



(51) International Patent Classification:

C12N 15/85 (2006.01) C07K 14/435 (2006.01)

C12N 5/0793 (2010.01) A61K 48/00 (2006.01)

C12N 15/65 (2006.01) A61P 27/16 (2006.01)

C12N 15/10 (2006.01)

(21) International Application Number:

PCT/US2019/041625

(22) International Filing Date:

12 July 2019 (12.07.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/697,652 13 July 2018 (13.07.2018) US

(71) Applicant: AKOUOS, INC. [US/US]; 7 Stetson Street, Brookline, Massachusetts 02446 (US).

(72) Inventor; and

(71) Applicant: NG, Robert [—/US]; 47 Lincoln Rd, Unit 1, Newton, Massachusetts 02458 (US).

(72) Inventor: SIMONS, Emmanuel J.; 7 Stetson Street, Brookline, Massachusetts 02446 (US).

(74) Agent: WHYTE, Jasmine et al.; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

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

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

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

(54) Title: METHODS OF TREATING NON-SYNDROMIC SENSORINEURAL HEARING LOSS

AAV-dSTRC Transduction in Explants

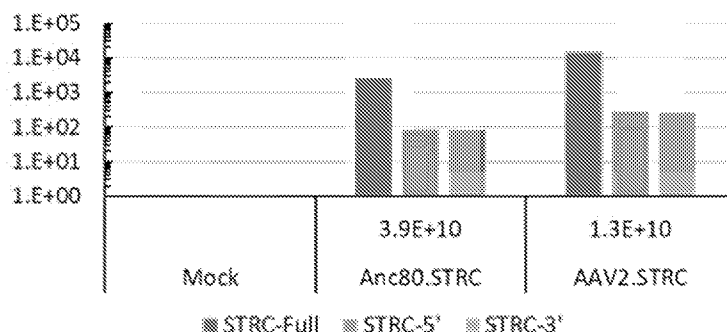


Figure 10

(57) Abstract: Provided herein are compositions that include at least two different nucleic acid vectors, where each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, and the use of these compositions to treat non-syndromic sensorineural hearing loss in a subject.



WO 2020/014625 A1

METHODS OF TREATING NON-SYNDROMIC SENSORINEURAL HEARING LOSS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Patent Application Serial
5 No. 62/697,652, filed July 13, 2018; the entire contents of which are herein
incorporated by reference.

TECHNICAL FIELD

The present disclosure relates generally to the use of nucleic acids to treat
hearing loss in a human subject.

BACKGROUND OF THE INVENTION

10 Non-syndromic deafness, in contrast to syndromic deafness, is hearing loss
that is not associated with other signs and symptoms. Seventy to eighty percent of
cases of genetic deafness are non-syndromic. While the causes of non-syndromic
deafness are complex, researchers have identified more than 30 genes that, when
15 altered, are associated with non-syndromic deafness (e.g., stereocilin (STRC)).
Current treatments consist mainly of hearing amplification for mild to severe hearing
loss and cochlear implants for severe to profound hearing loss; however, a long-felt
need remains for agents and methods for preventing or reversing non-syndromic
deafness.

20 Hearing loss can be conductive (arising from the ear canal or middle ear),
sensorineural (arising from the inner ear or auditory nerve), or mixed. Most forms of
non-syndromic deafness are associated with permanent hearing loss caused by
damage to structures in the inner ear (sensorineural deafness), although some forms
may involve changes in the middle ear (conductive hearing loss). The great majority
25 of human sensorineural hearing loss is caused by abnormalities in the hair cells of the
organ of Corti in the cochlea (poor hair cell function). The hair cells may be
abnormal at birth, or may be damaged during the lifetime of an individual (e.g., as a
result of noise trauma or infection).

SUMMARY

30 The present invention is based on the discovery that a composition including
at least two different nucleic acid vectors, where each of the at least two different

vectors includes a coding sequence that encodes a different portion of a stereocilin protein, can be used to generate a sequence encoding an active stereocilin protein (e.g., a full-length stereocilin protein) in a mammalian cell, and thereby treat non-syndromic sensorineural hearing loss in a subject in need thereof.

5 Provided herein are compositions that include at least two different nucleic acid vectors, where: each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions being at least 30 amino acid residues in length, wherein the amino acid sequence of each of the encoded portions may optionally partially overlap with the
10 amino acid sequence of a different one of the encoded portions; no single vector of the at least two different vectors encodes a full-length stereocilin protein; at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons; and when introduced into a mammalian cell the at least two
15 different vectors undergo homologous recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein. In some embodiments of any of the compositions described herein, each of the at least two different vectors is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector. In some embodiments of any of the compositions described herein,
20 each of the at least two different vectors is a human artificial chromosome (HAC), yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC). In some embodiments of any of the compositions described herein, each of the at least two different vectors is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a
25 lentivirus vector, or a retrovirus vector. In some embodiments of any of the compositions described herein, the viral vector is an AAV vector.

 In some embodiments of any of the compositions described herein, the amino acid sequence of none of the encoded portions overlaps with the amino acid sequence of a different one of the encoded portions. In some embodiments of any of the
30 compositions described herein, the amino acid sequence of each of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions. In some embodiments of any of the compositions described herein, the overlapping amino acid sequence is between about 30 amino acid residues to about 1000 amino acid residues in length.

In some embodiments of any of the compositions described herein, the vectors include two different vectors, each of which includes a different segment of an intron, wherein the intron includes the nucleotide sequence of an intron that is present in stereocilin genomic DNA, and wherein the two different segments overlap in

5 sequence by at least 100 nucleotides. In some embodiments of any of the compositions described herein, the two different segments overlap in sequence by about 100 nucleotides to about 800 nucleotides. In some embodiments of any of the compositions described herein, the nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 10,000 nucleotides in
10 length (e.g., between about 500 nucleotides to about 5,000 nucleotides in length).

In some embodiments of any of the compositions described herein, the number of different vectors in the composition is two. In some embodiments, one of the two vectors includes SEQ ID NO: 13 and the second of the two vectors includes SEQ ID NO: 17.

15 In some embodiments of any of the compositions described herein, one of the at least two different vectors includes a sequence encoding a stereocilin protein. In some embodiments, the sequence encoding a stereocilin protein is at least 90% (e.g., at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%) identical to SEQ ID NO: 11.

20 In some embodiments of any of the compositions described herein, a first of the two different vectors includes a coding sequence that encodes an N-terminal portion of the stereocilin protein. In some embodiments of any of the compositions described herein, the N-terminal portion of the stereocilin protein is between 30 amino acids to 1600 amino acids in length (e.g., between about 200 amino acids to
25 about 1500 amino acids in length). In some embodiments of any of the compositions described herein, the first vector further includes one or both of a promoter and a Kozak sequence. In some embodiments of any of the compositions described herein, the first vector includes a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter.

30 In some embodiments of any of the compositions described herein, the second of the two different vectors includes a coding sequence that encodes a C-terminal portion of the stereocilin protein. In some embodiments of any of the compositions described herein, the C-terminal portion of the stereocilin protein is between 30 amino acids to 1600 amino acids in length (e.g., between 200 amino acids to 1500 amino

acids in length). In some embodiments of any of the compositions described herein, the second vector further includes a poly(dA) sequence. Some embodiments of any of the compositions described herein further include a pharmaceutically acceptable excipient.

5 Also provided herein are kits that include any of the compositions described herein. Some embodiments of any of the kits described herein further include a pre-loaded syringe containing the composition.

 Also provided herein are methods that include introducing into a cochlea of a mammal a therapeutically effective amount of any of the compositions described
10 herein. In some embodiments of any of the methods described herein, the mammal is a human. In some embodiments of any of the methods described herein, the mammal has been previously identified as having a defective stereocilin gene.

 Also provided herein are methods of increasing expression of a full-length stereocilin protein in a mammalian cell that include introducing any of the
15 compositions described herein into the mammalian cell. In some embodiments of any of the methods described herein, the mammalian cell is a cochlear outer hair cell. In some embodiments of any of the methods described herein, the mammalian cell is a human cell. In some embodiments of any of the methods described herein, the mammalian cell has previously been determined to have a defective stereocilin gene.

20 Also provided herein are methods of increasing expression of a full-length stereocilin protein in an outer hair cell in a cochlea of a mammal that include introducing into the cochlea of the mammal a therapeutically effective amount of any of the compositions described herein. In some embodiments of any of the methods described herein, the mammal has been previously identified as having a defective
25 stereocilin gene. In some embodiments of any of the methods described herein, the mammal is a human.

 Also provided herein are methods of treating non-symptomatic sensorineural hearing loss in a subject identified as having a defective stereocilin gene that include administering a therapeutically effective amount of any of the compositions described
30 herein into the cochlea of the subject. In some embodiments of any of the methods described herein, the subject is a human. Some embodiments of any of the methods described herein further include, prior to the administering step, determining that the subject has a defective stereocilin gene.

Also provided herein are compositions that include two different nucleic acid vectors, where: a first nucleic acid vector of the two different nucleic acid vectors includes a promoter, a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter, and a splicing donor signal sequence positioned at the 3' end of the first coding sequence; and a second nucleic acid vector of the two different nucleic acid vectors including a splicing acceptor signal sequence, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the amino acid sequence of each of the encoded portions do not overlap; where no single vector of the two different vectors encodes a full-length stereocilin protein; and when introduced into a mammalian cell, splicing occurs between the splicing donor signal sequence and the splicing acceptor signal sequence, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein. In some embodiments of any of the compositions described herein, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons.

Also provided are compositions that include two different nucleic acid vectors, where: a first nucleic acid vector of the two different nucleic acid vectors includes a promoter, a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splicing donor signal sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a second detectable marker gene, a splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the amino acid sequence of each of the encoded portions do not overlap; where no single vector of the two different vectors encodes a full-length stereocilin protein; and when introduced into a mammalian cell, splicing occurs between the splicing donor signal and the splicing acceptor signal, thereby forming a recombined nucleic acid that encodes a full-length

stereocilin protein. In some embodiments of any of the compositions described herein, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons. In some embodiments of any of the compositions described herein, the first or second detectable marker gene is alkaline phosphatase. In some embodiments of any of the compositions described herein, the first and second detectable marker genes are the same.

Also provided herein are compositions that include two different nucleic acid vectors, where: a first nucleic acid vector of the two different nucleic acid vectors includes a promoter, a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' to the promoter, a splicing donor signal sequence
5 positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splicing donor signal sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a F1 phage recombinogenic region, a splicing acceptor signal sequence positioned 3' of the F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion
10 of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence; where each of the encoded portions is at least 30 amino acid residues in length, where the amino acid sequence of each of the encoded portions do not overlap; where no single vector of the two different vectors encodes a full-length
15 stereocilin protein; and when introduced into a mammalian cell, splicing occurs between the splicing donor signal and the splicing acceptor signal, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein. In some embodiments of any of the compositions described herein, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of
20 stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons.

Also provided herein are compositions that include: a Cas9 nuclease; a guide RNA that includes at the 5' end of the guide RNA a complementary region consisting of 20 nucleotides that are complementary to 20 consecutive nucleotides within
25 positions 13955-14151 of SEQ ID NO:5, that can be used in CRISPR/Cas9 RNA-guided genome editing to remove a nonsense mutation at a predetermined site in the endogenous stereocilin pseudogene sequence of SEQ ID NO: 5; and when introduced

into a mammalian cell, a nucleic acid encoding a full-length stereocilin protein is reconstituted at the locus of the stereocilin pseudogene.

Also provided are compositions that include at least one nucleic acid vector, wherein the at least one nucleic acid vector includes an adeno-associated virus (AAV) vector that includes an antisense oligonucleotide that is at least 80% complementary to a contiguous nucleotide sequence of SEQ ID NO: 5 at positions 13955-14151, where the antisense oligonucleotide is between 15-30 nucleotides in length; and when introduced into a mammalian cell, a nucleic acid encoding a full-length stereocilin protein is generated at the locus of the stereocilin pseudogene.

Also provided herein are kits that include any of the compositions described herein. Some embodiments of any of the kits described herein further include a pre-loaded syringe containing the composition.

Also provided are methods that include introducing into a cochlea of a mammal a therapeutically effective amount of any of the compositions described herein. In some embodiments of any of the methods described herein, the mammal is a human. In some embodiments of any of the methods described herein, the mammal has been previously identified as having a defective stereocilin gene.

Also provided herein are methods of increasing expression of a full-length stereocilin protein in a mammalian cell that include introducing any of the compositions described herein into the mammalian cell. In some embodiments of any of the methods described herein, the mammalian cell is a cochlear outer hair cell. In some embodiments of any of the methods described herein, the mammalian cell is a human cell. In some embodiments of any of the methods described herein, the mammalian cell has previously been determined to have a defective stereocilin gene.

Also provided herein are methods of increasing expression of a full-length stereocilin protein in an outer hair cell in a cochlea of a mammal that include introducing into the cochlea of the mammal a therapeutically effective amount of any of the compositions described herein. In some embodiments of any of the methods described herein, the mammal has been previously identified as having a defective stereocilin gene. In some embodiments of any of the methods described herein, the mammal is a human.

Also provided herein are methods of treating non-symptomatic sensorineural hearing loss in a subject identified as having a defective stereocilin gene that include administering a therapeutically effective amount of any of the compositions described

herein into the cochlea of the subject. In some embodiments of any of the methods described herein, the subject is a human. Some embodiments of any of the methods described herein further include, prior to the administering step, determining that the subject has a defective stereocilin gene.

5 The term “a” and “an” refers to one or to more than one (i.e., at least one) of the grammatical object of the article. By way of example, “an element” encompasses one element and more than one element.

 The term “mutation in a stereocilin gene” refers to a modification in a wildtype stereocilin gene that results in the production of a stereocilin protein having
10 one or more of: a deletion in one or more amino acids, one or more amino acid substitutions, and one or more amino acid insertions as compared to the wildtype stereocilin protein, and/or results in a decrease in the expressed level of the encoded stereocilin protein in a mammalian cell as compared to the expressed level of the encoded stereocilin protein in a mammalian cell not having a mutation. In some
15 embodiments, a mutation can result in the production of a stereocilin protein having a deletion in one or more amino acids (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids). In some embodiments, the mutation can result in a frameshift in the stereocilin gene. The term “frameshift” is known in the art to encompass any mutation in a coding sequence that results in a shift in the reading
20 frame of the coding sequence. In some embodiments, a frameshift can result in a nonfunctional protein. In some embodiments, a point mutation can be a nonsense mutation (i.e., result in a premature stop codon in an exon of the gene). A nonsense mutation can result in the production of a truncated protein (as compared to a corresponding wildtype protein) that may or may not be functional. In some
25 embodiments, the mutation can result in the loss (or a decrease in the level) of expression of stereocilin mRNA or stereocilin protein or both the mRNA and protein. In some embodiments, the mutation can result in the production of an altered stereocilin protein having a loss or decrease in one or more biological activities (functions) as compared to a wildtype stereocilin protein.

30 In some embodiments, the mutation is an insertion of one or more nucleotides into a stereocilin gene. In some embodiments, the mutation is in a regulatory sequence of the stereocilin gene, i.e., a portion of the gene that is not coding sequence. In some embodiments, a mutation in a regulatory sequence may be in a promoter or enhancer region and prevent or reduce the proper transcription of the stereocilin gene.

The term “conservative mutation” refers to a mutation that does not change the amino acid encoded at the site of the mutation (due to codon degeneracy).

Modifications can be introduced into a nucleotide sequence by standard techniques known in the art, such as site-directed mutagenesis and PCR-mediated mutagenesis. Conservative amino acid substitutions are ones in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, and histidine), acidic side chains (e.g., aspartic acid and glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, and tryptophan), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, and methionine), beta-branched side chains (e.g., threonine, valine, and isoleucine), and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, and histidine).

Unless otherwise specified, a “nucleotide sequence encoding an amino acid sequence” includes all nucleotide sequences that are degenerate versions of each other and thus encode the same amino acid sequence.

The term “endogenous” refers to any material originating from within an organism, cell, or tissue.

The term “exogenous” refers to any material introduced from or originating from outside an organism, cell, or tissue that is not produced or does not originate from the same organism, cell, or tissue in which it is being introduced.

The term “isolated” means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not “isolated,” but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is “isolated.” An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

The term “transfected,” “transformed,” or “transduced” refers to a process by which exogenous nucleic acid is transferred or introduced into a cell. A “transfected,” “transformed,” or “transduced” mammalian cell is one that has been transfected, transformed or transduced with exogenous nucleic acid.

The term “expression” refers to the transcription and/or translation of a particular nucleotide sequence encoding a protein.

The term “transient expression” refers to the expression of a non-integrated coding sequence for a short period of time (e.g., hours or days). The coding sequence that is transiently expressed in a cell (e.g., a mammalian cell) is lost upon multiple rounds of cell division.

5 The term “subject” is intended to include any mammal. In some embodiments, the subject is a rodent (e.g., a rat or mouse), a rabbit, a non-human primate, or a human. In some embodiments, the subject has or is at risk of developing non-syndromic deafness. In some embodiments, the subject has been previously identified as having a mutation in a stereocilin gene. In some embodiments, the
10 subject has been identified as having a mutation in a stereocilin gene and has been diagnosed with non-symptomatic sensorineural hearing loss. In some embodiments, the subject has been identified as having non-symptomatic sensorineural hearing loss.

 A treatment is “therapeutically effective” when it results in a reduction in one or more of the number, severity, and frequency of one or more symptoms of a disease
15 state (e.g., non-symptomatic sensorineural hearing loss) in a subject (e.g., a human). In some embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of an active stereocilin protein (e.g., a wildtype, full-length stereocilin protein or of a variant of a stereocilin protein that has the desired activity) (e.g., as compared to the expression level prior to treatment with the
20 composition). In some embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of an active stereocilin protein (e.g., a wildtype, full-length stereocilin protein or active variant) in a target cell (e.g., a cochlear outer hair cell). In some embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of
25 an active stereocilin protein (e.g., a wildtype, full-length stereocilin protein or active variant), and/or an increase in one or more activities of a stereocilin protein in a target cell (e.g., as compared to a reference level, such as the level(s) in a subject prior to treatment, the level(s) in a subject having a mutation in a stereocilin gene, or the level(s) in a subject or a population of subjects having non-symptomatic sensorineural
30 hearing loss).

 The term “nucleic acid” or “polynucleotide” refers to deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), or a combination thereof, in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding

properties as the reference nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses complementary sequences as well as the sequence explicitly indicated. In some embodiments of any of the nucleic acids described herein, the nucleic acid is DNA. In some embodiments of any of the

5 nucleic acids described herein, the nucleic acid is RNA.

The term “active stereocilin protein” means a protein encoded by DNA that, if substituted for both wildtype alleles encoding full-length stereocilin protein in auditory hair cells of what is otherwise a wildtype mammal, and if expressed in the auditory hair cells of that mammal, results in that mammal’s having a level of hearing
10 approximating the normal level of hearing of a similar mammal that is entirely wildtype. Non-limiting examples of active stereocilin proteins are full-length stereocilin proteins (e.g., any of the full-length stereocilin proteins described herein).

For example, an active stereocilin protein can include a sequence of a wildtype, full-length stereocilin protein (e.g., a wildtype, human, full-length
15 stereocilin protein) including 1 amino acid substitution to about 160 amino acid substitutions, 1 amino acid substitution to about 155 amino acid substitutions, 1 amino acid substitution to about 150 amino acid substitutions, 1 amino acid substitution to about 145 amino acid substitutions, 1 amino acid substitution to about 140 amino acid substitutions, 1 amino acid substitution to about 135 amino acid
20 substitutions, 1 amino acid substitution to about 130 amino acid substitutions, 1 amino acid substitution to about 125 amino acid substitutions, 1 amino acid substitution to about 120 amino acid substitutions, 1 amino acid substitution to about 115 amino acid substitutions, 1 amino acid substitution to about 110 amino acid substitutions, 1 amino acid substitution to about 105 amino acid substitutions, 1
25 amino acid substitution to about 100 amino acid substitutions, 1 amino acid substitution to about 95 amino acid substitutions, 1 amino acid substitution to about 90 amino acid substitutions, 1 amino acid substitution to about 85 amino acid substitutions, 1 amino acid substitution to about 80 amino acid substitutions, 1 amino acid substitution to about 75 amino acid substitutions, 1 amino acid substitution to
30 about 70 amino acid substitutions, 1 amino acid substitution to about 65 amino acid substitutions, 1 amino acid substitution to about 60 amino acid substitutions, 1 amino acid substitution to about 55 amino acid substitutions, 1 amino acid substitution to about 50 amino acid substitutions, 1 amino acid substitution to about 45 amino acid substitutions, 1 amino acid substitution to about 40 amino acid substitutions, 1 amino

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

acid substitutions to about 155 amino acid substitutions, about 100 amino acid
substitutions to about 150 amino acid substitutions, about 100 amino acid
substitutions to about 145 amino acid substitutions, about 100 amino acid
substitutions to about 140 amino acid substitutions, about 100 amino acid
5 substitutions to about 135 amino acid substitutions, about 100 amino acid
substitutions to about 130 amino acid substitutions, about 100 amino acid
substitutions to about 125 amino acid substitutions, about 100 amino acid
substitutions to about 120 amino acid substitutions, about 100 amino acid
substitutions to about 115 amino acid substitutions, about 100 amino acid
10 substitutions to about 110 amino acid substitutions, about 100 amino acid
substitutions to about 105 amino acid substitutions, between about 110 amino acid
substitutions to about 160 amino acid substitutions, about 110 amino acid
substitutions to about 155 amino acid substitutions, about 110 amino acid
substitutions to about 150 amino acid substitutions, about 110 amino acid
15 substitutions to about 145 amino acid substitutions, about 110 amino acid
substitutions to about 140 amino acid substitutions, about 110 amino acid
substitutions to about 135 amino acid substitutions, about 110 amino acid
substitutions to about 130 amino acid substitutions, about 110 amino acid
substitutions to about 125 amino acid substitutions, about 110 amino acid
20 substitutions to about 120 amino acid substitutions, about 110 amino acid
substitutions to about 115 amino acid substitutions, between about 120 amino acid
substitutions to about 160 amino acid substitutions, about 120 amino acid
substitutions to about 155 amino acid substitutions, about 120 amino acid
substitutions to about 150 amino acid substitutions, about 120 amino acid
25 substitutions to about 145 amino acid substitutions, about 120 amino acid
substitutions to about 140 amino acid substitutions, about 120 amino acid
substitutions to about 135 amino acid substitutions, about 120 amino acid
substitutions to about 130 amino acid substitutions, about 120 amino acid
substitutions to about 125 amino acid substitutions, between about 130 amino acid
30 substitutions to about 160 amino acid substitutions, about 130 amino acid
substitutions to about 155 amino acid substitutions, about 130 amino acid
substitutions to about 150 amino acid substitutions, about 130 amino acid
substitutions to about 145 amino acid substitutions, about 130 amino acid
substitutions to about 140 amino acid substitutions, about 130 amino acid

substitutions to about 135 amino acid substitutions, between about 140 amino acid substitutions to about 160 amino acid substitutions, about 140 amino acid substitutions to about 155 amino acid substitutions, about 140 amino acid substitutions to about 150 amino acid substitutions, about 140 amino acid substitutions to about 145 amino acid substitutions, between about 150 amino acid substitutions to about 160 amino acid substitutions, or about 150 amino acid substitutions to about 155 amino acid substitutions. One skilled in the art would appreciate that amino acids that are conserved between wildtype stereocilin proteins from different species can be mutated without losing activity, while those amino acids that are not conserved between wildtype stereocilin proteins from different species should not be mutated as they are more likely (than amino acids that are not conserved between different species) to be involved in activity.

An active stereocilin protein can include, e.g., a sequence of a wildtype, full-length stereocilin protein (e.g., a wildtype, human, full-length stereocilin protein) that has 1 amino acid to about 50 amino acids, 1 amino acid to about 45 amino acids, 1 amino acid to about 40 amino acids, 1 amino acid to about 35 amino acids, 1 amino acid to about 30 amino acids, 1 amino acid to about 25 amino acids, 1 amino acid to about 20 amino acids, 1 amino acid to about 15 amino acids, 1 amino acid to about 10 amino acids, 1 amino acid to about 9 amino acids, 1 amino acid to about 8 amino acids, 1 amino acid to about 7 amino acids, 1 amino acid to about 6 amino acids, 1 amino acid to about 5 amino acids, 1 amino acid to about 4 amino acids, 1 amino acid to about 3 amino acids, about 2 amino acids to about 50 amino acids, about 2 amino acids to about 45 amino acids, about 2 amino acids to about 40 amino acids, about 2 amino acids to about 35 amino acids, about 2 amino acids to about 30 amino acids, about 2 amino acids to about 25 amino acids, about 2 amino acids to about 20 amino acids, about 2 amino acids to about 15 amino acids, about 2 amino acids to about 10 amino acids, about 2 amino acids to about 9 amino acids, about 2 amino acids to about 8 amino acids, about 2 amino acids to about 7 amino acids, about 2 amino acids to about 6 amino acids, about 2 amino acids to about 5 amino acids, about 2 amino acids to about 4 amino acids, about 3 amino acids to about 50 amino acids, about 3 amino acids to about 45 amino acids, about 3 amino acids to about 40 amino acids, about 3 amino acids to about 35 amino acids, about 3 amino acids to about 30 amino acids, about 3 amino acids to about 25 amino acids, about 3 amino acids to about 20 amino acids, about 3 amino acids to about 15 amino acids, about 3 amino acids to

[illegible]

amino acids, about 10 amino acids to about 40 amino acids, about 10 amino acids to about 35 amino acids, about 10 amino acids to about 30 amino acids, about 10 amino acids to about 25 amino acids, about 10 amino acids to about 20 amino acids, about 10 amino acids to about 15 amino acids, about 15 amino acids to about 50 amino acids, about 15 amino acids to about 45 amino acids, about 15 amino acids to about 40 amino acids, about 15 amino acids to about 35 amino acids, about 15 amino acids to about 30 amino acids, about 15 amino acids to about 25 amino acids, about 15 amino acids to about 20 amino acids, about 20 amino acids to about 50 amino acids, about 20 amino acids to about 45 amino acids, about 20 amino acids to about 40 amino acids, about 20 amino acids to about 35 amino acids, about 20 amino acids to about 30 amino acids, about 20 amino acids to about 25 amino acids, about 25 amino acids to about 50 amino acids, about 25 amino acids to about 45 amino acids, about 25 amino acids to about 40 amino acids, about 25 amino acids to about 35 amino acids, about 25 amino acids to about 30 amino acids, about 30 amino acids to about 50 amino acids, about 30 amino acids to about 45 amino acids, about 30 amino acids to about 40 amino acids, about 30 amino acids to about 35 amino acids, about 35 amino acids to about 50 amino acids, about 35 amino acids to about 45 amino acids, about 35 amino acids to about 40 amino acids, about 40 amino acids to about 50 amino acids, about 40 amino acids to about 45 amino acids, about 45 amino acids to about 50 amino acids, deleted. In some embodiments where two or more amino acids are deleted from the sequence of a wildtype, full-length stereocilin protein, the two or more deleted amino acids can be contiguous in the sequence of the wildtype, full-length protein. In other examples where two or more amino acids are deleted from the sequence of a wildtype, full-length stereocilin protein, the two or more deleted amino acids are not contiguous in the sequence of the wildtype, full-length protein. One skilled in the art would appreciate that amino acids that are not conserved between wildtype, full-length stereocilin proteins from different species can be deleted without losing activity, while those amino acids that are conserved between wildtype, full-length stereocilin proteins from different species should not be deleted as they are more likely (than amino acids that are not conserved between different species) to be involved in activity.

In some examples, an active stereocilin protein can, e.g., include a sequence of a wildtype, full-length stereocilin protein that has between 1 amino acid to about 100 amino acids, 1 amino acid to about 95 amino acids, 1 amino acid to about 90 amino

[illegible]

[illegible]

[illegible]

[illegible]

acids to about 80 amino acids, about 60 amino acids to about 75 amino acids, about 60 amino acids to about 70 amino acids, about 60 amino acids to about 65 amino acids, about 70 amino acids to about 100 amino acids, about 70 amino acid to about 95 amino acids, about 70 amino acids to about 90 amino acids, about 70 amino acids to about 85 amino acids, about 70 amino acids to about 80 amino acids, about 70 amino acids to about 75 amino acids, about 80 amino acids to about 100 amino acids, about 80 amino acid to about 95 amino acids, about 80 amino acids to about 90 amino acids, about 80 amino acids to about 85 amino acids, about 90 amino acids to about 100 amino acids, about 90 amino acids to about 95 amino acids, or about 95 amino acids to about 100 amino acids, removed from its N-terminus and/or 1 amino acid to 100 amino acids (or any of the subranges of this range described herein) removed from its C-terminus.

In some embodiments, an active stereocilin protein can, e.g., include the sequence of a wildtype, full-length stereocilin protein where 1 amino acid to 50 amino acids, 1 amino acid to 45 amino acids, 1 amino acid to 40 amino acids, 1 amino acid to 35 amino acids, 1 amino acid to 30 amino acids, 1 amino acid to 25 amino acids, 1 amino acid to 20 amino acids, 1 amino acid to 15 amino acids, 1 amino acid to 10 amino acids, 1 amino acid to 9 amino acids, 1 amino acid to 8 amino acids, 1 amino acid to 7 amino acids, 1 amino acid to 6 amino acids, 1 amino acid to 5 amino acids, 1 amino acid to 4 amino acids, 1 amino acid to 3 amino acids, about 2 amino acids to 50 amino acids, about 2 amino acids to 45 amino acids, about 2 amino acids to 40 amino acids, about 2 amino acids to 35 amino acids, about 2 amino acids to 30 amino acids, about 2 amino acids to 25 amino acids, about 2 amino acids to 20 amino acids, about 2 amino acids to 15 amino acids, about 2 amino acids to 10 amino acids, about 2 amino acids to 9 amino acids, about 2 amino acids to 8 amino acids, about 2 amino acids to 7 amino acids, about 2 amino acids to 6 amino acids, about 2 amino acids to 5 amino acids, about 2 amino acids to 4 amino acids, about 3 amino acids to 50 amino acids, about 3 amino acids to 45 amino acids, about 3 amino acids to 40 amino acids, about 3 amino acids to 35 amino acids, about 3 amino acids to 30 amino acids, about 3 amino acids to 25 amino acids, about 3 amino acids to 20 amino acids, about 3 amino acids to 15 amino acids, about 3 amino acids to 10 amino acids, about 3 amino acids to 9 amino acids, about 3 amino acids to 8 amino acids, about 3 amino acids to 7 amino acids, about 3 amino acids to 6 amino acids, about 3 amino acids to 5 amino acids, about 4 amino acids to 50 amino acids, about 4 amino acids to 45 amino acids,

about 4 amino acids to 40 amino acids, about 4 amino acids to 35 amino acids, about
4 amino acids to 30 amino acids, about 4 amino acids to 25 amino acids, about 4
amino acids to 20 amino acids, about 4 amino acids to 15 amino acids, about 4 amino
acids to 10 amino acids, about 4 amino acids to 9 amino acids, about 4 amino acids to
5 8 amino acids, about 4 amino acids to 7 amino acids, about 4 amino acids to 6 amino
acids, about 5 amino acids to 50 amino acids, about 5 amino acids to 45 amino acids,
about 5 amino acids to 40 amino acids, about 5 amino acids to 35 amino acids, about
5 amino acids to 30 amino acids, about 5 amino acids to 25 amino acids, about 5
amino acids to 20 amino acids, about 5 amino acids to 15 amino acids, about 5 amino
10 acids to 10 amino acids, about 5 amino acids to 9 amino acids, about 5 amino acids to
8 amino acids, about 5 amino acids to 7 amino acids, about 6 amino acids to 50 amino
acids, about 6 amino acids to 45 amino acids, about 6 amino acids to 40 amino acids,
about 6 amino acids to 35 amino acids, about 6 amino acids to 30 amino acids, about
6 amino acids to 25 amino acids, about 6 amino acids to 20 amino acids, about 6
15 amino acids to 15 amino acids, about 6 amino acids to 10 amino acids, about 6 amino
acids to 9 amino acids, about 6 amino acids to 8 amino acids, about 7 amino acids to
50 amino acids, about 7 amino acids to 45 amino acids, about 7 amino acids to 40
amino acids, about 7 amino acids to 35 amino acids, about 7 amino acids to 30 amino
acids, about 7 amino acids to 25 amino acids, about 7 amino acids to 20 amino acids,
20 about 7 amino acids to 15 amino acids, about 7 amino acids to 10 amino acids, about
7 amino acids to 9 amino acids, about 8 amino acids to 50 amino acids, about 8 amino
acids to 45 amino acids, about 8 amino acids to 40 amino acids, about 8 amino acids
to 35 amino acids, about 8 amino acids to 30 amino acids, about 8 amino acids to 25
amino acids, about 8 amino acids to 20 amino acids, about 8 amino acids to 15 amino
25 acids, about 8 amino acids to 10 amino acids, about 10 amino acids to 50 amino acids,
about 10 amino acids to 45 amino acids, about 10 amino acids to 40 amino acids,
about 10 amino acids to 35 amino acids, about 10 amino acids to 30 amino acids,
about 10 amino acids to 25 amino acids, about 10 amino acids to 20 amino acids,
about 10 amino acids to 15 amino acids, about 15 amino acids to 50 amino acids,
30 about 15 amino acids to 45 amino acids, about 15 amino acids to 40 amino acids,
about 15 amino acids to 35 amino acids, about 15 amino acids to 30 amino acids,
about 15 amino acids to 25 amino acids, about 15 amino acids to 20 amino acids,
about 20 amino acids to 50 amino acids, about 20 amino acids to 45 amino acids,
about 20 amino acids to 40 amino acids, about 20 amino acids to 35 amino acids,

about 20 amino acids to 30 amino acids, about 20 amino acids to 25 amino acids, about 25 amino acids to 50 amino acids, about 25 amino acids to 45 amino acids, about 25 amino acids to 40 amino acids, about 25 amino acids to 35 amino acids, about 25 amino acids to 30 amino acids, about 30 amino acids to 50 amino acids, about 30 amino acids to 45 amino acids, about 30 amino acids to 40 amino acids, about 30 amino acids to 35 amino acids, about 35 amino acids to 50 amino acids, about 35 amino acids to 45 amino acids, about 35 amino acids to 40 amino acids, about 40 amino acids to 50 amino acids, about 40 amino acids to 45 amino acids, or about 45 amino acids to about 50 amino acids, are inserted. In some examples, the 1 amino acid to 50 amino acids (or any subrange thereof) can be inserted as a contiguous sequence into the sequence of a wildtype, full-length protein. In some examples, the 1 amino acid to 50 amino acids (or any subrange thereof) are not inserted as a contiguous sequence into the sequence of a wildtype, full-length protein. As can be appreciated in the art, the 1 amino acid to 50 amino acids can be inserted into a portion of the sequence of a wildtype, full-length protein that is not well-conserved between species.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Methods and materials are described herein for use in the present invention; other suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 is an exemplary schematic representation of a genetic map of a STRC vector (pITR-201; SEQ ID NO: 12; 4324 basepairs (bp)) that can be used in any of the present methods described herein. The vector includes an inverted terminal repeat (ITR) sequence (SEQ ID NO: 18), a 5' FVIII stuffer sequence (SEQ ID NO: 19), a CMV enhancer (SEQ ID NO: 20), a CMV promoter (SEQ ID NO: 21), a 5' untranslated region (UTR) sequence (SEQ ID NO: 24), a 5' STRC coding sequence (SEQ ID NO: 25), a splicing donor signal sequence (SD) (SEQ ID NO: 6), a highly

recombinogenic sequence from F1 phage (AK; SEQ ID NO: 26), and an ITR sequence (SEQ ID NO: 36).

Figure 2 is an exemplary schematic representation of a genetic map of a STRC vector (pITR-202; SEQ ID NO: 13; 4434 bp) that can be used in any of the present methods described herein. The vector includes an ITR sequence (SEQ ID NO: 18), an AK sequence (SEQ ID NO: 26), a SD sequence (SEQ ID NO: 6), a 3' STRC coding sequence (SEQ ID NO: 28), a 3' UTR sequence (SEQ ID NO: 33), a bovine growth hormone poly-adenylation signal (bGHpA) (SEQ ID NO: 34), a 3' FVIII stuffer sequence (SEQ ID NO: 35), and an ITR sequence (SEQ ID NO: 36).

Figure 3 is an exemplary schematic representation of a genetic map of a STRC vector (pITR-202GFP; SEQ ID NO: 14; 4693 bp) that can be used in any of the present methods described herein. The vector includes an ITR sequence (SEQ ID NO: 18), an AK sequence (SEQ ID NO: 26), a SD sequence (SEQ ID NO: 6), a 3' STRC coding sequence (SEQ ID NO: 28), a T2A sequence (SEQ ID NO: 29), a tGFP sequence (SEQ ID NO: 31), a 3' UTR sequence (SEQ ID NO: 33), a bGHpA sequence (SEQ ID NO: 34), and an ITR sequence (SEQ ID NO: 36).

Figure 4 is an exemplary schematic representation of a genetic map of a STRC vector (pITR-203; SEQ ID NO: 15; 4619 bp) that can be used in any of the present methods described herein. The vector includes an ITR sequence (SEQ ID NO: 18), a 5' FVIII stuffer_500 sequence (SEQ ID NO: 37), a CMV enhancer (SEQ ID NO: 20), a CMV promoter sequence (SEQ ID NO: 21), a 5' UTR sequence (SEQ ID NO: 24), a 5' STRC coding sequence (SEQ ID NO: 25), a SD sequence (SEQ ID NO: 6), an AP sequence (AP) (SEQ ID NO: 27), and an ITR sequence (SEQ ID NO: 36).

Figure 5 is an exemplary schematic representation of a genetic map of a STRC vector (pITR-204; SEQ ID NO: 16; 4729 bp) that can be used in any of the present methods described herein. The vector includes an ITR sequence (SEQ ID NO: 18), an AP sequence (SEQ ID NO: 27), a splicing acceptor signal (SA) sequence (SEQ ID NO: 7), a 3' STRC coding sequence (SEQ ID NO: 28), a 3' UTR sequence (SEQ ID NO: 33), a bGHpA sequence (SEQ ID NO: 34), and an ITR sequence (SEQ ID NO: 36).

Figure 6 is an exemplary schematic representation of a genetic map of a STRC vector (pITR-205; SEQ ID NO: 17; 4460 bp) that can be used in any of the present methods described herein. The vector includes an ITR sequence (SEQ ID

NO: 18), a CMV enhancer (SEQ ID NO: 20), a chicken β -actin (CBA) promoter (SEQ ID NO: 22), a chimeric intron sequence (SEQ ID NO: 23), a 5' UTR sequence (SEQ ID NO: 24), a 5' STRC coding sequence (SEQ ID NO: 25), a SD sequence (SEQ ID NO: 6), an AK sequence (SEQ ID NO: 26), and an ITR sequence (SEQ ID NO: 36).

Figure 7 is an image of an immunoblot of STRC protein levels from transfected HEK293FT cells with single or dual STRC vectors 72-hours post-transfection using a polyclonal STRC antibody. β -actin (ACTB) was used as a loading control. Lane 1 – Prestrained Page Rule; Lane 2 – 400 ng pITR-201 plus 400 ng pITR-202; Lane 3 – 400 ng pITR-201 plus 400 ng pITR-202GFP; Lane 4 – 400 ng pITR-203 plus 400 ng pITR-204; Lane 5 – 400 ng pITR-205 plus 400 ng pITR-202; Lane 6 – 400 ng pITR-205 plus 400 ng pITR-202GFP; Lane 7 – 800 ng pITR-202GFP; and Lane 8 – untransfected control cell lysate.

Figure 8 is an image of an immunoblot of STRC protein levels from transduced HEK293FT cells with dual STRC vectors (AAV2.STRC and Anc80.STRC) at different MOIs (2.6E5, 8.6E4, 2.6E5, and 8.6E4) 72-hours post-transfection using a polyclonal STRC antibody. β -actin (ACTB) was used as a loading control.

Figure 9 is a bar graph showing the relative RNA expression of STRC (5' STRC, 3' STRC and full-length STRC) relative to GAPDH in HEK293FT cells transduced with AAV2.STRC (at MOI 8.6E4 and MOI 2.6E5) and Anc80.STRC (at MOI 8.6E4 and 2.6E5), respectively.

