTTAAACGGCCTTCAT

SEQ ID NO: 16 – Sequence 5' of Region A in REGN-huABC

TATGAGTATTCAACATTTCCGTGTGCGCCCTTATCCCTTTTTTGCGGCATTTTGCCTTC
CTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGG
GTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTT
TTCGCCCCGAAGAACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCG
CGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATT
CTCAGAATGACTTGGTTGAGTACTACCAAGTCACAGAAAAGCATCTTACGGATGGCA
TGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
AACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAAC

ATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCAT
ACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCA
AACTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGA
TGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGT
TTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCAC
TGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAA
GCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTT
CATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACAAA
ATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAA
GGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAC
CACCGCTACCAGCGGTGGTTTTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGA
AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGT
AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAA
TCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT
CAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTGGGGCTGAACGGGGGGTTCGTGC
ACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGA
GCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAA
GCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTG
GTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGA
TGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACG
GTTCTGCGCTTTTGCTGGCCTTTTGCTCACATGTCCTGCAGGCAGCTGCGCGCTCGC
TCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTGCACC
GGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGG
GTTCTGCGGCCGACGCGT

SEQ ID NO: 17 - Sequence 3' of Region C in REGN-huABC

GCTAGCCCTCCTCGCGAGTAACCACGTGCGCGGACCGAGCGGCCGAGGAACCCCT
AGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCG
ACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAG
CGCGCAGCTGCCTGCAGGGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCG
GTATTTACACCCGCATACGTCAAAGCAACCATAGTACGCGCCCTGTAGCGGCGCATT
AAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCC
TAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTCTCGCCACGTTCGCCGGCTTTCCC
CGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCAC
CTCGACCCCAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGA
TAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGT
TCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGA
TTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACG
CGAATTTTAACAAAATATTAACGTTTACAATTTTATGGTGCCTCTCAGTACAATCTG
CTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCC

CTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGG
GAGCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGGAGACGAAAGG
GCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGAC
GTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTATTTTTCTAA
ATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAA
TATTGAAAAAGGAAGAG

SEQ ID NO: 18 - REGN-huAB

TTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG
GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTCCCCCTGGAAGCTCCCT
CGTGCGTCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT
CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGC
CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACCTCACGTAAAGGATTTTGGTTCATGAGGTACCCTGCGCGCTCGCTCGCTC
ACTGAGGCCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTTCGCCCGGCCCTC
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGCCAACTCCATCACTAGGGGTTCCT
GCGGCCGCACGCGTATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTA
AAGGGAAGCCACATTTTTCCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTT
GATAGGCCACCGCAGCTTGGGTACACAGATGCGCGGATAATGCCATTTGCTGCGA
TGATTGAGTCCACAAGGCTGAGTGACTTGGGGAAAATATTCCCGATTGGACTAACCT
GTTGGCGGCCTGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGC
TTTTAATGATGGCAGCTCCAGAAATGGTGAGTAAAATGACCGGGCACCGGACGGCTC
GGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGAGCCAGGAAAGGGCTGTGTTTA
TGGAAGCCAGTAACACTGTGGCCTACTATCTTCCGTGGTGCCATCTACATTTTTG
GGACTCGGGAATTATGTGAGTACCGAAACTACTTATGCTTATGGTAGGTGTACCACAC
GCACAGGGAAAGAATTGCGTTTATGTGGGACAGGTGAAAACAATCGCAAAAAAGCA
ATGGAAAGGGCTTTGAGAGTAATTTATCTTCTGAACAATAGCAATATGGCAACTTCTA
AATGGTGAGAGGGAGTCCCTCTAAAGCAATGGATTGAAGATTGGTTGGACAAACAA
TGGGAAAGTCATTGTCTTAGCAGAATTAAGTTCATACTTTTTTTTTTTTTTTTTTGC
TAATTCAAGAAGCTTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACC
GATGACTGTTTTTGGTTAATGGATACTGAATGCGCTTTGTGTGGACTGTGCAATTTCAA
AGATTTACCGTATGACCAAGATGCACCTGATGCTACAAGTATAAATAGGGGAACAAAT
GCTTTCTGTTCTTCCCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGA
GGTCTGAAATAGTCACATCGATTGAATTCACCATGGGACCAGAGCCAGCGAAGTCT
GCTCCCGCCCCGAAAAAGGGCTCCAAGAAGGCGGTGACTAAGGCGCAGAAGAAAG
GCGGCAAGAAGCGCAAGCGCAGCCGCAAGGAGAGCTATTCCATCTATGTGTACAAG

TTCTGAAGCAGGTCCACCCTGACACCGGCATTTTCGTCCAAGGCCATGGGCATCATGA
ACTCGTTTGTGAACGACATTTTCGAGCGCATCGCAGGTGAGGCTTCCCGCCTGGCGC
ATTACAACAAGCGCTCGACCATCACCTCCAGGGAGATCCAGACGGCCGTGCGCCTGC
TGCTGCCTGGGGAGTTGGCCAAGCACGCCGTGTCCGAGGGTACTAAGGCCATCACCA
AGTACACCAGCGCTAAGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAG
CTGTTACACGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCAC
AAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
GAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAG
GACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC
ACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGA
AGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTG
CAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAA
GCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAAGTTCGAGAGATCTACGGGTGGCATCCCTGTGACCCCTCC
CCAGTGCCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCCACCAGCCTTGTCTTA
ATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTTCTATAATATTATGGGGTGG
AGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGG
GGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAATCTC
CGCCTCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAG
GCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTGGTAGAGACGGGGTTTCACCA
TATTGGCCAGGCTGGTCTCCAACCTCAATCTCAGGTGATCTACCCACCTTGGCCTCC
CAAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCCTTCTGATTTTG
TAGGTGACTTCTTGGTACCGCTAGCCCTCCTCGCGAGTAACCACGTGCGCGGACCG
AGCGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCT
CGCTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCCGGGCG
GCCTCAGTGAGCGAGCGAGCGCGCAGAAGCTTGATTATCAAAAAGGATCTTCACCTA
GATCCTTTTAAATTAAAAATGAAGTTTAAATCAAGCCCAATCTGAATAATGTTACAAC
CAATTAAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAACTGCAATTTATT
CATATCAGGATTATCAATAACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAA
ACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGAC
TCGTCCAACATCAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGTTATCAAGTG
AGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTT
CTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATC
AACCAAACCGTTATTCATTGATGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTG
TTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAG
CGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTT
TCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTG
ATGGTCCGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAA
CATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTT

CCCATAACAAGCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTAT
ACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTCGACGTTTCCCGTTGA
ATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCAT
GATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACGGGCCAGAG
CTGCATCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGA
CGTCATTGTCGATATCATTGTCACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAAGGCCGCGTTGCTGGCGTT

SEQ ID NO: 19 - REGN-huABC

TTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG
GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTT
CGGGAAGCGTGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT
CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGCGC
CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACTCACGTAAAGGGATTTTGGTTCATGAGGTACCCTGCGCGCTCGCTCGCTC
ACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCCTC
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCT
GCGGCCGCACGCGTATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTA
AAGGGAAGCCACATTTTTCCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTT
GATAGGCCACCGGCAGCTTGGGTACACAGATGCCCGCGATAATGCCATTTGCTGCGA
TGATTGAGTCCACAAGGCTGAGTGACTTGGGGAAAAATATTCCCCGATTGGACTAACCT
GTTGGCGGCTGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATATTTCTGC
TTTAATGATGGCAGCTCCAGAAATGGTGAGTAAATGACCGGGCACCGGACGGCTC
GGGTACTTTCTGTTCTTAATTAGGTCATGCCCGTGTGAGCCAGGAAAGGGCTGTGTTTA
TGGGAAGCCAGTAACACTGTGGCCTACTATCTTCCGTGGTGCCATCTACATTTTTG
GGACTCGGGAATTATGTGAGTACCGAACTACTTATGCTTATGGTAGGTGTACCACAC
GCACAGGGAAAGAATTGCGTTTATGTGGGACAGGTGAAAACAATCGCAAAAAAGCA
ATGGAAAGGGCTTTGAGAGTAATTTATCTTCTGAACAATAGCAATATGGCAACTTCTA
AATGGTGAGAGGGAGTCCCTCTAAAGCAATGGATTTGAAGATTGGTTGGACAAACAA
TGGGAAAGTCATTGTCTTAGCAGAATTAAGTTCATACTTTTTTTTTTTTTTTTTTGC
TAATTCAAGAAGCTTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCA
GATGACTGTTTTTGGTTAATGGATACTGAATGCGCTTTGTGTGGACTGTGCAATTTCAA
AGATTTACCGTATGACCAAGATGCACCTGATGCTACAAGTATAAATAGGGGAACAAAT
GCTTTCTGTTCTTCCCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGA
GGTCTGTAAATAGTCACATCGATTGAATTCACCATGGGACCAGAGCCAGCGAAGTCT
GCTCCCGCCCCGAAAAAGGGCTCCAAGAAGGCGGTGACTAAGGCGCAGAAAGAAAG
GCGGCAAGAAGCGCAAGCGCAGCCGCAAGGAGAGCTATTCCATCTATGTGTACAAGG
TTCTGAAGCAGGTCCACCTGACACCGGCATTTTCGTCCAAGGCCATGGGCATCATGA
ACTCGTTTGTGAACGACATTTTCGAGCGCATCGCAGGTGAGGCTTCCCGCCTGGCGC
ATTACAACAAGCGCTCGACCATCACCTCCAGGGAGATCCAGACGGCCGTGCGCCTGC
TGCTGCCTGGGGAGTTGGCCAAGCACGCCGTGTCCGAGGGTACTAAGGCCATCACCA
AGTACACCAGCGCTAAGGATCCACCGGTGCGCCACCATGGTGAGCAAGGGCGAGGAG
CTGTTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCAC
AAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT

GAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAAGTGTTCAGCCGCTACCCCGACCACATGAAGCAGCACGA
CTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAG
GACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC
ACAAGCTGGAGTACAACATAACAGCCACAACGTCTATATCATGGCCGACAAGCAGA
AGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTG
CAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAA
GCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAACTCGAGAGATCTACGGGTGGCATCCCTGTGACCCCTCC
CCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCCAACGACCTTGTCTTA
ATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTTCTATAATATTATGGGGTGG
AGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGG
GGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAATCTC
CGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAG
GCATGCATGACCAGGCTCAGCTAATTTTGTTTTTTTTGGTAGAGACGGGGTTTCACCA
TATTGGCCAGGCTGGTCTCCAACCTCTAATCTCAGGTGATCTACCCACCTTGGCCTCC
CAAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCTTCTGATTTTG
TAGGTCGACAATATCACCTTACCTTTCTATGGCATTAAATAATTAATATTTTCTAAAT
GTTGACAACACTATGGCTGATGGGTCAAATCTGGCCCAACCCCTTCTTTGTAAATAAA
GTTTTIATTGGAACACAGACATGCTCATTTGTTACATACTGTCTGTGGCTGTCTTACG
CCATAACAGCAGAGTGGGGTCGTTAGGACAGACCGCATGGCCATAAAGTTAAGTG
CTGTATTTACTGTCTGGTCTTTACAGAAAATCTCTCCAACCGTTTTCTATATCATCTTG
TCCTCCTATAGGTATTCCATTATATGTGTACCTCATTGTTATTTACCCGTTCTTTTTTG
TGGGTAATTTGGTCCGTTTTCTGTGCATCAATATTATAAGTGCCATTCTTCTAAGTGTC
TTGGGGGTTCGATATTTGCGTATATCTATGGTGACATTTIAGGATAAATTCCTAGAAGTG
GAACTGCTAGGTCAAAACTTTTCATCTGTTTTTTGAGCTTTCCGGTGTCTAAGTTTACAC
AGCTGGGAAGTTTACACAGTTTCTACCCCTGTGACTGTAGGTGGGGGACCATTTCCTT
TCACTTCATGTGAATCATTGTGAAGTTGGCATGTAGTTATGTCTCATTGTCATTGTTTT
ATTTCTCAAGAAGTTGAAAGCTTTTTTTTTTTCTTTTTCTTGAGACGAGGTCTCGCTCT
GTCGCCCAGGCTGGAGTGCAATGGCGTGATCTCGGCTCACTGCAACCTCCAACCGCC
AGGGTTCAAGCGATTCTCTGCCTCGGCCTCCTGAGTTGCTGGAATTACAGGCACATG
CCCCACACCCAGCTAATTTTGTATTTTITAGTAGAGGCGGGTTTCACCGTGTGGGC
AGGCTGGTCAAGAAGTTGAAAGCTTTTTCATACTTGTGCGCCATTGTCATTCTTCTTT
GGAGAAATGCTGTACATTTCAATTGTTTATTCCAGGCAATGTCTCAAGGACCAAGCAC
AGAGGAGGCGTGGGGCACCTGCAGTGCCCTCAGGGCAGCCTGACAGAGAAGCAGG
GCAAACGGGAAGTACACAGCCTTTCGATTAAATGTCAATCTTTGATTTGACAATTTG
GGGTCAATTTTCTAAATCCAACCTTAAATACAGAGCAGCTGGAATGTGAAGATGTCATT
CTGGGGTTCTCTGCAGATTTCTAAGTATAAGACGAAGGGTGGGCATCCTTATAGAGGAG
AATTGTTAGTGTGCGGGGGATCGAGGTGTGAGGGTGGAGAGAGAATCCTTTCTGCCA
GCTCACCACCGCAGATGGGCTGTCCTGGAAGGACTTGGTCTCAGACCCTGAGTTAAA
ACGGCCTTCATGCTAGCCCTCCTCGCGAGTAACCACGTGCGCGGACCGAGCGGCCGC
AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTG
AGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCCGGGCGGCCTCAGTG
AGCGAGCGAGCGCGCAGAAGCTTGATTATCAAAAAGGATCTTCACCTAGATCCTTTTA
AATTAATAATGAAGTTTTAAATCAAGCCCAATCTGAATAATGTTACAACCAATTAACC
AATTCTGATTAGAAAACTCATCGAGCATCAATGAAACTGCAATTTATTCATATCAGG
ATTATCAATAACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCG
AGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAA
CATCAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCA
CCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGA

CTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACC
GTTATTCATTTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGA
CAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAAC
AATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCGGGGAT
CGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGA
AGAGGCATAAATCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGC
AACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCATACAAG
CGATAGATTGTGCGACCTGATTGCCCCGACATTATCGCGAGCCCATTATACCCATATAA
ATCAGCATCCATGTTGGAATTTAATCGCGGCCCTCGACGTTTCCCGTTGAATATGGCTCA
TAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGATGATATAT
TTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACGGGCCAGAGCTGCATCGC
GCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGACGTCATTGT
CGATATCATTGTACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
GGCCGCGTTGCTGGCGTT