Figure 10 is a bar graph showing the relative RNA expression of STRC (5' STRC, 3' STRC and full-length STRC) relative to GAPDH in P1-3 cochlear explants from WT mice infected 16-hours with AAV2.STRC (at MOI 1.3E10) and Anc80.STRC (at MOI 3.9E10), respectively.

Figure 11A is an exemplary fluorescent image of P1-3 cochlear explants from mock-infected WT mice showing Myo7a and phalloidin staining.

Figure 11B is an exemplary fluorescent image of P1-3 cochlear explants from mock-infected WT mice showing Myo7a and phalloidin staining. Magnification 20x.

Figure 11C is an exemplary fluorescent image of P1-3 cochlear explants from AAV.Anc80.STRC-infected (at MOI 7.8E8 VG/cochlea) WT mice showing Myo7a and phalloidin staining.

Figure 11D is an exemplary fluorescent image of P1-3 cochlear explants from AAVanc80.dSTRC-infected (at MOI 2.6E8 VG/cochlea) WT mice showing Myo7a and phalloidin staining. Magnification 20x.

5 **Figure 11E** is an exemplary fluorescent image of P1-3 cochlear explants from AAV.Anc80.STRC-infected (at MOI 7.8E8 VG/cochlea) WT mice showing Myo7a and phalloidin staining. Magnification 20x.

Figure 12A is an exemplary fluorescent image of P1-3 cochlear explants from AAV2.STRC-infected (at MOI 7.8E8 VG/cochlea) WT mice showing Myo7a and phalloidin staining.

10 **Figure 12B** is an exemplary fluorescent image of P1-3 cochlear explants from AAV2.STRC-infected (at MOI 2.6E8 VG/cochlea) WT mice showing Myo7a and phalloidin staining. Magnification 20x.

Figure 12C is an exemplary fluorescent image of P1-3 cochlear explants from AAV2.STRC-infected (at MOI 7.8E8 VG/cochlea) WT mice showing Myo7a and
15 phalloidin staining. Magnification 20x.

DETAILED DESCRIPTION

Provided herein are compositions that include at least two different nucleic acid vectors, where: each of the at least two different vectors includes a coding
20 sequence that encodes a different portion of a stereocilin protein, each of the encoded portions being at least 30 amino acid residues in length, where the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions; no single vector of the at least two different vectors encodes an active stereocilin protein (e.g., a full-
25 length stereocilin protein); at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons; and, when introduced into a mammalian cell comprising chromosomal DNA, the at least two different vectors undergo homologous recombination with each other and with the chromosomal DNA
30 of the cell, thereby forming a recombined nucleic acid inserted into the chromosomal DNA, where the recombined nucleic acid encodes an active stereocilin protein (e.g., a full-length stereocilin protein). Also provided are kits that include any of the compositions described herein. Also provided herein are methods that include

introducing into a cochlea of a mammal a therapeutically effective amount of any of the compositions described herein.

Also provided herein are methods of increasing expression of an active stereocilin protein (e.g., a full-length stereocilin protein) in a mammalian cell that include introducing any of the compositions described herein into the mammalian cell. Also provided herein are methods of increasing expression of an active stereocilin protein (e.g., a full-length stereocilin protein) in an outer hair cell in a cochlea of a mammal that include introducing into the cochlea of the mammal a therapeutically effective amount of any of the compositions described herein. Also provided herein are methods of treating non-symptomatic sensorineural hearing loss in a subject identified as having a defective stereocilin gene that include: administering a therapeutically effective amount of any of the compositions described herein into the cochlea of the subject.

Additional non-limiting aspects of the compositions, kits, and methods are described herein and can be used in any combination without limitation.

Stereocilin

The STRC gene encodes stereocilin, a protein that is normally expressed in the inner ear and is associated with the stereocilia of specialized hair cells in the inner ear (see, e.g., Verpy et al., *Nature* 456: 255-258, 2008; and Zhang et al., *J. Med. Genet.* 44: 233-240, 2007). Stereocilia, which are about 10-50 μ m in length, are important for hearing and balance. Stereocilia play a mechanosensing role in hearing. By bending in response to sound, they form a structure for mechanoreception of sound stimulation.

The human STRC gene is located on chromosome 15q15. It contains 29 exons encompassing ~19 kilobases (kb) (Verpy et al., *Nature Genetics* 29(3): 345-349, 2001; NCBI Accession No. NG011636.1). The gene is tandemly duplicated on chromosome 15q15 in a telomere-to-centromere orientation, with the tandem copies located less than 100kb apart on the chromosome. The coding sequence of the second copy is interrupted by a stop codon in exon 20, meaning that it is a “pseudogene” that encodes a nonfunctional truncated protein. (Hereinafter, references to “the STRC gene” mean the gene that does not have that stop codon in exon 20, while references to the “pseudogene” or the “STRC pseudogene” denote the gene that does have that

stop codon in exon 20.) In some examples, the full-length STRC protein is a full-length wildtype STRC protein. The full-length wildtype STRC protein expressed from the human STRC gene is 1775 residues in length, and contains a signal peptide and several hydrophobic segments. When the signal peptide is cleaved off during processing of the wildtype protein, the resulting mature wildtype stereocilin protein is 1719 residues in length. In some embodiments, a full-length STRC protein can include a heterologous signal peptide positioned N-terminal to an amino acid sequence of a mature wildtype stereocilin protein.

Various mutations in the STRC gene have been associated with hearing loss (e.g., non-symptomatic sensorineural hearing loss). For example, a homozygous 1-bp insertion was identified in exon 13 in a consanguineous Pakistani family with autosomal recessive non-syndromic sensorineural deafness-16 (DFNB16) (Verpy et al. (2001). Another example found in a family with DFNB16 includes both a 4-bp deletion in exon 5 and a larger deletion encompassing exons 17-29. The two deletions were inherited from the father and mother, respectively. The 4-bp deletion was predicted to result in the translation of 5 out-of-frame amino acids and a premature stop codon in exon 5.

Additional exemplary mutations in a stereocilin gene detected in subjects having non-symptomatic sensorineural hearing loss and methods of sequencing a nucleic acid encoding stereocilin are described in, e.g., Verpy et al., *Nature* 456:255-259, 2008; Verpy et al., *Res. Systems Neurosci.* 519:194-210, 2011; Hofrichter et al., *Clinical Genetics* 87:49-55, 2015; and Mandelker et al., *J. Mol. Diagnostics* 16:639-647, 2014. Methods of detecting mutations in a gene are well-known in the art. Non-limiting examples of such techniques include: real-time polymerase chain reaction (RT-PCR), PCR, sequencing, Southern blotting, and Northern blotting.

An exemplary human wildtype stereocilin protein is or includes the sequence of SEQ ID NO: 1. Non-limiting examples of nucleic acid encoding a wildtype stereocilin protein are or include SEQ ID NO: 2 and SEQ ID NO: 3. As can be appreciated in the art, at least some or all of the codons in SEQ ID NO: 2 can be codon-optimized to allow for optimal expression in a non-human mammal or in a human.

In some embodiments of any of the compositions described herein, the stereocilin protein comprises a sequence that is at least 75% (e.g., at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least

96%, at least 97%, at least 98%, at least 99%, or 100%) identical to SEQ ID NO: 1 or SEQ ID NO: 11. In some embodiments, a stereocilin protein can include a sequence that is identical to SEQ ID NO: 1, except that it includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid substitutions and/or deletions.

5

Human Full-length Wildtype Stereocilin Protein (SEQ ID NO: 1) (Signal sequence shown in bold)

MALSLWPLLLLLLLLLLLLLLSFAVTLAPTGFHSLDPGLSFLKSLLSTLDQAPQGSLSRSRFFTFLANISSS
FEPGRMGEGPVGEPPLQPPALRLHDFLVTLRGSPDWEPMGLLGDMALLGQEQTPrDFLVHQAGVLG
10 GLVEVLLGALVPGGPPTPRPPCTRDGPSDCVLAADWLPSSLLLLLEGTWQALVQVQPSVDPTNATGLD
GREAAPHFLLQGLLGLLTPTGELGSKEALWGGLLRTVGAPLYAAEQEGLLRVTHSLQDEVFSILGQPEPD
TNGQCQGVFFLTLSLLGNLQQLLLWGVNRHNSWDVQALGFLSGSPPPPPALLHCLSTGVPLPRASQPSA
HISPRQRRAITVEALCENHLGPAPPYSISNFSIHLLCQHTKPATPQPHPSTTAICQTAVWYAVSWAPGA
QGWLQACHDQFPDEFDAICSNLSFSALSGSNRRLLVKRLCAGLLPPPTSCPEGLPPVPLTPDIFWGCFL
15 ENETLWAERLCGEASLQAVPPSNQAWVQHVCGQPTPDVTASPPCHIGPCGERCPDGGSFVMVCANDTM
YEVLVPFWPWLQAGQCRISRGNDTCFLEGLLGLPLPSLPPLGPSPLCLTPGPFLGLMLSQLPQCQSSVP
ALAHPTRLHYLLRLTLFLLGPGAGGAEAQGMLGRALLSSLPDNCSEWDAFRPEGRSVLRTIGEYLEQ
DEEQPTPSGFEPVNPSSGISKMELLACFSPVLWDLQREKSVWALQILVQAYLHMPENLQQLVLSAE
REAAQGFLLTMLQGLQGLQVPPSEEQALGRLTALLQRYPRLTSQLFIDLSPILPFLAVSDLMRFPF
20 SLLANDSVQQGYRGSEENSLEEDKGLRPMTPRSLAAIRDYSPGMRPEQKEALAKRLLAPELFGEVPAWP
QELLWAVLPLPLPHLPLENFLQLSPHQIQALEDSWPAAGLPGHARHVLRSVLNQSVQDGEEQVRRGLPL
ACFLSPEELQSLVPLSDPTGPVERGLLECAANGTSLPEGRVAYELLGVLRSSGGAVLSPRELRVWAPLF
SQLGLRFLQELSEPQLRAMLPVLQGTSVTPAQAVLLLGRLLPRHDLSEELCSLHLLLPGLSPTQLQAI
PRRVLVGACSLAPELSRLSACQTAALLQTFRVKDGVKNMGTTGAGPAVCIPGQQPIPTTWPDCLLPLL
25 PLKLLQLDLSLALLANRRRYWELPWSEQQAQFLWKKMQVPTNLTLRNLQALGTLAGGMSCEFLQQINSMV
DFLEVVMHIYQLPTRVRGSLRACIWAELQRRMAMPEPEWTTVGPELNGLDKLLLDLPIQLMDRLSNES
IMLVVELVQRAPEQLLALTPLHQAALAERLQNLAPKETPVSGEVLETLGPLVGLGTSTRQIPLQIL
LSHLSQLQGFCGETFATELGWLLLQESVLGKPELWSQDEVEQAGRLVFTLSTEAISLIPREALGPETL
ERLLEKQQSWEQSRVGQLCREPQLAAKKAALVAGVVRPAAEDLPEPVPNCADVGRTFPAAWSATQIAEM
30 ELSDFEDCLTLFAGDPGLGPEELRAAMGKAKQLWGPPRGRPEQILQLGRLLIGLDRELQELILVDWG
VLSTLGQIDGWSTTQLRIVVSSFLRQSGRHVSHLDFVHLTALGYTLCGLRPEELQHISWEFSQAALFL
GTLHLQCSEEQLEVLALLVLPGGFGPISNWGPETFTIGTIAAGIPDLALSALLRGQIQGVTPLAISV
IPPPKFAVVFSPILQSSLTSAQAVAVTPEQMAFLSPEQRRAVAWAQHEGKESPEQQGRSTAWGLQDWSR
PSWSLVLTISFLGHLL

35

Human Stereocilin Signal Sequence (SEQ ID NO: 10)

MALSLWPLLLLLLLLLLLLLLSFAV

Human Full-length Wildtype Stereocilin Protein (Signal Sequence in Bold) (SEQ ID NO: 11)

5 **MALS**WPLLLLLLLLLLLLLLSFAVTLAPTGPSLSDPGLSFLKSLSTLDQAPQGSLSRSRFFTFLANISS
FEPGRMGEGPVGEPPPLQPPALRLHDFLVTLRGSPDWEPMGLLGMDLALLGQEQTPrDFLVHQAGVLG
GLVEVLLGALVPGGPPTPTRPPCTRDGPSDCVLAADWLPSLLLLLEGTWQALVQVQPSVDPPTNATGLD
GREAAAPHFLQGLLGLLTPTGELGSKEALWGGLLRTVGAPLYAAAFQEGLLRVTHSLQDEVFSILGQPEPD
TNGQCQGGNLQQLLLWGVRHNLSWDVQALGFSLSGSPPPFALLHCLSTGVPLPRASQPSAHISPRQRR
ITVEALCENHLGPAPPYSISNFSIHLLCQHTKPATPQPHPSTTAICQTAVWYAVSWAPGAQGWLQACHD
QFPDEFDLDAICSNLSFSALSGSNRRLVKRLCAGLLPPPTSCPEGLPPVPLTPDIFWGCFLNETLWAER
10 LCGEASLQAVPPSNQAWVQHVCQGPTPDVTASPPCHIGPCGERCPDGGSLFVMVCANDTMYEVLVPFWP
WLAGQCRI SRGGNDTCFLEGLLGPLLSPPLPGSPCLCTPGFPLLGMSLQPRCQSSVPALAHPTRLH
YLLRLTLFLGPGAGAEAGMLGRALLSSLPDNCSFWDFAFRPEGRRSVLRTIGEYLEQDEEQPTPSG
FEPTVNPSSGISKMELLACFSPVLWDLLQREKSVWALQILVQAYLHMPPENLQQLVLASAEREAQGFLT
LMLQGKLQGKLQVPPSEEQALGRLTALLLQRYPRLTSQLFIDLSPILPFLAVSDLMRFPPSLLANDSVL
15 AAIRDYSPGMRPEQKEALAKRLLAPELFGVEPAWPQELLWAVLPLPHLPLENFLQLSPHQIQALED SW
PAAGLGPGHARHVLRLSVNQSVQDGEQVRRGLGLACFLSPEELQSLVPLSDPTGPVERGLLECAANGT
LSPEGRVAYELLGVLRSSGGAVLSPRELRVWAPLFSQLGLRFLQELSEPQLRAMLPVLQGTSVTPAQAV
LLLGRLLRPHDLSLEELCSLHLLLPGLSPTLQAI PRVLVGACSCIAPELSRLSACQTAALLQTFRVK
DGVKNMGTTGAGPAVCIPGQPIPTTWPDCLLPLPLKLLQLDSALLANRRRYWELPWSEQQAQFLWKK
20 MQVPTNLTLRNLQALGTLAGGMSCEFLQQINSMVDFLEVVHMIYQLPTRVRGSLRACIWAELQRRMAM
EPEWTTVGPELNGLDSKLLLDLPIQLMDRLSNESIMLVVELVQRAPEQLLALTPLHQAAALAERALQNLA
PKETPVSGEVLETGLPLVGFLGTSTRQIPLQIILLSHLSQLQGFCIGETFATELGWLLQESVLGKPEL
WSQDEVEQAGRLVFTLSTEAI SLIPREALGPETLERLLEKQQSWEQSRVGQLCREPQLAAKKAALVAGV
VRPAAEDLPEPVPCADVRGTFPAAWSATQIAEMELSDFEDCLTLFAGDPGLGPEELRAAMGKAKQLWG
25 PPRGFRPEQILQLGRLLIGLGDRELQELIQLVDWGLVSTLGQIDGWSTQLRIVVSSFLRQSGRHVSHLD
FVHLTAIGYTLCLGRPEELQHSISWEFSQAALFLGLTLHLQCSSEQLLEVLAHLVLPGGFGPI SNWGPEI
FTEIGTIAAGIPDLALSALLRGQIQGVTPLAISVIPPKFAVVSPIQLSSLTSQAQAVAVTPEQMAFLS
PEQRRAVAWAQHEGKESPEQOGRSTAWGLQDWSRPSWSLVLTISFLGHLL

Human Wildtype Stereocilin cDNA (SEQ ID NO: 2)

[illegible]

CAAGGCTGGCTACAGGCCTGCCACGACCAGTTTCTGATGAGTTTTTGGATGCGATCTGCAGTAACCTC
TCCTTTTCAGCCCTGTCTGGCTCCAACCGCCGCTGGTGAAGCGGCTCTGTGCTGGCCTGCTCCCACCC
CCTACCAGCTGCCCTGAAGGCCTGCCCCCTGTTCCCCCAGACATCTTTGGGGCTGCTTCTTG
GAGAATGAGACTCTGTGGGCTGAGCGACTGTGTGGGGAGGCAAGTCTACAGGCTGTGCCCCCAGCAAC
5 CAGGCTTGGGTCCAGCATGTGTGCCAGGGCCCCACCCAGATGTCACTGCCTCCCCACCATGCCACATT
GGACCCTGTGGGGAACGCTGCCCCGATGGGGGCAGCTTCTGGTGATGGTCTGTGCCAATGACACCATG
TATGAGGTCCTGGTGCCCTTCTGGCCTTGGCTAGCAGGCCAATGCAGGATAAGTCGTGGGGGCAATGAC
ACTTGCTTCTAGAAAGGGCTGCTGGGCCCCCTTCTGCCCTCTCTGCCACCACTGGGACCATCCCCACTC
TGTCTGACCCCTGGCCCCCTTCTCCTTGGCATGCTATCCCAGTTGCCACGCTGTCACTCCTCTGTCCCA
10 GCTCTTGCTCACCCACACGCCTACACTATCTCCTCCGCCTGCTGACCTTCTCTTGGGTCCAGGGGCT
GGGGGCGCTGAGGCCCAGGGGATGCTGGGTGCGGCCCTACTGCTCTCCAGTCTCCCAGACAACCTGCTCC
TTCTGGGATGCCTTTTCGCCCAGAGGGCCGGCGCAGTGTGCTACGGACGATTGGGGAATACCTGGAACAA
GATGAGGAGCAGCCAACCCCATCAGGCTTTGAACCCACTGTCAACCCAGCTCTGGTATAAGCAAGATG
GAGCTGCTGGCCTGCTTTAGTCTGTGCTGTGGGATCTGCTCCAGAGGAAAAGAGTGTGGGGCCCTG
15 CAGATTCTAGTGACGGCGTACCTGCATATGCCCCCAGAAAACCTCCAGCAGCTGGTGCTTTCAGCAGAG
AGGGAGGCTGCACAGGGCTTCTGACACTCATGCTGCAGGGGAAGCTGCAGGGGAAGCTGCAGGTACCA
CCATCCGAGGAGCAGGCCCTGGGTGCGCTGACAGCCCTGCTGCTCCAGCGGTACCCACGCCTCACCTCC
CAGCTCTTCATTGACCTGTCAACCACTCATCCCTTTCTTGGCTGTCTCTGACCTGATGCGCTTCCCACCA
TCCCTGTTAGCCAACGACAGTGTACAGCAGGGCTACAGAGGGTCAGAGGAAAACAGTTTGGAGGAAGAC
20 AAAGGGTTAAGACCCATGACTCCTCGCAGCCTGGCTGCCATCCGGGATTACAGCCCAGGAATGAGGCCT
GAACAGAAGGAGGCTCTGGCAAAGCGACTGCTGGCCCCCTGAACTGTTTGGGGAAGTGCCTGCCTGGCCC
CAGGAGCTGCTGTGGGCAGTGCTGCCCCCTGCTCCCCCACCTCCCTCTGGAGAACTTTTTGCAGCTCAGC
CCTCACAGATCCAGGCCCTGGAGGATAGCTGGCCAGCAGCAGGTCTGGGGCCAGGGCATGCCCGCCAT
GTGCTGCGCAGCCTGGTAAACCAGAGTGTCCAGGATGGTGAGGAGCAGGTACGCAGGCTTGGGCCCCTC
25 GCCTGTTTCTGAGCCCTGAGGAGCTGCAGAGCCTAGTGCCCCCTGAGTGATCCAACGGGGCCAGTAGAA
CGGGGGCTGCTGGAATGTGCAGCCAATGGGACCCCTCAGCCCAGAAGGACGGGTGGCATATGAACTTCTG
GGTGTGTTGCGCTCATCTGGAGGAGCGGTGCTGAGCCCCCGGGAGCTGCGGGTCTGGGCCCCCTCTCTTC
TCTCAGCTGGGCCTCCGCTTCTTTCAGGAGCTGTGAGAGCCCCAGCTTAGAGCCATGCTTCTGTCTCTG
CAGGGAAGTAGTGTTACACCTGCTCAGGCTGTCTGCTGCTTGGACGGCTCCTTCTTAGGCACGATCTA
30 TCCCTGGAGGAACCTCTGCTCCTTGACCTTCTGCTACCAGGCCTCAGCCCCCAGACACTCCAGGCCATC
CCTAGGCGAGTCTGGTTCGGGGCTTGTTCCTGCCTGGCCCCCTGAACTGTCAGCCTCTCAGCCTGCCAG
ACCGCAGCACTGCTGCAGACCTTTTCGGGTAAAGATGGTGTTAAAAATATGGGTACAACAGGTGCTGGT
CCAGCTGTGTGTATCCCTGGTCAGCAGCCTATTCCCACCACCTGGCCAGACTGCCTGCTTCCCCTGCTC
CCATTAAAGCTGCTACAACTGGATTCTTGGCTCTTCTGGCAAATCGAAGACGCTACTGGGAGCTGCCC
35 TGGTCTGAGCAGCAGGCACAGTTTCTCTGGAAGAAGATGCAAGTACCCACCAACCTTACCCTCAGGAAT
CTGCAGGCTCTGGGCACCCCTGGCAGGAGGCATGTCTGTGAGTTTCTGCAGCAGATCAACTCCATGGTA
GACTTCTTGAAGTGGTGCACATGATCTATCAGCTGCCCACTAGAGTTCGAGGGAGCCTGAGGGCCTGT
ATCTGGGCAGAGCTACAGCGGAGGATGGCAATGCCAGAACCAGAATGGACAACCTGTAGGGCCAGAAGTG
AACGGGCTGGATAGCAAGCTACTCCTGGACTTACCAGTCCAGTTGATGGACAGACTATCCAATGAATCC
40 ATTATGTTGGTGGTGGAGCTGGTGCAAAGAGCTCCAGAGCAGCTGCTGGCACTGACCCCCCTCCACCAG
GCAGCCCTGGCAGAGAGGGCACTACAAAACCTGGCTCCAAAGGAGACTCCAGTCTCAGGGGAAGTGCTG

GAGACCTTAGGCCCTTTGGTTGGATTCTTGGGGACAGAGAGCACACGACAGATCCCCCTACAGATCCTG
 CTGTCCCATCTCAGTCAGCTGCAAGGCTTCTGCCTAGGAGAGACATTTGCCACAGAGCTGGGATGGCTG
 CTATTGCAGGAGTCTGTTCTTGGGAAACCAGAGTTGTGGAGCCAGGATGAAGTAGAGCAAGCTGGACGC
 CTAGTATTCACTCTGTCTACTGAGGCAATTTCTTGATCCCCAGGGAGGCCTTGGGTCCAGAGACCCTG
 5 GAGCGGCTTCTAGAAAAGCAGCAGAGCTGGGAGCAGAGCAGAGTTGGACAGCTGTGTAGGGAGCCACAG
 CTTGCTGCCAAGAAAGCAGCCCTGGTAGCAGGGGTGGTGCGACCAGCTGCTGAGGATCTTCCAGAACCT
 GTGCCAAATTGTGCAGATGTACGAGGGACATTCCCAGCAGCCTGGTCTGCAACCCAGATTGCAGAGATG
 GAGCTCTCAGACTTTGAGGACTGCCTGACATTATTTGCAGGAGACCCAGGACTTGGGCCTGAGGAACCTG
 CGGGCAGCCATGGGCAAAGCAAAACAGTTGTGGGGTCCCCCGGGGATTTCTGTCCTGAGCAGATCCTG
 10 CAGCTTGGTAGGCTCTTAATAGGTCTAGGAGATCGGGAACTACAGGAGCTGATCCTAGTGGACTGGGGA
 GTGCTGAGCACCCCTGGGGCAGATAGATGGCTGGAGCACCACTCAGCTCCGCATTGTGGTCTCCAGTTTC
 CTACGGCAGAGTGGTCGGCATGTGAGCCACCTGGACTTCGTTTCATCTGACAGCGCTGGGTTATACTCTC
 TGTGGACTGCGGCCAGAGGAGCTCCAGCACATCAGCAGTTGGGAGTTCAGCCAAGCAGCTCTCTTCCTC
 GGCACCCCTGCATCTCCAGTGCTCTGAGGAACAACCTGGAGGTTCTGGCCCACCTACTTGTACTGCCTGGT
 15 GGGTTTGGCCCAATCAGTAACTGGGGGCCTGAGATCTTCACTGAAATTGGCACCATAGCAGCTGGGATC
 CCAGACCTGGCTCTTTCAGCACTGCTGCGGGGACAGATCCAGGGCGTTACTCCTCTTGCCATTTCTGTCT
 ATCCCTCCTCCTAAATTTGCTGTGGTGTTTAGTCCCATCCAACCTATCTAGTCTCACCAGTGCTCAGGCT
 GTGGCTGTCACTCCTGAGCAAATGGCCTTTCTGAGTCTTGAGCAGCGACGAGCAGTTGCATGGGGCCAA
 CATGAGGGAAAGGAGAGCCCAGAACAGCAAGGTGGAAGTACAGCCTGGGGCCTCCAGGACTGGTCACGA
 20 CCTTCCTGGTCCCTGGTATTGACTATCAGCTTCCTTGGCCACCTGCTATGA

Human Wildtype Stereocilin cDNA – Codon Optimized (SEQ ID NO: 3)

ATGGCTCTGTCTCTGTGGCCTCTGCTGCTGCTCCTGTTGCTGCTTCTGCTGCTCAGCTTCGCCGTGACA
 CTGGCTCCAACAGGACCCCCACTCTCTGGATCCTGGCCTGAGCTTTCTGAAGTCCCTGCTGAGCACCCCTG
 25 GATCAGGCTCCTCAGGGCAGCCTGAGCAGATCCAGATTCTTCACCTTCCTGGCCAACATCAGCAGCAGC
 TTCGAGCCTGGCAGAATGGGAGAAGGACCTGTGGGAGAACCCTCCTCCACTGCAACCTCCAGCTCTGCGG
 CTGCACGATTTTCTGGTCACACTGAGAGGCAGCCCCGACTGGGAACCTATGCTGGGACTGCTGGGAGAT
 ATGCTGGCCCTGCTCGGACAAGAGCAGACCCCTAGAGATTTCTGGTGCATCAGGCTGGCGTGCTCGGA
 GGACTGGTTGAAGTTCTGCTTGGAGCACTGGTGCTGGCGGACCTCCTACACCTACAAGACCTCCATGC
 30 ACCAGAGATGGCCCCAGCGATTGTGTGCTGGCTGCTGATTGGCTGCCTAGCCTGCTCCTGCTTCTGGAA
 GGCACAAGATGGCAGGCCCTGGTGCAGGTTTCAGCCTTCTGTGGATCCTACCAATGCCACCGGCCCTGGAT
 GGAAGAGAAGCCGCTCCACACTTTCTGCAGGGACTGCTTGGACTGCTCACACCTACTGGCGAGCTGGGC
 TCTAAAGAAGCCCTTTGGGGAGGCCTGCTGAGAACAGTTGGAGCCCCCTGTACGCCGCCCTTTCAAGAG
 GGACTCCTGAGAGTGACACACAGCCTGCAGGACGAGGTGTTTCAGCATCCTGGGACAGCCAGAGCCTGAC
 35 ACCAATGGACAGTGTGAGGGCGTGTTCTTTCTGACCCTGAGCCTGCTGGGCAACCTGCAGCAACTGCTT
 CTGTGGGGCGTCAGACACAACCTGAGCTGGGATGTGCAGGCACTGGGCTTTCTGTCTGGAAGCCCTCCA
 CCTCCACCAGCACTGCTGCATTGTCTGTCTACTGGCGTGCCACTGCCTAGAGCCTCTCAGCCTAGCGCT
 CACATCAGCCCCAGACAGAGAAGGGCCATCACCGTGGAAGCTCTGTGCGAGAATCACCTGGGACCTGCT
 CCTCCATACAGCATCAGCAACTTCTCCATCCATCTGCTGTGTGTCAGCACACCAAGCCTGCCACACCTCAG
 40 CCTCATCCTAGCACACAGCCATCTGTGAGACCGCCGTTTGGTACGCCGTTTCTTGGGCTCCTGGTGCT

CAAGGATGGCTGCAGGCCTGTCACGATCAGTTCCCCGACGAGTTCCTGGACGCCATCTGCAGCAATCTG
AGCTTCTCTGCCCTGTCCGGCAGCAATCGGAGACTGGTCAAAAAGACTGTGCGCCGGACTGCTGCCCTCCT
CCAACATCTTGTCTGAGGGACTGCCTCCTGTGCCCTCTGACTCCCGATATCTTTTGGGGCTGCTTCCTG
GAAAACGAGACACTGTGGGCCGAGAGACTGTGTGGCGAAGCCTCTCTGCAAGCCGTGCCTCCATCTAAT
5 CAGGCCTGGGTGCAGCACGTTTGTGAGGGCCCTACACCTGACGTGACAGCCTCTCCTCCTTGTACATC
GGACCTTGCGGCGAGAGATGTCTGATGGCGGCAGCTTCTCGTGATGGTCTGCGCCAACGACACTATG
TACGAGGTGCTGGTGCCCTTCTGGCCTTGGCTGGCTGGCCAGTGTAGAATCTCCAGAGGCGGCAACGAT
ACCTGTTTCTGGAAGGACTCCTGGGACCACTGCTTCCATCTCTGCCTCCGCTTGGACCTCTCCTCTG
TGTCTTACCCCTGGACCTTTCCTGCTGGGGATGCTGTCTCAGCTGCCTAGATGCCAGTCTAGCGTGCCA
10 GCTCTGGCCCATCCAACCAGACTGCACTATCTGTTGCGGCTGCTGACCTTCCTCCTTGGACCTGGCGCT
GGCGGAGCTGAAGCTCAAGGCATGCTTGGCAGAGCCCTGCTGCTGTCATCCCTGCCTGACAACCTGCAGC
TTCTGGGACGCCTTCAGACCTGAAGGACGCAGAAGCGTGCTGAGGACCATCGGCGAGTACCTGGAACAG
GATGAGGAACAGCCTACACCAAGCGGCTTCAACCCACCGTGAATCCTAGCAGCGGCATCTCCAAGATG
GAACTGCTGGCCTGCTTCAGCCCCGTGCTGTGGGATCTGCTGCAGAGGGAAAAAGCGTGTTGGGCCCTG
15 CAGATCCTGGTCCAGGCCTATCTGCACATGCCTCCAGAGAACCTGCAACAGCTGGTGCTGTCTGCCGAG
AGAGAAGCTGCCCAGGGATTCTGACTCTGATGCTGCAGGGAAAAGCTGCAGGGCAAACCTCAGGTGCCA
CCTAGCGAAGAACAGGCCCTTGGCAGACTGACAGCACTCCTGCTCCAGAGATACCCAGACTGACCTCT
CAGCTGTTTATCGATCTGAGCCCTCTGATCCCATTCCTGGCCGTGTCGACCTGATGAGATTCCTCCT
AGTCTGCTGGCCAACGACTCCGTGCAGCAGGGATACAGAGGCAGCGAGGAAAACAGCCTGGAAGAGGAC
20 AAAGGCCTGCGGCCTATGACACCTAGATCTCTGCGCGCCATCCGGGATTACAGCCCTGGAATGAGGCC
GAGCAGAAAGAGGCCCTGGCTAAGAGACTGCTGGCTCCCGAGCTGTTTGGCGAAGTTCCAGCCTGGCCT
CAAGAACTGCTGTGGGCTGTTCTGCCCCCTGCTGCCACATCTGCCTCTGGAAAACCTTCCTGCAGCTGAGC
CCACACCAGATTACAGGCCCTCGAGGATTCTTGGCCTGCCGCTGGACTCGGACCTGGACATGCTAGACAC
GTGCTGAGAAGCCTGGTCAACCAGAGCGTTCAGGACGGCGAGGAACAAGTGGGAGACTTGGACCCCTG
25 GCCTGTTTTCTGAGCCCCGAGGAACTGCAGAGTCTGGTGCCCTCTGTCTGACCTACAGGCCCTGTTGAA
AGGGGCCTGCTGGAATGTGCCGCAATGGAACACTGAGCCCTGAGGGCAGAGTGGCCTATGAACTTCTG
GGAGTGCTGAGATCTAGCGGCGGAGCCGTTCTGTCCCTAGGGAACTTAGAGTGTTGGGCACCCCTGTTT
AGCCAGCTGGGCCTGAGATTCTGCAAGAGCTGTCTGAGCCACAGCTGAGGGCTATGCTGCCTGTGCTC
CAGGGCACATCTGTGACACCAGCTCAAGCCGTCTGCTGTTGGGCAGACTGCTCCCTAGACACGATCTG
30 TCCCTGGAAGAACTGTGCAGCCTGCACTTGCTGCTGCCTGGACTGTCTCCTCAGACACTGCAGGCCATT
CCTCGGAGAGTGCTTGTGGGAGCCTGTAGCTGTCTGGCCCCCTGAGCTGTCTAGACTGAGCGCCTGTCAA
ACAGCCGCTCTCCTGCAGACCTTCAGAGTGAAGGATGGCGTGAAGAACATGGGCACCACAGGCGCTGGA
CCTGCCGTGTGTATTCTGACAGCAGCCATTCTACCACCTGGCCAGATTGTCTGCTCCCACTGCTG
CCCCTGAAACTGCTGCAGCTGGATTCTCTGGCTCTGCTGGCTAACC GGCGGAGATATTGGGAACTGCCT
35 TGGAGCGAACAGCAGGCACAGTTCTGTGGAAGAAGATGCAGGTCCCCACCAATCTGACACTGCGGAAT
CTGCAGGCTCTCGGCACACTTGCTGGCGGAATGAGCTGCGAGTTCTCCAGCAGATCAACAGCATGGTG
GACTTCTGGAAGTGGTGCACATGATCTACCAGCTGCCAACCCAGAGTGCGGGGAAGCCTGAGAGCTTGT
ATTTGGGCTGAACTGCAGCGGCGGATGGCCATGCCTGAACCTGAATGGACAACAGTGGGCCCCGAGCTG
AACGGCCTGGACTCTAAACTTCTGCTGGATCTCCCCATCCAGCTGATGGACAGACTGAGCAACGAGAGC
40 ATCATGCTGGTGGTGGAACTGGTGCAGAGAGCCCCAGAACAGCTGCTGGCACTGACACCTCTGCATCAA
GCTGCTCTGGCCGAACGGGCCCTTCAGAATCTGGCTCCCAAAGAAACCCCTGTGTCCGGGGAAGTGCTG

GAAACACTGGGACCTCTTGTGGGCTTCCTGGGCACCGAGTCTACCAGACAGATTCTCTCCAGATCCTG
 CTGTCCCACCTGAGCCAGCTCCAGGGATTTTGTCTGGGCGAGACATTCGCCACCGAACTCGGATGGTTG
 CTGCTCCAAGAGAGCGTGCTGGGAAAGCCCGAACTGTGGTCACAGGACGAAGTGGAAACAGGCCGGCAGA
 CTGGTGTTTACCCTGTCTACCGAGGCCATCAGTCTGATCCCCAGAGAAGCACTGGGCCCTGAAACACTC
 5 GAGAGGCTGCTGGAAAAGCAGCAGTCTTGGGAACAGAGCAGAGTGGGCCAGCTGTGTAGAGAACCTCAG
 CTGGCCGCTAAAAAGGCCGCACTGGTTGCTGGCGTTGTCTCAGACCTGCTGCCGAGGATCTGCCAGAGCCA
 GTGCCTAATTGTGCCGATGTGCGGGGCACATTTCTGCCGCTTGGAGCGCTACACAGATCGCCGAAATG
 GAACTGAGCGACTTCGAGGACTGTCTGACTCTGTTTGCCGGCGATCCTGGACTGGGACCAGAAGAAGTCTG
 AGAGCCGCCATGGGCAAAGCCAAGCAACTTTGGGGACCTCCAAGAGGCTTCAGACCCGAACAGATTCTG
 10 CAGCTCGGCCGCTGCTTATCGGACTGGGCGATAGAGAGCTGCAAGAACTGATCCTGGTGGACTGGGGC
 GTGCTGTCTACTCTGGGACAAATCGACGGCTGGTCCACCACACAGCTGCGGATTGTGGTGTCCAGCTTC
 CTGAGGCAGTCTGGCAGACATGTGTCTCACCTGGACTTCGTGCACCTGACTGCCCTGGGCTACACACTG
 TGTGGACTGCGGCCTGAGGAACTTCAGCACATCAGCTCTTGGGAGTTCAGCCAGGCAGCCCTGTTTCTG
 GGAACCTTGCATCTGCAGTGCTCCGAAGAACAAGTGGAAAGTGTCTCGCCCATCTGCTCGTGCTGCCAGGC
 15 GGATTTGGCCCCATCTCTAATTGGGGACCCGAGATCTTCACCGAGATCGGCACCATTGCCGCTGGCATC
 CCTGATCTGGCCCTGTCTGCACTGCTGAGAGGACAGATCCAGGGCGTGACACCACTGGCCATTAGCGTG
 ATCCCTCCACCAAAGTTCGCCGTGGTGTTTAGCCCCATTGAGCTGTCCAGCCTGACATCTGCCCAAGCC
 GTGGCTGTGACCCCTGAACAGATGGCATTTCTGTCTCCCGAGCAGCGGAGAGCTGTTGCTTGGGCTCAA
 CACGAGGGCAAAGAGAGTCTGAGCAACAGGGAAGAAGCACCGCTTGGGGACTGCAGGATTGGTCCAGA
 20 CCTTCTTGGTCACTGGTGCTGACCATCAGCTTTCTGGGCCACCTCCTGTGA

A non-limiting example of a human wildtype stereocilin genomic DNA
 sequence is SEQ ID NO: 4. The exons in SEQ ID NO: 4 are: nucleotide positions 1-
 142 (exon 1), nucleotide positions 445-1230 (exon 2), nucleotide positions 1319-1343
 25 (exon 3), nucleotide positions 2111-3368 (exon 4), nucleotide positions 4325-4387
 (exon 5), nucleotide positions 4489-4605 (exon 6), nucleotide positions 4748-4914
 (exon 7), nucleotide positions 5570-5756 (exon 8), nucleotide positions 5895-6010
 (exon 9), nucleotide positions 6292-6424 (exon 10), nucleotide positions 6777-6959
 (exon 11), nucleotide positions 7264-7302 (exon 12), nucleotide positions 7486-7653
 30 (exon 13), nucleotide positions 7817-7882 (exon 14), nucleotide positions 8364-8489
 (exon 15), nucleotide positions 9467-9525 (exon 16), nucleotide positions 10598-
 10721 (exon 17), nucleotide positions 10826-10938 (exon 18), nucleotide positions
 13402-13537 (exon 19), nucleotide positions 13955-14151 (exon 20), nucleotide
 positions 14350-14440 (exon 21), nucleotide positions 14649-14805 (exon 22),
 35 nucleotide positions 15390-15559 (exon 23), nucleotide positions 17250-17405 (exon
 24), nucleotide positions 17787-17929 (exon 25), nucleotide positions 18119-18267
 (exon 26), nucleotide positions 18509-18605 (exon 27), nucleotide positions 18694-

18841 (exon 28), and nucleotide positions 19041-19238 (exon 29). The introns are located between each of these exons in SEQ ID NO: 4, i.e., at nucleotide positions 143-444 (intron 1), nucleotide positions 1231 to 1318 (intron 2), nucleotide positions 1344-2110 (intron 3), nucleotide positions 3369-4324 (intron 4), nucleotide positions 4388-4488 (intron 5), nucleotide positions 4606-4747 (intron 6), nucleotide positions 4915-5569 (intron 7), nucleotide positions 5757-5894 (intron 8), nucleotide positions 6011-6291 (intron 9), nucleotide positions 6425-6776 (intron 10), nucleotide position 6960-7263 (intron 11), nucleotide positions 7303-7485 (intron 12), nucleotide positions 7654-7816 (intron 13), nucleotide positions 7883-8363 (intron 14), nucleotide positions 8490-9466 (intron 15), nucleotide positions 9526-10597 (intron 16), nucleotide positions 10722-10825 (intron 17), nucleotide positions 10939-13401 (intron 18), nucleotide positions 13538-13954 (intron 19), nucleotide positions 14152-14349 (intron 20), nucleotide positions 14441-14648 (intron 21), nucleotide positions 14806-15389 (intron 22), nucleotide positions 15560-17249 (intron 23), nucleotide positions 17406-17786 (intron 24), nucleotide positions 17930-18118 (intron 25), nucleotide positions 18268-18508 (intron 26), nucleotide positions 18606-18693 (intron 27), and nucleotide positions 18842-19040 (intron 28).

Human Wildtype Stereocilin Gene (SEQ ID NO: 4)

```

20 gccctgccct cacctggcta tcccacacag gtgagaataa ccagaactca cctccggtac
   cagtgttcac ttggaaacat ggctctcagc ctctggcccc tgctgctgct gctgctgctg
   ctgctgctgc tgtcctttgc aggtaagaag aacagtgagc agaactgggg atgaggagga
   ggggtggctgg aaaaagactt taagaatatg gaggtgaacc tgtagatag aaggacaaag
   gagagaggca gagacttgtg caaaaggga aaatgagggt taagaaaagc aggccaagac
25 ttactgtagg ccagtgaag gggttcagct caccatcccc tcacctcatc ttagatcca
   ggtagggaaac tgtgctcagg ggcagggttg agtttgggct ctgtgttcct ctccttcagt
   gacctctggg ttctctcctt acagtgactc tggcccctac tgggcctcat tccttgacc
   ctgggtctctc cttcctgaag tcattgctct cactctgga ccaggctccc cagggtccc
   tgagccgctc acggttcttt acattcctgg ccaacatttc ttcttccttt gagcctggga
30 gaatggggga aggaccagta ggagagcccc cacctctcca gccgcctgct ctgcggctcc
   atgattttct agtgacactg agaggtagcc ccgactggga gccaatgcta gggctgctag
   gggatatgct ggcactgctg ggacaggagc agactccccg agatttcctg gtgcaccagg
   caggggtgct ggggtgactt gtggagggtg tgctgggagc ctagttcct gggggcccc
   ctaccccaac tcggcccca tgcaccctg atgggccgtc tgactgtgtc ctggctgctg
35 actggttgcc ttctctgctg ctgttggttag agggcacacg ctggcaagct ctggtgcagg
   tgcagcccag tgtggacccc accaatgcca caggcctcga tgggaggag gcagctcctc
   actttttgca ggggtctgtt ggtttgctta cccaacagg ggagctaggc tccaaggagg
   ctctttgggg cggctctgta cgcacagtgg gggccccct ctatgctgcc tttcaggagg

```

ggctgctccg tgtcactcac tccctgcagg atgaggtctt ctccatthttg gggcagccag
 agcctgatac caatgggcag tgccagggag gtgagtgtgg ccagggctgg gactgggatg
 tggcagggca aggaaagtga aattggggta gttttcttcc ttactctttc cctcctaggt
 aaccttcaac agctgctctt atggtaagta acaggagacc agttctgagg gattgggcct
 5 ggaaaatctg gaggtgaaga gctgaagacc tcagcctcta gagaggaaaa ctgatgggag
 gagtgtagtt tagtggtttt ggggtgtgac tgtctgggtt ggtgtcccag ctccacctct
 tcttagccat atgaccttga gcaggttaca tagtctttct atacctcagt tccccattt
 ataaaatgag aatgataata ttagttacca cagagttgtt gcacccggtt aaatgagttg
 atactgtgta tgcaaacgac ttaaaacgt gctggcacat agcgcttaat aatgtagct
 10 agtaaagatg ggatttgga aataaggaca cagctggatt cctctacccc ctactactt
 cagtacaaca atgccagaca gtagttagac atattgagtt gctgagcaga tttcctaaca
 tgaggcccg ctaggggttg gtttaagcta tctaaaagca tacgaagaaa ggagacagaa
 gggggccagg tggacagaaa gaattccaac tggggcttct cctaggtgat tttggacctt
 ggcagggcag ctttctcttt tttgccccgt tgcagcattt caaccagtaa cgcctaaact
 15 ctgagggacc tcgctttag aaaagcctat gcttgccatg ccccttgagg gctctgagtc
 agggctcagaa tcttcagctg gaggaatgt gaactgacca gatcctgcct gctcctccct
 ctgcacccag gggcgctccg cacaaccttt cctgggatgt ccaggcgctg ggctttctgt
 ctggatcacc acccccaccc cctgcctctc ttcactgcct gagcacgggc gtgcctctgc
 ccagagcttc tcagccgtca gccacatca gccacgcca acggcgagcc atcactgtgg
 20 aggcctctg tgagaaccac ttagggccag caccacccta cagcatttcc aacttctcca
 tccacttgct ctgccagcac accaagcctg cactccaca gcccacccc agcaccactg
 ccatctgcca gacagctgtg tggtatgcag tgcctgggc accaggtgcc caaggctggc
 tacaggcctg ccacgaccag tttcctgatg agtttttga tgcgatctgc agtaacctct
 ccttttcagc cctgtctggc tccaaccgcc gcctggtgaa gcggctctgt gctggcctgc
 25 tcccccccc taccagctgc cctgaaggcc tgccccctgt tccccacc ccagacatct
 tttggggctg cttcttgagg aatgagactc tgtgggctga gcgactgtgt ggggaggcaa
 gtctacaggc tgtgcccccc agcaaccagg cttgggtcca gcatgtgtgc cagggcccca
 cccagatgt cactgcctcc ccacatgcc acattggacc ctgtggggaa cgctgcccgg
 atgggggcag cttcctggtg atgggtctgt ccaatgacac catgtatgag gtctggtgc
 30 ccttctggcc ttggctagca ggccaatgca ggataagtcg tgggggcaat gacacttgct
 tctagaagg gctgctgggc ccccttctgc cctctctgcc accactggga ccatccccac
 tctgtctgac cctggcccc ttcctccttg gcatgctatc ccagttgcca cgctgtcagt
 cctctgtccc agctcttgct caccacacac gcctacacta tctcctccgc ctgctgacct
 tctcttggg tccaggggct gggggcgctg agggccagg gatgctgggt cgggccctac
 35 tgctctccag tctcccagac aactgctcct tctgggatgc ctttcgcca gagggccggc
 gcagtgtgct acggacgatt ggggaatacc tggaaacaaga tgaggagcag ccaaccccat
 caggctttga acccactgtc aacccagct ctggtataag caagatggag ctgctggcct
 gcttttagtgt gagtgtctct ccagagggaa agctcctaga acagtgagaa ggccctccag
 gggaattcct cgaataactca gaggcagtag tgtgggtag tagttgaagc acacagctct
 40 agagtccagc aggcttgat tcatacttg gttctgtgac cagccttgaa tgagttatth
 aacttctctg agcaataatth ttctcgtctc atthataaac tagggatgat aatggtatat
 gagataatac atgctgtggg cttagcacag tgcattgatac acaaacatgc aataaatatt
 acctgttat tcttttgggc tctttgactc tctcactthc tgcaccagaa agaaaaagga
 tcaagttaga ggactctaaa tthttccct agagagtgc aattggaggc tggcagaata

caggaagata aggtaggaat gagaaagatt cagggacact accaatcaga agacttttgg
 tctaggttca actgtgccac aaattagtgat gatcttaggc aagcaatttc atttagtttt
 tctgggcttc agtttttagt ctgtagaatg gaggggtgag aatatgttaa acaccataat
 taattcactg agtgcctatt atatgcaagg cacttttgcta ggttctgtag gatataataa
 5 gattttcttac tccatgttgg ggccaccttt ttcaaaccct gggcccagta aaatggaatt
 agatagtctc atagtatttg gttcagggtct acaagtatta attgagccaa ctatggacct
 ggcatgggag aggttacaag agaaattaga gatatgatcc cggacctaaa agagcttaat
 atctgaagaa tcacacttga gatgatggac aagcatccca gcaagtggag ctggaatgcc
 tgggggagct gcaggagaga cagagaagac agctctgttg gcatattgtc tttcttccca
 10 ccagcctgtg ctgtgggatc tgctccagag ggaaaagagt gtttgggccc tgcagattct
 agtgcaggta acagggtggag ggcacatggg tgggctgggt gacagccatg gctggaggtc
 cctgccccgt gaggtgaggc catacccacc atgacctcct attcgcaggc gtacctgcat
 atgccccag aaaacctcca gcagctgggtg ctttcagcag agagggaggc tgcacagggc
 ttcttgacac tcatgtctga ggggaagctg caggggaagc tgcagggtgag cactgagaaa
 15 ggggagcaag ggcacctgga gcctagtgtt cagagggctt gctttagtgg gaggaggaac
 tccagagagg aaatggcagg gatactgagc atctccagag gcagaatcca ttctgtgccc
 cctacaggta ccaccatccg aggagcaggc cctgggtcgc ctgacagccc tgctgtccca
 gcggtaccca cgcctcacct cccagctcct cattgacctg tcaccactca tccctttctt
 ggctgtctct gacctgatgc gcttcccacc atccctgtta gccaacgaca gtgtgtaagg
 20 ttcttgcaact actcctcctg ctctgtcac ggtcaggcca accgcatcca cctggagcag
 ccccttccgg agctcctctc tgtttttttc tttcatgcca gataggcaat gtgccaacat
 cgtagcaagg tttgagagag gcacatctca cgcctgagtg tgaaaaccca atcattatgc
 taatgaacta caaaaggatc agagagctcc tctctattaa aaccagggag aggatgggag
 tgggtggctca tgctgtaat cccagcacgt tgggagcccg aggcagggtg atcactaggt
 25 ccgcctagtg agttcgagac cagcctggcc aatatggtga aacccgtct ctattaaact
 acaaaaaatta gccaggcatg gtggtgggag cttgtagtcc cagttactct ggaggctgag
 gcaggaggat agcttgaacc tgggaggcag aggttgagc gaaccaagat cgtgccactg
 cactccagcc tgggtgacag agcgagactc cgtcttaaaa aaaacaaaaa aaaaaacaaa
 aaaaaaaaaac agggagagtc tccttcctat ctagacagca gggctacaga gggtcagagg
 30 aaaacagttt ggaggaagac aaagggttaa gacctatgac tcctcgagc ctggctgcca
 tccgggatta cagcccagga atgaggcctg aacagaagga ggctctggca aagcgactgc
 tggccccctga actgtttggg gaagtgcctg cctggcccca ggagctgctg tgggcagtgc
 tgccccctgt ccccccacct cctctggaga acttttttga gctcagccct caccaggtat
 gagaatcatc ttctttactt gactggccca tcttctgcta gtggggacaa agagtcaatg
 35 gcatgtctct cagtggcccc tccttgcaag aaccctatag tgaccccagt gcgagctaac
 ctcccccatc tcagatccag gccctggagg atagctggcc agcagcagg ctggggccag
 ggcatgccc ccatgtgctg cgcagcctgg taaaccagag tgtccaggat ggtgaggagc
 aggtacgcag gtgagttgtt gtgggatcag taaccaaggc aagagtggaa gaggtagaga
 gaggaaggca cagctgtcac gctgggtcgg tgttctagga agaaagggc aagagagtag
 40 gcagtggcct caggcagcat agagttccag gagagaggtc tatagatggg gccctgtgt
 agtgggtgtag tgtcagagtg cccagtgtat gtaccatac catctgctgc caggcctgcc
 ttagtgctag tcttggggac cacacaaagg tcagcttcat gccctcctca ggcttgggccc
 cctgcgctgt ttctgagcc ctgaggagct gcagagccta gtgccccctga gtgatccaac
 ggggccagta gaacgggggc tgctggaatg tgcagccaat gggaccctca gcccagaagg

acgggtgagc ccctcagcac aagcctacaa gacttttaggc ttcccctggg tctgtgtgga
 tggctttccc attgtgtcaa cttgagcaca gtggtgccag ccccatccc acttttgcaa
 cctccattcc ttactccatg gccattctta cctgttacca cctcttcctg gcccttctct
 atctgggtctg tagcacccca aacataccct ttgccatttt gaacctaatc tactccagtc
 5 caatccctag ttccaaaccc tagcccaggc cctgggaaat tcagatgtgg gattagagag
 gaagttcaag gttcatctgt cttttctctc cagtccataa ccttcttttg ttacagggtg
 catatgaact tctgggtgtg ttgcgctcat ctggaggagc ggtgctgagc ccccgaggc
 tgcgggtctg ggccccctctc ttctctcagc tgggcctccg cttccttcag gagctgtcag
 agccccagct tagagccatg cttcctgtcc tgcagggaac tagtgttaca cctgctcagg
 10 tttgcctgtc tcaactccctg gcatgtaccc tccatccccg cttgagcccc agtcaagaga
 atcccattca gggataaaag cagccccctc tttccctggg tgaacagtag aggtaaactc
 tgtctgcagg aggacgcctt cattcccttt cctcagatca agaagggacc tgagtcactg
 aggatgggta ctagggatgg ttaagaggca gcgggaagtt ttggagggtt tgccttagga
 acccacttag gacctggctg ctgggtcctg agagctgttg ttttcggtcc catcccaaca
 15 caggctgtcc tgctgcttg acggctcctt cctaggcacg atgtgagtag cagcaacttc
 tcagcctccc gccagaggtc tctatcctct tttaacctgg ctccctgcac tgccccctct
 ctctctccgc tccccctata cttactgcct tgctgcattg tgattgttgt cttccccaac
 acccttccct tcttcttcag gcctcttgct tctcttgctc tttagctatc cctggaggaa
 ctctgctcct tgcaccttct gctaccaggc ctcagccccc agacactcca ggccatccct
 20 aggcgagtc tggtcggggc ttgttcctgc ctggcccctg aactgtcacg cctctcagcc
 tgccagaccg cagcactgct gcagaccttt cgggtatgag agtggcaagg aggatgagat
 aatcagggat accggtctt tctggttggg aggaaggcat cttccctgag gccagggaag
 gcctttcata cctccccact tacacacaca cacacacaca cacacacaca cacacacaca
 accaattctc atgcagggtta aagatggtgt taaaaatatg ggtacaacag gtgctggtcc
 25 agctgtgtgt atccctggtc aggtaagtgt gagatctccc aactgagctc ctctcccat
 tctggggcag tttcatatgg ctggtgctac ctcccacact accctgcagt ggccctgaga
 gttctgggta gctctgtgcc cattagcagc cctccccagt gccagatgca ggacagcatg
 atccactcac attgtcctag actaatgtca aagctggaag ggccctgagaa atcttccagg
 ccacccaccc tgctttcaga tgaagagacc aaggctggga gaagctaagg gactttgttt
 30 gcctgggtgcc taactagcag caacacttga ccacagcagc ctgcagtgtg aggctcttag
 gcgtttattg ctacagtggc aaatgccatt ccacttctgt cctagctttg gtccctttcc
 acccccatgg ttccctttct ctgagtgtca agtacagact ctctcaccta tcaactacact
 gctataccca tcaccgccag cagcctattc ccaccacctg gccagactgc ctgcttcccc
 tgctcccat aaagctgcta caactggatt ccttggtctt tctggcaaata cgaagacgt
 35 actgggagct gccctggtct gagcagcagg taattctccc cacttaattt cagaacttcc
 tccctcaatg tagtctacct tctttacctc tcccttagcc ctatttggtc agcttatccc
 tactatcctt tatttgattg tttgagatac agtctcactc tgttgcccag gctgcagtgc
 agtggcatga tcagagtctg ctgtaacctc aaactcctga gctcaggcaa tctttctgcc
 tcagcctcct gaatagctag gacgacagg gtgtaccacc atgcctgggt aatttttaaa
 40 tttttttttt gttttttgag atgaagtctt gctctgtcac ccaggcttga gtacagtggc
 acaagcttgg ctactgcaa cctctgtctc ccgggttcaa gcgattctcc tgcctcagcc
 tcccagtag ctgggactac aggcaactccc cacaatgcct ggctaatttt tttttgttt
 tagtagagac agggtttcac catattggcc aggctggtct cgaactgctg accttgatg
 ctgcctgcct ctgcctctca aagtgtctggg attacaggtg tgagccacca tgcccgcca

atttttaaat tttttgtaga gacagacaat acaaaaatgt ggacactatg tggagacact
 atgttgagggt actatgctgt ccagattgggt cttgaaactcc tggcctcaag caatcctcct
 gccttggcct cccaaagtgc tgggattaca gacctgagcc actgcaccca gccccctagt
 atctcttata atgtgacttg cttttctttt tctttctcct tcccttttct ttcatttctt
 5 tctcactctc gagagaagag tgggcatctg ggagagtggg aggctggtgg gtccacaga
 gtgaggaggc aggactgggt ccaaggcagt cctgcctctc cactctaggg ggtatccttg
 gacagtgtct cttctgggaa ggggctcgtc tttctttctc ttgtaggcac agtttctctg
 gaagaagatg caagtacca ccaaccttac cctcaggaat ctgcagtgag taacttgtgt
 tgagcagtgc gctgaattcg accaacattt ttttgagtgc ttactatgtg ccaggcacca
 10 tgtgatatgg aatgggggat atagggatga atgatgcata gtccctgcct cgtggacgtt
 ctctagcac ctccctttgc cctcctttcc ttccacagtg ccatgcctat cctgactaga
 gccaaaggac tcagaaaacc tggattcagg ttccagtcct gtcacctact tgtcctcttg
 ggcaagtcat ttaacgtccc tgtgtcagtt ttcccttctt taaatgagaa ttacaatggc
 accagcctca taggtagtta ctgtgaagat taaatgaggt aggtcatgta agatatttaa
 15 cacagtgttt ggtccattgt aaagtcccag tagtcatttg ctactgttag tttacttcag
 gatgacttca gaggcactgg ccaagcaaga ataaatagga ataagaagg atcactttac
 ttacaccac attagaagaa caatgggctt cagaatcttt tttttttttt ttttttcgag
 acagtcttgc tctgttggcc aggctggagt gcagtggcgc gatttcggct cactgcaacc
 tctgcctccc aggttcaagc gattctcctg tctcagcctc tggagttagct gggattacag
 20 gaatgtgcca ccatacccag ctaatttttg tatttttagt agagatgggg tttcaccatt
 ttggccaggc tgggtctcaa ctctgacct caggatgatcc accgcctca gcctccaag
 ggcttcagaa tctaagacat ggctctagtt tcagtttacc acatttctag cagaatgatg
 ttgggaatgt cacctgactt ccataaatcc ttattttctc ctctgataaa cagcagtgat
 gttatgggga gctgatgaga tatctatgta aaaacatttc tcaaaccata aattacgggtg
 25 gatgaacatc tgtacttgtg ttgagagtac tgatatcaag gagcaaacag gctgttgtat
 gtgttgaatg agcctctccc cactcacaca cccacagggc tctgggcacc ctggcaggag
 gcatgtcctg tgagtttctg cagcagatca actccatggg agacttcctt gaagtgggtg
 acatgatcta tcagctgccc actagagttc gagggagcct ggtgagaggg ggtgcctgga
 ctttagtggt agcagggagg ctgggacctt aggtatagaa cccagctcct atgttctgct
 30 ctggcctcac actgcttccc tacagagggc ctgtatctgg gcagagctac agcggaggat
 ggcaatgcca gaaccagaat ggacaactgt agggccagaa ctgaacgggc tggatagcaa
 gctactcctg gacttaccgt aagtactgca gctagagata ttggccccctc agaaagctca
 atctgggggtg aagatctgcc cttagggaat gccctggagg aggtagtttt tctgtctggt
 agttccctga cataatttat agcccaaagc agaggatttt attcaaagtt gctctatgta
 35 ttgactgggt cccagaatat gctccagcac agggcagctg aggggtggca cactgtattg
 aagcctgcca agtaatctta caataaccta gtccacatta attgagattg agacagagca
 tctgaagtga gggaggcaat gctccaaatc tgccccagag gattgtagtt tgctcagggc
 actgtgttct tagtgcattc agaggagtag atcgagagaa aaatatatga aaaatgtgat
 aaataccttc aaatacctga ggggctatca agtagaaatt agattgtcat atttatgagt
 40 ggccccattg ggcaagacta agagtagtta acggagatca gattttttaca tagtataaga
 aaaactaagg tagtgagttc ctggctcctg gagctgttct agcctaagcc agatggcccc
 atggcaggaa tgtttagtag cacgttcata tacagggtgt gggaagaaaa ggctatagga
 acccaaggct cctccctacc catggagaaa tttattagta tgttactcat atgctgcttt
 tctcatttta cccctaccac caccctgttg ccatccgcac tgtaagtcag gataggaaaa

tgctggtggt acagtcttcc tggggaatat ggagctgaag tggagtaaaa gcagttgact
 tcatctctac ttttttcttt tttttctttt tttttttttt tgagacagag ttttgctgtg
 tcaccaaggc tggagtgcag tgacgtgatc tcggctcact gcaacctcca tcttcagggt
 tcaagcaatt ctctgcctc agtctcccga gtagctggga ctgtaggtgt gcaccaccat
 5 gccaggctaa tttttgtatt gttagtaggg acgagctttc accatgttgg ccaggctggt
 cttgaactcc tggcttcaag tgatctgccc acctcggctt cccaacattc ttatatTTTT
 ataggccttt ccacagattt cagctcttgt atgacttagc ccagttccag aactggtaat
 cctaggtagg gtacagggtta tcacctctga tttcgggtaa aagggattta tttatttatt
 tgttttattta tttatatTTT tgagacagag tctcgtctctg tcaccaggc tggagtgc aa
 10 tgggtgccatc tcggctcact gcaacctctc cctctgggggt tcaagcaatt ctctgcctc
 agcctgctga gtagctggga ttacaggcgc gtgccaccac acccggctaa tttttgcatt
 tttagtagag acgggggttt accatgttgc tcagggtggt ctggaatttc tgacctgtg
 atctgcctgc ctggcctcc caaagtgtctg ggattacagg catgagccac tgcgtccggc
 ctgtttttac ttttttttaa tgccattcag atctgtttta atatgtgggt tctgtgagat
 15 aatttagaat cccaagggtta cagatgagggt gaaagatcct agaccatgca tcaaaaaact
 tgagtttctc atttgtgaaa gaaggataag agaaacacct attttgtctg ggtgcagtgg
 ctcatgccta taatcccagc atttggggag gccaagggtg gtggatcacg gaggtcagggt
 gttcaagacc agactggcca acatggcaaa acaccatctc tactaaaaat acaaaagtta
 gctgggogtg gtggcacgtg cgtgtaattc cagctattcg ggaggctgag gcacgagaat
 20 tgcttgaacc tgggaggtgc ggggtgcagt gaactgagat cgcagcacca ctgtgctcca
 gcctgagtga tggagtgagg ccaggctctg ttgtaggatc aaatgagata acacctgaaa
 gaactttgta aattgtatag cacgtacaaa caagaaggga cctcttcaca agcagaggaa
 ggggtggtcct gtggaaaaaa acgggaattg ggagtgagag acctcaacat ttgatctctg
 tgaacctcag ttttttaatc tataaaatgg ggaaatgtta atggtactta atatttgag
 25 cttttgagtc cattagatca ggtaggattg ttcgttattt ttttttttta ggaagactag
 aaatatgttg ctcccttttt ctccccact caagcttgat ggtgggaatt ggccctggag
 ctgtttacta tcagttcctg tccagcttca ctaaaatttg tctgggggtca catcttagct
 gcggactgtg gggttttgtg gtcccttctc gacttggccc agctccacct gaatcctgtt
 gttgtcaaat tgctgtaata ggatccagtt gatggacaga ctatccaatg aatccattat
 30 gttggtggtg gagctggtgc aaagagctcc agagcagctg ctggcactga cccccctcca
 ccaggcagcc ctggcagaga gggcactaca aaacctggta agagtccacc ctaccagact
 cagatttgct gccctgggca attcttgctc ctcagacaat gctctctgac tgtcccccaa
 ccctctactt cttgctttct tgctgcaaaa cagattcctg tctacaaggc ctggccctgt
 ttttgctct gggttctgtt ccttgataat atgcttcacg ttacttgtcc atacctcttg
 35 gagtccgaga aatctcttg agtccacctc tcagtctttc tgctgtctcc tatctgggct
 cattgcttaa ggaagtgaac aaaggtagtg agcatcatag ggtgctgagc tgggagcagg
 agggagggaa ggttagggg cttggtgtct tgatcaagggt gtctggtatt ctgagtcaga
 agtgcattgt ccaagttctg atgctcttct ccaggctcca aaggagactc cagtctcagg
 ggaagtgtct gagaccttag gccctttggt tggattcctg gggacagaga gcacacgaca
 40 gatcccccta cagatcctgc tgtcccatct cagtcagctg caaggcttct gcctaggaga
 gacatttgcc acagagctgg gatggctgct attgcaggag tctgttcttg ggtatggacc
 ttcgagaact tcagattcta actcattcta taccaggtcc ctccagccacc atcatcagtg
 gcagcctgtt ccatattctt aaggtoctt ggagccctgt gtccgaaatc ctagcatgtc
 ctcttttccc cttccttttc ctcacagttc cctcagctcc ccagcccccg attttcttcc

tgtccccagg aaaccagagt tgtggagcca ggatgaagta gagcaagctg gacgcctagt
 attcactctg tctactgagg caatttcctt gatccccagg gtgagatgaa ggaagaaggg
 aagggagtaa atgcatagag gggactggtg agctgggttat ggggacccgt ggccaaagag
 ggcaaaggat atgaagccta gatctggggg gagactgcaa aacagagaca ggactttgga
 5 cttagagcta tagcagcagg tcctgatctg tccagatctc cccactctcc ttctaccttc
 tcatgcagga ggccttggtt ccagagaccc tggagcggct tctagaaaag cagcagagct
 gggagcagag cagagttgga cagctgtgta gggagccaca gcttgctgcc aagaaagcag
 ccctggtagc aggggtggtg cgaccagctg ctgaggatct tccaggtgaa actacccaaa
 tacttatatg tccagcagga tgtacagga gtatcaaacg gtctgggttc tacatgtgct
 10 cttccctggg actgggtttt ctaatttata aagcaaagag tttagaggga tgatcttcaa
 gcctcttgta gttctagaat tctgtagttc tgggagtttg taaactatta agttttcttt
 tagcccagaa cttccatttt cctgctctct cgtgtctgct ctagactcag ctctagctcg
 gctaagtgtg gagctctctg ctggggagat ccctagaagc tttgaaggag acattgtgag
 gctggagaac tgggttcaaa ttcagtgcta ccattaaatc tctgaataac atcctcagtc
 15 ttccatctat aaaagtcttg gcatctccaa tcacttcttg ttctattatc tcctaagccc
 tatacatatt actctgtaat actcctttga tccctatttc tcacagtgtc ctatcctcca
 aaggttgga gactcactct atctacagat atctctctgg gcatatttta ttactgcgct
 gacctcctgg ccctgccttc ccccttcaga acctgtgcca aattgtgcag atgtacgagg
 gacattccca gcagcctggt ctgcaaccca gattgcagag atggagctct cagactttga
 20 ggactgcctg acattatttg caggagaccc aggacttggg cctgaggaac tgcgggcagc
 catgggcaaa gcaaaacagg ttagggatgg agagccaact ggggttgcc atgaggaagc
 tatttggtg tgatgtagga cacaaagaga atggagagtt ggatgagagg tgggggaagc
 aagagataga agagttagaa gatttggttc acaagtagga ggtgaaggga gataaatatt
 gaggaagag agctagtata atgaatagag ggacgaaagc agtggttacc aaattttaat
 25 gcatatcacg atcatcaagg gaacagattt ttttctttat ttttttttct ttcttaaaaa
 aataatggca tgcttcggct ggggtgcagcg gctcacgcct ataatctcag aactttggga
 ggccaaggcg ggcagatcac gaggtcagga gatcaagacc atcctgtcta acacggcgaa
 acacggtctc tactaaaaat acaaaaaagt tagccgggca tgggtgtgca cacttgttgt
 ccagctact tgggaggctg aggcaggaga atggcgtgaa cctgggaggg ggagcttgca
 30 gtgagccgaa gtcaagccaa tgcactccat cctgggtgac agagcaagac tccatctcaa
 aaaaaaaaaa aaaaaaaaaa ggcattgcttc atgaatttgc gtgttatcct tgcacaggcg
 ccatgcaaat ctctgtatca ttccaatttt ttggggtagt tgctgctgaa ctgagcatgg
 gaacagtgcc agtgccagat taccatgctt cactgactta ataaaaacct ttggggaggc
 tgggcgcagt gactcatgcc tgtaatcaca gcaactttggg aggcggaggc aggtggattg
 35 cttgagccca ggagttagag accagactgg gcaacatggt gaaaccctgt ctctactaaa
 aatagaaaaa acattagctg ggtgtggcgg cacatgcctg taatcccagc tactcaggag
 gctggggtag gagaatccca tgagtgcagg aggtggaggg tgcaatgtgc caagatcgca
 ccactgccct ccagcctggg tgtcagagca agaccctgtc tcataaatta aaaaataagc
 ctctggggga aagagtctag acatctgcat ctccctttttt tttttttttt tttttttttg
 40 agacagagtc tcaactctgtc acccagcatc caggctggag tgcagtgggtg tgatcttggc
 tcaactgtaac ctctacatcc tgggttcaaa cgatcctcct gcctcagcct ctcaagtagc
 tgggactaca ggtgcaccac acctggctaa tttttgtatc tttggtagag atggggtttc
 actatgttgc ccaggatggt ctogaacttc tgggctcaag caatcctccc acctcagcct
 cccaaagtgc tgggattaca gctgttagcc actgtgctgg gccctaggca tctgttttaa

taagcgtctc tgtgtctgat gcacataaaa gtgtggaact catggactag agttagtttg
 ctcttctttt ccactgattg taatgtcttt caaaacacct tagaggaact gtaaggcaac
 ggtctcattt tatagtggag gaaactaaag aaaaggcaaa tgatttacct agagttatac
 agctaagggc agaggcaaga cttaaaaccc agcagtatga ctcccaatcc actgcttttc
 5 cactcacatt gttcctgtct ttctcctagt tgtgggggtcc cccccgggga tttcgtcctg
 agcagatcct gcagcttggt aggctcttaa taggtctagg agatcgggaa ctacaggagc
 tgatcctagt ggactgggga gtgctgagca ccctggggca gatagatggc tggagcacca
 ctcaggtaac acttttcttc ctccctacgg cttcccaaac acccatccca cagaccagc
 cctatagatc atctaaagcc caaggaattt ttttctgtg accctacctg gtcttctttt
 10 ctatcttttg ttgatacccc atactagtga ccttcaggac tctgatttat tcaactctgag
 gccctggaca cataatactg tctcctacct cttttcctgg aggcttcctc tttttctttc
 cttttctttt ctgagtcctc agccttcccc atgactcctt aggtcttaat agtaacagaa
 tataaccag taacacctat cacttccctg tccattaatt ctccataact ttctccttc
 cctctttctc ccacccccca cccagctcc gcattgtggt ctccagtttc ctacggcaga
 15 gtggctggca tgtgagccac ctggacttcg ttcactctgac agcgtgggt tatactctct
 gtggactgag gccagaggag ctccagcaca tcagcagttg ggagttcagg tcatttctga
 aggggctgag ggtgggtggt ctgaggtaaa ggtggactta ctggggaaag aaggatcatg
 aaggctctgt cccatggagg aagggaactc atttgaagcc atctcttcct ttgtctcatg
 accacagccc ctttactga agccgaattc ttcttccttc cttcctactg ttctacagcc
 20 aagcagctct cttcctcgcc accctgcac tccagtgtc tgaggaacaa ctggagggtc
 tggccacact acttgtactg cctggtgggt ttggcccaat cagtaactgg gggcctgaga
 tcttactga aattggcacc atagcagggt gggagctggg ccaactgtgg tgcaagttgg
 tttggtttct ataccatggg tggactggat ggaagactgc cctgcaattc ttaaggtggg
 ggccctgagg tgtttaaata aggggctaga gacatattgg ggaaggtcta tgatagggca
 25 ctttgggagt agttagagaa ggtctatagg tttgaagaga gggaaggtca gtctaagaca
 atgtttggat gccacttgct tcaacagctg ggatcccaga cctggctctt tcagcactgc
 tgcggggaca gatccagggc gttactcctc ttgccatttc tgtcatccct cctcctaaat
 ttgctgtaag tattaatgga ctggggtgac cacaggagag ccagggccca atggggacta
 catgcatgca ctgattccta cccctgccct cagggtggtgt ttagtcccat ccaactatct
 30 agtctcacca gtgtcaggc tgtggctgtc actcctgagc aaatggcctt tctgagtcct
 gagcagcgac gagcagttgc atgggcccac catgagggaa aggagagccc agaacagcaa
 ggtgagttcc cagctgcaca gcttgatcct ccatctcctg acccagaatc aaaccctaa
 tttggtgctg tctggctctt agagtgcacc caggagatc cctggagtga aggagtctac
 aggcagagcg ctaatttcca agtatcaatg ctctggaga gctgagttgt gatattactc
 35 ccattccctg tctattatag gtcgaagtac agcctggggc ctccaggact ggtcacgacc
 ttctgggtcc ctggtattga ctatcagctt ccttggccac ctgctatgag cctgtctcta
 cagtagaagg agattgtggg gagagaaatc ttaagtcata atgaataaag tgcaaacaga
 agtgcacct gattattttc agaagctgat gaggaata

40 The nucleotide sequence of the human STRC pseudogene is shown in SEQ ID NO: 5.

Human Stereocilin Pseudogene 1 (STRCP1) (SEQ ID NO: 5)

gccctgccct cacctggcta tcccacacag gtgagaataa ccagaactca cctccgttac
 cagtgttcac ttggaaacat ggctctcagc ctctggcccc tgctgctgct gctgctgctg

ctgctgctgc tgtcctttgc aggtaagaag aacagtgagc agaactgggg atgaggagga
 ggggtggctgg aaaaagactt taagaatatg gaggtgaacc tgtagatag aaggacaaag
 gagagaggca gagacttggt caaaaggga aaatgagggt taagaaaagc aggccaaagc
 ttactgtagg ccagtgaag gggttcagct caccatcccc tcacctcatc ttagatcca
 5 ggtaggaac tgtgctcagg ggcagggttg agtttgggct ctgtgttcct ctcttcagt
 gacctctggt ttctctcctt acagtgaactc tggcccctac tgggcctcat tccctggacc
 ctggtctctc ctctctgaag tcattgctct ccactctgga ccaggctccc cagggtctcc
 tgagccgctc acggttcttt acattcctgg ccaacatttc ttcttccttt gagcctggga
 gaatggggga aggaccagta ggagagcccc cactctcca gccgcctgct ctgcggtcc
 10 atgattttct agtgacactg agaggtagcc ccgactggga gccaatgcta gggctgctag
 gggatatgct ggcactgctg ggacaggagc agactccccg agatttcctg gtgcaccagg
 caggggtgct ggggtggactt gtggagggtg tgcctgggagc cttagttcct gggggcccc
 ctaccccaac tcagcccca tgcaccgctg atgggcccgtc tgactgtgtc ctggctgctg
 actggttgcc ttctctgctg ctgttggttag agggcacacg ctggcaagct ctggtgcagg
 15 tgcagcccag tgtggacccc accaatgcca caggcctcga tgggaggag gcagctctc
 actttttgca gggctctgttg ggtttgctta cccaacagg ggagctaggc tccaaggagg
 ctctttgggg cggctctgcta cgcacagtgg gggccccct ctatgctgcc tttcaggagg
 ggctgctccg tgtcactcac tccctgcagg atgaggctct ctccattttg gggcagccag
 agcctgatac caatgggcag tgccaggag gtgagtgtgg ccagggtggt gactgggatg
 20 tggcagggca aggaaagtga aattggggta gttttcttcc ttactctttc ctcctaggt
 aaccttcaac agctgctctt atggtaagta acaggagacc agttctgagg gattgggcct
 ggaaaatctg gaggtgaaga gctgaagacc tcagcctcta gagaggaaaa ctgatgggag
 gagtgtagtt tagtggtttt ggggtgtgac tgtctgggtt ggtgtcccag ctccacctct
 tcttagccat atgacctga gcagggtaca tagtctttct atacctcagt tccccattt
 25 ataaaatgag aatgataata ttagttacca cagagttggt gcacccggtt aaatgagttg
 atactgtgta tgcaaacgac ttaaaaccgt gctggcacat agcgcttaat aatgtagct
 agtaaagatg ggatttgga aataaggaca cagctggatt cctctacccc ctactactt
 cagtacaaca atgccagaca gtagttagac atattgagtt gctgagcaga tttcctaaca
 tgaggcccgct tgagggttgt gtttaagcta tctaaaagca tatgaagaaa ggagacagaa
 30 gggggccagg tggacagaaa gaattccaac tggggcttct cctaggtgat tttggacctt
 ggcagggcag ctttctcttt tttgccccgt tgcagcattt caaccagtaa cgcctaaact
 ctcagggacc tcgcttgtag aaaagcctat gcttgccatg ccccttgagg gctctgagtc
 agggctcagaa tcttcagctg gaggaatgt gaactgacca gatcctgcct gctcctccct
 ctgcacccag gggcgctccg cacaaccttt cctgggatgt ccaggcgctg ggctttctgt
 35 ctggatcacc acccccaccc cctgccctcc ttactgcct gagcacgggc gtgcctctgc
 ccagagcttc tcagccgtca gccacatca gccacgcca acggcgagcc atcactgtgg
 aggcctctg tgagaaccac ttaggccag caccacccta cagcatttcc aacttctcca
 tccacttgct ctgccagcac accaagcctg cactccaca gcccacccc agcaccactg
 ccatctgcca gacagctgtg tggtagcag tgcctgggc accagggtgc caaggctggc
 40 tacaggcctg ccacgaccag tttcctgatg agtttttga tgcatctgc agtaacctct
 ctttttcagc cctgtctggc tccaaccgcc gcctggtgaa gcggctctgt gctggcctgc
 tcccccccc taccagctgc cctgaaggcc tgccccctgt tcccctcacc ccagacatct
 tttggggctg cttcttgag aatgagactc tgtgggctga gcgactgtgt ggggaggcaa
 gtctacaggc tgtgcccccc agcaaccagg cttgggtcca gcatgtgtgc cagggcccca