SEQ ID NO: 20 - REGN-huAB2B3

TTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG
GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT
CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCGCTGCGC
CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGTGAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACTCACGTAAAGGGATTTTGGTTCATGAGGTACCCTGCGCGCTCGCTCGCTC
ACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCCGCCGGCCTC
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCT
GCGGCCGCACGCGTATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTA
AAGGGAAGCCACATTTTCCCTCCCTGGTCTCTGCAGCCATACCAAGGGCTATTT
GATAGGCCACCGCAGCTTGGGTACACAGATGCCCGCGATAATGCCATTTGCTGCGA
TGATTGAGTCCACAAGGCTGAGTGACTTGGGGAAAAATATTCCCCGATTGGACTAACCT
GTTGGCGCCTGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGC
TTTAATGATGGCAGCTCCAGAAATGGTGAGTAAATGACCGGGCACCGGACGGCTC
GGGTACTTTCTTTAATTAGGTCATGCCCGTGTGAGCCAGGAAAGGGCTGTGTTTA
TGGGAAGCCAGTAACACTGTGGCCTACTATCTTCCGTGGTGCCATCTACATTTTGTG
GACTCGGGAATTATGTGAGTACCATTCAAGAAGCTTTTCTGTTATCTGCTGTAGCTC
AGACGAAAATGCATTCTCACCAGATGACTGTTTTTGGTTAATGGATACTGAATGCGCT
TTGTGTGGAAGTGTGCAATTTCAAAGATTTACCGTATGACCAAGATGCACCTGATGCTA
CAAGTATAAATAGGGGAACAAATGCTTTCTGTTCTTCCCTCGGCTAAGGAGGTAGAGGT
GGAGGCGGAGCCGGATGTCAGAGGTCTGAAATAGTCACATCGATTGAATTCACCAT
GGGACCAGAGCCAGCAAGTCTGCTCCCGCCCCGAAAAAGGGCTCCAAGAAGGCG
GTGACTAAGGCGCAGAAGAAAGGCGGCAAGAAGCGCAAGCGCAGCCGCAAGGAGA
GCTATTCCATCTATGTGTACAAGGTTCTGAAGCAGGTCCACCCTGACACCGGCATTTTC
GTCCAAGGCCATGGGCATCATGAACTCGTTTGTGAACGACATTTTCGAGCGCATCGCA
GGTGAGGCTTCCCGCCTGGCGCATTACAACAAGCGCTCGACCATCACCTCCAGGGAG
ATCCAGACGGCCGTGCGCCTGCTGCTGCTGGGGAGTTGGCCAAGCACGCCGTGTCC
GAGGGTACTAAGGCCATCACCAAGTACACCAGCGCTAAGGATCCACCGGTGCGCCACC

ATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCC
CTGGCCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAAGTGCTTCAGCCGCTACCC
CGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCA
GGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA
AGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG
GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGT
CTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGATCCGCCA
CAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCA
TCGGCGCAGGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACCTCGAGAGATCTACG
GGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCC
AGTGCCCAACAGCCTTGTCTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGT
CCTTCTATAATATTATGGGGTGGAGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGA
AGACAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCA
CAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGC
CTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTT
TGGTAGAGACGGGGTTTCACCATATTGGCCAGGCTGGTCTCCAACCTCAATCTCAGG
TGATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCC
CTTCCCTGTCTTCTGATTTTGTAGTCTGACTTCTTGGTACCGCTAGCCCTCTCGCGA
GTAACCACGTGCGCGGACCGAGCGGCCGAGGAACCCCTAGTGATGGAGTTGGCCA
CTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGAC
GCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAAGCTTGATT
ATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAAGC
CCAATCTGAATAATGTTACAACCAATTAACCAATTCTGATTAGAAAACTCATCGAGC
ATCAAATGAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGC
CGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCC
TGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCTC
GTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGA
GAATGGCAAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGC
TCGTATCAAAAATCACTCGCATCAACCAAAACCGTTATTCAATCGTGATTGCGCCTGAG
CGAGACGAAATACGCGATCGCTGTAAAAGGACAATTACAAACAGGAATCGAATGCA
ACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTC
TTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCAT
CAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATCCGTCAGCCAGT
TTAGTCTGACCATCTCATCTGTAAATCATTTGGCAACGCTACCTTTGCCATGTTTCAGA
AACTACTTGGCGCATCGGGCTTCCCATACAAGCGATAGATTGTGCGACCTGATTGCC
CGACATTATCGCGAGCCCATTTATACCATATAAATCAGCATCCATGTTGGAATTTAATC
GCGGCCTCGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTTACTGTTTATG
TAAGCAGACAGTTTTATTGTTTCATGATGATATATTTTATCTTGTGCAATGTAACATCAG
AGATTTTGAGACACGGGCCAGAGCTGCATCGCGCGTTTCGGTGATGACGGTGAAAAC
CTCTGACACATGCAGCTCCCGGACGTCATTGTGATATCATTGTCACATGTGAGCAAA
AGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT

SEQ ID NO: 21 - REGN-huAB2B3C

TTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAG
GTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTCTCCCTT

CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT
CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCCACCGCTGCGC
CTTATCCGGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CTCTGCTGAAGCCAGTTACCTTCGGA AAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACCTCACGTAAAGGGATTTTGGTTCATGAGGTACCCTGCGCGCTCGCTCGCTC
ACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCGGGCGACCTTTGGTTCGCCCGGCCCTC
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGCCAACTCCATCACTAGGGGTTCCCT
GCGGCCGCACGCGTATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTA
AAGGGAAGCCACATTTTCCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTT
GATAGGCCACCGGCAGCTTGGGTACACAGATGCCCGCGATAATGCCATTTGCTGCGA
TGATTGAGTCCACAAGGCTGAGTGACTTGGGGAAAATATTCCCCGATTGGACTAACCT
GTTGGCGGCCTGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGC
TTTAATGATGGCAGCTCCAGAAATGGTGAGTAAATGACCGGGCACCGGACGGCTC
GGGTACTTTCGTTCTTAATTAGGTCATGCCCGTGTGAGCCAGGAAAGGGCTGTGTTTA
TGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTG
GGACTCGGAATTATGTGAGTACCATTCAAGAAGCTTTTCTGTTATCTGCTGTAGCTC
AGACGAAAATGCAATCTCACCAGATGACTGTTTTTGGTTAATGGATACTGAATGCGCT
TTGTGTGACTGTGCAATTTCAAAGATTACCGTATGACCAAGATGCACCTGATGCTA
CAAGTATAAATAGGGGAACAAATGCTTTCTGTTCTTCCCTCGGCTAAGGAGGTAGAGGT
GGAGGCGGAGCCGGATGTGAGAGGTCTGAAATAGTCACATCGATTGAATTCACCAT
GGGACCAGAGCCAGCGAAGTCTGCTCCCGCCCCGAAAAAGGGCTCCAAGAAGGCG
GTGACTAAGGCGCAGAAAGAAAGGCGGCAAGAAGCGCAAGCGCAGCCGCAAGGAGA
GCTATTCCATCTATGTGTACAAGGTTCTGAAGCAGGTCCACCCTGACACCGGCATTTCT
GTCCAAGGCCATGGGCATCATGAACTCGTTTGTGAACGACATTTTCGAGCGCATCGCA
GGTGAGGCTTCCCGCCTGGCGCATTACAACAAGCGCTCGACCATCACCTCCAGGGAG
ATCCAGACGGCCGTGCGCCTGCTGCTGCCTGGGGAGTTGGCCAAGCACGCCGTGTCC
GAGGGTACTAAGGCCATCACCAAGTACACCAGCGCTAAGGATCCACCGGTGCGCCACC
ATGGTGAGCAAGGGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCC
CTGGCCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAAGTCTCAGCCGCTACCC
CGACCACATGAAGCAGCAGGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCA
GGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGA
AGTTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG
GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGT
CTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTTCAGATCCGCCA
CAACATCGAGGACGGCAGCGTGAGCTCGCCGACCACTACCAGCAGAACACCCCCA
TCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACC
GCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACTCGAGAGATCTACG
GGTGGCATCCCTGTGACCCCTCCCCAGTGCTCTCTGGCCCTGGAAGTTGCCACTCC
AGTGCCCAACAGCCTTGTCTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGT
CCTTCTATAATATTATGGGGTGGAGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGA
AGACAACCTGTAGGGCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAAGTGCCA
CAATCTTGGCTCACTGCAATCTCCGCCTCTGGGTTCAAGCGATTCTCCTGCCTCAGC
CTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTT
TGGTAGAGACGGGGTTTACCATATTGGCCAGGCTGGTCTCCAACCTCTAATCTCAGG
TGATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCC

CTTCCCTGTCCTTCTGATTTTGTAGGTCGACAATATCACCTTCACCTTTCTATGGCATTA
AATAATTAAATATTTTCTAAATGTTGACAACTATGGCTGATGGGTCAAATCTGGCCCA
CCCCCTTCTTTTGTAAATAAAGTTTATTGGAACACAGACATGCTCATTTGTTACATA
CTGTCTGTGGCTGCTTTACGCCATAACAGCAGAGTGGGGTCGTTAGGACAGACCGC
ATGGCCCATAAAGTTTAAAGTGCTGTATTTACTGTCTGGTCCTTTACAGAAAATCTCTCC
AACCGTTTTCTATATCATCTTGTCTCTATAGGTATTCCATTATATGTGTACCTCATTTG
TTATTTACCCGTTCTTTTTTGTGGGTAATTTGGTCCGTTTCTGTGCATCAATATTATA
AGTGCCATTCTTCTAAGTGTCCTTGGGGGTCGATATTGCGTATATCTATGGTGACATT
TAGGATAAATTCCTAGAAAGTGGAAGTGTAGGTCAAACTTTTCATCTGTTTTTGAGCT
TTCCGGTGTCTAAGTTTACACAGCTGGGAAGTTTACACAGTTTCTACCCCTGTGACTG
TAGGTGGGGGACCATTTCCTTTCATTTCTCAAGAAGTTGAAAGCTTTTTTTTTCTTTT
ATGCTCATTTGCAATTGTTTTCTATTCTCAAGAAGTTGAAAGCTTTTTTTTTCTTTT
CTTGAGACGAGGTCTCGCTCTGTCGCCAGGCTGGAGTGCAATGGCGTGATCTCGGC
TCACTGCAACCTCCAACCGCCAGGGTTCAAGCGATTCTCTTGCCCTCGGCCCTCTGAG
TTGCTGGAATTACAGGCACATGCCCCACACCCAGCTAATTTTTGTATTTTAGTAGAG
GCGGGTTTACCGTGTTGGGCAGGCTGGTCAAGAAGTTGAAAGCTTTTTCTACTTG
TGCGCCATTGTCATTCTTCTTTGGAGAAATGCTGTACATTTCATTGTTATTCAGGC
AATGTCTCAAGGACCAAGCACAGAGGAGGCGTGGGGCACCTGCAGTGCCCTCAGG
GCAGCTGACAGAGAAGCAGGGCAAACGGGAAGTGCACAGCCTTTCGATTAAATG
TCAATCTTTGATTGACAATTTGGGGTCATTTCTAAATCCAAGTTAAATACAGAGCAG
CTGGAATGTGAAGATGTCATTCTGGGGTTCCTGCAGATTCTAAGTATAAGACGAAG
GGTGGGCATCCTTATAGAGGAGAAATTGTTAGTGTGCGGGGGATCGAGGTGTGAGGT
GGAGAGAGAATCCTTTCTGCCAGCTCACCACCGCAGATGGGCTGTCTGGAAGGACT
TGGTCTCAGACCCTGAGTTAAAACGGCCTTCATGCTAGCCCTCCTCGCGAGTAACCAC
GTGCGCGGACCGAGCGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTC
TGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGC
TTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAAGCTTGATTATCAAAAAG
GATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAAGCCCAATCTGAA
TAATGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCATCGAGCATCAAATGAA
ACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTGAATAAGCCGTTTCTGTA
ATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGT
CTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATA
AGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAA
AGTTTATGCATTCTTTCCAGACTTGTTC AACAGGCCAGCCATTACGCTCGTCATCAA
AATCACTCGCATCAACCAAACCGTTATTCATTCTGATTGCGCCTGAGCGAGACGAAA
TACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAG
GAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATACCT
GGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACG
GATAAAATGCTTGATGGTTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACC
ATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTG
GCGCATCGGGCTTCCCATACAAGCGATAGATTGTCGCACCTGATTGCCCCGACATTATC
GCGAGCCCATTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTCG
ACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGAC
AGTTTTATTGTTTCATGATGATATATTTTATCTTGTGCAATGTAACATCAGAGATTTTGA
GACACGGGCCAGAGCTGCATCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACAC
ATGCAGCTCCCGGACGTCATTGTGATATCATTGTACATGTGAGCAAAAAGGCCAGCA
AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT

SEQ ID NO: 22 - Modified Region E

GTGAGTACAATTCATTATTTAGTAGGATCTCAAATACCTAGCCAAGTAGTTAAACTTTTC
ATGAGCGCCCCCTCTCATCCCAGCATTTTGCTGCTGCCCAGCCTTAATTACTGGAGGC
AAACATCCTTTTGGTTTGCTGACTTGAAGTTTTATTTGCCACATGAATGTAAAGTCA
GTCTGGGGTGCCGCAAGGGCTGCTGGCCATGCTCATGGCTTTGTACCAAGGAACCT
GCCCCGTTAGGAGTGGCCAAAGGCAGCTGATGAAATAGAATCCTGCAGGTGGGAGG
GAGGCACTGACCGGGGCAGCGGGGAGAAGACAAAGAGACCTATTCATGCAATGCAG
GCGCACCAGCTGTTTTTACAATAAGGTGCATGTGGCTTTTGAGAAAGTGGGTCAATTTG
GATGGAGGGTCTCTAATAGGGAACAATTTTTCTTTCAAATACCTGAATCACAGGTCAG
TTCAATTCGCCACCTGCCCCACTGTGTGGTCCACACTTCGTTCACTTTCTGCCTTCGG
TAAGAGAAACCAGGAAGGCCGAGCAGTTCTCAGTCAGGACTCCTTGCTCGGCTGGC
TCAGCAGGTGCGCCCCCGCCATGCAGTTTCTCTTCTTGACCTTGAGACCTCCAGAGT
CTAAATTCACCTTGCACTGTGGTTCTGCGCCAGGCCACATAGCGAGATAAAGTGGG
GCACAGCAGCCCCACCCCTTGGATTTCAAAAACAGAAACTGGGCTGCAGGAACAGG
CTGCTGACAGGGACGCAATTTCAATACAGGTTTCATTATTACGGGGCCAGCTCCGCAC
TGTGTGGCAGGGGGACAGTTGTAGTGTGCACTCTGAAAGAGCTGTTGGTTAATTC
CTGCCTTCCTCCCTTTTCCCAG