cccagatgt cactgcctcc ccaccatgcc acattggacc ctgtggggaa cgctgcccgg
 atgggggag cttcctggtg atgggtctgtg ccaatgacac catgtatgag gtctggtgac
 ccttctggcc ttggctagca ggccaatgca ggataagtcg tgggggcaat gacacttgct
 tccagaagg gctgctgggc ccccttctgc cctctctgcc accactggga ccatccccac
 5 tctgtctgac ccttggtccc ttctctcttg gcatgctatc ccagttgcca cgctgtcagt
 cctctgtccc agctcttgct caccacacac gcctacacta tctcctccgc ctgctgacct
 tctcttggg tccaggggct gggggcgctg agggccagg gatgctgggt cgggccctac
 tgctctccag tctcccagac aactgctcct tctgggatgc ctttcgcca gagggccggc
 gcagtgtgct acggacgatt ggggaatacc tggaacaaga tgaggagcag ccaaccccat
 10 caggctttga accactgtc aacccagct ctggtataag caagatggag ctgctggcct
 gctttagtgt gagtgtctctg ccagagggaa agctcctaga acagtgagaa ggccctccag
 gggaattcct cgaatactca gaggcagtag tgtgggtag tagttgaagc acacagctct
 agagtccagc aggttggtat tcatactctg gttctgtgac cagccttgaa tgagttatct
 aacttctctg agcaatatct ttctcgtctc atttataaac tagggatgat aatggtatat
 15 gagataatac atgctgtggg cttagcacag tgcatgatac acaaacatgc aataaatatt
 accttggtat tcttttgggc tctttgactc tctcactttc tgcaccagaa agaaaaagga
 tcaagttaga ggactctaaa tttttccct agagagttag aattggaggc tggcagaata
 caggaagata aggtaggaat gagaaagatt caggacact accaatcaga agactttggt
 tctaggttca actgtgccac aaattagtgt gatcttaggc aagcaatttc atttagtttt
 20 tctgggcttc agtttttagt ctgtagaatg gaggggtgag aatatgttaa acaccataat
 taattcactg agtgcctatt atatgcaagg cactttgcta ggttctgtag gatataaaa
 gatttcttac tccatgttgg ggccaccttt ttcaaaccct gggcccagta aaatggaatt
 agatagtctc atagtatttg gttcaggtct acaagtatta attgagccaa ctatggacct
 ggcatgggag aggttacaag agaaattaga gatatgatcc cggacctaaa agagcttaat
 25 atctgaagaa tcacacttga gatgatggac aagcatcca gcaagtggag ttggaatgcc
 tgggggagct gcaggagaga cagagaagac agctctgttg gcatattgtc tttcttccca
 ccagcctgtg ctgtgggatc tgctccagag ggaaaagagt gtttggggcc tgcagattct
 agtgcaggta acaggtggag ggcacatggg tgggctgggt gacagccatg gctggaggtc
 cctgccccgt gaggtgaggc catacccacc atgacctcct attcgcaggc gtacctgcat
 30 atgccccag aaaacctcca gcagctggtg ctttcagcag agagggaggc tgcacagggc
 ttcttgacac tcatgtgca ggggaagctg caggggaagc tgcaggtgag cactgagaaa
 ggggagcaag ggcacctgga gcctagtgtt cagagggctt gctttagtgg gaggaggaac
 tccagagagg aaatggcagg gatactgagc atctccagag gcagaatcca ttctgtgccc
 cctacaggta ccaccatccg aggagcaggc cctgggtcgc ctgacagccc tgctgtccca
 35 gcggtaccca cgcctcacct cccagctctt cattgacctg tcaccactca tccctttctt
 ggctgtctct gacctgatgc gcttcccacc atccctgtta gccaacgaca gtgtgtaagg
 ttcttgcaact actcctcctg ctctgtcac ggtcaggcca accgcatcca cctggagcag
 ccccttccgg agctcctctc tgtttttttc tttcatgcca gataggcaat gtgccaacat
 cgtagcaagg tttgagagag gcacatctca cgcctgagtg tgaaaacca atcattatgc
 40 taatgaacta caaaaggatc agagagctcc tctctattaa aaccagggag aggatgggag
 tgggtggctca tgctgtaat ccagcacgt tgggagccc aggcagggtg atcactaggt
 ccgctagtgt agttcgagac cagcctggcc aatatggtga aaccctgtct ctattaaact
 acaaaaatta gccaggcatg gtgggtggcg cttgtagtcc cagttactct ggaggctgag
 gcaggaggat agcttgaacc tgggaggcag aggttgagc gagccaagat cgtgccactg

cactccagcc tgggtgacag agcgagactc cgtcttaaaa aaaacaaaaa acaaaacaaa
 acaaaaaaac agggagagtc tccttcctat ctagacagca gggctacaga gggtcagagg
 aaaacagttt ggaggaagac aaagggttaa gaccatgac tcctcgagc ctggctgcc
 tccgggatta cagcccagga atgaggcctg aacagaagga ggctctggca aagcgactgc
 5 tggcccctga actgtttggg gaagtgcctg cctggcccca ggagctgctg tgggcagtgc
 tgcccctgct ccccccactc cctctggaga actttttgca gctcagccct caccaggtat
 gagaatcatc ttctttactt gactggccca tcttctgcta gtggggacaa agagtcaatg
 gcatgtctct cagtggcccc tccttgcaag aaccctatag tgaccccagt gcgagctaac
 cttccccatc tcagatccag gccctggagg atagctggcc agcagcaggt ctggggccag
 10 ggcagtcctg ccatgtgctg cgcagcctgg taaaccagag tgtccaggat ggtgaggagc
 aggtacgcag gtgagttgtt gtgggatcag taaccaaggc aagagtggaa gaggtagaga
 gaggaaggca cagctgtcac gctgggtcgg tgttctagga agaaaggggc aagagagtag
 gcagtggcct caggcagcat agagtccag gagagaggtc tatagatggt gccctgtgt
 agtgggtgtag tgtcagagtg cccagtgtat gtaccatac catctgctgc caggcctgcc
 15 ttagtgctag tcttggggac cacacaaagg tcagcttcat gccctcctca ggcttggggc
 cctgcgctgt ttctgagcc ctgaggagct gcagagccta gtgcccctga gtgatccaac
 gggggcagta gaacgggggc tgctggaatg tgcagccaat gggaccctca gcccagaagg
 acgggtgagc ccctcagcac aagcctacaa gactttaggc ttcccctggg tctgtgtgga
 tggctttccc attgtgtcaa cttgagcaca gtggtgccag ccccatccc acttttgcaa
 20 cctccattcc ttactccatg gccattctta cctgttacca cctcttcctg gcccttctct
 atctgggtctg tagcacccca aacataccct ttgccatttt gaacctaatc tactccagtc
 caatccctag ttccaaaccc tagcccaggc cctgggaaat tcagatgtgg gattagagag
 gaagttcaag gttcatctgt cttttctctc cagtccataa ccttctttgg ttacagggtg
 catatgaact tctgggtgtg ttgcgctcat ctggaggagc ggtgctgagc cccggggagc
 25 tgccgggtctg gggccctctc ttctctcagc tgggcctccg cttccttcag gagctgtcag
 agccccagct tagagccatg cttcctgtcc tgcagggaac tagtgttaca cctgctcagg
 tttgcctgtc tcaactccctg gcatgtaccc tccatccccg cttgagcccc agtcaagaga
 atcccattca gggataaaag cagcccctcc tttccctggg tgaacagtag aggtaaactc
 tgtctgcagg aggacgcctt cattcccttt cctcagatca agaagggacc tgagtcactg
 30 aggatgggta ctggggatgg ttaagaggca gcgggaagtt ttggagggtt tgccttagga
 acccacttag gacctggctg ctgggtcctg agagctgttg ttttcggtcc catccaaca
 caggctgtcc tgctgcttg acggctcctt cctaggcacg atgtgagtag cagcaacttc
 tcagcctccc gccagaggtc tctatcctct tttaacctgg ctccctgcac tgcccctcct
 ctctctccgc tcccctcata cttactgcct tgctgcattg tgattgttgt cttccccaac
 35 acccttccct tcttcttcag gcctcttgct tctcttgctc tttagctatc cctggaggaa
 ctctgtctct tgcaccttct gctaccaggc ctcagcccc agacactcca ggccatccct
 agggaggtcc tggtcggggc ttgttcctgc ctggcccctg aactgtcacg cctctcagcc
 tgccagaccg cagcactgct gcagaccttt cgggtatgag agtggcaagg aggatgagat
 aatcagggat accggctctt tctggttggg aggaaggcat cttccctgag gccagggaag
 40 gcctttcata cctccccact tacacacaca cacacacaca cacacacaca accaattctc
 atgcagggta aagatgggtg taaaaatatg ggtacaacag gtgctggtcc agctgtgtgc
 atccctgggc aggtaaagtg gagatctccc aactgagctc ctctcccat tctggggcag
 tttcatatgg ctggtgctac ctcccacact accctgcagt ggccctgaga gttctgggta
 gctctgtgcc cattagcagc cctccccagt gccagatgca ggacagcatg atccactcac

attgtcctag actaatgtca aagctggaag ggcttgagaa atcttccagg ccaccacccc
 tgctttcaga tgaaaagacc aaggctggga gaagctaagg gactttgttt gcctgggtgc
 taactagcag caacacttga ccacagcagc ctgcagtgtg aggctcttag gcgtttattg
 ctacagtggc aaatgccatt ccacttctgt cctagctttg gtccctttcc acccccatgg
 5 ttccttttct ctgagtgtca agtacagact ctctcaccta tctactacact gctataccca
 tcaccgccag cagcctattc ccaccacctg gccagactgc ctgcttcccc tgctcccatt
 aaagctgcta caactggatt ccttggtctt tctggcaaat cgaagacgct actgggagct
 gccctggtct gagcagcagg taattctccc cacttaattt cagaacttcc tccctcaatg
 tagtctacct tctttacctt tcccttagcc ctatttggtc agcttatccc tactatcctt
 10 tatttgattg tttgagatac agtctcactc tggtgcccag gctgcagtgc agtggcatga
 tcagagttcg ctgtaacctc aaactcatga gctcaggcaa tctttctgcc tcagcctcct
 gaatagctag gatgacaggt ggttaccacc atgcctggct aatttttaaa tttttttttt
 gttttttgag atgaagtctt gctctgtcac ccaggcttga gtacagtggc acaagcttgg
 ctactgcaa cctctgtctc ccgggttcaa gcgattctcc tgcctcagcc tcccgagtag
 15 ctgggactac aggcactccc cacaatgccc ggctaatttt ttttttgttt tagtagagac
 agggtttcac catattggcc aggcctggtct cgaactgctg acctgtgat ctgcctgcct
 ctgcctctca aagtgtctgg attacaggtg tgagccacca tgcccgcca atttttaaat
 tttttgtaga gacagacaat acaaaaatgt ggacactatg tggagacact atgttgaggt
 actatgctgt ccagattggc cttgaactcc tggcctcaag caatcctcct gccttggcct
 20 cccaaagtgc tgggattaca gacctgagcc actgcaccca gcccctagt atcccttata
 atgtgacttg cttttctttt tctttctcct tcccttttct ttcatttctt tctactctc
 gagagaagag tgggcatctg ggagagtggg aggcctggtg gtcccacaga gtgaggaggc
 aggactgggt ccaaggcagt cctgcctctc cactctaggg ggtatccttg gacagtgtct
 cttctgggaa ggggctcgtc tttctttctc ttgtaggcac agtttctctg gaagaagatg
 25 caagtaccca ccaacctgac cctcaggaat ctacagttag taacttgtgt tgagcagtgc
 gctgaattcg accaacattt ttttgagtgc ttactatgtg ccaggcacca tatgatattg
 aatgggggat atagggatga atgatgcata gtccctgcct cgtggacgtt ctctagcac
 ctccctttgc cctcctttcc ttccacagtg ccatgcctat cctgactaga gccaaaggac
 tcagaaaacc tggattcagg ttccagtcct gtcacctact tgtcctcttg ggcaagtcct
 30 ttaacgtccc tgtgtcagtt ttcccttctt taaatgagaa ttacaatggc accagcctca
 taggtagtta ctgtgaagat taaatgaggt aggtcatgta agatatttaa cacagtgttt
 ggtccattgt aaagtccag tagtcatttg ctactgttag tttacttcag gatgacttca
 gaggcactgg ccaagcaaga ataaatagga atagaagggt atcactttac ttataccac
 attagaagaa caatgggctt cagaatcttt tttttttttt ttttcgagac agtcttgctc
 35 tggtgcccag gctggagtgc agtggcgca tttcggctca ctgcaacctc tgctcccag
 gttcaagcga ttctcctgtc tcagcctctg gagtagctgg gattacagga atgtgccacc
 ataccagct aatttttgta ttttttagtag agatgggggt tcaccatttt ggccaggctg
 gtctcaaaact cctgacctca ggtgatccac ccgcctcagc ctccaagggt cttcagaatc
 taagacatgt ctctagtttc agtttaccac atttctagca gaatgatgtt gggaatgtca
 40 cctgacttcc ataaatcctt attttctcct ctgataaaca gcagtgatgt tatggggagc
 tgatgagata tctatgtaaa aacatttctc aaaccataaa ttacgggtgga tgaacatctg
 tacttggtgt gagagtactg atatcaagga gcaaacaggc tggtgtatgt gttgaatgag
 cctctcccca ctcacacacc cacagggtct tgggcaccct ggcaggaggc atgtcctgtg
 agtttctgca gcagatcaac tccatggtag acttccttga agtgggtgcac atgatctatc

agctgcccac tagagttcga gggagcctgg tgagagggga tgcctggact ttagtgggag
 cagggaggct gggaccctag gtatagaacc cagctcctat gttctgctct ggccctcactc
 tgcttcccta cagagggcct gtatctgggc agagctacag cggaggatgg caatgccaga
 accagaatgg acaactgtag ggccagaact gaacgggctg gatagcaagc tactcctgga
 5 cttaccgtaa gtactgcagc tagagatatt ggcccctcag aaagctcaat ctgggggtgaa
 gatctgccct tagggaatgc cctggaggag gtagtttttc tgtctggtag ttccctgaca
 taatttatag cccaaagcag aggattttat tcaaagttgc tctatgtatt gactggttcc
 cagaatatgc tccagcacag ggcagctgag ggtggcaaca ctgtattgaa gcctgccaaag
 taatcttaca ataacctagt ccacattaat tgagattgag acagagcatc tgaagtgagg
 10 gaggcaatgc tccaaatctg cccagagga ttgtagtttg ctccagggcac tgtgttctta
 gtgcattcag aggagtggat cgagagaaaa atatatgaaa aatgtgataa ataccttcaa
 atacctgagg ggctatcaag tagaaattag attgtcatat ttatgagtgg cccattggg
 caagactaag agtagttaac ggagatcaga tttttacata gtataagaaa aactaaggta
 gtgagttcct ggtccttgga gctgttcgag cctaagccag atggcccat ggcaggaaatg
 15 ttgtagagca cgttcatata caggttgtgg gaagaaaagg ctataggaac ccaaggctcc
 tccctacca tggagaaatt tattagtatg ttactcatat gctgcttttc tcatthtacc
 cctaccacca ccccggttgc atccgcaactg taagtcagga taggaaaatg ctgggtgttac
 agtcttccctg gggaaatatgg agctgaagtg gagtaaaagc agttgacttc attcttactt
 ttttcttttt cttttttttt ttttttttga gacagagttt tgctgtgtca ccaaggctgg
 20 agtgcagtga cgtgatctcg gctcactgca acctccatct tccaggttca agcaattctc
 ctgcctcagt ctcccagta gctgggactg taggtgtgca ccaccatgcc aggctaattt
 ttgtattttgt tgtagggacg agctttcacc atggttgcca ggctggctct gaactcctgg
 cttcaagtga tctgcccacc tcggcttccc aacattctta tattttttata ggcctttcca
 cagatttcag ctcttgtagt acttagccca gttccagaac tggtaatcct aggtagggtta
 25 caggttatca cctctgattt cgggtaaaag ggattttattt atttattttgt ttattttattt
 atattttttga gacagagtct cgctctgtca cccaggctgg agtgcaatgg tgccatctcg
 gctcactgca acctctccct ctgggggttca agcaacttctg cctcagcctc ctgagtagct
 gggattacag gcgcgtgcca ccacacccgg ctaatttttg catttttagt agagacgggg
 tttcaccatg ttgctcaggg tgggtctcgaa tttctgacct tgtgatctgc ctgcctcggc
 30 ctcccaaagt gctgggatta caggcatgag cactgctgc cggcctgttt ttactttttt
 ttaatgccat tcagatctgt ttaaataatgt gggttctgtg agataattta gaatcccaag
 gttacagatg aggtgaaaga tcctagacca tgcatacaaaa aacttgagtt tctcatttgt
 gaaagaagga taagagaaac acctattttt tctgggtgca gtggctcatg cctataatcc
 cagcattttgg ggaggccaag gtgggtggat cacggaggtc aggtgttcaa gaccagactg
 35 gccaacatgg caaaacacca tctctactaa aaatacaaaa gttagctggg cgtgggtggca
 cgtgcgtgta attccagcta ttcgggaggc tgaggcacga gaattgcttg aacctgggag
 gtgcgggttg cagtgaactg agatcgagc accactgtgc tccagcctga gtgatggagt
 gaggccaggt cttgtttagt gatcaaatga gataacacct gaaagaactt tgtaaattgt
 atagcacgta caaacaagaa gggacctctt cacaagcaga ggaaggggtg tcctgtggaa
 40 aaaaacggga attgggagtg agagacctca acatttgatc tctgtgaacc tcagtttttt
 aatctataaa atggggaaat gttaatggta cttaatatth ggagcttttg agtccattag
 atcaggtagg attgttcgtt attttttttt ttttaggaaga ctagaaatat gttgctccct
 ttttctcccc cactcaagct tgatgggtgg aattggccct ggagctgttt actgtcagtt
 cctgtccagc ttcactaaat ttgggtctgg gtcacatctt agctggggac tgtgggggtc

tgtgggtccct tctcgacttg gccagctcc acttgaatga tattgttgtc aaattgctat
 aatagaatcc agttgatgga cagactatcc aatgaatcca ttatgttggg ggtggagctg
 gtgcaaagag ctccagagca gctgctggca ctgaccccc tccaccaggc agccctggca
 gagagggcac tacaaaacct ggtaagagtc caccctacca gactcagatt tgctgccctg
 5 ggcaattcct gctcctcaga caatgctctc tgactgtcct ccaaccctct acttcttgct
 ttcttgctgc caaacagggt cctgtctaca aggcctggcc cgttttgcct ctgggttctg
 ttcttgata atatgcttca cgttacttgt ccatacctct tggagtccga gaaatctctt
 ggagtccacc tctcagtctt tctgctgct cctatctggt ctcatgtctt aagaaagtga
 acaaaggtag tgagcatcat aggggtgctga gttgggagca ggaggaggagg aaggtcaggg
 10 ggcttgggtg cttgatcaag gtgtctggta ttctgagtca gaagtgcatt gtccaagtct
 tgatgctctt ctccaggctc caaaggagac tccagtctca ggggaagtgc tggagacctt
 aggccctttg gttggattcc tggggacaga gagcacacga cagatcccc tacagatcct
 gctgtcccat ctcagtcagc tgtaaggctt ctgcctagga gagacatttg ccacagagct
 gggatggctg ctattgcagg agtctgttct tgggtatgga tcttcgagaa cttcagattc
 15 taactcattc tataccaggt ccctcagcca ccatcatcag tggcagcctg ttccatattc
 ctaagggccc ttggagccct gtgtccgaaa tcctagcatg tcctcttttc cccttccttt
 tcctcacagt tccctcagct cccagcccc cgattttctt cctgtcccca ggaaaccaga
 gttgtggagc caggatgaag tagagcaagc tggacgccta gtattcactc tgtctactga
 ggcaatttcc ttgatcccca gggtgagatg aaggaagaag ggaaggagg aaatgcatag
 20 aggggactgg tgagctggtt atggggaccc gtggccaacg agggcaaagg atatgaagcc
 tagatctggg gggatactgc gaaacagaga caggactttg gacttatagc tatagcagca
 ggtcctgatc tgtccagatc tccccactct ccttctacct tctcatgcag gaggccttgg
 gtccagagac cctggagcgg cttctagaaa agcagcagag ctgggagcag agcagagttg
 gacagctgtg tagggggcca cagcttgctg ccaagaaaagc agccctggta gcaggggtg
 25 tgcgaccagc tgctgaggat cttccagggt aaactacca aatacttata tgtccagcag
 gatgtacagg gagtatcaaa cgggtctgggt tctacatgtg cttttctctg ggactgggtt
 ttctaattta taaagcaaag agtttagagg gatgatcttc aagtctctgc tagttctaaa
 attctatagc tctgggagtg tgtaaaactat taagttttcc tttagcccag aacttccatt
 ttctgctgt cttgtctctt taaacagctg actcagctct agctcggcta agtgtggagc
 30 tctctgctgg ggagatccct cgaaggagac attgtgaggc tagagaactg gtttcaaatt
 cagctctgcc atgaaatccc tgaccaacat cctcagtctt ccatctataa aatgaggggac
 ttggcatctc caatcacttc ttgttcaatt cacttctaag ccctatacat cttaacctgt
 aatactccct cgatccctac ttcccacagt gctccatcct ccaaagggtg gaagagtcatt
 tccatctaca gatacctcct ggccctgcct tcccccttca gaacctgtgc caaattgtgc
 35 agatgtacga gggacattcc cagcagcctg ctctgcaacc cagattgctg agatggagct
 ctcagacttt aaggactgcc tgacactatt tgcaggagac ccaggacttg ggcctgagga
 accacgggca gccatgggca aagcaaaatg ggtcaggggc ggagagccaa ctgaggttgg
 ccatgaggaa gctgtgtggg tgtgatgtag gacacgaaga gaatagagag ttggatgaga
 ggtgggggaa gcagaagata aaagagttag aagatttggg tcacaagtag aaggtgagga
 40 gagataaata ttgaggaaag agagctacta taatgaatag agggactaaa gcagtgatta
 ccaaatttta atgcatatca caatcatcaa gggaacagat ttttttcttt tttctttttt
 tttctttctt aaaaaataat ggcatacttc atgaatttgt gtgttatcct tgcgcaggcg
 ccatgctaatt ctctgtatca ttccaatttt tttttgtata tgtgctgctg aaccgagcac
 gggaacagtg ccaatgctgg attaccatgc tagactgact taataaaaaac ctttgcggag

gctggacgga gtggctcgtg cctgtaatca cagcactttg ggaggctgag gcagggtggat
 tgcttgagcc caggagtttg agaccagact gggcaacatg gtgaagccct gtctctacta
 aaaatagaaa aaaaattagc tgggtgtggc ggcacatgcc tgtaatccca gctactcagg
 aggctggggg aggagaatcc cctgagcgca ggagggtggag ggtgcaatga gccaagatca
 5 caccactgca ctccagcctg ggtgtcagag caagaccctg tctcataaat taaaaataa
 gcctctggga gaaagtctag acatctgcat ctgctttttt tttttttttt ttttttgaga
 cacagtctca ctctgtcacc cagcatccag gctggagtgc ggtgggtgtga tcttggtctca
 ctgtaacctc tacatcctgg gctcaaacga tcctcctgcc tcagcctctc gagaagctgg
 gactacagggt gcacaccact acacctggct aattttttgta ttttttgtag agatgagggt
 10 tcgctgtgtt gcccaggatg gtcttgaact cctgggctca agcattcctc ccacctcagc
 ctcccaaagt gctgggatta cagctgtgag ccactgtgct gggccccagg catctgcatt
 ttaataagcg tctctgtgtc tgatgcacat aaaagtgtgg aactcatgga ctagagttag
 tttgctcttc ttttccactg attgtaatgt ctttcaaaac accttagagg aactgtaagg
 caacggtctc attttatagt ggaggaaacc aaagaaaagg caagtgactt gcctagagtt
 15 atacagctaa ggtcagaggc aagacttaaa acccagcatt ctgactccca atctactggt
 tttccactca cattgttcct ctctttctcc tagttgtggg gtcccccccg gggatttggt
 cctgagcaga tcctgcagct cggtagactc ttaataggtc tgggagatca ggaactacag
 gagctgatcc tagtggactg gggagtgtg agcaccctgg ggcagataga tggctggagc
 tccactcagg taacactttt actcctccct accgcttccc aaacacccat cccacagacc
 20 cagccctata gatcatctaa agcccaagga attttttctt gtgaccctac ctgggtcattc
 tttctgtctt ttgttaatat ccatactag tgaccttcag gactcttgat ttattcactc
 tgaggccctg gacacataat actgtctcct acctcttttc ctggaggctt cctctttttc
 tttccttttc ttttctgagt cctcagcctt ccccatgact ccttaggtct taatagtaac
 agaataatac ccagtaacac ctatcacttc cctgtccatt aattctccat aactttcctc
 25 ctccccctct tctcccaccc cccaccccag ctccgcattg tggctctccag tttcctacgg
 cagagtgggc ggcagtgtgag ccacctggac ttcgttcacg tgacagcgct ggggtatact
 ctctgtggac tgcggccaga ggagctccag cacatcagca gttgggagtt taggtcattt
 gtgaaggggc tgagggtggt cgtgttcagg taaagggtga cttgctgggg aaaggaggat
 catgaagtta tagtcccatg gaggaaggga actcatttga agccatctct tcctttgtct
 30 catgaccaca gtccctttca ctgaagccga attcttcttc cttccttccc actgttctac
 agccaagcag ctctcttcct gggcaccctg catctgcagt gctctgagga acaactggag
 tttctggccc acctctttgt actgcctggg gggtttgccc caatcagtaa ctgggggctt
 gagatcttca ctgaaattgg caccatagca ggtggggagc tgggccactg ctgggtgcaag
 ttggtttggt ttctatacca tgggtggact ggatggaaga ctgccctgca attcttcagg
 35 tgtgggcctg agagggtgtt taaataaggg gctagagaca tattggggaa ggtctatgat
 agggcacttt gggagtagtt agagaagggt tataggtttg aagagaggga aggtcagtct
 aagacaatgt ttggatgcc cttgcttcaa cagctgggat cccagacctg gctctttcag
 cactgctgcg gggacagatc cagggcgtta ctctcttgc catttctgtc atccctcctc
 ctaaatattg tgtaagtatt aatggactgg ggtgaccaca ggagagccag ggcccaatgg
 40 ggactacatg catgcaactga ttctacccc tgccctcagg tgggtgttag tcccatccaa
 ctatctagtc tcgccagtgc tcaggctgtg gctgtcactc ctgagcaaag ggcctttctg
 agtctgagc agcgacgagc agttgcatgg gcccaacatg agggaaaagg gagcccagaa
 cagcaagggt agttcccagc tgcacagctt gatcctccat ctctgaccc agaatacaac
 ccctaatttg gtgctgtctg gctcttagag tgcacccagg gagatccctg gagtgaagga

gtctacaggc agagcgctaa tttccaagta tcaatgctcc tggagagctg agttgtgata
 ttactcccat tccctgtcta ttataggtcg aagtacagcc tggggcctcc aggactggtc
 acgaccttcc tgggtccctgg tattgactat cagcttcctt ggccacctgc tatgagcctg
 tctctacagt agaaggagat tgtggggaga gaaatcttaa gtcataatga ataaagtgca
 5 aacagaagtg catcctgatt attttcagaa gctgatgagg aata

In some embodiments, a stereocilin can be encoded by a sequence that is at least 70%, at least 72%, at least 74%, at least 75%, at least 76%, at least 78%, at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%,
 10 at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%, or 100% identical to SEQ ID NO: 8 or SEQ ID NO: 9.

One skilled in the art would appreciate that mutation of amino acids that are not conserved between the same protein from different species is less likely to have an effect on the function of a protein and therefore, these amino acids should be selected
 15 for mutation. Amino acids that are conserved between the same protein from different species should not be mutated, as these mutations are more likely to result in a change in the function of the protein. Non-limiting examples of stereocilin from other mammalian species are shown below.

20 **Mouse Stereocilin Protein (SEQ ID NO: 8)**

malslqpql lllsllpqev tsaptgpqsl daglsllksf vatldqapqr slsqsrfsaf
 lanisssfq grmgegpvge ppplqppalr lhdfvlvtrg spdwepmgl lgdvlallgq
 eqtprdfvlh qagvlggive allgalvpgg ppaptrppct rdgpsdcvla adwlpslml
 25 legtrwqalv qlqpsvdptn atgldgrepa phflqgllgl ltpagelgse ealwggllrt
 vgaplyaaafq egllrvthsl qdevfsimgq pepdasgqcq ggnlqqlllw gmrnnlswwa
 ralgflsgsp ppppallhcl srgvplpras qpaahisprq rraisvealc enhsgpeppy
 sisnfsiyll cqhikpatpr pppttprppp ttpqpppttt qpipdttqpp pvtprppptt
 pqpppstavi cqtavwyavs wapgargwlq achdqfpdqf ldmicgnlsf salsgpnrrl
 30 vkqlcagllp pptscppgli pvpltpeifw gcflenetlw aerlcvedsl qavpprnqaw
 vqhvcrgptl datdfppcrv gpcgercpdg gsflmvmcan dtlyealvpf wawlagqcri
 srggndtcfl egmlgpplls llplgspplc lapgpflgm lsqprcqss vpalahptrl
 hyllrlltfl lgpgtggaet qgmigqalll sslpdncsfw dafrpegrrs vlrtvgeylq
 reeptppgld ssllsgsgms kmellscfsp vlwdllqrek svwalrtlvk aylrmpedl
 35 qqlvlsaeme aaqgfltlml rswaklkvqp seeqamgrlt alllqryprl tsqfildmsp
 lipflavpdl mrfppsllan dsvlaairdh ssgmkpeqke alakrllape lfgevpdpwq
 ellwaalpll phlplesflq lsphqiqale dswpvaglgp gharhvlrsl vnqsmedgee
 qvrlrgslac flspeelqsl vplsdpmgpv eqglllecaan gtlspegrva yellgvlrss
 ggtvlsprel rvwaplfpql glrflqelse tqqlramlpal qgasvtpaqa vllfgrllpk
 40 hdlsleelcs lhpplpglsp qtlqaipkrv lvgacsclgp elsrlsacqi aallqtfrvk

dgvknmgaag agsavcipgq pttwpcdllp llplkllqld aaallanrrl yrqlpwseqq
 aqflwkkmqv ptnlslnrlq algnlaggmt ceflqqissm vdfldvvhml yqlptgvres
 lraciwtelq rrmtmpepel ttlgpelsel dtklldldpi qlmdrlsnds imlvvemvqg
 apeqllaltp lhqtalaera lknlapketp iskevletlg plvgflgies trriplpill
 5 shlsqlqgfc lgetfatelg wlllqepvlg kpelwsqdei eqagrlvftl saeaissipr
 ealgpeter llgkhqsw eq srvg hlcges qlahkkaalv agivhpaag lqepvpncad
 irgtfpaaws atqisemels dfedclslfa gdp glg peel raamgkakql wgp prg frpe
 qilqlgrlli glgerelqel tlvdwgvls lgqidgwssm qlravvssfl rqsgrhvshl
 dfiyltalgy tlcglrpeel qhisswefsq aalflgslhl pcseeqlevl ayllvlpggf
 10 gpvsnwgpei fteigtiaag ipdlalsall rgqiqltpl aisvipapkf avvfnpigls
 sltrgqavav tpeqlaylsp eqrravawaq hegkeipeql grnsawglyd wfqaswalal
 pvsifghll

Dog Stereocilin Protein (SEQ ID NO: 9)

15 malslwpvll lltcaatlvs sgiqslldpdl sllksllstm dqapqgslsr sqfsaflani
 sssfesgrmg egpvgepppl qspalrlhdf lvtlrgspdw epmlgllgdv lallgqeftp
 rdflghqagv lgglavellg alvpigp ptp trppcirdgp sdcvlvadwl pslllllegt
 rwqalvqvqp svdpntatgl dgrepaphfl qgllglltpv gelgseealw gllrtvgap
 20 lyaa fgegl rvt dslrnev fsilgqpepd angqcqggnl rqlllwgirh nlswdvqalg
 flsgsppppp allhclstgv plprasqpsa hinprqrrai svealcenhs gpappysisn
 fsihllcgha qpatpqp pps taavcq tamw yavswapgaq gwlqachdqf pdqfleaic
 nlsfsalsgp nrrlvkqlca gllppptncp eglppaplt p evfwgcflen etlwaerlcg
 eaglgavpps nqawvqhvcq gtpdvtafp pchvgpcger cpdggsflmm vcandtmyea
 25 lvpfwplag qcrisrggnd tcflegllgp llpslpplgp splclapapf llgmlsq lpr
 cqssvpalah strlhyllrl ltfillgpgag gtea qgm lqg almlsslpdn csfwdaf rpe
 grrsvlrvtg eylereeq lt ppgfeptasp ssgitkmell acfspvlwdl lqreksvwal
 qilvqaylhm ppenlqq l vl saereaaqgf ltlmhrswaq lqvppseeqa lgrltalllq
 ryprltsqlf idlsplipfl avsdmlrfpp sllandsvla airdyspgmr peqkealakr
 30 llapel fge v pawsgellwa vlpllp hpl enflqlsphq iqaledswpa aglgpgharh
 vlrslnqsv qdgee qvrrl gplacflspe elqslvplsd pmgpvergll ecaangtlsp
 qgrvayellg vlrssggavl splelrwap lfpqlglrfl qelsepqlra mlpalqgtsv
 tpaqavlllg rllprhdsl eelcslhpl pglssqtlqa iprrvligac sclapelsrl
 sacqtaallq tfrvkdgvkn igttgasaav cipgqqpipt twpcdllpl plklldsa
 35 allanrrrryr dlpwseqqag flwkkmqvpt nltlrnlqal gtlaggmsce flqqinlmad
 flevvhmiyq lptgvrgrsl aciwaelqrk mtmpepelat lgselsgldt klldspiyl
 mdrlnesim lmvelvrrap eqllaltp lh rvalaeralq nlapkettvs revletlgpl
 vgflgiestr ripllillah lnqlqgfclg epfatelgw llsqepilgkp elwsegeveq
 agrlvltlst eaisliprea lgpeterll ekqqsweqsr vgqlcgtpl apkkaalvag
 40 vvrptaedls epvpncadvr gt fpaawsat qiaemelsdf edclalfagd pglg peelra
 amgkakqlwg pprgfrpeqi qlslrlligl gerelqelil vdwgvlstlg qidgwssiq l
 rvvssflrq sgrhvshldf lhl talgytl cglrpeelqh isswefsqaa lflgnlhlqc
 seeqlevlaq llvlpggfgp vsnwgpeift eigtiaagip dlalsallqe qiqgltplai

svipapkfav vfsptqlssl tsvqamavtp eqmaflspeq rravvwaqhe gkespeqqgr
stawglqdws qpswamalti cflgnli

Vectors

- 5 The compositions provided herein include at least two (e.g., two, three, four, five, or six) nucleic acid vectors, where: each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions being at least 30 amino acids (e.g., between about 30 amino acids to about 1600 amino acids, about 30 amino acids to about 1550 amino acids, about 30 amino acids to about 1500 amino acids, about 30 amino acids to about 1450 amino acids, about 30 amino acids to about 1400 amino acids, about 30 amino acids to about 1350 amino acids, about 30 amino acids to about 1300 amino acids, about 30 amino acids to about 1250 amino acids, about 30 amino acids to about 1200 amino acids, about 30 amino acids to about 1150 amino acids, about 30 amino acids to about 1100 amino acids, about 30 amino acids to about 1050 amino acids, about 30 amino acids to about 1000 amino acids, about 30 amino acids to about 950 amino acids, about 30 amino acids to about 900 amino acids, about 30 amino acids to about 850 amino acids, about 30 amino acids to about 800 amino acids, about 30 amino acids to about 750 amino acids, about 30 amino acids to about 700 amino acids, about 30 amino acids to about 650 amino acids, about 30 amino acids to about 600 amino acids, about 30 amino acids to about 550 amino acids, about 30 amino acids to about 500 amino acids, about 30 amino acids to about 450 amino acids, about 30 amino acids to about 400 amino acids, about 30 amino acids to about 350 amino acids, about 30 amino acids to about 300 amino acids, about 30 amino acids to about 250 amino acids, about 30 amino acids to about 200 amino acids, about 30 amino acids to about 150 amino acids, about 30 amino acids to about 100 amino acids, about 30 amino acids to about 50 amino acids, about 50 amino acids to about 1600 amino acids, about 50 amino acids to about 1550 amino acids, about 50 amino acids to about 1500 amino acids, about 50 amino acids to about 1450 amino acids, about 50 amino acids to about 1400 amino acids, about 50 amino acids to about 1350 amino acids, about 50 amino acids to about 1300 amino acids, about 50 amino acids to about 1250 amino acids, about 50 amino acids to about 1200 amino acids, about 50 amino acids to about 1150 amino acids, about 50 amino acids to about 1100 amino acids, about 50 amino acids to about 1050 amino acids, about 50 amino acids to about 1000 amino acids, about 50

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

amino acids to about 1250 amino acids, about 1250 amino acids to about 1600 amino acids, about 1250 amino acids to about 1550 amino acids, about 1250 amino acids to about 1500 amino acids, about 1250 amino acids to about 1450 amino acids, about 1250 amino acids to about 1400 amino acids, about 1250 amino acids to about 1350 amino acids, about 1250 amino acids to about 1300 amino acids, about 1300 amino acids to about 1600 amino acids, about 1300 amino acids to about 1550 amino acids, about 1300 amino acids to about 1500 amino acids, about 1300 amino acids to about 1450 amino acids, about 1300 amino acids to about 1400 amino acids, about 1300 amino acids to about 1350 amino acids, about 1350 amino acids to about 1600 amino acids, about 1350 amino acids to about 1550 amino acids, about 1350 amino acids to about 1500 amino acids, about 1350 amino acids to about 1450 amino acids, about 1350 amino acids to about 1400 amino acids, about 1400 amino acids to about 1600 amino acids, about 1400 amino acids to about 1550 amino acids, about 1400 amino acids to about 1500 amino acids, about 1400 amino acids to about 1450 amino acids, about 1450 amino acids to about 1600 amino acids, about 1450 amino acids to about 1550 amino acids, about 1450 amino acids to about 1500 amino acids, about 1500 amino acids to about 1600 amino acids, about 1500 amino acids to about 1550 amino acids, or about 1550 amino acids to about 1600 amino acids), wherein the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions; no single vector of the at least two different vectors encodes an active stereocilin protein (e.g., a full-length stereocilin protein (e.g., a full-length wildtype stereocilin protein)); and, when introduced into a mammalian cell that contains chromosomal DNA, the at least two different vectors undergo homologous recombination with each other and with the chromosomal DNA of the cell, thereby forming a recombined nucleic acid inserted into the chromosomal DNA, where the recombined nucleic acid encodes an active stereocilin protein (e.g., a full-length stereocilin protein). In some embodiments, one of the nucleic acid vectors can include a coding sequence that encodes a portion of a stereocilin protein, where the encoded portion is, e.g., about 900 amino acids to about 1600 amino acids, about 900 amino acids to about 1550 amino acids, about 900 amino acids to about 1500 amino acids, about 900 amino acids to about 1450 amino acids, about 900 amino acids to about 1400 amino acids, about 900 amino acids to about 1350 amino acids, about 900 amino acids to about 1300 amino acids, about 900 amino acids to about 1250 amino acids, about 900 amino acids to about 1200 amino acids,

[illegible]

1400 amino acids, about 1150 amino acids to about 1350 amino acids, about 1150 amino acids to about 1300 amino acids, about 1150 amino acids to about 1250 amino acids, about 1150 amino acids to about 1200 amino acids, about 1200 amino acids to about 1600 amino acids, about 1200 amino acids to about 1550 amino acids, about
5 1200 amino acids to about 1500 amino acids, about 1200 amino acids to about 1450 amino acids, about 1200 amino acids to about 1400 amino acids, about 1200 amino acids to about 1350 amino acids, about 1200 amino acids to about 1300 amino acids, about 1200 amino acids to about 1250 amino acids, about 1250 amino acids to about 1600 amino acids, about 1250 amino acids to about 1550 amino acids, about 1250
10 amino acids to about 1500 amino acids, about 1250 amino acids to about 1450 amino acids, about 1250 amino acids to about 1400 amino acids, about 1250 amino acids to about 1350 amino acids, about 1250 amino acids to about 1300 amino acids, about 1300 amino acids to about 1600 amino acids, about 1300 amino acids to about 1550 amino acids, about 1300 amino acids to about 1500 amino acids, about 1300 amino acids to about 1450 amino acids, about 1300 amino acids to about 1400 amino acids, about 1300 amino acids to about 1350 amino acids, about 1350 amino acids to about 1600 amino acids, about 1350 amino acids to about 1550 amino acids, about 1350 amino acids to about 1500 amino acids, about 1350 amino acids to about 1450 amino acids, about 1350 amino acids to about 1400 amino acids, about 1400 amino acids to about 1600 amino acids, about 1400 amino acids to about 1550 amino acids, about 1400 amino acids to about 1500 amino acids, about 1400 amino acids to about 1450 amino acids, about 1450 amino acids to about 1600 amino acids, about 1450 amino acids to about 1550 amino acids, about 1450 amino acids to about 1500 amino acids, about 1500 amino acids to about 1600 amino acids, about 1500 amino acids to about
20 1550 amino acids, or about 1550 amino acids to about 1600 amino acids in length.

In some embodiments of these compositions, at least one of the coding sequences includes a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks the intronic sequence that naturally occurs between the two neighboring exons.

30 In some embodiments, the amino acid sequence of none of the encoded portions overlaps even in part with the amino acid sequence of a different one of the encoded portions. In some embodiments, the amino acid sequence of one or more of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions. In some embodiments, the amino acid sequence of each

of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions.

In some embodiments, the overlapping amino acid sequence is between about 30 amino acid residues to about 1000 amino acids (e.g., or any of the subranges of this range described herein) in length.

In some examples, the vectors include two different vectors, each of which comprises a different segment of an intron, wherein the intron includes the nucleotide sequence of an intron that is present in a stereocilin genomic DNA (e.g., any of the exemplary introns in SEQ ID NO: 4 described herein), and wherein the two different segments overlap in sequence by at least 100 nucleotides (e.g., about 100 nucleotides to about 5,000 nucleotides, about 100 nucleotides to about 4,500 nucleotides, about 100 nucleotides to about 4,000 nucleotides, about 100 nucleotides to about 3,500 nucleotides, about 100 nucleotides to about 3,000 nucleotides, about 100 nucleotides to about 2,500 nucleotides, about 100 nucleotides to about 2,000 nucleotides, about 100 nucleotides to about 1,500 nucleotides, about 100 nucleotides to about 1,000 nucleotides, about 100 nucleotides to about 800 nucleotides, about 100 nucleotides to about 600 nucleotides, about 100 nucleotides to about 400 nucleotides, about 100 nucleotides to about 200 nucleotides, about 200 nucleotides to about 5,000 nucleotides, about 200 nucleotides to about 4,500 nucleotides, about 200 nucleotides to about 4,000 nucleotides, about 200 nucleotides to about 3,500 nucleotides, about 200 nucleotides to about 3,000 nucleotides, about 200 nucleotides to about 2,500 nucleotides, about 200 nucleotides to about 2,000 nucleotides, about 200 nucleotides to about 1,500 nucleotides, about 200 nucleotides to about 1,000 nucleotides, about 200 nucleotides to about 800 nucleotides, about 200 nucleotides to about 600 nucleotides, about 200 nucleotides to about 400 nucleotides, about 400 nucleotides to about 5,000 nucleotides, about 400 nucleotides to about 4,500 nucleotides, about 400 nucleotides to about 4,000 nucleotides, about 400 nucleotides to about 3,500 nucleotides, about 400 nucleotides to about 3,000 nucleotides, about 400 nucleotides to about 2,500 nucleotides, about 400 nucleotides to about 2,000 nucleotides, about 400 nucleotides to about 1,500 nucleotides, about 400 nucleotides to about 1,000 nucleotides, about 400 nucleotides to about 800 nucleotides, about 400 nucleotides to about 600 nucleotides, about 600 nucleotides to about 5,000 nucleotides, about 600 nucleotides to about 4,500 nucleotides, about 600 nucleotides to about 4,000 nucleotides, about 600 nucleotides to about 3,500 nucleotides, about 600 nucleotides

to about 3,000 nucleotides, about 600 nucleotides to about 2,500 nucleotides, about 600 nucleotides to about 2,000 nucleotides, about 600 nucleotides to about 1,500 nucleotides, about 600 nucleotides to about 1,000 nucleotides, about 600 nucleotides to about 800 nucleotides, about 800 nucleotides to about 5,000 nucleotides, about 800 nucleotides to about 4,500 nucleotides, about 800 nucleotides to about 4,000 nucleotides, about 800 nucleotides to about 3,500 nucleotides, about 800 nucleotides to about 3,000 nucleotides, about 800 nucleotides to about 2,500 nucleotides, about 800 nucleotides to about 2,000 nucleotides, about 800 nucleotides to about 1,500 nucleotides, about 800 nucleotides to about 1,000 nucleotides, about 1,000 nucleotides to about 5,000 nucleotides, about 1,000 nucleotides to about 4,500 nucleotides, about 1,000 nucleotides to about 4,000 nucleotides, about 1,000 nucleotides to about 3,500 nucleotides, about 1,000 nucleotides to about 3,000 nucleotides, about 1,000 nucleotides to about 2,500 nucleotides, about 1,000 nucleotides to about 2,000 nucleotides, about 1,000 nucleotides to about 1,500 nucleotides, about 1,500 nucleotides to about 5,000 nucleotides, about 1,500 nucleotides to about 4,500 nucleotides, about 1,500 nucleotides to about 4,000 nucleotides, about 1,500 nucleotides to about 3,500 nucleotides, about 1,500 nucleotides to about 3,000 nucleotides, about 1,500 nucleotides to about 2,500 nucleotides, about 1,500 nucleotides to about 2,000 nucleotides, about 2,000 nucleotides to about 5,000 nucleotides, about 2,000 nucleotides to about 4,500 nucleotides, about 2,000 nucleotides to about 4,000 nucleotides, about 2,000 nucleotides to about 3,500 nucleotides, about 2,000 nucleotides to about 3,000 nucleotides, about 2,000 nucleotides to about 2,500 nucleotides, about 2,500 nucleotides to about 5,000 nucleotides, about 2,500 nucleotides to about 4,500 nucleotides, about 2,500 nucleotides to about 4,000 nucleotides, about 2,500 nucleotides to about 3,500 nucleotides, about 2,500 nucleotides to about 3,000 nucleotides, about 3,000 nucleotides to about 5,000 nucleotides, about 3,000 nucleotides to about 4,500 nucleotides, about 3,000 nucleotides to about 4,000 nucleotides, about 3,000 nucleotides to about 3,500 nucleotides, about 3,500 nucleotides to about 5,000 nucleotides, about 3,500 nucleotides to about 4,500 nucleotides, about 3,500 nucleotides to about 4,000 nucleotides, about 4,000 nucleotides to about 5,000 nucleotides, about 4,000 nucleotides to about 4,500 nucleotides, about 4,500 nucleotides to about 5,000 nucleotides), in length.

The overlapping nucleotide sequence in any two of the different vectors can include part or all of one or more exons of a stereocilin gene (e.g., any one or more of the exemplary exons in SEQ ID NO: 4 described herein).

In some embodiments, the number of different vectors in the composition is
5 two, three, four, or five. In compositions where the number of different vectors in the composition is two, the first of the two different vectors can include a coding sequence that encodes an N-terminal portion of the stereocilin protein. In some examples, the N-terminal portion of the stereocilin gene is between about 30 amino acids to about 1780 amino acids (or any of the subranges of this range described
10 above) in length. In some examples, the first vector further includes one or both of a promoter (e.g., any of the promoters described herein or known in the art) and a Kozak sequence (e.g., any of the exemplary Kozak sequences described herein or known in the art). In some examples, the first vector includes a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter. In some
15 examples, the second of the two different vectors includes a coding sequence that encodes a C-terminal portion of the stereocilin protein. In some examples, the C-terminal portion of the stereocilin protein is between 30 amino acids to about 1780 amino acids (or any of the subranges of this range described above) in length. In some examples, the second vector further includes a poly(A) sequence.

20 In some examples where the number of different vectors in the composition is two, the N-terminal portion encoded by one of the two vectors can include a portion comprising amino acid position 1 to about amino acid position 1,500, about amino acid position 1,490, about amino acid position 1,480, about amino acid position 1,470, about amino acid position 1,460, about amino acid position 1,450, about amino acid
25 position 1,440, about amino acid position 1,430, about amino acid position 1,420, about amino acid position 1,410, about amino acid position 1,400, about amino acid position 1,390, about amino acid position 1,380, about amino acid position 1,370, about amino acid position 1,360, about amino acid position 1,350, about amino acid position 1,340, about amino acid position 1,330, about amino acid position 1,320,
30 about amino acid position 1,310, about amino acid position 1,300, about amino acid position 1,290, about amino acid position 1,280, about amino acid position 1,270, about amino acid position 1,260, about amino acid position 1,250, about amino acid position 1,240, about amino acid position 1,230, about amino acid position 1,220, about amino acid position 1,210, about amino acid position 1,200, about amino acid

position 1,190, about amino acid position 1,180, about amino acid position 1,170,
about amino acid position 1,160, about amino acid position 1,150, about amino acid
position 1,140, about amino acid position 1,130, about amino acid position 1,120,
about amino acid position 1,110, about amino acid position 1,100, about amino acid
5 position 1,090, about amino acid position 1,080, about amino acid position 1,070,
about amino acid position 1,060, about amino acid position 1,050, about amino acid
position 1,040, about amino acid position 1,030, about amino acid position 1,020,
about amino acid position 1,010, about amino acid position 1,000, about amino acid
position 990, about amino acid position 980, about amino acid position 970, about
10 amino acid position 960, about amino acid position 950, about amino acid position
940, about amino acid position 930, about amino acid position 920, about amino acid
position 910, about amino acid position 900, about amino acid position 890, about
amino acid position 880, about amino acid position 870, about amino acid position
860, about amino acid position 850, about amino acid position 840, about amino acid
15 position 830, about amino acid position 820, about amino acid position 810, about
amino acid position 800, about amino acid position 790, about amino acid position
780, about amino acid position 770, about amino acid position 760, about amino acid
position 750, about amino acid position 740, about amino acid position 730, about
amino acid position 720, about amino acid position 710, about amino acid position
20 700, about amino acid position 690, about amino acid position 680, about amino acid
position 670, about amino acid position 660, about amino acid position 650, about
amino acid position 640, about amino acid position 630, about amino acid position
620, about amino acid position 610, about amino acid position 600, about amino acid
position 590, about amino acid position 580, about amino acid position 570, about
25 amino acid position 560, about amino acid position 550, about amino acid position
540, about amino acid position 530, about amino acid position 520, about amino acid
position 510, about amino acid position 500, about amino acid position 490, about
amino acid position 480, about amino acid position 470, about amino acid position
460, about amino acid position 450, about amino acid position 440, about amino acid
30 position 430, about amino acid position 420, about amino acid position 410, about
amino acid position 400, about amino acid position 390, about amino acid position
380, about amino acid position 370, about amino acid position 360, about amino acid
position 350, about amino acid position 340, about amino acid position 330, about
amino acid position 320, about amino acid position 310, about amino acid position

300, about amino acid position 290, about amino acid position 280, about amino acid position 270, about amino acid position 260, about amino acid position 250, about amino acid position 240, about amino acid position 230, about amino acid position 220, about amino acid position 210, about amino acid position 200, about amino acid position 190, about amino acid position 180, about amino acid position 170, about amino acid position 160, about amino acid position 150, about amino acid position 140, about amino acid position 130, about amino acid position 120, about amino acid position 110, about amino acid position 100, about amino acid position 90, about amino acid position 80, about amino acid position 70, about amino acid position 60, about amino acid position 50, or about amino acid position 40 of a wildtype stereocilin protein (e.g., SEQ ID NO: 1).

In some examples where the number of different vectors in the composition is two, the N-terminal portion of the precursor stereocilin protein can include a portion comprising amino acid position 1 to amino acid position 310, amino acid position 1 to about amino acid position 320, amino acid position 1 to about amino acid position 330, amino acid position 1 to about amino acid position 340, amino acid position 1 to about amino acid position 350, amino acid position 1 to about amino acid position 360, amino acid position 1 to about amino acid position 370, amino acid position 1 to about amino acid position 380, amino acid position 1 to about amino acid position 390, amino acid position 1 to about amino acid position 400, amino acid position 1 to about amino acid position 410, amino acid position 1 to about amino acid position 420, amino acid position 1 to about amino acid position 430, amino acid position 1 to about amino acid position 440, amino acid position 1 to about amino acid position 450, amino acid position 1 to about amino acid position 460, amino acid position 1 to about amino acid position 470, amino acid position 1 to about amino acid position 480, amino acid position 1 to about amino acid position 490, amino acid position 1 to about amino acid position 500, amino acid position 1 to about amino acid position 510, amino acid position 1 to about amino acid position 520, amino acid position 1 to about amino acid position 530, amino acid position 1 to about amino acid position 540, amino acid position 1 to about amino acid position 550, amino acid position 1 to about amino acid position 560, amino acid position 1 to about amino acid position 570, amino acid position 1 to about amino acid position 580, amino acid position 1 to about amino acid position 590, amino acid position 1 to about amino acid position 600, amino acid position 1 to about amino acid position 610, amino acid

[illegible]

[illegible]

position 1 to about amino acid position 1,600 of a wildtype stereocilin protein (e.g., SEQ ID NO: 1).

As used herein, the term “vector” means a composition including a polynucleotide capable of carrying at least one exogenous nucleic acid fragment, e.g.,
5 a plasmid vector, a transposon, a cosmid, an artificial chromosome (e.g., a human artificial chromosome (HAC), a yeast artificial chromosome (YAC), a bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC)) or a viral vector (e.g., any adenoviral vectors (e.g., pSV or pCMV vectors), any retroviral vectors as described herein) and any Gateway® vectors. A vector can, e.g., include
10 sufficient cis-acting elements for expression; other elements for expression can be supplied by the host mammalian cell or in an *in vitro* expression system. The term “vector” includes any genetic element (e.g., a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector, etc.) that is capable of replicating when associated with the proper control elements. Thus, the term includes cloning and
15 expression vectors, as well as viral vectors (e.g., an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector).

Vectors include all those known in the art, including cosmids, plasmids (e.g., naked or contained in liposomes) and viruses (e.g., lentiviruses, retroviruses, adenoviruses, and adeno-associated viruses) that incorporate the recombinant
20 polynucleotide. Skilled practitioners will be capable of selecting suitable vectors and mammalian cells for making any of the nucleic acids described herein. In some embodiments, the vector is a plasmid (i.e. a circular DNA molecule that can autonomously replicate inside a cell). In some embodiments, the vector can be a cosmid (e.g., pWE and sCos series (Wahl et al. (1987), Evans et al. (1989)).

25 In some embodiments, the vector(s) is an artificial chromosome. An artificial chromosome is a genetically engineered chromosome that can be used as a vector to carry large DNA inserts. In some embodiments, the artificial chromosome is human artificial chromosome (HAC) (see, e.g., Kouprina et al., *Expert Opin. Drug Deliv* 11(4): 517-535, 2014; Basu et al., *Pediatr. Clin. North Am.* 53: 843-853, 2006; Ren et al., *Stem. Cell Rev.* 2(1):43-50, 2006; Kazuki et al., *Mol. Ther.* 19(9):1591-1601,
30 2011; Kazuki et al., *Gen. Ther.* 18: 384-393, 2011; and Katoh et al., *Biochem. Biophys. Res. Commun.* 321:280-290, 2004).

In some embodiments, the vector(s) is a yeast artificial chromosome (YAC) (see, e.g., Murray et al., *Nature* 305: 189-193, 1983; Ikeno et al. (1998) *Nat. Biotech.*

16:431-439, 1998). In some embodiments, the vector(s) is a bacterial artificial chromosome (BAC) (e.g., pBeloBAC11, pECBAC1, and pBAC108L). In some embodiments, the vector(s) is a P1-derived artificial chromosome (PAC). Examples of artificial chromosome are known in the art.

5 In some embodiments, the vector(s) is a viral vector (e.g., adeno-associated virus, adenovirus, lentivirus, and retrovirus). Non-limiting examples of viral vectors are described herein. In some embodiments, the vector(s) is an adeno-associated viral vector (AAV) (see, e.g., Asokan et al., *Mol. Ther.* 20: 699-7080, 2012).

“Recombinant AAV vectors” or “rAAVs” are typically composed of, at a minimum, a
10 transgene or a portion thereof and a regulatory sequence, and optionally 5' and 3' AAV inverted terminal repeats (ITRs). Such a recombinant AAV vector is packaged into a capsid and delivered to a selected target cell (e.g., an outer hair cell).

The AAV sequences of the vector typically comprise the cis-acting 5' and 3' ITR sequences (See, e.g., B. J. Carter, in "Handbook of Parvoviruses", ed., P. Tijsser, CRC Press, pp. 155-168, 1990). Typical AAV ITR sequences are about 145
15 nucleotides in length. In some embodiments, at least 75% of a typical ITR sequence (e.g., at least 80%, at least 85%, at least 90%, or at least 95%) is incorporated into the AAV vector. The ability to modify these ITR sequences is within the skill of the art. (See, e.g., texts such as Sambrook et al., "Molecular Cloning. A Laboratory Manual",
20 2d ed., Cold Spring Harbor Laboratory, New York, 1989; and K. Fisher et al., *J Virol.* 70:520-532, 1996). In some embodiments, any of the coding sequences described herein are flanked by 5' and 3' AAV ITR sequences in the AAV vectors. The AAV ITR sequences may be obtained from any known AAV, including presently identified AAV types.

25 A non-limiting example of a 5' AAV ITR sequence is SEQ ID NO: 18. A non-limiting example of a 3' AAV ITR sequence is SEQ ID NO: 36.

AAV vectors as described herein may include any of the regulatory elements described herein (e.g., one or more of a promoter, a polyA sequence, and an IRES).

In some embodiments, the AAV vector is selected from the group consisting
30 of: an AAV1 vector, an AAV2 vector, an AAV3 vector, an AAV4 vector, an AAV5 vector, an AAV6 vector, an AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV2.7m8 vector, an AAV8BP2 vector, and an AAV293 vector. Additional exemplary AAV vectors that can be used herein are known in the art. See, e.g., Kanaan et al., *Mol. Ther. Nucleic Acids* 8:184-197, 2017; Li et al., *Mol. Ther.* 16(7):

1252-1260; Adachi et al., *Nat. Commun.* 5: 3075, 2014; Isgrig et al., *Nat. Commun.* 10(1): 427, 2019; and Gao et al., *J. Virol.* 78(12): 6381-6388.

In some embodiments, an AAV vector provided herein includes or consists of a sequence that is at least 80% identical (e.g., at least 82%, at least 84%, at least 85%,
 5 at least 86%, at least 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 98%, at least 99%, or 100% identical) to SEQ ID NOs: 12-17.

pITR-201 (SEQ ID NO: 12)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggcgctcgggcgacctttggtcgcccg
 10 gcctcagtgagcgagcgagcgcgagagaggagtggtgccaactccatcactaggggttcctgcggccgc
 acgcgtagacttgccaaaaactctggcaagttgaattgcttccaaaagtccaactttatcagaagga
 cctattccctacggaaactagctgaaggtctcctggcactctggatctcgtggaaggaggcctttctca
 gggaacagaggggagcgattaagtggtgaaaagcaaacagacctggaaaagtccctttctgagagtagc
 aacagaaaagctctgcaaagactccctccaagctattggatcctcttgcttgggataaccacttgagtac
 15 tcagataccaaaagaagagtggaaatcccaagagaagtcaccagaaaaaacagcttttaagaaaaagga
 tacacttttgtccctgaacgcttgtgaaagcaatctgacaatagcagcaattgaaaagggacaaaataa
 gcccgaaatagaagtcacctgggcaagcaaggtaggactgaaaggctgtgctctcaaaaccaccagt
 cttgaaacgcactcaacgggtgaaagtagtaggaaagggtgaattttacaaaggacgtaggactcaaaga
 gtgagtttttccaagcagcagaaacctatttcttactaacttggataatttactgaaaaataatacaca
 20 caatcaagaaaaaaaaattcaggaagaaatagaaaagaaggaaaaacttaatccaagagtgaatagtttt
 gcctcagataactacagtgactggcactaagaattttctgaaagaaccttttcttactgagcactaggct
 gaaaatagaaggttacttgaacggggacttgactccagtacttcaagatttttaggtactttgaaaattc
 aacaaatagaacaaagaaacacacagctactttctcaaaaaaggaggaggaagaaaacttggaaggctt
 gggaaatcaaaccaagcaaattgtagagaaattgactgacaccacaaggatatctcctaatacaagcca
 25 gcagaattttgtcacgcaacgtagtaagagagctttgaaactagatcccatatatggagttccgcgtta
 cataacttacggtaaatggcccgctggctgaccgccaacgacccccgcccattgacgtcaataatga
 cgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtattttacggtaaa
 ctgcccacttggcagtagcatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggta
 aatggcccgctggcattatgccagtagcatgaccttatgggactttcctacttggcagtagcatctacg
 30 tattagtcacgctattaccatgggtgatgcggttttgggtgatgcggttttggcagtagcatcaatgggcg
 tggatagcgggtttgactcacggggatttccaagtctccacccattgacgtcaatgggagtttgttttg
 gcacccaaaatcaacgggactttccaaaatgtcgtaacaaactccgccccattgacgcaaatgggcggtag
 gcgtgtacgggtgggaggtctatataagcagagctcggttagtgaaccgtcaccgggtggtgccttgccct
 cacctggctatccacacaggtgagaataaccagaactcacctccggtaccagtggttacttggccacc
 35 atggctctcagcctctggccccctgctgctgctgctgctgctgctgctgctgctgctgctcctttgcagtgact
 ctggccccctactgggacctcattccctggaccctgggtctctccttccctgaagtcattgctctccactctg
 gaccaggctccccagggtccctgagccgctcacgggttctttacattcctggccaacattttcttcttcc
 tttgagcctgggagaatgggggaaggaccagtaggagagccccacctctccagccgctgctctgcgg
 ctccatgattttctagtgacactgagaggttagccccgactgggagccaatgctagggctgctaggggat

atgctggcactgctgggacaggagcagactccccgagatttcctggtgcaccaggcaggggtgctgggt
 ggacttgtggaggtgctgctgggagccttagttcctgggggccccctaccccaactcggcccccattgc
 acccgtgatgggcccgtctgactgtgtcctggctgctgactgggtgccttctctgctgctgtttagag
 ggcacacgctggcaagctctggtgcaggtgcagcccagtggtgacccccaccaatgccacaggcctcgat
 5 gggagggaggcagctcctcactttttgcaggggtctgttgggtttgcttaccccaacaggggagctaggc
 tccaaggaggctctttggggcggtctgctacgcacagtgggggccccctctatgctgcctttcaggag
 gggctgctccgtgtcactcactccttgccaggtgaggtcttctccattttggggcagccagagcctgat
 accaatgggcagtgccaggggaggtaaccttcaacagctgctcttatggggcgctccggcacaacctttcc
 tgggatgtccaggcgctgggctttctgtctggatcaccacccccacccctgcctccttactgcctg
 10 agcacgggctgctctgcccagagcttctcagccgtcagcccacatcagcccacgccaacggcgagcc
 atcactgtggaggccctctgtgagaaccacttagggccagcaccacccctacagcatttccaacttctcc
 atccacttgcctctgccagcacaccaagcctgccactccacagccccatcccagcaccactgccatctgc
 cagacagctgtgtggtatgcagtgtcctgggacccaggtgcccaaggctggctacaggcctgccacgac
 cagtttctctgatgagtttttgatgcatctgcagtaacctctccttttcagccctgtctggctccaac
 15 cgcgcctggtgaagcggtctgtgctggcctgctccacccccctaccagctgcctgaaggcctgcc
 cctgttccccctcaccccagacatcttttggggctgcttcttgagaatgagactctgtgggctgagcga
 ctgtgtggggaggcaagtctacaggctgtgccccccagcaaccaggcttggtccagcatgtgtgccag
 ggccccaccccagatgtcactgcctccccaccatgccacattggaccctgtggggaacgctgcccggat
 gggggcagcttctggtgatggtctgtgccaatgacaccatgtatgaggtcctggtgcccttctggcct
 20 tggctagcaggccaatgcaggataagtcgtgggggcaatgacacttgcttcttagaagggtgctgggc
 ccccttctgcctctctgccaccactgggacccatccccactctgtctgacccctggcccttctcctt
 ggcattgctatcccagttgccacgctgtcagtcctctgtcccagctcttgctcaccacacagcctacac
 tatctcctccgctgctgaccttctcttgggtccaggggtggtggcgctgaggcccaggggatgctg
 ggtcgggcccactactgctctccagctctccagacaaactgctccttctgggatgcctttcgcccagagggc
 25 cggcgagtgctgctacggacgattggggaatacctggaacaagatgaggagcagccaaccccatcaggc
 tttgaacccactgtcaaccccagctctggtataagcaagatggagctgctggcctgcttttagtcctgtg
 ctgtgggatctgctccagagggaaaagagtgtttgggcccctgcagattctagtgcaggttaagtatcaag
 gttacaagacaggtttaaggagaccaatagaaaactgggcttgtcgagacagagaagactcttgctttc
 tgggattttgcccatttccgctatttggttaaaaaatgagctgatttaacaaaaatttaacgcgaattt
 30 taacaaaataagcttgaattcagctgacgtgcctcgacccgctaggaacccctagtgtgaggttgcc
 actccctctctgcgcgctcgctcgctcactgaggccgggacgaccaaaggctgcccagcggcggttt
 gcccgggcgccctcagtgagcgagcgagcgcgagctgcctgcagg

pITR-202 (SEQ ID NO: 13)

35 cctgcaggcagctgcgcgctcgctcgctcactgaggccgccccggcgctcggcgacctttggtcgcccg
 gcctcagtgagcgagcgagcgcgagagaggagtgccaaactccatcactaggggttctcgggccgc
 acgcgtgggattttgccgatttccgctatttggttaaaaaatgagctgatttaacaaaaatttaacgcg
 aattttaacaaaaatgataggcacctattgggtcttactgacatccactttgcctttctctccacaggcgt
 acctgcataatgccccagaaaacctccagcagctgggtgctttcagcagagagggaggctgcacagggt
 40 tcttgacactcatgctgcaggggaagctgcaggggaagctgcaggtaccaccatccgaggagcagggcc

tgggtcgccctgacagccctgctgctccagcggtagccacgcctcacctcccagctcttcattgacctgt
caccactcatccctttcttggtgtctctgacctgatgcgcttcccaccatccctgttagccaacgaca
gtgtcctgggtgccatccgggattacagcccaggaatgaggcctgaacagaaggaggctctggcaaagc
gactgctggccctgaactgtttggggaagtgcctgcctggccccaggagctgctgtgggcagtgtgc
5 cctgctccccacctccctctgggagaactttttgcagctcagccctcaccagatccaggccctggagg
atagctggccagcagcaggtctggggccagggcagctgcccgcctatgtgctgcgcagcctggtaaaccaga
gtgtccaggatggtgaggagcaggtacgcaggcttggggccctcgctgtttcctgagccctgaggagc
tgcagagcctagtgtccctgagtgatccaacggggccagtagaacgggggctgctggaatgtgcagcca
atgggaccctcagcccagaaggacgggtggcatatgaacttctgggtgtgttgcgctcatctggaggag
10 cgggtgctgagccccgggagctgcgggtctggggccctctcttctctcagctgggcctccgcttccttc
aggagctgtcagagccccagcttagagccatgcttctgtcctccagggaactagtgttacacctgtctc
aggctgtcctgctgcttggacggctccttcttaggcacgatctatccctggaggaaactctgtccttgc
accttctgtaccaggcctcagccccagacactccaggccatccctaggcgagctcctgggtcggggctt
gttctctgctggccccctgaactgtcacgcctctcagcctgccagaccgcagcactgctgcagaccttct
15 ggggttaaagatggtgttaaaaaatatgggtacaacaggtgctgggtccagctgtgtgtatccctggtcagc
ctattcccaccacctggccagactgcctgcttcccctgtctccattaaagctgtacaactggattcct
tgggtcttctggcaaatacgaagacgctactgggagctgcccctggtctgagcagcaggcacagtttctct
ggaagaagatgcaagtacccaccaaccttaccctcaggaatctgcaggctctgggcaccctggcaggag
gcatgtcctgtgagtttctgcagcagatcaactccatggtagacttccctgaagtgggtgcacatgatct
20 atcagctgcccactagagttcgagggagcctgagggcctgtatctgggcagagctacagcggaggatgg
caatgccagaaccagaatggacaactgtagggccagaactgaacgggctggatagcaagctactcctgg
acttaccgatccagttgatggacagactatccaatgaatccattatgttggtgggtggagctggtgcaaa
gagctccagagcagctgctggcactgacccccctccaccaggcagccctggcagagaggggcactacaaa
acctggctccaaaggagactccagtctcaggggaagtgtgtagaccttaggccctttggttgattcc
25 tggggacagagagcacacgacagatccccctacagatcctgctgtcccattctcagtcagctgcaaggct
tctgcctaggagagacatttgccacagagctgggatggctgctattgcaggagtctgttcttgggaaac
cagagttgtggagccaggatgaagtagagcaagctggacgcctagtattcactctgtctactgaggcaa
tttcttctgatccccagggagggccttgggtccagagaccctggagcggcttctagaaaagcagcagagct
gggagcagagcagagttggacagctgtgtagggagccacagcttgcctgccaagaaagcagccctggtag
30 caggggtggtgagcagcagctgctgaggatcttccagaacctgtgccaattgtgcagatgtacgaggga
cattcccagcagcctggtctgcaaccagattgcagagatggagctctcagactttgaggactgcctga
cattatattgcaggagaccaggacttgggcctgaggaaactgcgggcagccatgggcaaaagcaaaacagt
tgtggggtcccccccggggatttctgctcctgagcagatcctgcagcttggtaggctcttaataaggtctag
gagatcgggaactacaggagctgatcctagtggactggggagtgtgagcaccctggggcagatagatg
35 gctggagcaccactcagctccgcattgtgggtctccagttttctacggcagagtgggtcggcatgtgagcc
acctggacttctgttcatctgacagcgtgggttatactctctgtggactgcggccagaggagctccagc
acatcagcagttgggagttcagccaagcagctctcttctcctggcaccctgcattctccagtgtctgagg
aacaactggagggttctggcccacctacttgtactgcctgggtgggtttggccaatcagtaactgggggc
ctgagatcttactgaaattggcaccatagcagctgggatcccagacctggctctttcagcactgctgc
40 ggggacagatccaggggcttactcctcttgcatttctgtcatccctcctcctaaatttctgtgtggtgt
ttagtcccatccaactatctagtctcaccagtgtcaggctgtggctgtcactcctgagcaaatggcct

ttctgagtcctgagcagcgacgagcagttgcatgggccaacatgagggaaaggagagcccagaacagc
 aagggtcgaagtacagcctggggcctccaggactggtcacgaccttcctgggtccctgggtattgactatca
 gcttccttgccacctgctatgagcctgtctctacagtagaaggagattgtggggagagaaatcttaag
 tcataatgaataaagtgcacaacagaagtgcacccctgattatcttcagaagctgatgaggaataactagt
 5 gctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgccccctccccgtgccttcct
 tgaccctggaaggtgccactcccactgtcctttcctaataaaaatgaggaaattgcatcgcattgtctga
 gtaggtgtcattctattctgggggtgggggtggggcaggacagcaagggggaggattgggaagacaata
 gcaggcatgctggggatgcggtgggctctatggcaattcagactcccactagaagaaacagaacttgaa
 aaaaggataattgtggtgaacacctcaaccagtggtccaaaaactgaaaaactttgaccccgagcacc
 10 ctcacacagatagactactgaaagaaggagaaaggggcaacttactcagtctcccttatcagattgcctt
 acgaggagtactagactccctcaagcaaatagatctcacttaccacttgcaaaggtataactactttcac
 tctattagacctatatactgaccagggtcctattccaagacaactcttctactcttccagcagactct
 tatagaaagaaagattctgggggtccaagaaagcagtagctttcttacaaggagccaaaaaaaataacctt
 tcttttagcacttctaacttgagtgaaactgggtgatcaaagagaggttggctccctggggacaagtgcc
 15 acaaatcagtcaactacaagaaagttgagaacactgttctcccgaaaccaagcttgaattcagctgac
 gtgcctcggaccgctaggaacccctagtgatggagttggccactccctctctgcgcgctcgtcgtca
 ctgaggccggggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgag
 cgcgagctgcctgcagg

20 pITR-202GFP (SEQ ID NO: 14)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccgggctcggggcgacctttgggtcgcccg
 gcctcagtgagcgagcgagcgcgagagaggagtgccaaactccatcactaggggttcctgcggccgc
 acgcgtgggattttgccgatttcggcctattgggttaaaaaatgagctgatttaacaaaaatttaacgcg
 aattttaacaaaaatgataggcacctattgggtcttactgacatccactttgcctttctctccacaggcgt
 25 acctgcatatgccccagaaaacctccagcagctgggtgctttcagcagagagggaggctgcacagggtc
 tcttgacactcatgctgcaggggaagctgcaggggaagctgcaggtaccaccatccgaggagcaggccc
 tgggtcgctgacagccctgctgctccagcggtaccacgcctcacctcccagctcttcattgacctgt
 caccactcatccctttcttgggtgtctctgacctgatgcgcttcccaccatccctgttagccaacgaca
 gtgtcctgggtgccatccgggattacagcccaggaatgaggcctgaacagaaggaggctctggcaaagc
 30 gactgctggccctgaactgtttggggaagtgcctgcctggccccaggagctgctgtgggcagtgtgc
 cctgctccccacctccctctggagaactttttgcagctcagccctcaccagatccaggccctggagg
 atagctggccagcagcaggtctggggccaggcatgcccggcatgtgctgcgcagcctggtaaaccaga
 gtgtccaggatggtgaggagcaggtacgcaggcttggggccctcgctgtttcctgagccctgaggagc
 tgcagagcctagtggccctgagtgatccaacggggccagtagaacgggggctgctggaatgtgcagcca
 35 atggggacctcagcccagaaggacgggtggcatatgaacttctgggtgtgttgcgctcatctggaggag
 cgggtgctgagccccgggagctgcgggtctggggccctctcttctctcagctgggcctccgcttctctc
 aggagctgtcagagccccagcttagagccatgcttcctgtcctccagggaactagtgttacacctgtc
 aggtgtcctgctgcttggacggctccttcttaggcagcatctatccctggaggaaactctgctccttgc
 accttctgctaccaggcctcagccccagacactccaggccatccctaggcgagtcctgggtcggggctt
 40 gttcctgcctggccccctgaactgtcacgcctctcagcctgccagaccgcagcactgctgcagaccttctc

gggttaaagatggtgttaaaaatatgggtacaacaggtgctggtccagctgtgtgtatccctggtcagc
ctattcccaccacctggccagactgctgtctccctgctccattaaagctgtacaactggattcct
tggctcttctggcaaactcgaagacgctactgggagctgcccgtgtgagcagcaggcacagtttctct
ggaagaagatgcaagtacccaccaaccttacccctcaggaatctgcaggctctgggcacctggcaggag
5 gcatgtcctgtgagtttctgcagcagatcaactccatggtagacttccttgaagtgggtgcacatgatct
atcagctgcccactagagttcagaggagcctgagggcctgtatctgggcagagctacagcggaggatgg
caatgccagaaccagaatggacaactgtagggccagaactgaacgggctggatagcaagctactcctgg
acttacccgatccagttgatggacagactatccaatgaatccattatgttggtggtggagctgggtgcaaa
gagctccagagcagctgctggcactgacccccctccaccaggcagccctggcagagagggcactacaaa
10 acctggctccaaaggagactccagttctcaggggaagtgtctggagaccttaggccctttggttgattcc
tggggacagagagcacacgacagatccccctacagatcctgctgtcccatctcagtcagctgcaaggct
tctgcctagggagagacatttgccacagagctgggatggctgctattgcaggagtctgttcttgggaaac
cagagttgtggagccaggatgaagtagagcaagctggacgcctagtattcactctgtctactgaggcaa
tttccttgatccccaggaggccttgggtccagagacctggagcggcttctagaaaagcagcagagct
15 gggagcagagcagagttggacagctgtgtagggagccacagcttgcctgccaagaaagcagccctggtag
caggggtggtgcgaccagctgctgaggatcttcagaacctgtgccaattgtgcagatgtacgaggga
cattcccagcagcctggtctgcaaccagattgcagagatggagctctcagactttgaggactgctga
cattatttgaggagaccaggaacttgggcctgaggaaactgcgggcagccatgggcaaagcaaacagt
tgtggggtcccccccgggatttcgtcctgagcagatcctgcagcttggtaggctcttaataggtctag
20 gagatcgggaactacaggagctgatcctagtggactggggagtgtgtagcaccctggggcagatagatg
gctggagcaccactcagctccgcattgtggtctccagtttctacggcagagtggctcgcatgtgagcc
acctggacttcgttcatctgacagcgtgggttatactctctgtggactgcggccagaggagctccagc
acatcagcagttgggagttcagccaagcagctctcttccctcgccaccctgcactctccagtgtctgagg
aacaactggaggttctggcccacctacttgtaactgcctgggtgggtttggccaatcagtaactgggggc
25 ctgagatcttactgaaattggcaccatagcagctgggatcccagacctggctctttcagcactgctgc
ggggacagatccaggcggttactcctcttgccatttctgtcatccctcctcctaaatttgctgtggtgt
ttagtcccatccaactatctagtctcaccagtgtcaggctgtggctgtcactcctgagcaaatggcct
ttctgagtcctgagcagcagcagcagttgcatgggccaacatgagggaaaggagagccagaacagc
aaggctgaagtacagcctggggcctccaggactggtcacgaccttccctggtccctggtattgactatca
30 gcttcccttgccacctgctaggtccggagagggcagaggaaagtctgctaacatgcggtgacgtcgagg
agaatcctggccaatggagagcgacgagagcggcctgcccgccatggagatcgagtgccgcatcacgg
gcacctgaacggcggtggagtccgagctgggtggcgggcggagagggcaccctcgagcaggggccgcatga
ccaacaagatgaagagcaccaaaggcgccctgaccttcagcccctacctgctgagccacgtgatgggt
acggcttctaccacttcggcacctaccccagcggctacgagaaccccttctgacgccatcaacaacg
35 gcggctacaccaacaccccgcatcgagaagtacgaggacggcggtgctgacgtgagcttcagctacc
gctacgaggccggccgctgatcggcgacttcaagggtgatgggcaccggcttccccgaggacagcgtga
tcttcaccgacaagatcatccgcagcaacgccaccgtggagcacctgcaccccatgggcgataacgatc
tggatggcagcttcacccgcaccttcagcctgcgcgacggcggtactacagctccgtgggtggacagcc
acatgcacttcaagagcggccatccaccccagcatcctgcagaacggggggcccatgttcgccttccgcc
40 gcgtggaggaggatcacagcaacaccgagctgggcatcgtggagtaccagcacgccttcaagacccgg
atgcagatgccggtgaagaataagcctgtctctacagtagaaggagattgtggggagagaaatcttaag