SEQ ID NO: 23 – Unmodified Region E

GTGAGTACAATTCATTATTTAGTAGGATCTCAAATACCTAGCCAAGTAGTTAAACTTTTC
ATGAGCGCCCCCTCTCATCCCAGCATTTTGCTGCTGCCCAGCCTTAATTACTGGAGGC
AAACATCCTTTTGGTTTGCTGACTTGAAGTTTTATTTGCCACATGAATGTAAAGTCA
GTCTGGGGTGCCGCAAGGGCTGCTGGCCATGCTCATGGCTTTGTACCAAGGAACCT
GCCCCGTTAGGAGTGGCCAAAGGCAGCTGATGAAATAGAATCCTGCAGGTGGGAGG
GAGGCACTGACCGGGGCAGCGGGGAGAAGACAAAGAGACCTATTCATGCAATGCAG
GCGCACCAGCTGTTTTTACAATAAGGTGCATGTGGCTTTTGAGAAAGTGGGTCAATTTG
GATGGAGGGTCTCTAATAGGGAACAATTTTTCTTTCAAATACCTGAATCACAGGTCAG
TTCAATTCGCCACCTGCCCCACTGTGTGGTCCACACTTCGTTCACTTTCTGCCTTCGG
TAAGAGAAACCAGGAAGGCCGAGCAGTTCTCAGTCAGGACTCCTTGCTCGGCTGGC
TCAGCAGGTGCGCCCCCGCCATGCAGTTTCTCTTCTTGACCTTGAGACCTCCAGAGT
CTAAATTCACCTTGCACTGTGGTTCTGCGCCAGGCCACATAGCGAGATAAAGTGGG
GCACAGCAGCCCCACCCCTTGGATTTCAAAAACAGAAACTGGGCTGCAGGAACAGG
CTGCTGACAGGGACGCAATTTCAATACAGGTTTCATTATTACGGGGCCAGCTCCGCAC
TGTGTGGCAGGGGGACAGTTGTAGTGTGCACTCTGAAAGAGCTGTTGGTTAATTC
CTGCCTTCCTCCCTTTTCCCAG

SEQ ID NO: 24 – H2-Region E-eGFP

ATGGGACCAGAGCCAGCGAAGTCTGCTCCCCGCCCCGAAAAAGGGCTCCAAGAAGGC
GGTGACTAAGGCGCAGAAGAAAGGCGGCAAGAAGCGCAAGCGCAGCCGCAAGGAG
AGCTATTCCATCTATGTGTACAAGGTTCTGAAGCAGGTCCACCCTGACACCGGCATTT
CGTCCAAGGCCATGGGCATCATGAACCTCGTTTGTGAACGACATTTTCGAGCGCATCGC
AGGTGAGGGCTTCCCGCCTGGCGCATTACAACAAGCGCTCGACCATCAGCTCCAGGGA
GATCCAGACGGCCGTGCGCCTGCTGCTGCCTGGGGAGTTGGCCAAGCACGCCGTGTC
CGAGGGTACTAAGGCCATCACCAAGTACACCAGCGCTAAGGATCCAGGTGAGTACAA
TTCATTATTTAGTAGGATCTCAAATACCTAGCCAAGTAGTTAAACTTTTCATGAGCGCCC
CCTCTCATCCCAGCATTTTGCTGCTGCCCAGCCTTAATTACTGGAGGCAAACATCCTTT
TGGTTTGCTGACTTGAAGTTTTATTTGCCACATGAATGTAAAGTCAGTCTGGGGTGC
CGCAAGGGCTGCTGGCCATGCTCATGGCTTTGTACCAAGGAACCTGCCCCGTTAGG
AGTGGCCAAAGGCAGCTGATGAAATAGAATCCTGCAGGTGGGAGGGAGGCACTGAC

CGGGGCAGCGGGGAGAAGACAAAGAGACCTATTTCATGCAATGCAGGCGCACCAGCT
GTTTTTACAATAAGGTGCATGTGGCTTTTGAGAAAGTGGGTCATTTGGATGGAGGGTC
TCTAATAGGGAACAATTTTTCTTTCAAATACCTGAATCACAGGTCAGTTCAATTCGCCA
CCTGCCCCACTGTGTGGTCCACACTTCGTTCACTTTCTGCCTTCGGTAAGAGAAACCA
GGAAGGCCGAGCAGTTCTCAGTCAGGACTCCTTGCTCGGCTGGCTCAGCAGGTCCG
CCCCCGCCATGCAGTTTCTCTTTGACCTTGAGACCTCCAGAGTCTAAATTCACCTT
GCACTGTGGTTTCTGGGCCAGGCCACATAGCGAGATAAAGTGGGGCACAGCAGCCCC
ACCCCTTGATTTCAAACAGAACTGGGCTGCAGGAAACAGGCTGCTGACAGGGA
CGCAATTTCAATACAGGTTTCATTATTCACGGGCCAGCTCCGCACTGTGTGGCAGGGG
GACAGTTGTTAGTGTTCGACTCTGAAAGAGCTGTTGGTTAATTCTCGCTTCTCCCT
TTTCCCAGGTCGGTCCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGT
GGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTTCAGCGTGC
CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCTGAAGTTCATCTGCAC
CACCGGCAAGCTGCCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGC
CATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC
TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC
AACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAG
GTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCA
CTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCACAACCACTA
CCTGAGCACCCAGTCCGCCCTGAGCAAGACCCCAACGAGAAGCGCGATCATATGG
TCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAA

SEQ ID NO: 25 – REGN-huABE

ATGCAAAGCTTCTGAAAAGGTTTGCTCACCTTACTGACTAATTAAAGGGAAGCCACA
TTTTTCCCTCCCTGGTCTCTGCAGCCATACCACCAAGGGCTATTTGATAGGCCACG
GCAGCTTGGGTTACACAGATGCCCGCGATAATGCCATTTGCTGCGATGATTGAGTCC
ACAAGGCTGAGTGACTTGGGGAAAATATTCCCCGATTGGACTAACCTGTTGGCGGCC
TGCTTACCATCACACTTGTTTTTATAATGAAAAAGGAAATTATTTCTGCTTTTAATGA
TGGCAGCTCCAGAAATGGTGAGTAAATGACCGGGCACCGGACGGCTCGGGTACTT
TCGTTCTTAATTAGGTCATGCCCGTGTGAGCCAGGAAAGGGCTGTGTTTATGGGAAG
CCAGTAACACTGTGGCCTACTATCTTCCGTGGTGCCATCTACATTTTTTGGGACTCG
GGAATTATGTGAGTACCGAACTACTTATGCTTATGGTAGGTGTACCACACGCACAG
GGAAAGAATTGCGTTTATGTGGGACAGGTGAAAACAATCGCAAAAAAGCAATGGAA
AGGGCTTTGAGAGTAATTTATCTTCTGAACAATAGCAATATGGCAACTTCTAAATGG
TGAGAGGGGAGTCCCTCTAAAGCAATGGATTTGAAGATTGGTTGGACAAACAATGGG
AAAGTCATTGTCTTAGCAGAATTAAGTTCATACTTTTTTTTTTTTTTTTTTTGCTAA
TTCAAGAAGCTTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCAGA
TGACTGTTTTTGGTTAATGGATACTGAATGCGCTTTGTGTGGACTGTCAATTTCAA
GATTTACCGTATGACCAAGATGCACCTGATGCTACAAGTATAAATAGGGGAACAAA
TGCTTTCTGTTCTTCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAG
AGGTCCTGAAATAGTCACATCGATTGAATTCACCATGGGACCAGAGCCAGCGAAGT
CTGCTCCCGCCCCGAAAAAGGGCTCCAAGAAGGCGGTGACTAAGGCGCAGAAGAA
AGGCGGCAAGAAGCGCAAGCGCAGCCGCAAGGAGAGCTATTCCATCTATGTGTACA