5

10

15

20

25

30

35

40

tcactccttgccaggatgaggtcttctccatTTTTGGGGCAGCCAGAGCCTGATACCAATGGGCAGTGCCA
GGGAGGTAACCTTCAACAGCTGCTCTTATGGGGCGTCCGGCACAACCTTCTGGGATGTCCAGGCGCT
GGGCTTCTGTCTGGATCACCACCCCCACCCCTGCCCTCCTTCACTGCCTGAGCACGGGCGTGCCCTCT
GCCAGAGCTTCTCAGCCGTGAGCCACATCAGCCACGCCAACGGCGAGCCATCACTGTGGAGGCCCT
5 CTGTGAGAACCCTTAGGCCAGCACCACCCTACAGCATTTCACACTTCTCCATCCACTTGCTCTGCCA
GCACACCAAGCCTGCCACTCCACAGCCCCATCCAGCACCCTGCCATCTGCCAGACAGCTGTGTGGTA
TGCAGTGTCTGGGCACCAGGTGCCAAGGCTGGCTACAGGCCTGCCACGACCAGTTTCTGATGAGTT
TTTGGATGCGATCTGCAGTAACCTCTCCTTTTCAAGCCCTGTCTGGCTCCAACCGCCGCTGGTGAAGCG
GCTCTGTGCTGGCCTGCTCCCCCCCCCTACCAGCTGCCCTGAAGGCTGCCCCCTGTTCCCTCACCCC
10 AGACATCTTTTGGGGCTGCTTCTTGGAGAATGAGACTCTGTGGGCTGAGCGACTGTGTGGGAGGCAAG
TCTACAGGCTGTGCCCCCAGCAACCAGGCTTGGGTCCAGCATGTGTGCCAGGGCCCCCCCCAGATGT
CACTGCCTCCCCACCATGCCACATTGGACCCTGTGGGGAACGCTGCCCGGATGGGGGCAGCTTCTGGT
GATGGTCTGTGCCAATGACACCATGTATGAGGTCTGGTGCCTTCTGGCCTTGGCTAGCAGGCCAATG
CAGGATAAGTCGTGGGGCAATGACACTTGCTTCTAGAAGGGCTGTGGGCCCCCTTCTGCCTCTCT
15 GCCACCACTGGGACCATCCCCACTCTGTCTGACCCTGGCCCCCTCCTCCTTGGCATGCTATCCCAGTT
GCCACGCTGTCACTCCTCTGTCCAGCTCTTGTCCACCCACACGCCTACACTATCTCCTCCGCTGCT
GACCTTCTCTTGGGTCCAGGGGCTGGGGGCGCTGAGGCCAGGGGATGCTGGGTGGGCCCCCTACTGCT
CTCCAGTCTCCAGACAACCTGCTCCTTCTGGGATGCCTTTCGCCAGAGGGCCGGCGCAGTGTGCTACG
GACGATTGGGGAATACCTGGAACAAGATGAGGAGCAGCCAACCCCATCAGGCTTTGAACCCACTGTCAA
20 CCCCAGCTCTGGTATAAGCAAGATGGAGCTGCTGGCCTGCTTTAGTCCTGTGCTGTGGGATCTGCTCCA
GAGGGAAAAGAGTGTTTGGGCCCTGCAGATTCTAGTGCAGGTAAGTATCAAGGTTACAAGACAGGTTTA
AGGAGACCAATAGAACTGGGCTTGTGCAGACAGAGAAGACTCTTGCGTTTCTCTAGGTGGAGGCCGAA
AGTACATGTTTGCATGGGAACCCAGACCCTGAGTACCAGATGACTACAGCCAAGGTGGGACCAGGC
TGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCGAAGCGCCAGGGTGCCCGGTATGTGTGGAACCGCA
25 CTGAGCTCATGCAGGCTTCCCTGGACCCGTCTGTGACCCATCTCATGGGTCTCTTTGAGCCTGGAGACA
TGAAATACGAGATCCACCGAGACTCCACACTGGACCCCTCCTGATGGAGATGACAGAGGCTGCCTGC
GCCTGCTGAGCAGGAACCCCCGCGGCTTCTTCTCTCTCGTGGAGGGTGGTCGCATCGACCATGGTCATC
ATGAAAGCAGGGCTTACCAGGGCACTGACTGAGACGATCATGTTGACGACGCCATTGAGAGGGCGGGCC
AGCTCACCAGCGAGGAGGACACGCTGAGCCTCGTCACTGCCGACCACTCCCAGCTTCTCTCTCTCGGAG
30 GCTACCCCTGCGAGGGAGCTCCATCTTCGGGTGGCCCCCTGGCAAGGCCCGGACAGGAAGGCCTACA
CGGTCTCTCTATACGGAAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGCCGGATGTTACCGAGA
GCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCTGGACGAAGAGACCCACGCAGGCG
AGGACGTGGCGGTGTTGCGCGCGGGCCCGCAGGCGCACCTGGTTCACGGCGTGCAAGGAGCAGACCTTCA
TAGCGCACTCATGGCTTCTCGCGCCTGCCTGGAGCCCTACACCGCTGCGACCTGGCGCCCCCGCCG
35 GCACCAACCGACGCGCGCACCCGGGGCGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGGAACC
CCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGT
CGCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

pITR-204 (SEQ ID NO: 16)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggggtcgggcgacctttggtcgcccg
 gcctcagtgagcgagcgagcgcgagagaggagtgccaaactccatcactaggggttctctgcgccgc
 acgcgctctaggtggaggccgaaagtacatgtttcgcatgggaaccccagaccctgagtaccagatgac
 5 tacagccaaggtgggaccaggctggacgggaagaatctggtgcaggaatggctggcgaagcgccaggg
 gcccggtatgtgtggaaccgcactgagctcatgcaggcttccctggaccgctctgtgacccatctcatg
 ggtctcttttgagcctggagacatgaaatacgagatccaccgagactccacactggaccctccctgatg
 gagatgacagaggctgccctgcgctgctgagcaggaacccccgcggcttcttctctctctgagggt
 ggtcgcatcgaccatggtcatcatgaaagcagggttaccgggcaactgactgagacgatcatgttcgac
 10 gacgccattgagagggcgggccagctcaccagcgaggaggacacgctgagcctcgctcactgcccaccac
 tcccacgtcttctctctcggaggctacccctgcgaggagctccatcttcgggctggcccctggcaag
 gcccgggacaggaaggcctacacggtcctcctatacggaaacgggtccaggctatgtgctcaaggacggc
 gcccgccggatgttaccgagagcgagagcgggagccccgagtatcggcagcagtcagcagtgcccctg
 gacgaagagacccacgcaggcgaggacgtggcggtgttcgcgcgcgcccgaggcgccacctggttcac
 15 ggcgtgcaggagcagaccttcatagcgacgtcatggccttcgcccgtgctggagccctacaccgcc
 tgcgacctggcgccccccgcccgcaccaccgacgcccgcacccggggcggtataggcacctattggtct
 tactgacatccactttgctttctctccacaggcgtaacctgcataatgccccagaaaacctccagcagc
 tgggtgctttcagcagagaggaggctgcacagggttctctgacactcatgctgcaggggaagctgcagg
 ggaagctgcaggataccaccatccgaggagcaggccctgggtcgctgacagccctgctgctccagcgg
 20 acccagcctcacctcccagctcttcatgacctgtcaccactcatcccttcttgggtgtctctgacc
 tgatgcgcttcccaccatccctgttagccaacgacagtgctcctggctgccatccgggattacagcccag
 gaatgaggcctgaacagaaggaggctctggcaaagcgactgctggcccctgaactgtttggggaagtgc
 ctgcttggccccaggagctgctgtgggcagtgtgcccctgctccccacctccctctggagaactttt
 tgcagctcagccctcaccagatccaggccctggaggatagctggccagcagcaggtctggggccagggc
 25 atgcccgcctatgtgctgcgagcctggtaaacagagtgctccaggatggtgaggagcaggtacgcaggc
 ttggggccctcgctgtttcctgagccctgaggagctgcagagcctagtgtcccctgagtgtccaacgg
 ggccagtagaacgggggctgctggaatgtgcagccaatgggaccctcagcccagaaggacgggtggcat
 atgaacttctgggtgtgttgcgctcatctggaggagcgggtgctgagccccgggagctgcggtctggg
 cccctctcttctctcagctgggcctccgcttcccttcaggagctgtcagagccccagcttagagccatgc
 30 ttctgtcctccagggaactagtgttacacctgctcaggctgtcctgctgcttgacgggtccttccta
 ggcaogatctatccctggaggaactctgctccttgacaccttctgctaccaggcctcagccccagacac
 tccaggccatccctaggcgagtcctggtcggggcttgttctgctgctggccctgaactgtcacgcctct
 cagcctgccagaccgcagcactgctgcagacctttcgggttaaagatggtgttaaaaatatgggtacaa
 cagggtgctgggtccagctgtgtgtatccctggtcagcctattcccaccacctggccagactgctgcttc
 35 cctgctccattaaagctgctacaactggattccttggctcttctggcaaatcgaagacgctactggg
 agctgccctgggtctgagcagcaggcacagtttctctggaagaagatgcaagtaccaccaaccttacc
 tcaggaatctgcaggctctgggcacccctggcaggaggcatgtcctgtgagtttctgcagcagatcaact
 ccatggtagacttccctgaagtgggtgcacatgatctatcagctgccactagagttcgaggggagcctga
 gggcctgtatctgggcagagctacagcggaggatggcaatgccagaaccagaatggacaactgtagggc
 40 cagaactgaacgggctggatagcaagctactcctggacttaccgatccagttgatggacagactatcca
 atgaatccattatgttgggtgggagctgggtgcaaagagctccagagcagctgctggcactgaccccc

tccaccaggcagccctggcagagagggcactacaaaacctggctccaaaggagactccagtctcagggg
 aagtgtctggagaccttaggcccttgggttgattcctggggacagagagcacacgacagatccccctac
 agatcctgtctgtcccatctcagtcagctgcaaggcttctgcctaggagagacatttgccacagagctgg
 gatggctgtctattgcaggagctctgttcttgggaaaccagagttgtggagccaggatgaagtagagcaag
 5 ctggacgcctagtagtattcactctgtctactgaggcaatttccttgatccccagggaggccttgggtccag
 agaccctggagcggcttctagaaaagcagcagagctgggagcagagcagagttggacagctgtgtaggg
 agccacagcttgtgtccaagaaagcagccctggtagcaggggtggtgcgaccagctgtgtaggatcttc
 cagaacctgtgccaaattgtgcagatgtacgagggacattcccagcagcctggtctgcaaccagattg
 cagagatggagctctcagactttgaggactgcctgacattatttgaggagaccaggacttgggcctg
 10 aggaactgcgggcagccatgggcaaagcaaacagttgtgggggtccccccggggatttcgtcctgagc
 agatcctgcagcttggtaggctcttaataaggctcaggagatcgggaaactacaggagctgatcctagtgg
 actggggagtgctgagcaccctggggcagatagatggctggagcaccactcagctccgcattgtggtct
 ccagtttctctacggcagagtgggtcgcatgtgagccactggacttcgttcatctgacagcgtgggtt
 atactctctgtggactgcggccagaggagctccagcacatcagcagttgggagttcagccaagcagctc
 15 tcttcctcggcaccctgcactctccagtgtctgaggaacaactggaggttctggccacactactgttac
 tgcctggtgggttttggcccaatcagtaactgggggcctgagatcttcactgaaattggcaccatagcag
 ctgggatcccagacctggctctttcagcactgtctgcggggacagatccagggcggttactcctcttgcca
 tttctgtcatccctcctcctaaatttgctgtggtgttttagtcccatccaactatctagtctcaccagtg
 ctccaggtgtggctgtcactcctgagcaaattggcctttctgagtcctgagcagcgacgagcagttgcat
 20 gggcccaacatgaggggaaaggagagcccagaacagcaagggtcgaagtacagcctggggcctccaggact
 ggtcacgaccttcctgggtccctgggtattgactatcagcttccttggccacctgctatgagcctgtctct
 acagtagaaggagattgtggggagagaaatcttaagtataatgaataaagtgcacacagaagtgcac
 ctgattattttcagaagctgatgaggaataactagtgtctgatcagcctcgactgtgccttctagtggcc
 agccatctgttgtttgccccctccccgtgccttccttgaccctggaagggtgccactcccactgtccttt
 25 cctaataaaaatgaggaaattgcacgcattgtctgagtaggtgtcattctattctgggggggtgggtg
 ggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggtctatgg
 aagcttgaattcagctgacgtgcctcggaaccgctaggaacccctagtgtgagttggccactccctct
 ctgcgcgctcgctcgctcactgaggccgggcgaccaaagggtcgcccgacgcccgggctttgcccggg
 gcctcagtgtgagcgagcgagcgcgagctgcctgcagg

30

pITR-205 (SEQ ID NO: 17)

cctgcaggcagctgcgcgctcgctcgctcactgaggccgcccggggtcgggcgacctttggtcgcccc
 gcctcagtgtgagcgagcgagcgcgagagagggagtgccaaactccatcactaggggttcctgcggccgc
 acgcgtgacattgattattgactagttattaatagtaataatcaattacgggggtcattagttcatagccat
 35 atatggagttccgcgttacataacttacggtaaatggcccgctggctgaccgccaacgacccccgcc
 cattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatggg
 tggactatttacggtaaaactgcccacttggcagtagcatcaagtgtatcatatgccaaagtacgcccccta
 ttgacgtcaatgacggtaaatggcccgctggcattatgccagtagcatgaccttatgggactttccta
 cttggcagtagcatctacgtatttagtcatcgctattaccatgggtcgaggtgagccccacgttctgtcttc
 40 actctccccatctccccccccctcccccccccaattttgtattttattttttaattttttgtgca

95

cccacccccagatgtcactgcctccccaccatgccacattggaccctgtggggaacgctgcccggatgg
 gggcagcttctctggtgatggtctgtgccaatgacaccatgtatgaggtcctggtgcccttctggccttg
 gctagcaggccaatgcaggataagtcgtgggggcaatgacacttgcttcctagaagggctgctggggcc
 ccttctgccctctctgccaccactgggaccatccccactctgtctgaccctggccccttctccttg
 5 catgctatcccagttgccacgctgtcagtcctctgtcccagctcttgctcaccacacgcctacacta
 tctcctccgctgctgaccttctcttgggtccagggctggggcgctgaggcccaggggatgctggg
 tcgggcccactactgctctccagtctcccagacaactgctccttctgggatgcctttcgcccagagggccg
 gcgcagtgtgctacggacgattggggaatacctggaacaagatgaggagcagccaaccccatcaggctt
 tgaaccactgtcaaccccagctctggtataagcaagatggagctgctggcctgctttagtctgtgct
 10 gtgggatctgctccagagggaaaagagtgtttgggcccctgcagattctagtgcaggttaagtatcaaggt
 tacaagacaggtttaaggagaccaatagaaaactgggcttgcgagacagagaagactcttgctttctg
 ggattttgcccatttccgctatttggttaaaaaatgagctgatttaacaaaaatttaacgcgaatttta
 aaaaaataagcttgaattcagctgacgtgcctcggaccgctaggaacccctagtgtatggagttggccac
 tcctctctgctgcgcgtcgtcgtcactgaggccggcgaccaaaggctcgcccgacgcccgggctttgc
 15 ccggggcgccctcagtgagcgagcgagcgcgagctgcctgcagg

In some embodiments, the vector(s) is an adenovirus (see, e.g., Dmitriev et al.
 (1998) *J. Virol.* 72: 9706-9713; and Poulin et al., *J. Virol* 8: 10074-10086, 2010). In
 some embodiments, the vector(s) is a retrovirus (see, e.g., Maier et al. (2010) *Future*
 20 *Microbiol* 5: 1507-23).

In some embodiments, the vector(s) is a lentivirus (see, e.g., Matrai et al.
 (2010) *Mol Ther.* 18: 477-490; Banasik et al. (2010) *Gene Ther.* 17:150-7; and
 Wanisch et al. (2009) *Mol. Ther.* 17: 1316-32). A lentiviral vector refers to a vector
 derived from at least a portion of a lentivirus genome, including especially a self-
 25 inactivating lentiviral vector as provided in Milone et al., *Mol. Ther.* 17(8): 1453-
 1464 (2009). Non-limiting lentivirus vectors that may be used in the clinic include
 the LENTIVECTOR® gene delivery technology from Oxford BioMedica, the
 LENTIMAX™ vector system from Lentigen, and the like. Other types of lentiviral
 vectors are also available and would be known to one skilled in the art.

30 The vectors provided herein can be of different sizes. The choice of vector that
 is used in any of the compositions, kits, and methods described herein may depend on
 the size of the vector.

In some embodiments, the vector(s) is a plasmid and can include a total length
 of up to about 1 kb, up to about 2 kb, up to about 3 kb, up to about 4 kb, up to about 5
 35 kb, up to about 6 kb, up to about 7 kb, up to about 8kb, up to about 9 kb, up to about
 10 kb, up to about 11 kb, up to about 12 kb, up to about 13 kb, up to about 14 kb, or

up to about 15 kb. In some embodiments, the vector(s) is a plasmid and can have a total length in a range of about 1 kb to about 2 kb, about 1 kb to about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about 1 kb to about 6 kb, about 1 kb to about 7 kb, about 1 kb to about 8 kb, about 1 kb to about 9 kb, about 1 kb to about 10 kb, about 1 kb to about 11 kb, about 1 kb to about 12 kb, about 1 kb to about 13 kb, about 1 kb to about 14 kb, or about 1 kb to about 15 kb.

In some embodiments, the vector(s) is a transposon (e.g., PiggyBac transposon) and can include greater than 200 kb. In some examples, the vector(s) is a transposon having a total length in the range of about 1 kb to about 10 kb, about 1 kb to about 20 kb, about 1 kb to about 30 kb, about 1 kb to about 40 kb, about 1 kb to about 50 kb, about 1 kb to about 60 kb, about 1 kb to about 70 kb, about 1 kb to about 80 kb, about 1 kb to about 90 kb, about 10 kb to about 20 kb, about 10 kb to about 30 kb, about 10 kb to about 40 kb, about 10 kb to about 50 kb, about 10 kb to about 60 kb, about 10 kb to about 70 kb, about 10 kb to about 90 kb, about 10 kb to about 100 kb, about 20 kb to about 30 kb, about 20 kb to about 40 kb, about 20 kb to about 50 kb, about 20 kb to about 60 kb, about 20 kb to about 70 kb, about 20 kb to about 80 kb, about 20 kb to about 90 kb, about 20 kb to about 100 kb, about 30 kb to about 40 kb, about 30 kb to about 50 kb, about 30 kb to about 60 kb, about 30 kb to about 70 kb, about 30 kb to about 80 kb, about 30 kb to about 90 kb, about 30 kb to about 100 kb, about 40 kb to about 50 kb, about 40 kb to about 60 kb, about 40 kb to about 70 kb, about 40 kb to about 80 kb, about 40 kb to about 90 kb, about 40 kb to about 100 kb, about 50 kb to about 60 kb, about 50 kb to about 70 kb, about 50 kb to about 80 kb, about 50 kb to about 90 kb, about 50 kb to about 100 kb, about 60 kb to about 70 kb, about 60 kb to about 80 kb, about 60 kb to about 90 kb, about 60 kb to about 100 kb, about 70 kb to about 80 kb, about 70 kb to about 90 kb, about 70 kb to about 100 kb, about 80 kb to about 90 kb, about 80 kb to about 100 kb, about 90 kb to about 100 kb, about 1 kb to about 100 kb, about 100 kb to about 200 kb, about 100 kb to about 300 kb, about 100 kb to about 400 kb, or about 100 kb to about 500 kb.

In some embodiments, the vector is a cosmid and can have a total length of up to 55 kb. In some examples, the vector is a cosmid and has a total number of nucleotides of about 1 kb to about 10 kb, about 1 kb to about 20 kb, about 1 kb to about 30 kb, about 1 kb to about 40 kb, about 1 kb to about 50 kb, about 1 kb to about 55 kb, about 10 kb to about 20 kb, about 10 kb to about 30 kb, about 10 kb to about 40 kb, about 10 kb to about 50 kb, about 10 kb to about 55 kb, about 15 kb to about

55 kb, about 15 kb to about 50 kb, about 15 kb to about 40 kb, about 15 kb to about 30 kb, about 15 kb to about 20 kb, about 20 kb to about 55 kb, about 20 kb to about 50 kb, about 20 kb to about 40 kb, about 20 kb to about 30 kb, about 25 kb to about 55 kb, about 25 kb to about 50 kb, about 25 kb to about 40 kb, about 25 kb to about 30 kb, about 30 kb to about 55 kb, about 30 kb to about 50 kb, about 30 kb to about 40 kb, about 35 kb to about 55 kb, about 40 kb to about 55 kb, about 40 kb to about 50 kb, or about 45 kb to about 55 kb.

In some embodiments, the vector(s) is an artificial chromosome and can have a total number of nucleotides of about 100 kb to about 2000 kb. In some
10 embodiments, the artificial chromosome(s) is a human artificial chromosome (HAC) and can have a total number of nucleotides in the range of about 1 kb to about 10 kb, 1 kb to about 20 kb, about 1 kb to about 30 kb, about 1 kb to about 40 kb, about 1 kb to about 50 kb, about 1 kb to about 60 kb, about 10 kb to about 20 kb, about 10 kb to about 30 kb, about 10 kb to about 40 kb, about 10 kb to about 50 kb, about 10 kb to
15 about 60 kb, about 20 kb to about 30 kb, about 20 kb to about 40 kb, about 20 kb to about 50 kb, about 20 kb to about 60 kb, about 30 kb to about 40 kb, about 30 kb to about 50 kb, about 30 kb to about 60 kb, about 40 kb to about 50 kb, about 40 kb to about 60 kb, or about 50 kb to about 60 kb.

In some embodiments, the artificial chromosome(s) is a yeast artificial
20 chromosome (YAC) and can have a total number of nucleotides up to 1000 kb. In some embodiments, the artificial chromosome(s) is a YAC having a total number of nucleotides in the range of about 100 kb to about 1,000 kb, about 100 kb to about 900 kb, about 100 kb to about 800 kb, about 100 kb to about 700 kb, about 100 kb to about 600 kb, about 100 kb to about 500 kb, about 100 kb to about 400 kb, about 100
25 kb to about 300 kb, about 100 kb to about 200 kb, about 200 kb to about 1,000 kb, about 200 kb to about 900 kb, about 200 kb to about 800 kb, about 200 kb to about 700 kb, about 200 kb to about 600 kb, about 200 kb to about 500 kb, about 200 kb to about 400 kb, about 200 kb to about 300 kb, about 300 kb to about 1,000 kb, about 300 kb to about 900 kb, about 300 kb to about 800 kb, about 300 kb to about 700 kb,
30 about 300 kb to about 600 kb, about 300 kb to about 500 kb, about 300 kb to about 400 kb, about 400 kb to about 1,000 kb, about 400 kb to about 900 kb, about 400 kb to about 800 kb, about 400 kb to about 700 kb, about 400 kb to about 600 kb, about 400 kb to about 500 kb, about 500 kb to about 1,000 kb, about 500 kb to about 900 kb, about 500 kb to about 800 kb, about 500 kb to about 700 kb, about 500 kb to

about 600 kb, about 600 kb to about 1,000 kb, about 600 kb to about 900 kb, about 600 kb to about 800 kb, about 600 kb to about 700 kb, about 700 kb to about 1,000 kb, about 700 kb to about 900 kb, about 700 kb to about 800 kb, about 800 kb to about 1,000 kb, about 800 kb to about 900 kb, or about 900 kb to about 1,000 kb.

5 In some embodiments, the artificial chromosome(s) is a bacterial artificial chromosome (BAC) and can have a total number of nucleotides of up to 750 kb. In some embodiments, the artificial chromosome(s) is a BAC and can have a total number of nucleotides in the range of about 100 kb to about 750 kb, about 100 kb to about 700 kb, about 100 kb to about 600 kb, about 100 kb to about 500 kb, about 100
10 kb to about 400 kb, about 100 kb to about 300 kb, about 100 kb to about 200 kb, about 150 kb to about 750 kb, about 150 kb to about 700 kb, about 150 kb to about 600 kb, about 150 kb to about 500 kb, about 150 kb to about 400 kb, about 150 kb to about 300 kb, about 150 kb to about 200 kb, about 200 kb to about 750 kb, about 200 kb to about 700 kb, about 200 kb to about 600 kb, about 200 kb to about 500 kb,
15 about 200 kb to about 400 kb, about 200 kb to about 300 kb, about 250 kb to about 750 kb, about 250 kb to about 700 kb, about 250 kb to about 600 kb, about 250 kb to about 500 kb, about 250 kb to about 400 kb, about 250 kb to about 300 kb, about 300 kb to about 750 kb, about 300 kb to about 700 kb, about 300 kb to about 600 kb, about 300 kb to about 500 kb, about 300 kb to about 400 kb, about 350 kb to about
20 750 kb, about 350 kb to about 700 kb, about 350 kb to about 600 kb, about 350 kb to about 500 kb, about 350 kb to about 400 kb, about 400 kb to about 750 kb, about 400 kb to about 700 kb, about 450 kb to about 600 kb, about 450 kb to about 500 kb, about 500 kb to about 750 kb, about 500 kb to about 700 kb, about 500 kb to about 600 kb, about 550 kb to about 750 kb, about 550 kb to about 700 kb, about 550 kb to
25 about 600 kb, about 600 kb to about 750 kb, about 600 kb to about 700 kb, or about 650 kb to about 750 kb.

In some embodiments, the artificial chromosome(s) is a P1-derived artificial chromosome (PAC) and can have a total number of nucleotides of up to 300 kb. In some embodiments, the P1-derived artificial chromosome(s) can have a total number
30 of nucleotides in the range of about 100 kb to about 300 kb, about 100 kb to about 200 kb, or about 200 kb to about 300 kb.

In some embodiments, the vector(s) is a viral vector and can have a total number of nucleotides of up to 10 kb. In some embodiments, the viral vector(s) can have a total number of nucleotides in the range of about 1 kb to about 2 kb, 1 kb to

about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about 1 kb to about 6 kb, about 1 kb to about 7 kb, about 1 kb to about 8 kb, about 1 kb to about 9 kb, about 1 kb to about 10 kb, about 2 kb to about 3 kb, about 2 kb to about 4 kb, about 2 kb to about 5 kb, about 2 kb to about 6 kb, about 2 kb to about 7 kb, about 2 kb to about 8 kb, about 2 kb to about 9 kb, about 2 kb to about 10 kb, about 3 kb to about 4 kb, about 3 kb to about 5 kb, about 3 kb to about 6 kb, about 3 kb to about 7 kb, about 3 kb to about 8 kb, about 3 kb to about 9 kb, about 3 kb to about 10 kb, about 4 kb to about 5 kb, about 4 kb to about 6 kb, about 4 kb to about 7 kb, about 4 kb to about 8 kb, about 4 kb to about 9 kb, about 4 kb to about 10 kb, about 5 kb to about 6 kb, about 5 kb to about 7 kb, about 5 kb to about 8 kb, about 5 kb to about 9 kb, about 5 kb to about 10 kb, about 6 kb to about 7 kb, about 6 kb to about 8 kb, about 6 kb to about 9 kb, about 6 kb to about 10 kb, about 7 kb to about 8 kb, about 7 kb to about 9 kb, about 7 kb to about 10 kb, about 8 kb to about 9 kb, about 8 kb to about 10 kb, or about 9 kb to about 10 kb.

In some embodiments, the vector(s) is a lentivirus and can have a total number of nucleotides of up to 8 kb. In some examples, the lentivirus(es) can have a total number of nucleotides of about 1 kb to about 2 kb, about 1 kb to about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about 1 kb to about 6 kb, about 1 kb to about 7 kb, about 1 kb to about 8 kb, about 2 kb to about 3 kb, about 2 kb to about 4 kb, about 2 kb to about 5 kb, about 2 kb to about 6 kb, about 2 kb to about 7 kb, about 2 kb to about 8 kb, about 3 kb to about 4 kb, about 3 kb to about 5 kb, about 3 kb to about 6 kb, about 3 kb to about 7 kb, about 3 kb to about 8 kb, about 4 kb to about 5 kb, about 4 kb to about 6 kb, about 4 kb to about 7 kb, about 4 kb to about 8 kb, about 5 kb to about 6 kb, about 5 kb to about 7 kb, about 5 kb to about 8 kb, about 6 kb to about 8 kb, about 6 kb to about 7 kb, or about 7 kb to about 8 kb.

In some embodiments, the vector(s) is an adenovirus and can have a total number of nucleotides of up to 8 kb. In some embodiments, the adenovirus(es) can have a total number of nucleotides in the range of about 1 kb to about 2 kb, about 1 kb to about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about 1 kb to about 6 kb, about 1 kb to about 7 kb, about 1 kb to about 8 kb, about 2 kb to about 3 kb, about 2 kb to about 4 kb, about 2 kb to about 5 kb, about 2 kb to about 6 kb, about 2 kb to about 7 kb, about 2 kb to about 8 kb, about 3 kb to about 4 kb, about 3 kb to about 5 kb, about 3 kb to about 6 kb, about 3 kb to about 7 kb, about 3 kb to about 8 kb, about 4 kb to about 5 kb, about 4 kb to about 6 kb, about 4 kb to about 7 kb, about

4 kb to about 8 kb, about 5 kb to about 6 kb, about 5 kb to about 7 kb, about 5 kb to about 8 kb, about 6 kb to about 7 kb, about 6 kb to about 8 kb, or about 7 kb to about 8 kb.

In some embodiments, the vector(s) is an adeno-associated virus (AAV vector) and can include a total number of nucleotides of up to 5 kb. In some
5 embodiments, the AAV vector(s) can include a total number of nucleotides in the range of about 1 kb to about 2 kb, about 1 kb to about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about 2 kb to about 3 kb, about 2 kb to about 4 kb, about 2 kb to about 5 kb, about 3 kb to about 4 kb, about 3 kb to about 5 kb, or about 4 kb to
10 about 5 kb.

In some embodiments, the vector(s) is a Gateway® vector and can include a total number of nucleotides of up to 5 kb. In some embodiments, each Gateway®
vector(s) includes a total number of nucleotides in the range of about 1 kb to about 2 kb, about 1 kb to about 3 kb, about 1 kb to about 4 kb, about 1 kb to about 5 kb, about
15 2 kb to about 3 kb, about 2 kb to about 4 kb, about 2 kb to about 5 kb, about 3 kb to about 4 kb, about 3 kb to about 5 kb, or about 4 kb to about 5 kb.

In some embodiments of any of the compositions, kits, and methods provided herein, the at least two different vectors can be substantially the same type of vector and may differ in size. In some embodiments, the at least two different vectors can be
20 different types of vector, and may have substantially the same size or have different sizes.

In some embodiments, any of the at least two vectors can have a total number of nucleotides in the range of about 500 nucleotides to about 10,000 nucleotides, about 500 nucleotides to about 9,500 nucleotides, about 500 nucleotides to about
25 9,000 nucleotides, about 500 nucleotides to about 8,500 nucleotides, about 500 nucleotides to about 8,000 nucleotides, about 500 nucleotides to about 7,800 nucleotides, about 500 nucleotides to about 7,600 nucleotides, about 500 nucleotides to about 7,400 nucleotides, about 500 nucleotides to about 7,200 nucleotides, about 500 nucleotides to about 7,000 nucleotides, about 500 nucleotides to about 6,800
30 nucleotides, about 500 nucleotides to about 6,600 nucleotides, about 500 nucleotides to about 6,400 nucleotides, about 500 nucleotides to about 6,200 nucleotides, about 500 nucleotides to about 6,000 nucleotides, about 500 nucleotides to about 5,800 nucleotides, about 500 nucleotides to about 5,600 nucleotides, about 500 nucleotides to about 5,400 nucleotides, about 500 nucleotides to about 5,200 nucleotides, about

[illegible]

[illegible]

nucleotides, about 1,200 nucleotides to about 6,800 nucleotides, about 1,200
nucleotides to about 6,600 nucleotides, about 1,200 nucleotides to about 6,400
nucleotides, about 1,200 nucleotides to about 6,200 nucleotides, about 1,200
nucleotides to about 6,000 nucleotides, about 1,200 nucleotides to about 5,800
5 nucleotides, about 1,200 nucleotides to about 5,600 nucleotides, about 1,200
nucleotides to about 5,400 nucleotides, about 1,200 nucleotides to about 5,000
nucleotides, about 1,200 nucleotides to about 4,800 nucleotides, about 1,200
nucleotides to about 4,600 nucleotides, about 1,200 nucleotides to about 4,400
nucleotides, about 1,200 nucleotides to about 4,200 nucleotides, about 1,200
10 nucleotides to about 4,000 nucleotides, about 1,200 nucleotides to about 3,800
nucleotides, about 1,200 nucleotides to about 3,600 nucleotides, about 1,200
nucleotides to about 3,400 nucleotides, about 1,200 nucleotides to about 3,200
nucleotides, about 1,200 nucleotides to about 3,000 nucleotides, about 1,200
nucleotides to about 2,800 nucleotides, about 1,200 nucleotides to about 2,600
15 nucleotides, about 1,200 nucleotides to about 2,400 nucleotides, about 1,200
nucleotides to about 2,200 nucleotides, about 1,200 nucleotides to about 2,000
nucleotides, about 1,200 nucleotides to about 1,800 nucleotides, about 1,200
nucleotides to about 1,600 nucleotides, about 1,200 nucleotides to about 1,400
nucleotides, about 1,400 nucleotides to about 10,000 nucleotides, about 1,400
20 nucleotides to about 9,500 nucleotides, about 1,400 nucleotides to about 9,000
nucleotides, about 1,400 nucleotides to about 8,500 nucleotides, about 1,400
nucleotides to about 8,000 nucleotides, about 1,400 nucleotides to about 7,800
nucleotides, about 1,400 nucleotides to about 7,600 nucleotides, about 1,400
nucleotides to about 7,400 nucleotides, about 1,400 nucleotides to about 7,200
25 nucleotides, about 1,400 nucleotides to about 7,000 nucleotides, about 1,400
nucleotides to about 6,800 nucleotides, about 1,400 nucleotides to about 6,600
nucleotides, about 1,400 nucleotides to about 6,400 nucleotides, about 1,400
nucleotides to about 6,200 nucleotides, about 1,400 nucleotides to about 6,000
nucleotides, about 1,400 nucleotides to about 5,800 nucleotides, about 1,400
30 nucleotides to about 5,600 nucleotides, about 1,400 nucleotides to about 5,400
nucleotides, about 1,400 nucleotides to about 5,200 nucleotides, about 1,400
nucleotides to about 5,000 nucleotides, about 1,400 nucleotides to about 4,800
nucleotides, about 1,400 nucleotides to about 4,600 nucleotides, about 1,400
nucleotides to about 4,400 nucleotides, about 1,400 nucleotides to about 4,200

nucleotides, about 1,400 nucleotides to about 4,000 nucleotides, about 1,400
nucleotides to about 3,800 nucleotides, about 1,400 nucleotides to about 3,600
nucleotides, about 1,400 nucleotides to about 3,400 nucleotides, about 1,400
nucleotides to about 3,200 nucleotides, about 1,400 nucleotides to about 3,000
5 nucleotides, about 1,400 nucleotides to about 2,600 nucleotides, about 1,400
nucleotides to about 2,400 nucleotides, about 1,400 nucleotides to about 2,200
nucleotides, about 1,400 nucleotides to about 2,000 nucleotides, about 1,400
nucleotides to about 1,800 nucleotides, about 1,400 nucleotides to about 1,600
nucleotides, about 1,600 nucleotides to about 10,000 nucleotides, about 1,600
10 nucleotides to about 9,500 nucleotides, about 1,600 nucleotides to about 9,000
nucleotides, about 1,600 nucleotides to about 8,500 nucleotides, about 1,600
nucleotides to about 8,000 nucleotides, about 1,600 nucleotides to about 7,800
nucleotides, about 1,600 nucleotides to about 7,600 nucleotides, about 1,600
nucleotides to about 7,400 nucleotides, about 1,600 nucleotides to about 7,200
15 nucleotides, about 1,600 nucleotides to about 7,000 nucleotides, about 1,600
nucleotides to about 6,800 nucleotides, about 1,600 nucleotides to about 6,400
nucleotides, about 1,600 nucleotides to about 6,200 nucleotides, about 1,600
nucleotides to about 6,000 nucleotides, about 1,600 nucleotides to about 5,800
nucleotides, about 1,600 nucleotides to about 5,600 nucleotides, about 1,600
20 nucleotides to about 5,400 nucleotides, about 1,600 nucleotides to about 5,200
nucleotides, about 1,600 nucleotides to about 5,000 nucleotides, about 1,600
nucleotides to about 4,800 nucleotides, about 1,600 nucleotides to about 4,600
nucleotides, about 1,600 nucleotides to about 4,400 nucleotides, about 1,600
nucleotides to about 4,200 nucleotides, about 1,600 nucleotides to about 4,000
25 nucleotides, about 1,600 nucleotides to about 3,800 nucleotides, about 1,600
nucleotides to about 3,600 nucleotides, about 1,600 nucleotides to about 3,400
nucleotides, about 1,600 nucleotides to about 3,200 nucleotides, about 1,600
nucleotides to about 3,000 nucleotides, about 1,600 nucleotides to about 2,800
nucleotides, about 1,600 nucleotides to about 2,600 nucleotides, about 1,600
30 nucleotides to about 2,400 nucleotides, about 1,600 nucleotides to about 2,200
nucleotides, about 1,600 nucleotides to about 2,000 nucleotides, about 1,600
nucleotides to about 1,800 nucleotides, about 1,800 nucleotides to about 10,000
nucleotides, about 1,800 nucleotides to about 9,500 nucleotides, about 1,800
nucleotides to about 9,000 nucleotides, about 1,800 nucleotides to about 8,500

nucleotides, about 1,800 nucleotides to about 8,000 nucleotides, about 1,800
nucleotides to about 7,800 nucleotides, about 1,800 nucleotides to about 7,600
nucleotides, about 1,800 nucleotides to about 7,400 nucleotides, about 1,800
nucleotides to about 7,200 nucleotides, about 1,800 nucleotides to about 7,000
5 nucleotides, about 1,800 nucleotides to about 6,800 nucleotides, about 1,800
nucleotides to about 6,600 nucleotides, about 1,800 nucleotides to about 6,400
nucleotides, about 1,800 nucleotides to about 6,200 nucleotides, about 1,800
nucleotides to about 6,000 nucleotides, about 1,800 nucleotides to about 5,800
nucleotides, about 1,800 nucleotides to about 5,600 nucleotides, about 1,800
10 nucleotides to about 5,400 nucleotides, about 1,800 nucleotides to about 5,200
nucleotides, about 1,800 nucleotides to about 5,000 nucleotides, about 1,800
nucleotides to about 4,800 nucleotides, about 1,800 nucleotides to about 4,600
nucleotides, about 1,800 nucleotides to about 4,400 nucleotides, about 1,800
nucleotides to about 4,200 nucleotides, about 1,800 nucleotides to about 4,000
15 nucleotides, about 1,800 nucleotides to about 3,800 nucleotides, about 1,800
nucleotides to about 3,600 nucleotides, about 1,800 nucleotides to about 3,400
nucleotides, about 1,800 nucleotides to about 3,200 nucleotides, about 1,800
nucleotides to about 3,000 nucleotides, about 1,800 nucleotides to about 2,800
nucleotides, about 1,800 nucleotides to about 2,600 nucleotides, about 1,800
20 nucleotides to about 2,400 nucleotides, about 1,800 nucleotides to about 2,200
nucleotides, about 1,800 nucleotides to about 2,000 nucleotides, about 2,000
nucleotides to about 10,000 nucleotides, about 2,000 nucleotides to about 9,500
nucleotides, about 2,000 nucleotides to about 9,000 nucleotides, about 2,000
nucleotides to about 8,500 nucleotides, about 2,000 nucleotides to about 8,000
25 nucleotides, about 2,000 nucleotides to about 7,800 nucleotides, about 2,000
nucleotides to about 7,600 nucleotides, about 2,000 nucleotides to about 7,400
nucleotides, about 2,000 nucleotides to about 7,200 nucleotides, about 2,000
nucleotides to about 7,000 nucleotides, about 2,000 nucleotides to about 6,800
nucleotides, about 2,000 nucleotides to about 6,600 nucleotides, about 2,000
30 nucleotides to about 6,400 nucleotides, about 2,000 nucleotides to about 6,200
nucleotides, about 2,000 nucleotides to about 6,000 nucleotides, about 2,000
nucleotides to about 5,800 nucleotides, about 2,000 nucleotides to about 5,600
nucleotides, about 2,000 nucleotides to about 5,400 nucleotides, about 2,000
nucleotides to about 5,200 nucleotides, about 2,000 nucleotides to about 5,000

nucleotides, about 2,000 nucleotides to about 4,800 nucleotides, about 2,000
nucleotides to about 4,600 nucleotides, about 2,000 nucleotides to about 4,400
nucleotides, about 2,000 nucleotides to about 4,200 nucleotides, about 2,000
nucleotides to about 4,000 nucleotides, about 2,000 nucleotides to about 3,800
5 nucleotides, about 2,000 nucleotides to about 3,600 nucleotides, about 2,000
nucleotides to about 3,400 nucleotides, about 2,000 nucleotides to about 3,200
nucleotides, about 2,000 nucleotides to about 3,000 nucleotides, about 2,000
nucleotides to about 2,800 nucleotides, about 2,000 nucleotides to about 2,600
nucleotides, about 2,000 nucleotides to about 2,400 nucleotides, about 2,000
10 nucleotides to about 2,200 nucleotides, about 2,200 nucleotides to about 10,000
nucleotides, about 9,500 nucleotides, about 9,000 nucleotides, about 8,500
nucleotides, about 8,000 nucleotides, about 7,800 nucleotides, about 7,600
nucleotides, about 7,400 nucleotides, about 7,200 nucleotides, about 7,000
nucleotides, about 6,800 nucleotides, about 6,600 nucleotides, about 6,400
15 nucleotides, about 6,200 nucleotides, about 6,000 nucleotides, about 5,800
nucleotides, about 5,600 nucleotides, about 5,400 nucleotides, about 5,200
nucleotides, about 5,000 nucleotides, about 4,800 nucleotides, about 4,600
nucleotides, about 4,400 nucleotides, about 4,200 nucleotides, about 4,000
nucleotides, about 3,800 nucleotides, about 3,600 nucleotides, about 3,400
20 nucleotides, about 3,200 nucleotides, about 3,000 nucleotides, about 2,800
nucleotides, about 2,600 nucleotides, about 2,400 nucleotides, about 2,400
nucleotides to about 10,000 nucleotides, about 2,400 nucleotides to about 9,500
nucleotides, about 2,400 nucleotides to about 9,000 nucleotides, about 2,400
nucleotides to about 8,500 nucleotides, about 2,400 nucleotides to about 8,000
25 nucleotides, about 2,400 nucleotides to about 7,800 nucleotides, about 2,400
nucleotides to about 7,600 nucleotides, about 2,400 nucleotides to about 7,400
nucleotides, about 2,400 nucleotides to about 7,200 nucleotides, about 2,400
nucleotides to about 7,000 nucleotides, about 2,400 nucleotides to about 6,800
nucleotides, about 2,400 nucleotides to about 6,600 nucleotides, about 2,400
30 nucleotides to about 6,400 nucleotides, about 2,400 nucleotides to about 6,200
nucleotides, about 2,400 nucleotides to about 6,000 nucleotides, about 2,400
nucleotides to about 5,800 nucleotides, about 2,400 nucleotides to about 5,600
nucleotides, about 2,400 nucleotides to about 5,400 nucleotides, about 2,400
nucleotides to about 5,200 nucleotides, about 2,400 nucleotides to about 5,000