AGGTTCTGAAGCAGGTCCACCCTGACACCGGCATTTTCGTCCAAGGCCATGGGCATCA
TGAATCGTTTTGTGAACGACATTTTCGAGCGCATCGCAGGTGAGGCTTCCCCGCTGG
CGCATTACAACAAGCGCTCGACCATCACCTCCAGGGAGATCCAGACGGCCGTGCGC
CTGCTGCTGCCTGGGGAGTTGGCCAAGCACGCCGTGTCCGAGGGTACTAAGGCCAT
CACCAAGTACACCAGCGCTAAGGATCCAGGTGAGTACAATTCATTATTTAGTAGGAT
CTCAAATACCTAGCCAAGTAGTTAACTTTTCATGAGCGCCCCCTCTCATCCCAGCAT
TTTGCTGCTGCCCAGCCTTAATTACTGGAGGCAAACATCCTTTTGTTTGCTGACTTG
AAGTTTTATTTGCCACATGAATGTAAAGTCAGTCTGGGGTGCCGCAAGGGCTGCTG
GCCATGCTCATGGCTTTGTCACCAAGGAACCTGCCCCGTTAGGAGTGCCCAAAGGC
AGCTGATGAAATAGAATCCTGCAGGTGGGAGGGAGGCACTGACCGGGGCAGCGGG
GAGAAGACAAAGAGACCTATTCATGCAATGCAGGCGCACCAGCTGTTTTTACAATA
AGGTGCATGTGGCTTTTGAGAAAGTGGGTCAATTTGGATGGAGGGTCTCTAATAGGGA
ACAATTTTCTTTCAAATACCTGAATCACAGGTCAGTTCAATTCGCCACCTGCCCCAC
TGTGTGGTCCACACTTCGTTCACTTTCTGCCTTCGGTAAGAGAAACCAGGAAGGCCG
AGCAGTTCTCAGTCAGGACTCCTTGCTCGGCTGGCTCAGCAGGTGCGCCCCCGCCA
TGCAGTTTCTCTCTTGACCTTGAGACCTCCAGAGTCTAAATTCACCTTGCACTGTGGT
TCCTGGGCCAGGCCACATAGCGAGATAAAGTGGGGCACAGCAGCCCCACCCCTTGG
ATTTCAAAACAGAACTGGGCTGCAGGAAACAGGCTGCTGACAGGGACGCAATTC
AATACAGGTTTCATTATTACGGGCCAGCTCCGCACTGTGTGGCAGGGGGACAGTTG
TTAGTGTTGCACTCTGAAAGAGCTGTTGGTTAATTCCTGCCTTCCTCCCTTTTCCCAG
GTCCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCC
ATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGA
GGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG
GCAAGCTGCCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCACT
GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGC
CCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG
ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAA
GGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAAT
ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTG
AACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTA
CCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACC
TGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA
GTAATCGAGAGATCTACGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCCTCTCCTG
GCCCTGGAAGTTGCCACTCCAGTGCCCAACAGCCTTGTCTTAATAAAATTAAGTTGC
ATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTGGTA
TGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGGGA
ACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGT
TCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCATGCATGAC
CAGGCTCAGCTAATTTTTGTTTTTTGGTAGAGACGGGGTTTACCATATTGGCCAGG
CTGGTCTCCAATCCTAATCTCAGGTGATCTACCCACCTTGGCCTCCCAAATTGCTGG
GATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCTTCTGATTTTGTAGGTCGACTT
CTTGGTACCGCTAGCCCTCCTCGCGAGTAACCACGTGCGCGGACCGAGCGGCCGCA

GGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGA
GGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGA
GCGAGCGAGCGCGCAGAAGCTTGATTATCAAAAAGGATCTTCACCTAGATCCTTTTA
AATTAAAAATGAAGTTTTAAATCAAGCCCAATCTGAATAATGTTACAACCAATTAAC
CAATTCTGATTAGAAAACTCATCGAGCATCAAAATGAAACTGCAATTTATTCATATC
AGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTC
ACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCG
TCCAACATCAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGTTATCAAGTGA
GAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTT
CTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCAT
CAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGC
TGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCC
AGCGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATACCTGGAATGCT
GTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAA
ATGCTTGATGGTCGGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTC
ATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGC
ATCGGGCTTCCCATAACAAGCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCG
AGCCCATTATACCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTCGA
CGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGAC
AGTTTTTATTGTTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTT
GAGACACGGGCCAGAGCTGCATCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGA
CACATGCAGCTCCCGGACGTCATTGTGATATCATTGTACATGTGAGCAAAAGGCC
AGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC
CGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCC
GACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGATAGGTGCTTCGCTCCA
AGCTGGGCTGTGTGCACGAACCCCCGTTACCCCGACCGCTGCGCCTTATCCGGTA
ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCA
CTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAG
TGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCAC
CGCTGGTAGCGGTGGTTTTTTTGGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGG
ATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA
CTCACGTTAAGGGATTTTGGTCATGAGGTACCCTGCGCGCTCGCTCGCTCACTGAGG
CCGCCCCGGCAAAGCCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGCTCAGTGAGC
GAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTGCGGCCG
CACGCGT

SEQ ID NO: 26 – Second Modified Region B

GACCGGGCACCGGACGGCTCGGGTACTTTCGTTCTTAATTAGGTCATGCCCCGTGTGA
GCCAGGAAAGGGCTGTGTTTATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTC
CGTGGTGCCATCTACATTTTGGGACTCGGGAATTATGTGAGTACCGAACTACTTA
TGCTTATGGTAGGTGTACCACACGCACAGGGAAAGAATTGCGTTTATGTGGGACAG
GTGAAAACAATCGCAAAAAGCAATGGAAAGGGCTTTGAGAGTAATTTATCTTCTG
AACAAATAGCAATATGGCAACTTCTAAATGGTGAGAGGGAGTCCCTCTAAAGCAATT
GATTTGAAGATTGGTTGGACAAACAATGGGAAAGTCATTGTCTTAGCAGAATTAAGT
TCATACTTTTTTTTTTTTTTTTTTTTTGCTAATTC AAGAAGCTTTTCTGTTATCTGCTGT
AGCTCAGACGAAAATGCATTCTCACCAGATGACTGTTTTTGGTTAATCCATACTGAA
TGCGCTTTGTGTGGACTGTCGAATTTCAAAGATTTACCGTATGACCAAGATGCACCT
GATGCTACAAGTATAAATAGGGGAACAAATGCTTTCTGTTCTTCCTCGGCTAAGGAG
GTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC

SEQ ID NO: 27 – Second Modified Region B3

ATTCAAGAAGCTTTTCTGTTATCTGCTGTAGCTCAGACGAAAATGCATTCTCACCAG
ATGACTGTTTTTGGTTAATCCATACTGAATGCGCTTTGTGTGGACTGTCGAATTTCAA
AGATTTACCGTATGACCAAGATGCACCTGATGCTACAAGTATAAATAGGGGAACAA
ATGCTTTCTGTTCTTCCTCGGCTAAGGAGGTAGAGGTGGAGGCGGAGCCGGATGTCA
GAGGTCCTGAAATAGTCAC

CLAIMS**WHAT IS CLAIMED IS:**

1. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 1 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 2.
2. The polynucleotide of claim 1, wherein the polynucleotide comprises SEQ ID NO: 16 or a functional portion or derivative thereof.
3. The polynucleotide of claim 1, wherein the polynucleotide consists of SEQ NO: 1 or a functional portion or derivative thereof.
4. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 3 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 26 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 4.
5. The polynucleotide of claim 4, wherein the polynucleotide consists of SEQ NO: 3 or a functional portion or derivative thereof, or SEQ ID NO: 26 or a functional portion or derivative thereof.
6. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 5 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 6.
7. The polynucleotide of claim 6, wherein the polynucleotide comprises SEQ ID NO: 17 or a functional portion or derivative thereof.
8. The polynucleotide of claim 4, wherein the polynucleotide consists of SEQ NO: 5 or a functional portion or derivative thereof.
9. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 8 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 27 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 9.

10. The polynucleotide of claim 9, wherein the polynucleotide consists of SEQ NO: 8 or a functional portion or derivative thereof, or SEQ ID NO: 27 or a functional portion or derivative thereof.
11. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 10 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 11.
12. The polynucleotide of claim 11, wherein the polynucleotide consists of SEQ NO: 10 or a functional portion or derivative thereof.
13. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 12 or a functional portion or derivative thereof.
14. The polynucleotide of claim 13, wherein the polynucleotide consists of SEQ NO: 12 or a functional portion or derivative thereof.
15. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 13 or a functional portion or derivative thereof.
16. The polynucleotide of claim 15, wherein the polynucleotide consists of SEQ NO: 13 or a functional portion or derivative thereof.
17. A polynucleotide comprising a region having at least 85% sequence identity to SEQ ID NO: 22 or a functional portion or derivative thereof, wherein said region is not identical to SEQ ID NO: 23.
18. The polynucleotide of claim 17, wherein the polynucleotide consists of SEQ NO: 22 or a functional portion or derivative thereof.
19. A polynucleotide comprising in a 5' to 3' direction:
 - a) a first region having at least 85% sequence identity to SEQ ID NO: 1 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 2 or a functional portion or derivative thereof;
 - b) a second region having at least 85% sequence identity to SEQ ID NO: 3 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 26 or a functional portion or derivative thereof;

portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 4 or a functional portion or derivative thereof; and

c) a linker having a length from 0 to two thousand nucleotides between the first region and the second region,

wherein the first region is operably linked to the second region and the polynucleotide is not identical to SEQ ID NO: 14.

20. The polynucleotide of claim 19, wherein the polynucleotide comprises a third region in the 3' direction of the second region coding for a protein.

21. The polynucleotide of claim 19, wherein the linker has a length of 0 nucleotides.

22. A polynucleotide comprising in a 5' to 3' direction:

a) a first region having at least 85% sequence identity to SEQ ID NO: 1 or a functional portion, or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 2 or a functional portion or derivative thereof;

b) a second region having at least 85% sequence identity to SEQ ID NO: 3 or a functional portion or derivative thereof, at least 85% sequence identity to SEQ ID NO: 26 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 4 or a functional portion or derivative thereof;

c) a third region coding for a protein;

d) a fourth region having at least 85% sequence identity to SEQ ID NO: 5 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 6 or a functional portion or derivative thereof;

e) a first linker having a length from 0 to two thousand nucleotides between the first region and the second region;

f) a second linker having a length from 0 to 100 nucleotides between the second region and the third region;

g) a third linker between having a length from 0 to 100 nucleotides between the third region and the fourth region,

wherein the first region, the second region, third region and the fourth region are operably linked together and the polynucleotide is not identical to SEQ ID NO: 15.

23. The polynucleotide of claim 22, wherein the polynucleotide comprises SEQ ID NO: 16 or a functional portion or derivative thereof.
24. The polynucleotide of claim 22, wherein the polynucleotide comprises SEQ ID NO: 17 or a functional portion or derivative thereof.
25. The polynucleotide of claim 22, wherein the first linker has a length of 0 nucleotides.
26. The polynucleotide of claim 22, wherein the second linker has a length of 0 nucleotides.
27. The polynucleotide of claim 22, wherein the third linker has a length of 0 nucleotides.
28. The polynucleotide of claim 22, wherein the protein is human TMPRSS3.
29. The polynucleotide of claim 22, wherein the protein is a reporter gene.
30. The polynucleotide of claim 29, wherein the reporter gene is a green fluorescent protein.
31. The polynucleotide of claim 22, wherein the polynucleotide has a SEQ ID NO: 19.
32. A polynucleotide comprising in a 5' to 3' direction:
 - a) a first region having at least 85% sequence identity to SEQ ID NO: 1 or a functional portion, or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 2 or a functional portion or derivative thereof;
 - b) a second region having at least 85% sequence identity to SEQ ID NO: 7 or a functional portion or derivative thereof;
 - c) a third region having at least 85% sequence identity to SEQ ID NO: 8, at least 85% sequence identity to SEQ ID NO: 27 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 9 or a functional portion or derivative thereof;
 - d) a fourth region coding for a protein;

e) a fifth region having at least 85% sequence identity to SEQ ID NO: 5 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 6 or a functional portion or derivative thereof;

f) a first linker having a length from 0 to two thousand nucleotides between the first region and the second region;

g) a second linker having a length from 0 to 100 nucleotides between the second region and the third region, wherein the third linker is not identical to SEQ ID NO: 11;

h) a third linker between having a length from 0 to 100 nucleotides between the third region and the fourth region;

i) a fourth linker having a length from 0 to 100 nucleotides between the fourth region and the fifth region,

wherein the first region, second region, third region, fourth region and fifth region are operably linked together.

33. The polynucleotide of claim 32, wherein the first linker has a length of 0 nucleotides.

34. The polynucleotide of claim 32, wherein the second linker has a length of 0 nucleotides.

35. The polynucleotide of claim 32, wherein the third linker has a length of 0 nucleotides.

36. The polynucleotide of claim 32, wherein the fourth linker has a length of 0 nucleotides.

37. The polynucleotide of claim 32, wherein the protein is human TMPRSS3.

38. The polynucleotide of claim 32, wherein the protein is a reporter gene.

39. The polynucleotide of claim 38, wherein the reporter gene is a green fluorescent protein.

40. The polynucleotide of claim 32, wherein the polynucleotide has a SEQ ID NO: 21.

41. A polynucleotide comprising in a 5' to 3' direction:

a) a first region having at least 85% sequence identity to SEQ ID NO: 1 or a functional portion, or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 2 or a functional portion or derivative thereof;

b) a second region having at least 85% sequence identity to SEQ ID NO: 3 or a functional portion or derivative thereof, at least 85% sequence identity to SEQ ID NO: 26 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 4 or a functional portion or derivative thereof;

c) a third region coding for a protein or a part thereof;

d) a fourth region having at least 85% sequence identity to SEQ ID NO: 22 or a functional portion or derivative thereof, or at least 85% sequence identity to SEQ ID NO: 23 or a functional portion or derivative thereof;

e) a first linker having a length from 0 to two thousand nucleotides between the first region and the second region;

f) a second linker having a length from 0 to 100 nucleotides between the second region and the third region;

g) a third linker between having a length from 0 to 100 nucleotides between the third region and the fourth region,

wherein the first region, the second region, third region and the fourth region are operably linked together and the polynucleotide is not identical to SEQ ID NO: 14 or SEQ ID NO: 15.