nucleotides, about 2,400 nucleotides to about 4,800 nucleotides, about 2,400
nucleotides to about 4,600 nucleotides, about 2,400 nucleotides to about 4,400
nucleotides, about 2,400 nucleotides to about 4,200 nucleotides, about 2,400
nucleotides to about 4,000 nucleotides, about 2,400 nucleotides to about 3,800
5 nucleotides, about 2,400 nucleotides to about 3,600 nucleotides, about 2,400
nucleotides to about 3,400 nucleotides, about 2,400 nucleotides to about 3,200
nucleotides, about 2,400 nucleotides to about 3,000 nucleotides, about 2,400
nucleotides to about 2,800 nucleotides, about 2,400 nucleotides to about 2,600
nucleotides, about 2,600 nucleotides to about 10,000 nucleotides, about 2,600
10 nucleotides to about 9,500 nucleotides, about 2,600 nucleotides to about 9,000
nucleotides, about 2,600 nucleotides to about 8,500 nucleotides, about 2,600
nucleotides to about 8,000 nucleotides, about 2,600 nucleotides to about 7,800
nucleotides, about 2,600 nucleotides to about 7,600 nucleotides, about 2,600
nucleotides to about 7,400 nucleotides, about 2,600 nucleotides to about 7,200
15 nucleotides, about 2,600 nucleotides to about 7,000 nucleotides, about 2,600
nucleotides to about 6,800 nucleotides, about 2,600 nucleotides to about 6,600
nucleotides, about 2,600 nucleotides to about 6,400 nucleotides, about 2,600
nucleotides to about 6,200 nucleotides, about 2,600 nucleotides to about 6,000
nucleotides, about 2,600 nucleotides to about 5,800 nucleotides, about 2,600
20 nucleotides to about 5,600 nucleotides, about 2,600 nucleotides to about 5,400
nucleotides, about 2,600 nucleotides to about 5,200 nucleotides, about 2,600
nucleotides to about 5,000 nucleotides, about 2,600 nucleotides to about 4,800
nucleotides, about 2,600 nucleotides to about 4,600 nucleotides, about 2,600
nucleotides to about 4,400 nucleotides, about 2,600 nucleotides to about 4,200
25 nucleotides, about 2,600 nucleotides to about 4,000 nucleotides, about 2,600
nucleotides to about 3,800 nucleotides, about 2,600 nucleotides to about 3,600
nucleotides, about 2,600 nucleotides to about 3,400 nucleotides, about 2,600
nucleotides to about 3,200 nucleotides, about 2,600 nucleotides to about 3,000
nucleotides, about 2,600 nucleotides to about 2,800 nucleotides, about 2,800
30 nucleotides to about 10,000 nucleotides, about 2,800 nucleotides to about 9,500
nucleotides, about 2,800 nucleotides to about 9,000 nucleotides, about 2,800
nucleotides to about 8,500 nucleotides, about 2,800 nucleotides to about 8,000
nucleotides, about 2,800 nucleotides to about 7,800 nucleotides, about 2,800
nucleotides to about 7,600 nucleotides, about 2,800 nucleotides to about 7,400

nucleotides, about 2,800 nucleotides to about 7,200 nucleotides, about 2,800
nucleotides to about 7,000 nucleotides, about 2,800 nucleotides to about 6,800
nucleotides, about 2,800 nucleotides to about 6,600 nucleotides, about 2,800
nucleotides to about 6,400 nucleotides, about 2,800 nucleotides to about 6,200
5 nucleotides, about 2,800 nucleotides to about 6,000 nucleotides, about 2,800
nucleotides to about 5,800 nucleotides, about 2,800 nucleotides to about 5,600
nucleotides, about 2,800 nucleotides to about 5,400 nucleotides, about 2,800
nucleotides to about 5,200 nucleotides, about 2,800 nucleotides to about 5,000
nucleotides, about 2,800 nucleotides to about 4,800 nucleotides, about 2,800
10 nucleotides to about 4,600 nucleotides, about 2,800 nucleotides to about 4,400
nucleotides, about 2,800 nucleotides to about 4,200 nucleotides, about 2,800
nucleotides to about 4,000 nucleotides, about 2,800 nucleotides to about 3,800
nucleotides, about 2,800 nucleotides to about 3,600 nucleotides, about 2,800
nucleotides to about 3,400 nucleotides, about 2,800 nucleotides to about 3,200
15 nucleotides, about 2,800 nucleotides to about 3,000 nucleotides, about 3,000
nucleotides to about 10,000 nucleotides, about 3,000 nucleotides to about 9,500
nucleotides, about 3,000 nucleotides to about 9,000 nucleotides, about 3,000
nucleotides to about 8,500 nucleotides, about 3,000 nucleotides to about 8,000
nucleotides, about 3,000 nucleotides to about 7,800 nucleotides, about 3,000
20 nucleotides to about 7,600 nucleotides, about 3,000 nucleotides to about 7,400
nucleotides, about 3,000 nucleotides to about 7,200 nucleotides, about 3,000
nucleotides to about 7,000 nucleotides, about 3,000 nucleotides to about 6,800
nucleotides, about 3,000 nucleotides to about 6,600 nucleotides, about 3,000
nucleotides to about 6,400 nucleotides, about 3,000 nucleotides to about 6,200
25 nucleotides, about 3,000 nucleotides to about 6,000 nucleotides, about 3,000
nucleotides to about 5,800 nucleotides, about 3,000 nucleotides to about 5,600
nucleotides, about 3,000 nucleotides to about 5,400 nucleotides, about 3,000
nucleotides to about 5,200 nucleotides, about 3,000 nucleotides to about 5,000
nucleotides, about 3,000 nucleotides to about 4,800 nucleotides, about 3,000
30 nucleotides to about 4,600 nucleotides, about 3,000 nucleotides to about 4,400
nucleotides, about 3,000 nucleotides to about 4,200 nucleotides, about 3,000
nucleotides to about 4,000 nucleotides, about 3,000 nucleotides to about 3,800
nucleotides, about 3,000 nucleotides to about 3,600 nucleotides, about 3,000
nucleotides to about 3,400 nucleotides, about 3,000 nucleotides to about 3,200

[illegible]

nucleotides, about 3,400 nucleotides to about 4,200 nucleotides, about 3,400
nucleotides to about 4,000 nucleotides, about 3,400 nucleotides to about 3,800
nucleotides, about 3,400 nucleotides to about 3,600 nucleotides, about 3,600
nucleotides to about 10,000 nucleotides, about 3,600 nucleotides to about 9,500
5 nucleotides, about 3,600 nucleotides to about 9,000 nucleotides, about 3,600
nucleotides to about 8,500 nucleotides, about 3,600 nucleotides to about 8,000
nucleotides, about 3,600 nucleotides to about 7,800 nucleotides, about 3,600
nucleotides to about 7,600 nucleotides, about 3,600 nucleotides to about 7,400
nucleotides, about 3,600 nucleotides to about 7,200 nucleotides, about 3,600
10 nucleotides to about 7,000 nucleotides, about 3,600 nucleotides to about 6,800
nucleotides, about 3,600 nucleotides to about 6,600 nucleotides, about 3,600
nucleotides to about 6,400 nucleotides, about 3,600 nucleotides to about 6,200
nucleotides, about 3,600 nucleotides to about 6,000 nucleotides, about 3,600
nucleotides to about 5,800 nucleotides, about 3,600 nucleotides to about 5,600
15 nucleotides, about 3,600 nucleotides to about 5,400 nucleotides, about 3,600
nucleotides to about 5,200 nucleotides, about 3,600 nucleotides to about 5,000
nucleotides, about 3,600 nucleotides to about 4,800 nucleotides, about 3,600
nucleotides to about 4,600 nucleotides, about 3,600 nucleotides to about 4,400
nucleotides, about 3,600 nucleotides to about 4,200 nucleotides, about 3,600
20 nucleotides to about 4,000 nucleotides, about 3,600 nucleotides to about 3,800
nucleotides, about 3,800 nucleotides to about 10,000 nucleotides, about 3,800
nucleotides to about 9,500 nucleotides, about 3,800 nucleotides to about 9,000
nucleotides, about 3,800 nucleotides to about 8,500 nucleotides, about 3,800
nucleotides to about 8,000 nucleotides, about 3,800 nucleotides to about 7,800
25 nucleotides, about 3,800 nucleotides to about 7,600 nucleotides, about 3,800
nucleotides to about 7,400 nucleotides, about 3,800 nucleotides to about 7,200
nucleotides, about 3,800 nucleotides to about 7,000 nucleotides, about 3,800
nucleotides to about 6,800 nucleotides, about 3,800 nucleotides to about 6,600
nucleotides, about 3,800 nucleotides to about 6,400 nucleotides, about 3,800
30 nucleotides to about 6,200 nucleotides, about 3,800 nucleotides to about 6,000
nucleotides, about 3,800 nucleotides to about 5,800 nucleotides, about 3,800
nucleotides to about 5,600 nucleotides, about 3,800 nucleotides to about 5,400
nucleotides, about 3,800 nucleotides to about 5,200 nucleotides, about 3,800
nucleotides to about 5,000 nucleotides, about 3,800 nucleotides to about 4,800

nucleotides, about 3,800 nucleotides to about 4,600 nucleotides, about 3,800
nucleotides to about 4,200 nucleotides, about 3,800 nucleotides to about 4,000
nucleotides, about 4,000 nucleotides to about 10,000 nucleotides, about 4,000
nucleotides to about 9,500 nucleotides, about 4,000 nucleotides to about 9,000
5 nucleotides, about 4,000 nucleotides to about 8,500 nucleotides, about 4,000
nucleotides to about 8,000 nucleotides, about 4,000 nucleotides to about 7,800
nucleotides, about 4,000 nucleotides to about 7,600 nucleotides, about 4,000
nucleotides to about 7,400 nucleotides, about 4,000 nucleotides to about 7,200
nucleotides, about 4,000 nucleotides to about 7,000 nucleotides, about 4,000
10 nucleotides to about 6,800 nucleotides, about 4,000 nucleotides to about 6,600
nucleotides, about 4,000 nucleotides to about 6,400 nucleotides, about 4,000
nucleotides to about 6,200 nucleotides, about 4,000 nucleotides to about 6,000
nucleotides, about 4,000 nucleotides to about 5,800 nucleotides, about 4,000
nucleotides to about 5,600 nucleotides, about 4,000 nucleotides to about 5,400
15 nucleotides, about 4,000 nucleotides to about 5,200 nucleotides, about 4,000
nucleotides to about 5,000 nucleotides, about 4,000 nucleotides to about 4,800
nucleotides, about 4,000 nucleotides to about 4,600 nucleotides, about 4,000
nucleotides to about 4,400 nucleotides, about 4,000 nucleotides to about 4,200
nucleotides, about 4,200 nucleotides to about 10,000 nucleotides, about 4,200
20 nucleotides to about 9,500 nucleotides, about 4,200 nucleotides to about 9,000
nucleotides, about 4,200 nucleotides to about 8,500 nucleotides, about 4,200
nucleotides to about 8,000 nucleotides, about 4,200 nucleotides to about 7,800
nucleotides, about 4,200 nucleotides to about 7,600 nucleotides, about 4,200
nucleotides to about 7,400 nucleotides, about 4,200 nucleotides to about 7,200
25 nucleotides, about 4,200 nucleotides to about 7,000 nucleotides, about 4,200
nucleotides to about 6,800 nucleotides, about 4,200 nucleotides to about 6,600
nucleotides, about 4,200 nucleotides to about 6,400 nucleotides, about 4,200
nucleotides to about 6,200 nucleotides, about 4,200 nucleotides to about 6,000
nucleotides, about 4,200 nucleotides to about 5,800 nucleotides, about 4,200
30 nucleotides to about 5,600 nucleotides, about 4,200 nucleotides to about 5,400
nucleotides, about 4,200 nucleotides to about 5,200 nucleotides, about 4,200
nucleotides to about 5,000 nucleotides, about 4,200 nucleotides to about 4,800
nucleotides, about 4,200 nucleotides to about 4,600 nucleotides, about 4,200
nucleotides to about 4,400 nucleotides, about 4,400 nucleotides to about 10,000

nucleotides, about 4,400 nucleotides to about 9,500 nucleotides, about 4,400
nucleotides to about 9,000 nucleotides, about 4,400 nucleotides to about 8,500
nucleotides, about 4,400 nucleotides to about 8,000 nucleotides, about 4,400
nucleotides to about 7,800 nucleotides, about 4,400 nucleotides to about 7,600
5 nucleotides, about 4,400 nucleotides to about 7,400 nucleotides, about 4,400
nucleotides to about 7,200 nucleotides, about 4,400 nucleotides to about 7,000
nucleotides, about 4,400 nucleotides to about 6,800 nucleotides, about 4,400
nucleotides to about 6,600 nucleotides, about 4,400 nucleotides to about 6,400
nucleotides, about 4,400 nucleotides to about 6,200 nucleotides, about 4,400
10 nucleotides to about 6,000 nucleotides, about 4,400 nucleotides to about 5,800
nucleotides, about 4,400 nucleotides to about 5,600 nucleotides, about 4,400
nucleotides to about 5,400 nucleotides, about 4,400 nucleotides to about 5,200
nucleotides, about 4,400 nucleotides to about 5,000 nucleotides, about 4,400
nucleotides to about 4,800 nucleotides, about 4,400 nucleotides to about 4,600
15 nucleotides, about 4,600 nucleotides to about 10,000 nucleotides, about 4,600
nucleotides to about 9,500 nucleotides, about 4,600 nucleotides to about 9,000
nucleotides, about 4,600 nucleotides to about 8,500 nucleotides, about 4,600
nucleotides to about 8,000 nucleotides, about 4,600 nucleotides to about 7,800
nucleotides, about 4,600 nucleotides to about 7,600 nucleotides, about 4,600
20 nucleotides to about 7,400 nucleotides, about 4,600 nucleotides to about 7,200
nucleotides, about 4,600 nucleotides to about 7,000 nucleotides, about 4,600
nucleotides to about 6,800 nucleotides, about 4,600 nucleotides to about 6,600
nucleotides, about 4,600 nucleotides to about 6,400 nucleotides, about 4,600
nucleotides to about 6,200 nucleotides, about 4,600 nucleotides to about 6,000
25 nucleotides, about 4,600 nucleotides to about 5,800 nucleotides, about 4,600
nucleotides to about 5,600 nucleotides, about 4,600 nucleotides to about 5,400
nucleotides, about 4,600 nucleotides to about 5,200 nucleotides, about 4,600
nucleotides to about 5,000 nucleotides, about 4,600 nucleotides to about 4,800
nucleotides, about 4,800 nucleotides to about 10,000 nucleotides, about 4,800
30 nucleotides to about 9,500 nucleotides, about 4,800 nucleotides to about 9,000
nucleotides, about 4,800 nucleotides to about 8,500 nucleotides, about 4,800
nucleotides to about 8,000 nucleotides, about 4,800 nucleotides to about 7,800
nucleotides, about 4,800 nucleotides to about 7,600 nucleotides, about 4,800
nucleotides to about 7,400 nucleotides, about 4,800 nucleotides to about 7,200

nucleotides, about 4,800 nucleotides to about 7,000 nucleotides, about 4,800
nucleotides to about 6,800 nucleotides, about 4,800 nucleotides to about 6,600
nucleotides, about 4,800 nucleotides to about 6,400 nucleotides, about 4,800
nucleotides to about 6,200 nucleotides, about 4,800 nucleotides to about 6,000
5 nucleotides, about 4,800 nucleotides to about 5,800 nucleotides, about 4,800
nucleotides to about 5,600 nucleotides, about 4,800 nucleotides to about 5,400
nucleotides, about 4,800 nucleotides to about 5,200 nucleotides, about 4,800
nucleotides to about 5,000 nucleotides, about 5,000 nucleotides to about 10,000
nucleotides, about 5,000 nucleotides to about 9,500 nucleotides, about 5,000
10 nucleotides to about 9,000 nucleotides, about 5,000 nucleotides to about 8,500
nucleotides, about 5,000 nucleotides to about 8,000 nucleotides, about 5,000
nucleotides to about 7,800 nucleotides, about 5,000 nucleotides to about 7,600
nucleotides, about 5,000 nucleotides to about 7,400 nucleotides, about 5,000
nucleotides to about 7,200 nucleotides, about 5,000 nucleotides to about 7,000
15 nucleotides, about 5,000 nucleotides to about 6,800 nucleotides, about 5,000
nucleotides to about 6,600 nucleotides, about 5,000 nucleotides to about 6,400
nucleotides, about 5,000 nucleotides to about 6,200 nucleotides, about 5,000
nucleotides to about 6,000 nucleotides, about 5,000 nucleotides to about 5,800
nucleotides, about 5,000 nucleotides to about 5,600 nucleotides, about 5,000
20 nucleotides to about 5,400 nucleotides, about 5,000 nucleotides to about 5,200
nucleotides, about 5,200 nucleotides to about 10,000 nucleotides, about 5,200
nucleotides to about 9,500 nucleotides, about 5,200 nucleotides to about 9,000
nucleotides, about 5,200 nucleotides to about 8,500 nucleotides, about 5,200
nucleotides to about 8,000 nucleotides, about 5,200 nucleotides to about 7,800
25 nucleotides, about 5,200 nucleotides to about 7,600 nucleotides, about 5,200
nucleotides to about 7,400 nucleotides, about 5,200 nucleotides to about 7,200
nucleotides, about 5,200 nucleotides to about 7,000 nucleotides, about 5,200
nucleotides to about 6,800 nucleotides, about 5,200 nucleotides to about 6,600
nucleotides, about 5,200 nucleotides to about 6,400 nucleotides, about 5,200
30 nucleotides to about 6,200 nucleotides, about 5,200 nucleotides to about 6,000
nucleotides, about 5,200 nucleotides to about 5,800 nucleotides, about 5,200
nucleotides to about 5,600 nucleotides, about 5,200 nucleotides to about 5,400
nucleotides, about 5,400 nucleotides to about 10,000 nucleotides, about 5,400
nucleotides to about 9,500 nucleotides, about 5,400 nucleotides to about 9,000

nucleotides, about 5,400 nucleotides to about 8,500 nucleotides, about 5,400
nucleotides to about 8,000 nucleotides, about 5,400 nucleotides to about 7,800
nucleotides, about 5,400 nucleotides to about 7,600 nucleotides, about 5,400
nucleotides to about 7,400 nucleotides, about 5,400 nucleotides to about 7,200
5 nucleotides, about 5,400 nucleotides to about 7,000 nucleotides, about 5,400
nucleotides to about 6,800 nucleotides, about 5,400 nucleotides to about 6,600
nucleotides, about 5,400 nucleotides to about 6,400 nucleotides, about 5,400
nucleotides to about 6,200 nucleotides, about 5,400 nucleotides to about 6,000
nucleotides, about 5,400 nucleotides to about 5,800 nucleotides, about 5,400
10 nucleotides to about 5,600 nucleotides, about 5,600 nucleotides to about 10,000
nucleotides, about 5,600 nucleotides to about 9,500 nucleotides, about 5,600
nucleotides to about 9,000 nucleotides, about 5,600 nucleotides to about 8,500
nucleotides, about 5,600 nucleotides to about 8,000 nucleotides, about 5,600
nucleotides to about 7,800 nucleotides, about 5,600 nucleotides to about 7,600
15 nucleotides, about 5,600 nucleotides to about 7,400 nucleotides, about 5,600
nucleotides to about 7,200 nucleotides, about 5,600 nucleotides to about 7,000
nucleotides, about 5,600 nucleotides to about 6,800 nucleotides, about 5,600
nucleotides to about 6,600 nucleotides, about 5,600 nucleotides to about 6,400
nucleotides, about 5,600 nucleotides to about 6,200 nucleotides, about 5,600
20 nucleotides to about 6,000 nucleotides, about 5,600 nucleotides to about 5,800
nucleotides, about 5,800 nucleotides to about 10,000 nucleotides, about 5,800
nucleotides to about 9,500 nucleotides, about 5,800 nucleotides to about 9,000
nucleotides, about 5,800 nucleotides to about 8,500 nucleotides, about 5,800
nucleotides to about 8,000 nucleotides, about 5,800 nucleotides to about 7,800
25 nucleotides, about 5,800 nucleotides to about 7,600 nucleotides, about 5,800
nucleotides to about 7,400 nucleotides, about 5,800 nucleotides to about 7,200
nucleotides, about 5,800 nucleotides to about 7,000 nucleotides, about 5,800
nucleotides to about 6,800 nucleotides, about 5,800 nucleotides to about 6,600
nucleotides, about 5,800 nucleotides to about 6,400 nucleotides, about 5,800
30 nucleotides to about 6,200 nucleotides, about 5,800 nucleotides to about 6,000
nucleotides, about 6,000 nucleotides to about 10,000 nucleotides, about 6,000
nucleotides to about 9,500 nucleotides, about 6,000 nucleotides to about 9,000
nucleotides, about 6,000 nucleotides to about 8,500 nucleotides, about 6,000
nucleotides to about 8,000 nucleotides, about 6,000 nucleotides to about 7,800

nucleotides, about 6,000 nucleotides to about 7,600 nucleotides, about 6,000
nucleotides to about 7,400 nucleotides, about 6,000 nucleotides to about 7,200
nucleotides, about 6,000 nucleotides to about 7,000 nucleotides, about 6,000
nucleotides to about 6,800 nucleotides, about 6,000 nucleotides to about 6,600
5 nucleotides, about 6,000 nucleotides to about 6,400 nucleotides, about 6,000
nucleotides to about 6,200 nucleotides, about 6,200 nucleotides to about 10,000
nucleotides, about 6,200 nucleotides to about 9,000 nucleotides, about 6,200
nucleotides to about 8,500 nucleotides, about 6,200 nucleotides to about 8,000
nucleotides, about 6,200 nucleotides to about 7,800 nucleotides, about 6,200
10 nucleotides to about 7,600 nucleotides, about 6,200 nucleotides to about 7,400
nucleotides, about 6,200 nucleotides to about 7,200 nucleotides, about 6,200
nucleotides to about 7,000 nucleotides, about 6,200 nucleotides to about 6,800
nucleotides, about 6,200 nucleotides to about 6,600 nucleotides, about 6,200
nucleotides to about 6,400 nucleotides, about 6,400 nucleotides to about 10,000
15 nucleotides, about 6,400 nucleotides to about 9,500 nucleotides, about 6,400
nucleotides to about 9,000 nucleotides, about 6,400 nucleotides to about 8,500
nucleotides, about 6,400 nucleotides to about 8,000 nucleotides, about 6,400
nucleotides to about 7,800 nucleotides, about 6,400 nucleotides to about 7,600
nucleotides, about 6,400 nucleotides to about 7,400 nucleotides, about 6,400
20 nucleotides to about 7,200 nucleotides, about 6,400 nucleotides to about 7,000
nucleotides, about 6,400 nucleotides to about 6,800 nucleotides, about 6,400
nucleotides to about 6,600 nucleotides, about 6,600 nucleotides to about 10,000
nucleotides, about 6,600 nucleotides to about 9,500 nucleotides, about 6,600
nucleotides to about 9,000 nucleotides, about 6,600 nucleotides to about 8,500
25 nucleotides, about 6,600 nucleotides to about 8,000 nucleotides, about 6,600
nucleotides to about 7,800 nucleotides, about 6,600 nucleotides to about 7,600
nucleotides, about 6,600 nucleotides to about 7,400 nucleotides, about 6,600
nucleotides to about 7,200 nucleotides, about 6,600 nucleotides to about 7,000
nucleotides, about 6,600 nucleotides to about 6,800 nucleotides, about 6,800
30 nucleotides to about 10,000 nucleotides, about 6,800 nucleotides to about 9,500
nucleotides, about 6,800 nucleotides to about 9,000 nucleotides, about 6,800
nucleotides to about 8,500 nucleotides, about 6,800 nucleotides to about 8,000
nucleotides, about 6,800 nucleotides to about 7,800 nucleotides, about 6,800
nucleotides to about 7,600 nucleotides, about 6,800 nucleotides to about 7,400

nucleotides, about 6,800 nucleotides to about 7,200 nucleotides, about 6,800 nucleotides to about 7,000 nucleotides, about 7,000 nucleotides to about 10,000 nucleotides, about 7,000 nucleotides to about 9,500 nucleotides, about 7,000 nucleotides to about 9,000 nucleotides, about 7,000 nucleotides to about 8,500 nucleotides, about 7,000 nucleotides to about 8,000 nucleotides, about 7,000 nucleotides to about 7,800 nucleotides, about 7,000 nucleotides to about 7,600 nucleotides, about 7,000 nucleotides to about 7,400 nucleotides, about 7,000 nucleotides to about 7,200 nucleotides, about 7,200 nucleotides to about 10,000 nucleotides, about 7,200 nucleotides to about 9,500 nucleotides, about 7,200 nucleotides to about 9,000 nucleotides, about 7,200 nucleotides to about 8,500 nucleotides, about 7,200 nucleotides to about 8,000 nucleotides, about 7,200 nucleotides to about 7,800 nucleotides, about 7,200 nucleotides to about 7,600 nucleotides, about 7,200 nucleotides to about 7,400 nucleotides, about 7,400 nucleotides to about 10,000 nucleotides, about 7,400 nucleotides to about 9,500 nucleotides, about 7,400 nucleotides to about 9,000 nucleotides, about 7,400 nucleotides to about 8,500 nucleotides, about 7,400 nucleotides to about 8,000 nucleotides, about 7,400 nucleotides to about 7,800 nucleotides, about 7,400 nucleotides to about 7,600 nucleotides, about 7,600 nucleotides to about 10,000 nucleotides, about 7,600 nucleotides to about 9,500 nucleotides, about 7,600 nucleotides to about 9,000 nucleotides, about 7,600 nucleotides to about 8,500 nucleotides, about 7,600 nucleotides to about 8,000 nucleotides, about 7,600 nucleotides to about 7,800 nucleotides, about 7,800 nucleotides to about 10,000 nucleotides, about 7,800 nucleotides to about 9,500 nucleotides, about 7,800 nucleotides to about 9,000 nucleotides, about 7,800 nucleotides to about 8,500 nucleotides, about 7,800 nucleotides to about 8,000 nucleotides, about 8,000 nucleotides to about 10,000 nucleotides, about 8,000 nucleotides to about 9,500 nucleotides, about 8,000 nucleotides to about 9,000 nucleotides, about 8,000 nucleotides to about 8,500 nucleotides, about 8,500 nucleotides to about 10,000 nucleotides, about 8,500 nucleotides to about 9,500 nucleotides, about 8,500 nucleotides to about 9,000 nucleotides, about 9,000 nucleotides to about 10,000 nucleotides, about 9,000 nucleotides to about 9,500 nucleotides, or about 9,500 nucleotides to about 10,000 nucleotides (inclusive).

Provided herein are exemplary vectors that can be used in any of the compositions and methods described herein. See, e.g., Figures 1-6.

A variety of different methods known in the art can be used to introduce any of vectors disclosed herein into a mammalian cell (e.g., a cochlear outer hair cell). Non-limiting examples of methods for introducing nucleic acid into a mammalian cell include: lipofection, transfection (e.g., calcium phosphate transfection, transfection using highly branched organic compounds, transfection using cationic polymers, dendrimer-based transfection, optical transfection, particle-based transfection (e.g., nanoparticle transfection), or transfection using liposomes (e.g., cationic liposomes)), microinjection, electroporation, cell squeezing, sonoporation, protoplast fusion, impalefection, hydrodynamic delivery, gene gun, magnetofection, viral transfection, and nucleofection.

Skilled practitioners will appreciate that any of the vectors described herein can be introduced into a mammalian cell by, for example, lipofection, and can be stably integrated into an endogenous gene locus (e.g., a stereocilin gene locus or a stereocilin pseudogene 1 gene locus). In some embodiments, the vectors provided herein stably integrate into an endogenous defective stereocilin gene locus, and thereby replace the defective stereocilin gene with a nucleic acid encoding a functioning (e.g., wildtype) stereocilin protein.

Various molecular biology techniques that can be used to introduce a mutation(s) and/or a deletion(s) into an endogenous gene are also known in the art. Non-limiting examples of such techniques include site-directed mutagenesis, CRISPR (e.g., CRISPR/Cas9-induced knock-in mutations and CRISPR/Cas9-induced knock-out mutations), and TALENs. These methods can be used to correct the sequence of a defective endogenous gene present in a chromosome of a target cell.

Any of the vectors described herein can further include a control sequence, e.g., a control sequence selected from the group of a transcription initiation sequence, a transcription termination sequence, a promoter sequence, an enhancer sequence, an RNA splicing sequence, a polyadenylation (polyA) sequence, and a Kozak consensus sequence. Non-limiting examples of these control sequences are described herein. In some embodiments, a promoter can be a native promoter, a constitutive promoter, an inducible promoter, and/or a tissue-specific promoter.

Promoters

The term “promoter” means a DNA sequence recognized by enzymes/proteins in a mammalian cell required to initiate the transcription of a specific gene (e.g., a stereocilin gene). A promoter typically refers to, e.g., a nucleotide sequence to which an RNA polymerase and/or any associated factor binds and at which transcription is initiated. Non-limiting examples of promoters are described herein. Additional examples of promoters are known in the art.

In some embodiments, a vector encoding an N-terminal portion of a stereocilin protein (e.g., a human stereocilin protein) can include a promoter and/or an enhancer. The vector encoding the N-terminal portion of the stereocilin protein can include any of the promoters and/or enhancers described herein or known in the art.

In some embodiments, the promoter is an inducible promoter, a constitutive promoter, a mammalian cell promoter, a viral promoter, a chimeric promoter, an engineered promoter, a tissue-specific promoter, or any other type of promoter known in the art. In some embodiments, the promoter is a RNA polymerase II promoter, such as a mammalian RNA polymerase II promoter. In some embodiments, the promoter is a RNA polymerase III promoter, including, but not limited to, a H1 promoter, a human U6 promoter, a mouse U6 promoter, or a swine U6 promoter. The promoter will generally be one that is able to promote transcription in cochlear cells such as hair cells. In some examples, the promoter is a cochlea-specific promoter or a cochlea-oriented promoter.

A variety of promoters are known in the art that can be used herein. Non-limiting examples of promoters that can be used herein include: human EF1a, human cytomegalovirus (CMV) (US Patent No. 5,168,062), human ubiquitin C (UBC), mouse phosphoglycerate kinase 1, polyoma adenovirus, simian virus 40 (SV40), β -globin, β -actin, α -fetoprotein, γ -globin, β -interferon, γ -glutamyl transferase, mouse mammary tumor virus (MMTV), Rous sarcoma virus, rat insulin, glyceraldehyde-3-phosphate dehydrogenase, metallothionein II (MT II), amylase, cathepsin, MI muscarinic receptor, retroviral LTR (e.g. human T-cell leukemia virus HTLV), AAV ITR, interleukin-2, collagenase, platelet-derived growth factor, adenovirus 5 E2, stromelysin, murine MX gene, glucose regulated proteins (GRP78 and GRP94), α -2-macroglobulin, vimentin, MHC class I gene H-2 κ b, HSP70, proliferin, tumor necrosis factor, thyroid stimulating hormone α gene, immunoglobulin light chain, T-cell receptor, HLA DQ α and DQ β , interleukin-2 receptor, MHC class II, MHC class II

HLA-DR α , muscle creatine kinase, prealbumin (transthyretin), elastase I, albumin gene, c-fos, c-HA-ras, neural cell adhesion molecule (NCAM), H2B (TH2B) histone, rat growth hormone, human serum amyloid (SAA), troponin I (TN I), duchenne muscular dystrophy, human immunodeficiency virus, and Gibbon Ape Leukemia

5 Virus (GALV) promoters. Additional examples of promoters are known in the art. See, e.g., Lodish, *Molecular Cell Biology*, Freeman and Company, New York 2007. In some embodiments, the promoter is the CMV immediate early promoter. In some embodiments, the promoter is a CMV promoter, e.g., a CMV promoter comprising or consisting of SEQ ID NO: 21. In some embodiments, the promoter is a CAG
10 promoter or a CAG/CBA promoter. In some embodiments, the promoter is a CBA promoter, e.g., a CBA promoter comprising or consisting of SEQ ID NO: 22.

The term “constitutive” promoter refers to a nucleotide sequence that, when operably linked with a nucleic acid encoding a protein (e.g., a stereocilin protein), causes RNA to be transcribed from the nucleic acid in a mammalian cell under most
15 or all physiological conditions.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter, the cytomegalovirus (CMV) promoter (see, e.g., Boshart et al, *Cell* 41:521-530, 1985), the SV40 promoter, the dihydrofolate reductase promoter, the beta-actin promoter, the phosphoglycerol kinase (PGK)
20 promoter, and the EF1-alpha promoter (Invitrogen).

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and
25 inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech, and Ariad. Additional examples of inducible promoters are known in the art.

Examples of inducible promoters regulated by exogenously supplied compounds include the zinc-inducible sheep metallothioneine (MT) promoter, the
30 dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No et al, *Proc. Natl. Acad. Sci. U.S.A.* 93:3346-3351, 1996), the tetracycline-repressible system (Gossen et al, *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551, 1992), the tetracycline-inducible system (Gossen et al, *Science* 268:1766-1769, 1995, see also

Harvey et al, *Curr. Opin. Chem. Biol.* 2:512-518, 1998), the RU486-inducible system (Wang et al, *Nat. Biotech.* 15:239-243, 1997) and Wang et al, *Gene Ther.* 4:432-441, 1997), and the rapamycin-inducible system (Magari et al. *J. Clin. Invest.* 100:2865-2872, 1997).

5 The term “tissue-specific” promoter refers to a promoter that is active only in certain specific cell types and/or tissues (e.g., transcription of a specific gene occurs only within cells expressing transcription regulatory proteins that bind to the tissue-specific promoter).

10 In some embodiments, the regulatory sequences impart tissue-specific gene expression capabilities. In some cases, the tissue-specific regulatory sequences bind tissue-specific transcription factors that induce transcription in a tissue-specific manner.

Exemplary tissue-specific promoters include but are not limited to the following: a liver-specific thyroxin binding globulin (TBG) promoter, an insulin
15 promoter, a glucagon promoter, a somatostatin promoter, a pancreatic polypeptide (PPY) promoter, a synapsin-1 (Syn) promoter, a creatine kinase (MCK) promoter, a mammalian desmin (DES) promoter, an alpha-myosin heavy chain (a-MHC) promoter, and a cardiac Troponin T (cTnT) promoter. Additional exemplary promoters include Beta-actin promoter, hepatitis B virus core promoter (Sandig et al.,
20 *Gene Ther.* 3:1002-1009, 1996), alpha-fetoprotein (AFP) promoter (Arbuthnot et al., *Hum. Gene Ther.* 7:1503-1514, 1996), bone osteocalcin promoter (Stein et al., *Mol. Biol. Rep.* 24:185-196, 1997); bone sialoprotein promoter (Chen et al., *J. Bone Miner. Res.* 11:654-664, 1996), CD2 promoter (Hansal et al., *J. Immunol.* 161:1063-1068, 1998); immunoglobulin heavy chain promoter; T cell receptor alpha-chain promoter,
25 neuronal such as neuron-specific enolase (NSE) promoter (Andersen et al., *Cell. Mol. Neurobiol.* 13:503-515, 1993), neurofilament light-chain gene promoter (Piccioli et al., *Proc. Natl. Acad. Sci. U.S.A.* 88:5611-5615, 1991), and the neuron-specific vgf gene promoter (Piccioli et al., *Neuron* 15:373-384, 1995).

30 In some embodiments, the tissue-specific promoter is a cochlea-specific promoter. In some embodiments, the tissue-specific promoter is a cochlear hair cell-specific promoter. Non-limiting examples of cochlear hair cell-specific promoters include but are not limited to: a ATOH1 promoter, a POU4F3 promoter, a LHX3 promoter, a MYO7A promoter, a MYO6 promoter, a α 9ACHR promoter, and a α 10ACHR promoter. In some embodiments, the promoter is an outer hair cell-

specific promoter such as a PRESTIN promoter or an ONCOMOD promoter. See, e.g., Zheng et al., *Nature* 405:149-155, 2000; Tian et al. *Dev. Dyn.* 231:199-203, 2004; and Ryan et al., *Adv. Otorhinolaryngol.* 66: 99-115, 2009.

5 **Enhancers and 5' Cap**

In some instances, a vector can include a promoter sequence and/or an enhancer sequence. The term “enhancer” refers to a nucleotide sequence that can increase the level of transcription of a nucleic acid encoding a protein of interest (e.g., a stereocilin protein). Enhancer sequences (50-1500 basepairs in length) generally
10 increase the level of transcription by providing additional binding sites for transcription-associated proteins (e.g., transcription factors). In some embodiments, an enhancer sequence is found within an intronic sequence. Unlike promoter sequences, enhancer sequences can act at much larger distance away from the transcription start site (e.g., as compared to a promoter). Non-limiting examples of
15 enhancers include a RSV enhancer, a CMV enhancer, and a SV40 enhancer. In some embodiments, the CMV enhancer sequence comprises or consists of SEQ ID NO: 20.

Poly(A) Sequences

In some embodiments, any of the vectors provided herein can include a
20 poly(A) sequence. Most nascent eukaryotic mRNAs possess a poly(A) tail at their 3' end which is added during a complex process that includes cleavage of the primary transcript and a coupled polyadenylation reaction (see, e.g., Proudfoot et al., *Cell* 108:501-512, 2002). The poly(A) tail confers mRNA stability and transferability (Molecular Biology of the Cell, Third Edition by B. Alberts et al., Garland Publishing,
25 1994). In some embodiments, the poly(A) sequence is positioned 3' to the nucleic acid sequence encoding the C-terminus of the stereocilin protein.

As used herein, “polyadenylation” refers to the covalent linkage of a polyadenylyl moiety, or its modified variant, to a messenger RNA molecule. In eukaryotic organisms, most messenger RNA (mRNA) molecules are polyadenylated
30 at the 3' end. The 3' poly(A) tail is a long sequence of adenine nucleotides (e.g., 50, 60, 70, 100, 200, 500, 1000, 2000, 3000, 4000, or 5000) added to the pre-mRNA through the action of an enzyme, polyadenylate polymerase. In higher eukaryotes, the poly(A) tail is added onto transcripts that contain a specific sequence, the

polyadenylation signal or “poly(A) sequence.” The poly(A) tail and the protein bound to it aid in protecting mRNA from degradation by exonucleases.

Polyadenylation is also important for transcription termination, export of the mRNA from the nucleus, and translation. Polyadenylation occurs in the nucleus immediately after transcription of DNA into RNA, but additionally can also occur later in the cytoplasm. After transcription has been terminated, the mRNA chain is cleaved through the action of an endonuclease complex associated with RNA polymerase. The cleavage site is usually characterized by the presence of the base sequence AAUAAA near the cleavage site. After the mRNA has been cleaved, adenosine residues are added to the free 3' end at the cleavage site.

As used herein, a “poly(A) sequence” is a sequence that triggers the endonuclease cleavage of an mRNA and the additional of a series of adenosines to the 3' end of the cleaved mRNA.

There are several poly(A) sequences that can be used, including those derived from bovine growth hormone (bgh) (Woychik et al., *Proc. Natl. Acad. Sci. U.S.A.* 81(13):3944-3948, 1984; U.S. Patent No. 5,122,458), mouse- β -globin, mouse- α -globin (Orkin et al., *EMBO J.* 4(2):453-456, 1985; Thein et al., *Blood* 71(2):313-319, 1988), human collagen, polyoma virus (Batt et al., *Mol. Cell Biol.* 15(9):4783-4790, 1995), the Herpes simplex virus thymidine kinase gene (HSV TK), IgG heavy-chain gene polyadenylation signal (US 2006/0040354), human growth hormone (hGH) (Szymanski et al., *Mol. Therapy* 15(7):1340-1347, 2007), the group consisting of SV40 poly(A) site, such as the SV40 late and early poly(A) site (Schek et al., *Mol. Cell Biol.* 12(12):5386-5393, 1992). In some embodiments, the bGH polyA sequence comprises or consists of SEQ ID NO: 34.