42. The polynucleotide of claim 41, wherein the third region codes for a part of the protein and the polynucleotide further comprises a fifth region in the 3' direction of the fourth region coding for another part of the same protein coded by the third region.

43. The polynucleotide of claim 42, wherein the polynucleotide comprises a fourth linker having a length from 0 to 100 nucleotides between the fourth region and the fifth region.

44. The polynucleotide of claim 41, wherein the first linker has a length of 0 nucleotides.

45. The polynucleotide of claim 41, wherein the second linker has a length of 0 nucleotides.

46. The polynucleotide of claim 41, wherein the third linker has a length of 0 nucleotides.

47. The polynucleotide of claim 41, wherein the fourth linker has a length of 0 nucleotides.

48. The polynucleotide of claim 41, wherein the protein is human TMPRSS3.

49. The polynucleotide of claim 41, wherein the protein is a reporter gene.
50. The polynucleotide of claim 49, wherein the reporter gene is a green fluorescent protein.
51. A nucleic acid vector comprising the polynucleotide of any one of claims 1 -50.
52. The nucleic acid vector of claim 51, wherein the polynucleotide is operably linked to a transgene coding for a protein.
53. The nucleic acid vector of claim 52, wherein the polynucleotide is capable of directing cell- specific expression of the protein from the nucleic acid sequence in a mammalian inner ear cell.
54. The nucleic acid vector of claim 53, wherein the inner ear cell is a hair cell.
55. The nucleic acid vector of claim 54, wherein the hair cell is an inner hair cell, an outer hair cell and/or vestibular hair cell.
56. The nucleic acid vector of claim 53, wherein the inner ear cell is a sensory supporting cell.
57. The nucleic acid vector of any one of claims 52-56, wherein the protein is selected from the group consisting of: Actin, gamma 1; ADP-ribosyltransferase 1; Activating transcription factor 6; Atonal bHLH transcription factor 1; ATPase plasma membrane Ca²⁺ transporting 2; Brain-derived neurotrophic factor; Binding immunoglobulin protein; Calcium binding protein 2; Calcium channel, voltage-dependent, alpha 2/delta subunit 4; Coiled-Coil Domain Containing 92; Cadherin-related 23; Cholinergic receptor nicotinic alpha 9 subunit; Calcium and integrin binding family member 2; Calcium and integrin binding family member 3; Clarin 1; Clarin 2; Pejvakin; Dystonin; ELFN1 adhesion G protein-coupled receptor; Epidermal growth factor receptor kinase substrate 8; EPS8-like 2; Espin; Espin-like; Fc receptor-like B; Fascin actin-bundling protein 2; Growth factor independent 1 transcriptional repressor; Glutaredoxin domain-containing cysteine-rich protein 1; Glutaredoxin domain-containing cysteine-rich protein 2; Heat shock protein 70; Heat shock protein 90; Inositol-requiring enzyme 1; Potassium voltage-gated channel subfamily A member 10; Kielin/chordin-like protein; Kinocilin; Late cornified envelope protein 6A; LIM homeobox 3; Lipoxxygenase homology domains 1; Leucine-rich repeats and transmembrane domains 2; Leucine-rich transmembrane and O-methyltransferase domain

containing 1; Leucine-rich transmembrane and O-methyltransferase domain containing 2; Makorin ring finger protein 20; Myosin XV; Myosin IIIA; Myosin IIIB; Myosin VI; Myosin VIIA; Nescient helix-loop-helix 1; Neurotrophin 3; Otoferlin; Protocadherin-related 15; Protein kinase R (PKR)-like endoplasmic reticulum kinase; POU class 4 homeobox 3; Peripherin 2; Protein tyrosine phosphatase receptor type Q; Radixin; Serpin family E member 3; SKI family transcriptional corepressor 1; Solute carrier family 17, member 8 (vesicular glutamate transporter 3); Solute carrier family 8, member 2; Stereocilin; Tctex1 domain containing 1; Transmembrane channel-like 1; Transmembrane protease, serine 3; Taperin; TRIO and F-actin binding protein; Tetratricopeptide repeat domain 24; Usher syndrome 1C; Solute carrier family 26, member 4; Xin actin-binding repeat containing 2; and Zinc finger CCHC-type containing 12.

58. The nucleic acid vector of any one of claims 51-57, wherein the nucleic acid vector is a plasmid, cosmid, artificial chromosome, or viral vector.

59. The nucleic acid vector of claim 58, wherein the nucleic acid vector is a viral vector selected from the group consisting of an adeno-associated virus (AAV), an adenovirus, and a lentivirus.

60. The nucleic acid vector of claim 59, wherein the viral vector is an AAV vector.

61. The nucleic acid vector of claim 60, wherein the AAV vector has a serotype selected from the group consisting of AAV1, AAV2, AAV2quad(Y-F), AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, rh10, rh39, rh43, rh74, Anc80, Anc80L65, DJ/8, DJ/9, 7m8, PHP.B, PHP.eb, and PHP.S.

62. A composition comprising the nucleic acid vector of claim 51-61.

63. The composition of claim 62, further comprising a pharmaceutically acceptable excipient.

64. A method of increasing expression of a protein in a subpopulation of mammalian inner ear cells, comprising contacting the mammalian inner ear cells with the nucleic acid vector of any one of claim 51-61 or the composition of any one of claim 61-63.

65. The method of claim 64, wherein expression of the protein is increased in cochlear hair cells.

66. The method of claim 64, wherein expression of the protein is increased in vestibular hair cells.
67. The method of claim 64, wherein expression of the protein is increased in sensory support cells.
68. The method of claim 64, wherein expression of the protein is increased in inner hair cells and/or outer hair cells.
69. The method of any one of claims 64, wherein expression of the protein is increased in interdental hair cells.
70. A method of treating a subject in need thereof having or at risk of developing a condition, the method comprising administering to the subject an effective amount of the nucleic acid vector of any one of claims 51-61 or the composition of any of claims 62-63.
71. The method of claim 70, wherein the condition is hearing loss.
72. The method of claim 71, wherein the hearing loss is genetic hearing loss.
73. The method of claim 72, wherein the genetic hearing loss is autosomal dominant hearing loss, autosomal recessive hearing loss, or X-linked hearing loss.
74. The method of claim 71, wherein the hearing loss is acquired hearing loss.
75. The method of claim 74, wherein the acquired hearing loss is noise-induced hearing loss, age-related hearing loss, disease or infection-related hearing loss, head trauma-related hearing loss, or ototoxic drug-induced hearing loss.
76. The method of claim 70, wherein the condition is vestibular dysfunction.
77. The method of claim 76, wherein the vestibular dysfunction is vertigo, dizziness, or imbalance.
78. The method of claim 70, wherein the condition is tinnitus.
79. A kit comprising the nucleic acid vector of any one of claims 51-61 or the composition of any one of claims 62-63.