The poly(A) sequence can be a sequence of AATAAA. The AATAAA sequence may be substituted with other hexanucleotide sequences with homology to AATAAA which are capable of signaling polyadenylation, including ATTAAA, AGTAAA, CATAAA, TATAAA, GATAAA, ACTAAA, AATATA, AAGAAA, AATAAT, AAAAAA, AATGAA, AATCAA, AACAAA, AATCAA, AATAAC, AATAGA, AATTAA, or AATAAG (see, e.g., WO 06/12414).

In some embodiments, the poly(A) sequence can be a synthetic polyadenylation site (see, e.g., the pCl-neo expression vector of Promega which is based on Levitt et al., *Genes Dev.* 3(7):1019-1025, 1989). In some embodiments, the poly(A) sequence is the polyadenylation signal of soluble neuropilin-1 (sNRP)

(AAATAAAATACGAAATG, SEQ ID NO: 38) (see, e.g., WO 05/073384).

Additional examples of poly(A) sequences are known in the art.

Internal Ribosome Entry Site (IRES)

5 In some embodiments, a vector encoding the C-terminus of the stereocilin protein can include a polynucleotide internal ribosome entry site (IRES). An IRES sequence is used to produce more than one polypeptide from a single gene transcript. An IRES forms a complex secondary structure that allows translation initiation to occur from any position with an mRNA immediately downstream from where the
10 IRES is located (see, e.g., Pelletier and Sonenberg, *Mol. Cell. Biol.* 8(3):1103-1112, 1988).

There are several IRES sequences known to those in skilled in the art, including those from, e.g., foot and mouth disease virus (FMDV), encephalomyocarditis virus (EMCV), human rhinovirus (HRV), cricket paralysis
15 virus, human immunodeficiency virus (HIV), hepatitis A virus (HAV), hepatitis C virus (HCV), and poliovirus (PV). See e.g., Alberts, *Molecular Biology of the Cell*, Garland Science, 2002; and Hellen et al., *Genes Dev.* 15(13):1593-612, 2001.

In some embodiments, the IRES sequence that is incorporated into the vector that encodes the C-terminus of a stereocilin protein is the foot and mouth disease virus
20 (FMDV). The Foot and Mouth Disease Virus 2A sequence is a small peptide (approximately 18 amino acids in length) that has been shown to mediate the cleavage of polyproteins (Ryan, M D et al., *EMBO* 4:928-933, 1994; Mattion et al., *J. Virology* 70:8124-8127, 1996; Furler et al., *Gene Therapy* 8:864-873, 2001; and Halpin et al., *Plant Journal* 4:453-459, 1999). The cleavage activity of the 2A sequence has
25 previously been demonstrated in artificial systems including plasmids and gene therapy vectors (AAV and retroviruses) (Ryan et al., *EMBO* 4:928-933, 1994; Mattion et al., *J. Virology* 70:8124-8127, 1996; Furler et al., *Gene Therapy* 8:864-873, 2001; and Halpin et al., *Plant Journal* 4:453-459, 1999; de Felipe et al., *Gene Therapy* 6:198-208, 1999; de Felipe et al., *Human Gene Therapy* 11:1921-1931, 2000; and
30 Klump et al., *Gene Therapy* 8:811-817, 2001).

Reporter Sequences

Any of the vectors provided herein can optionally include a sequence encoding a reporter protein (“a reporter sequence”). Non-limiting examples of reporter sequences include DNA sequences encoding: a beta-lactamase, a beta-galactosidase (LacZ), an alkaline phosphatase, a thymidine kinase, a green fluorescent protein (GFP), a red fluorescent protein, an mCherry fluorescent protein, a yellow fluorescent protein, a chloramphenicol acetyltransferase (CAT), and a luciferase. Additional examples of reporter sequences are known in the art. When associated with regulatory elements which drive their expression, the reporter sequence can provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence, or other spectrographic assays; fluorescent activating cell sorting (FACS) assays; immunological assays (e.g., enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and immunohistochemistry).

In some embodiments, the reporter sequence is turbo green fluorescent protein (tGFP) (SEQ ID NO: 31).

tGFP cDNA (SEQ ID NO: 31)

```

ATGGAGAGCGACGAGAGCGGCCTGCCCCGCATGGAGATCGAGTGCCGCATCACCGGCACCCTGAACGGC
GTGGAGTTCGAGCTGGTGGGCGGCGGAGAGGGGCACCCCCGAGCAGGGCCGCATGACCAACAAGATGAAG
AGCACCAAGGCGCCCTGACCTTCAGCCCCTACCTGCTGAGCCACGTGATGGGCTACGGCTTCTACCAC
TTCGGCACCTACCCCAGCGGCTACGAGAACCCTTCTCTGCACGCCATCAACAACGGCGGCTACACCAAC
ACCCGCATCGAGAAGTACGAGGACGGCGGCGTGCTGCACGTGAGCTTCAGCTACCGCTACGAGGCCGGC
CGCGTGATCGGCGACTTCAAGGTGATGGGCACCGGCTTCCCCGAGGACAGCGTGATCTTCACCGACAAG
ATCATCCGCAGCAACGCCACCGTGGAGCACCTGCACCCCATGGGCGATAACGATCTGGATGGCAGCTTC
ACCCGCACCTTCAGCCTGCGCGACGGCGGCTACTACAGCTCCGTGGTGGACAGCCACATGCACCTTCAAG
AGCGCCATCCACCCCAGCATCCTGCAGAACGGGGGCCCCATGTTTCGCCTTCCGCCGCGTGGAGGAGGAT
CACAGCAACACCGAGCTGGGCATCGTGGAGTACCAGCACGCCTTCAAGACCCCGGATGCAGATGCCGGT
GAAGAA

```

tGFP Protein (SEQ ID NO: 32)

```

MESDESLPAMEIECRITGTLNGVEFELVGGEGTPEQGRMTNKMKSTKGALTFSPYLLSHVMGYGFYH
FGTYPSPGYENPFLHAINNGGYTNTRIEKYEDGGVLHVSFSYRYEAGRVI GDFKVMGTGFPEDSVIFTDK
IIRSNATVEHLHPMGDNDLDGSFTRTFSLRDGGYYSSVVD SHMHFKSAIHPSILQNGGPMFAFRVEED
HSNTELGIVEYQHAFKTPDADAGEE

```

In some embodiments, the reporter sequence is the LacZ gene, and the presence of a vector carrying the LacZ gene in a mammalian cell (e.g., a cochlear outer hair cell) is detected by assays for beta-galactosidase activity. In other embodiments, the reporter is a fluorescent protein (e.g., green fluorescent protein) or luciferase, the presence of a vector carrying the fluorescent protein or luciferase in a mammalian cell (e.g., a cochlear outer hair cell) may be measured by fluorescent techniques (e.g., fluorescent microscopy or FACS) or light production in a luminometer (e.g., a spectrophotometer or an IVIS imaging instrument). In some embodiments, the reporter sequence can be used to verify the tissue-specific targeting capabilities and tissue-specific promoter regulatory activity of any of the vectors described herein.

Flanking Regions Untranslated Regions (UTRs)

In some embodiments, any of the vectors described herein (e.g., any of the at least two different vectors) can include an untranslated region. In some embodiments, a vector can include a 5' UTR or a 3' UTR.

Untranslated regions (UTRs) of a gene are transcribed but not translated. The 5' UTR starts at the transcription start site and continues to the start codon but does not include the start codon. The 3' UTR starts immediately following the stop codon and continues until the transcriptional termination signal. There is growing body of evidence about the regulatory roles played by the UTRs in terms of stability of the nucleic acid molecule and translation. The regulatory features of a UTR can be incorporated into any of the vectors, compositions, kits, or methods as described herein to enhance the stability of a stereocilin protein.

Natural 5' UTRs include a sequence that plays a role in translation initiation. They harbor signatures like Kozak sequences, which are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus sequence CCR(A/G)CCAUGG, where R is a purine (A or G) three bases upstream of the start codon (AUG), which is followed by another "G". The 5' UTR have also been known, e.g., to form secondary structures that are involved in elongation factor binding.

For example, in some embodiments, a 5' UTR is included in any of the vectors described herein. Non-limiting examples of 5' UTRs including those from the

following genes: albumin, serum amyloid A, Apolipoprotein A/B/E, transferrin, alpha fetoprotein, erythropoietin, and Factor VIII, can be used to enhance expression of a nucleic acid molecule, such as a mRNA.

5 In some embodiments, a 5' UTR from a mRNA that is transcribed by a cell in the cochlea can be included in any of the vectors, compositions, kits, and methods described herein.

3' UTRs are known to have stretches of adenosines and uridines embedded in them. These AU-rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU-rich
10 elements (AREs) can be separated into three classes (Chen et al., *Mol. Cell. Biol.* 15:5777-5788, 1995; Chen et al., *Mol. Cell Biol.* 15:2010-2018, 1995): Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. For example, c-Myc and MyoD mRNAs contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A) (U/A) nonamers. GM-CSF and TNF-
15 alpha mRNAs are examples that contain class II AREs. Class III AREs are less well defined. These U-rich regions do not contain an AUUUA motif. Two well-studied examples of this class are c-Jun and myogenin mRNAs.

Most proteins binding to the AREs are known to destabilize the messenger, whereas members of the ELAV family, most notably HuR, have been documented to
20 increase the stability of mRNA. HuR binds to AREs of all the three classes. Engineering the HuR specific binding sites into the 3' UTR of nucleic acid molecules will lead to HuR binding and thus, stabilization of the message *in vivo*.

An exemplary human wildtype 5' UTR is or includes the sequence of SEQ ID NO: 24. An exemplary human wildtype 3' UTR is or includes the sequence of SEQ
25 ID NO: 33.

In some embodiments of any of the compositions described herein, a 5' UTR, a 3' UTR, or both are included in a vector (e.g., any of the vectors described herein). For example, any of the 5' UTRs described herein can be operatively linked to the start codon in any of the coding sequences described herein. For example, any of the
30 3' UTRs can be operatively linked to the 3'-terminal codon (last codon) in any of the coding sequences described herein.

In some embodiments of any of the compositions described herein, the 5' UTR comprises at least 10 contiguous (e.g., at least 15 contiguous, at least 20 contiguous, at least 25 contiguous, at least 30 contiguous, at least 35 contiguous, at

least 40 contiguous, at least 45 contiguous, at least 50 contiguous, at least 55 contiguous, at least 60 contiguous, at least 65 contiguous, or at least 70 contiguous) nucleotides from anywhere within SEQ ID NO: 24.

For example, a 5' UTR can include or consist of one or more of: nucleotide positions 1 to 70, nucleotide positions 1 to 60, nucleotide positions 1 to 50, nucleotide positions 1 to 40, nucleotide positions 1 to 30, nucleotide positions 1 to 20, nucleotide positions 1 to 10, nucleotide positions 10 to 70, nucleotide positions 10 to 60, nucleotide positions 10 to 50, nucleotide positions 10 to 40, nucleotide positions 10 to 30, nucleotide positions 10 to 20, nucleotide positions 20 to 70, nucleotide positions 20 to 60, nucleotide positions 20 to 50, nucleotide positions 20 to 40, nucleotide positions 20 to 30, nucleotide positions 30 to 70, nucleotide positions 30 to 60, nucleotide positions 30 to 50, nucleotide positions 30 to 40, nucleotide positions 40 to 70, nucleotide positions 40 to 60, nucleotide positions 40 to 50, nucleotide positions 50 to 70, nucleotide positions 50 to 60, nucleotide positions 60 to 70, , nucleotide positions 100 to 105, of SEQ ID NO: 24.

In some embodiments of any of the compositions described herein, the 5' UTR comprises a sequence that is at least 70% (e.g., at least 75%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%) identical to SEQ ID NO: 24.

In some embodiments of any of the compositions described herein, the 3' UTR comprises at least 10 contiguous (e.g., at least 15 contiguous, at least 20 contiguous, at least 25 contiguous, at least 30 contiguous, at least 35 contiguous, at least 40 contiguous, at least 45 contiguous, at least 50 contiguous, at least 55 contiguous, at least 60 contiguous, at least 65 contiguous, at least 70 contiguous, at least 75 contiguous, at least 80 contiguous, at least 85 contiguous, at least 90 contiguous, at least 95 contiguous, at least 100 contiguous, or at least 105 contiguous) nucleotides from anywhere within SEQ ID NO: 33.

For example, a 3' UTR can include or consist of one or more of: nucleotide positions 1 to 107, nucleotide positions 1 to 105, nucleotide positions 1 to 100, nucleotide positions 1 to 90, nucleotide positions 1 to 80, nucleotide positions 1 to 70, nucleotide positions 1 to 60, nucleotide positions 1 to 50, nucleotide positions 1 to 40, nucleotide positions 1 to 30, nucleotide positions 1 to 20, nucleotide positions 1 to 10, nucleotide positions 10 to 107, nucleotide positions 10 to 105, nucleotide positions 10 to 100, nucleotide positions 10 to 90, nucleotide positions 10 to 80, nucleotide

positions 10 to 70, nucleotide positions 10 to 60, nucleotide positions 10 to 50, nucleotide positions 10 to 40, nucleotide positions 10 to 30, nucleotide positions 10 to 20, nucleotide positions 20 to 107, nucleotide positions 20 to 105, nucleotide positions 20 to 100, nucleotide positions 20 to 90, nucleotide positions 20 to 80, nucleotide positions 20 to 70, nucleotide positions 20 to 60, nucleotide positions 20 to 50, nucleotide positions 20 to 40, nucleotide positions 20 to 30, nucleotide positions 30 to 107, nucleotide positions 30 to 105, nucleotide positions 30 to 100, nucleotide positions 30 to 90, nucleotide positions 30 to 80, nucleotide positions 30 to 70, nucleotide positions 30 to 60, nucleotide positions 30 to 50, nucleotide positions 30 to 40, nucleotide positions 40 to 107, nucleotide positions 40 to 105, nucleotide positions 40 to 100, nucleotide positions 40 to 90, nucleotide positions 40 to 80, nucleotide positions 40 to 70, nucleotide positions 40 to 60, nucleotide positions 40 to 50, nucleotide positions 50 to 107, nucleotide positions 50 to 105, nucleotide positions 50 to 100, nucleotide positions 50 to 90, nucleotide positions 50 to 80, nucleotide positions 50 to 70, nucleotide positions 50 to 60, nucleotide positions 60 to 107, nucleotide positions 60 to 105, nucleotide positions 60 to 100, nucleotide positions 60 to 90, nucleotide positions 60 to 80, nucleotide positions 60 to 70, nucleotide positions 70 to 107, nucleotide positions 70 to 105, nucleotide positions 70 to 100, nucleotide positions 70 to 90, nucleotide positions 70 to 80, nucleotide positions 80 to 107, nucleotide positions 80 to 105, nucleotide positions 80 to 100, nucleotide positions 80 to 90, nucleotide positions 90 to 107, nucleotide positions 90 to 105, nucleotide positions 90 to 100, nucleotide positions 100 to 107, nucleotide positions 100 to 105, of SEQ ID NO: 33.

In some embodiments of any of the compositions described herein, the 3' UTR comprises a sequence that is at least 70% (e.g., at least 75%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%) identical to SEQ ID NO: 33.

Human 5' UTR of STRC (SEQ ID NO: 24)

GCCCTGCCCTCACCTGGCTATCCACACAGGTGAGAATAACCAGAACTCACCTCCGGTACCAGTGTTCACTTG

Human 3' UTR of STRC (SEQ ID NO: 33)

GCCTGTCTCTACAGTAGAAGGAGATTGTGGGGAGAGAAAATCTTAAGTCATAATGAATAAAGTGCAAACA
GAAGTGCATCCTGATTATTTTCAGAAGCTGATGAGGAATA

5 In some embodiments, the introduction, removal, or modification of 3' UTR AREs can be used to modulate the stability of an mRNA encoding a stereocilin protein. In other embodiments, AREs can be removed or mutated to increase the intracellular stability and thus increase translation and production of a stereocilin protein.

10 In other embodiments, non-UTR sequences may be incorporated into the 5' or 3' UTRs. In some embodiments, introns or portions of intron sequences may be incorporated into the flanking regions of the polynucleotides in any of the vectors, compositions, kits, and methods provided herein. Incorporation of intronic sequences may increase protein production as well as mRNA levels. An intron can be an intron
15 from a stereocilin gene or can be an intron from a heterologous gene, e.g., a hybrid adenovirus/mouse immunoglobulin intron (Yew et al., *Human Gene Ther.* 8(5):575-584, 1997), an SV40 intron (Ostedgaard et al., *Proc. Natl. Acad. Sci. U.S.A.* 102(8):2952-2957, 2005) an MVM intron (Wu et al., *Mol. Ther.* 16(2):280-289, 2008), a factor IX truncated intron 1 (Wu et al., *Mol. Ther.* 16(2):280-289, 2008;
20 Kurachi et al., *J. Biol. Chem.* 270(10):5276-5281, 1995), a chimeric β -globulin splice donor/immunoglobulin heavy chain splice acceptor intron (Wu et al., *Mol. Ther.* 16(2):280-289, 2008; Choi et al., *Mol. Brain* 7:17, 2014) a SV40 late splice donor/splice acceptor intron (19S/16S) (Yew et al., *Human Gene Ther.* 8(5):575-584, 1997), a hybrid adenovirus splice donor/ IgG splice acceptor (Choi et al., *Mol. Brain*
25 7:17, 1991; Huang and Gorman, *Mol. Cell Biol.* 10(4):1805-1810, 1990).

Non-limiting examples of a splice donor and splice acceptor sequences are SEQ ID NOs: 6 and 7, respectively.

In some embodiments of any of the vectors described herein, the vector includes a chimeric intron sequence (SEQ ID NO: 23).

30 In some embodiments of any of the vectors described herein, the vector includes a T2A sequence (SEQ ID NO: 29).

T2A cDNA sequence (SEQ ID NO: 29)

GAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAAATCCTGGCCCA

T2A Protein (SEQ ID NO: 30)

GSGEGRGSLLTCTGDVEENPGP

Additional Sequences

- 5 Any of the vectors provided herein can optionally include additional nucleotide sequences (“a stuffer sequence”) in order to optimize the total number of basepairs in the vector. For example, in order to optimize packaging, each vector can be designed to contain a total of about 4,000 base pairs to about 4,700 base pairs, e.g., about 4,000 base pairs to about 4,650 base pairs, about 4,000 base pairs to about 4,600
- 10 base pairs, about 4,000 base pairs to about 4,550 base pairs, about 4,000 base pairs to about 4,500 base pairs, about 4,000 base pairs to about 4,450 base pairs, about 4,000 base pairs to about 4,400 base pairs, about 4,000 base pairs to about 4,350 base pairs, about 4,000 base pairs to about 4,300 base pairs, about 4,000 base pairs to about 4,250 base pairs, about 4,000 base pairs to about 4,200 base pairs, about 4,000 base pairs to
- 15 about 4,150 base pairs, about 4,000 base pairs to about 4,100 base pairs, about 4,000 base pairs to about 4,050 base pairs, about 4,050 base pairs to about 4,700 base pairs, about 4,050 base pairs to about 4,650 base pairs, about 4,050 base pairs to about 4,600 base pairs, about 4,050 base pairs to about 4,550 base pairs, about 4,050 base pairs to about 4,500 base pairs, about 4,050 base pairs to about 4,450 base pairs, about 4,050
- 20 base pairs to about 4,400 base pairs, about 4,050 base pairs to about 4,350 base pairs, about 4,050 base pairs to about 4,300 base pairs, about 4,050 base pairs to about 4,250 base pairs, about 4,050 base pairs to about 4,200 base pairs, about 4,050 base pairs to about 4,150 base pairs, about 4,050 base pairs to about 4,100 base pairs, about 4,100 base pairs to about 4,700 base pairs, about 4,100 base pairs to about 4,650 base pairs,
- 25 about 4,100 base pairs to about 4,600 base pairs, about 4,100 base pairs to about 4,550 base pairs, about 4,100 base pairs to about 4,500 base pairs, about 4,100 base pairs to about 4,450 base pairs, about 4,100 base pairs to about 4,400 base pairs, about 4,100 base pairs to about 4,350 base pairs, about 4,100 base pairs to about 4,300 base pairs, about 4,100 base pairs to about 4,250 base pairs, about 4,100 base pairs to about 4,200
- 30 base pairs, about 4,100 base pairs to about 4,150 base pairs, about 4,150 base pairs to about 4,700 base pairs, about 4,150 base pairs to about 4,650 base pairs, about 4,150 base pairs to about 4,600 base pairs, about 4,150 base pairs to about 4,550 base pairs, about 4,150 base pairs to about 4,500 base pairs, about 4,150 base pairs to about 4,450 base pairs, about 4,150 base pairs to about 4,400 base pairs, about 4,150 base pairs to

[illegible]

A stuffer sequence can be any nucleotide sequence, e.g., up to 1000 bp, that can be included in any of the vectors described herein that is not transcribed and that does not serve a regulatory function in order to achieve a desirable vector size (e.g., a vector size of about 4 kb to about 5 kb, or any of the vector sizes provided herein).

- 5 For example, a stuffer sequence can be any nucleotide sequences of about 100 bp to about 1000 bp (e.g., about 100 bp to about 900 bp, about 100 bp to about 850 bp, about 100 bp to about 800 bp, about 100 bp to about 750 bp, about 100 bp to about 700 bp, about 100 bp to about 650 bp, about 100 bp to about 600 bp, about 100 bp to about 550 bp, about 100 bp to about 500 bp, about 100 bp to about 450 bp, about 100 bp to about 400 bp, about 100 bp to about 350 bp, about 100 bp to about 300 bp, about 100 bp to about 250 bp, about 100 bp to about 200 bp, about 100 bp to about 150 bp, about 150 bp to about 1000 bp, about 150 bp to about 900 bp, about 150 bp to about 850 bp, about 150 bp to about 800 bp, about 150 bp to about 750 bp, about 150 bp to about 700 bp, about 150 bp to about 650 bp, about 150 bp to about 600 bp, about 150 bp to about 550 bp, about 150 bp to about 500 bp, about 150 bp to about 450 bp, about 150 bp to about 400 bp, about 150 bp to about 350 bp, about 150 bp to about 300 bp, about 150 bp to about 250 bp, about 150 bp to about 200 bp, about 200 bp to about 1000 bp, about 200 bp to about 900 bp, about 200 bp to about 850 bp, about 200 bp to about 800 bp, about 200 bp to about 750 bp, about 200 bp to about 700 bp, about 200 bp to about 650 bp, about 200 bp to about 600 bp, about 200 bp to about 550 bp, about 200 bp to about 500 bp, about 200 bp to about 450 bp, about 200 bp to about 400 bp, about 200 bp to about 350 bp, about 200 bp to about 300 bp, about 200 bp to about 250 bp, about 250 bp to about 1000 bp, about 250 bp to about 900 bp, about 250 bp to about 850 bp, about 250 bp to about 800 bp, about 250 bp to about 750 bp, about 250 bp to about 700 bp, about 250 bp to about 650 bp, about 250 bp to about 600 bp, about 250 bp to about 550 bp, about 250 bp to about 500 bp, about 250 bp to about 450 bp, about 250 bp to about 400 bp, about 250 bp to about 350 bp, about 250 bp to about 300 bp, about 300 bp to about 1000 bp, about 300 bp to about 900 bp, about 300 bp to about 850 bp, about 300 bp to about 800 bp, about 300 bp to about 750 bp, about 300 bp to about 700 bp, about 300 bp to about 650 bp, about 300 bp to about 600 bp, about 300 bp to about 550 bp, about 300 bp to about 500 bp, about 300 bp to about 450 bp, about 300 bp to about 400 bp, about 300 bp to about 350 bp, about 350 bp to about 1000 bp, about 350 bp to about 900 bp, about 350 bp to about 850 bp, about 350 bp to about 800 bp, about 350 bp to about 750 bp,

about 350 bp to about 700 bp, about 350 bp to about 650 bp, about 350 bp to about 600 bp, about 350 bp to about 550 bp, about 350 bp to about 500 bp, about 350 bp to about 450 bp, about 350 bp to about 400 bp, about 400 bp to about 1000 bp, about 400 bp to about 900 bp, about 400 bp to about 850 bp, about 400 bp to about 800 bp, about 400 bp to about 750 bp, about 400 bp to about 700 bp, about 400 bp to about 650 bp, about 400 bp to about 600 bp, about 400 bp to about 550 bp, about 400 bp to about 500 bp, about 400 bp to about 450 bp, about 450 bp to about 1000 bp, about 450 bp to about 900 bp, about 450 bp to about 850 bp, about 450 bp to about 800 bp, about 450 bp to about 750 bp, about 450 bp to about 700 bp, about 450 bp to about 650 bp, about 450 bp to about 600 bp, about 450 bp to about 550 bp, about 450 bp to about 500 bp, about 500 bp to about 1000 bp, about 500 bp to about 900 bp, about 500 bp to about 850 bp, about 500 bp to about 800 bp, about 500 bp to about 750 bp, about 500 bp to about 700 bp, about 500 bp to about 650 bp, about 500 bp to about 600 bp, about 500 bp to about 550 bp, about 550 bp to about 1000 bp, about 550 bp to about 900 bp, about 550 bp to about 850 bp, about 550 bp to about 800 bp, about 550 bp to about 750 bp, about 550 bp to about 700 bp, about 550 bp to about 650 bp, about 550 bp to about 600 bp, about 600 bp to about 1000 bp, about 600 bp to about 900 bp, about 600 bp to about 850 bp, about 600 bp to about 800 bp, about 600 bp to about 750 bp, about 600 bp to about 700 bp, about 600 bp to about 650 bp, about 650 bp to about 1000 bp, about 650 bp to about 900 bp, about 650 bp to about 850 bp, about 650 bp to about 800 bp, about 650 bp to about 750 bp, about 650 bp to about 700 bp, about 700 bp to about 1000 bp, about 700 bp to about 900 bp, about 700 bp to about 850 bp, about 700 bp to about 800 bp, about 700 bp to about 750 bp, about 750 bp to about 1000 bp, about 750 bp to about 900 bp, about 750 bp to about 850 bp, about 750 bp to about 800 bp, about 800 bp to about 1000 bp, about 800 bp to about 900 bp, about 800 bp to about 850 bp, about 850 bp to about 1000 bp, about 850 bp to about 900 bp, about 900 bp to about 1000 bp, about 900 to about 950 bp, or about 950 bp to about 1000 bp). SEQ ID NOs. 19, 35 and 37 are exemplary human factor FVIII stuffer sequences that can be used in any of the vectors described herein. Additional stuffer sequences are known in the art. Exemplary vectors that include stuffer sequences are shown in Figures 1, 2 and 4.

Mammalian Cells

Also provided herein is a cell (e.g., a mammalian cell) that includes any of the nucleic acids, vectors (e.g., at least two different vectors described herein), or compositions described herein. Skilled practitioners will appreciate that the nucleic acids and vectors described herein can be introduced into any mammalian cell. Non-limiting examples of vectors and methods for introducing vectors into mammalian cells are described herein.

In some embodiments, the cell is a human cell, a mouse cell, a porcine cell, a rabbit cell, a dog cell, a cat cell, a rat cell, or a non-human primate cell. In some embodiments, the cell is a specialized cell of the cochlea. In some embodiments, the cell is a cochlear inner hair cell or a cochlear outer hair cell. In some embodiments, the cell is a cochlear inner hair cell. In some embodiments, the cell is a cochlear outer hair cell.

In some embodiments, the mammalian cell is *in vitro*. In some embodiments, the mammalian cell is present in a mammal. In some embodiments, the mammalian cell is autologous cell obtained from a subject and cultured *ex vivo*.

Methods

Also provided herein is a method of introducing into a cochlea of a mammal (e.g., a human) a therapeutically effective amount of any of the compositions described herein. Also provided are methods of increasing expression of an active stereocilin protein (e.g., a full-length stereocilin protein) in an outer hair cell in a cochlea of a mammal (e.g., a human) that include introducing into the cochlea of the mammal a therapeutically effective amount of any of the compositions described herein. Also provided are methods of treating non-symptomatic sensorineural hearing loss in a subject (e.g., a human) identified as having a defective stereocilin gene, where the methods include administering a therapeutically effective amount of any of the compositions described herein into the cochlea of a subject.

In some embodiments of any of these methods, the mammal has been previously identified as having a defective stereocilin gene (e.g., a stereocilin gene having a mutation that results in a decrease in the expression and/or activity of a stereocilin protein encoded by the gene). Some embodiments of any of these methods further include, prior to the introducing or administering step, determining that the subject has a defective stereocilin gene. Some embodiments of any of these

methods can further include detecting a mutation in a stereocilin gene in a subject. Some embodiments of any of the methods can further include identifying or diagnosing a subject as having non-symptomatic sensorineural hearing loss.

In some embodiments of any of these methods, two or more doses of any of the compositions described herein are introduced or administered into the cochlea of the mammal or subject. Some embodiments of any of these methods can include introducing or administering a first dose of the composition into the cochlea of the mammal or subject, assessing hearing function of the mammal or subject following the introducing or the administering of the first dose, and administering an additional dose of the composition into the cochlea of the mammal or subject found not to have a hearing function within a normal range (e.g., as determined using any test for hearing known in the art).

In some embodiments of any of the methods described herein, the composition can be formulated for intra-cochlear administration. In some embodiments of any of the methods described herein, the compositions described herein can be administered via intra-cochlear administration or local administration. In some embodiments of any of the methods described herein, the compositions are administered through the use of a medical device (e.g., any of the exemplary medical devices described herein).

In some embodiments, intra-cochlear administration can be performed using any of the methods described herein or known in the art. For example, a composition can be administered or introduced into the cochlea using the following surgical technique: first using visualization with a 0 degree, 2.5-mm rigid endoscope, the external auditory canal is cleared and a round knife is used to sharply delineate an approximately 5-mm tympanomeatal flap. The tympanomeatal flap is then elevated and the middle ear is entered posteriorly. The chorda tympani nerve is identified and divided, and a curette is used to remove the scutal bone, exposing the round window membrane. To enhance apical distribution of the administered or introduced composition, a surgical laser may be used to make a small 2-mm fenestration in the oval window to allow for perilymph displacement during trans-round window membrane infusion of the composition. The microinfusion device is then primed and brought into the surgical field. The device is maneuvered to the round window, and the tip is seated within the bony round window overhang to allow for penetration of the membrane by the microneedle(s). The footpedal is engaged to allow for a

measured, steady infusion of the composition. The device is then withdrawn and the round window and stapes foot plate are sealed with a gelfoam patch.

In some embodiments of any of the methods described herein, the subject or mammal is a rodent, a non-human primate, or a human. In some embodiments of any of the methods described herein, the subject or mammal is an adult, a teenager, a juvenile, a child, a toddler, an infant, or a newborn. In some embodiments of any of the methods described herein, the subject or mammal is 1-5, 1-10, 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 1-110, 2-5, 2-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-110, 10-30, 10-40, 10-50, 10-60, 10-70, 10-80, 10-90, 10-100, 10-110, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 20-110, 30-50, 30-60, 30-70, 30-80, 30-90, 30-100, 40-60, 40-70, 40-80, 40-90, 40-100, 50-70, 50-80, 50-90, 50-100, 60-80, 60-90, 60-100, 70-90, 70-100, 70-110, 80-100, 80-110, or 90-110 years of age. In some embodiments of any of the methods described herein, the subject or mammal is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 months of age.

In some embodiments of any of the methods described herein, the subject or mammal has or is at risk of developing non-syndromic sensorineural hearing loss. In some embodiments of any of the methods described herein, the subject or mammal has been previously identified as having a mutation in a stereocilin gene. (All references here to “stereocilin gene” refer to the gene that is not the stereocilin pseudogene.) In some embodiments of any of the methods described herein, the subject or mammal has any of the mutations in a stereocilin gene that are described herein or are known in the art to be associated with non-symptomatic sensorineural hearing loss.

In some embodiments of any of the methods described herein, the subject or mammal has been identified as being a carrier of a mutation in a stereocilin gene (e.g., via genetic testing). In some embodiments of any of the methods described herein, the subject or human has been identified as having a mutation in a stereocilin gene and has been diagnosed with non-symptomatic sensorineural hearing loss. In some embodiments of any of the methods described herein, the subject or human has been identified as having non-symptomatic sensorineural hearing loss.

In some embodiments, successful treatment of non-symptomatic sensorineural hearing loss can be determined in a subject using any of the conventional functional hearing tests known in the art. Non-limiting examples of functional hearing tests are various types of audiometric assays (e.g., pure-tone testing, speech testing, test of the

middle ear, auditory brainstem response, and otoacoustic emissions).

Also provided herein are methods of increasing expression of an active stereocilin protein (e.g., a full-length stereocilin protein) in a mammalian cell that include introducing any of the compositions described herein into the mammalian
5 cell. In some embodiments of these methods, the mammalian cell is a cochlear outer hair cell. In some embodiments of these methods, the mammalian cell is a human cell (e.g., a human cochlear outer hair cell). In some embodiments of these methods, the mammalian cell is *in vitro*. In some embodiments of these methods, the mammalian cell is in a mammal. In some embodiments of these methods, the mammalian cell is
10 originally obtained from a mammal and is cultured *ex vivo*. In some embodiments, the mammalian cell has previously been determined to have a defective stereocilin gene.

Methods for introducing any of the compositions described herein into a mammalian cell are known in the art (e.g., via lipofection or through the use of a viral
15 vector, e.g., any of the viral vectors described herein).

An increase in expression of an active stereocilin protein (e.g., a full-length stereocilin protein) as described herein is, e.g., as compared to a control or to the level of expression of an active stereocilin protein (e.g., a full-length stereocilin protein) prior to the introduction of the vector(s).

20 Methods of detecting expression and/or activity of stereocilin are known in the art. In some embodiments, the level of expression of a stereocilin protein can be detected directly (e.g., detecting stereocilin protein or detecting stereocilin mRNA). Non-limiting examples of techniques that can be used to detect expression and/or activity of stereocilin directly include: real-time PCR, Western blotting,
25 immunoprecipitation, immunohistochemistry, or immunofluorescence. In some embodiments, expression of a stereocilin protein can be detected indirectly (e.g., through functional hearing tests).

Pharmaceutical Compositions and Kits

30 In some embodiments, any of the compositions described herein can further include one or more agents that promote the entry of a nucleic acid or any of the vectors described herein into a mammalian cell (e.g., a liposome or cationic lipid). In some embodiments, any of the vectors described herein can be formulated using

natural and/or synthetic polymers. Non-limiting examples of polymers that may be included in any of the compositions described herein can include, but are not limited to, DYNAMIC POLYCONJUGATE® (Arrowhead Research Corp., Pasadena, Calif.), formulations from Mirus Bio (Madison, Wis.) and Roche Madison (Madison, Wis.), PhaseRX polymer formulations such as, without limitation, SMARTT POLYMER TECHNOLOGY® (PhaseRX, Seattle, Wash.), DMRI/DOPE, poloxamer, VAXFECTIN® adjuvant from Vical (San Diego, Calif.), chitosan, cyclodextrin from Calando Pharmaceuticals (Pasadena, Calif.), dendrimers and poly (lactic-co-glycolic acid) (PLGA) polymers, RONDEL™ (RNAi/Oligonucleotide Nanoparticle Delivery) polymers (Arrowhead Research Corporation, Pasadena, Calif.), and pH responsive co-block polymers, such as, but not limited to, those produced by PhaseRX (Seattle, Wash.). Many of these polymers have demonstrated efficacy in delivering oligonucleotides *in vivo* into a mammalian cell (see, e.g., deFougerolles, *Human Gene Ther.* 19:125-132, 2008; Rozema et al., *Proc. Natl. Acad. Sci. U.S.A.* 104:12982-12887, 2007; Rozema et al., *Proc. Natl. Acad. Sci. U.S.A.* 104:12982-12887, 2007; Hu-Lieskovan et al., *Cancer Res.* 65:8984-8982, 2005; Heidel et al., *Proc. Natl. Acad. Sci. U.S.A.* 104:5715-5721, 2007).

Any of the compositions described herein can be, e.g., a pharmaceutical composition. A pharmaceutical composition can include any of the compositions described herein and one or more pharmaceutically or physiologically acceptable carriers, diluents, or excipients. Such compositions may comprise one or more buffers, such as neutral-buffered saline, phosphate-buffered saline, and the like; one or more carbohydrates, such as glucose, mannose, sucrose, and dextran; mannitol; one or more proteins, polypeptides, or amino acids, such as glycine; one or more antioxidants; one or more chelating agents, such as EDTA or glutathione; and/or one or more preservatives.

In some embodiments, the composition includes a pharmaceutically acceptable carrier (e.g., phosphate buffered saline, saline, or bacteriostatic water). Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, injectable gels, drug-release capsules, and the like.

As used herein, the term “pharmaceutically acceptable carrier” includes solvents, dispersion media, coatings, antibacterial agents, antifungal agents, and the

like that are compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into any of the compositions described herein.

In some embodiments, a single dose of any of the compositions described herein can include a total sum amount of the at least two different vectors of at least 1 ng, at least 2 ng, at least 4 ng, about 6 ng, about 8 ng, at least 10 ng, at least 20 ng, at least 30 ng, at least 40 ng, at least 50 ng, at least 60 ng, at least 70 ng, at least 80 ng, at least 90 ng, at least 100 ng, at least 200 ng, at least 300 ng, at least 400 ng, at least 500 ng, at least 1 μ g, at least 2 μ g, at least 4 μ g, at least 6 μ g, at least 8 μ g, at least 10 μ g, at least 12 μ g, at least 14 μ g, at least 16 μ g, at least 18 μ g, at least 20 μ g, at least 22 μ g, at least 24 μ g, at least 26 μ g, at least 28 μ g, at least 30 μ g, at least 32 μ g, at least 34 μ g, at least 36 μ g, at least 38 μ g, at least 40 μ g, at least 42 μ g, at least 44 μ g, at least 46 μ g, at least 48 μ g, at least 50 μ g, at least 52 μ g, at least 54 μ g, at least 56 μ g, at least 58 μ g, at least 60 μ g, at least 62 μ g, at least 64 μ g, at least 66 μ g, at least 68 μ g, at least 70 μ g, at least 72 μ g, at least 74 μ g, at least 76 μ g, at least 78 μ g, at least 80 μ g, at least 82 μ g, at least 84 μ g, at least 86 μ g, at least 88 μ g, at least 90 μ g, at least 92 μ g, at least 94 μ g, at least 96 μ g, at least 98 μ g, at least 100 μ g, at least 102 μ g, at least 104 μ g, at least 106 μ g, at least 108 μ g, at least 110 μ g, at least 112 μ g, at least 114 μ g, at least 116 μ g, at least 118 μ g, at least 120 μ g, at least 122 μ g, at least 124 μ g, at least 126 μ g, at least 128 μ g, at least 130 μ g, at least 132 μ g, at least 134 μ g, at least 136 μ g, at least 138 μ g, at least 140 μ g, at least 142 μ g, at least 144 μ g, at least 146 μ g, at least 148 μ g, at least 150 μ g, at least 152 μ g, at least 154 μ g, at least 156 μ g, at least 158 μ g, at least 160 μ g, at least 162 μ g, at least 164 μ g, at least 166 μ g, at least 168 μ g, at least 170 μ g, at least 172 μ g, at least 174 μ g, at least 176 μ g, at least 178 μ g, at least 180 μ g, at least 182 μ g, at least 184 μ g, at least 186 μ g, at least 188 μ g, at least 190 μ g, at least 192 μ g, at least 194 μ g, at least 196 μ g, at least 198 μ g, or at least 200 μ g, e.g., in a buffered solution.

The compositions provided herein can be, e.g., formulated to be compatible with their intended route of administration. A non-limiting example of an intended route of administration is local administration (e.g., intra-cochlear administration).

In some embodiments, the therapeutic compositions are formulated to include a lipid nanoparticle. In some embodiments, the therapeutic compositions are formulated to include a polymeric nanoparticle. In some embodiments, the therapeutic compositions are formulated to comprise a mini-circle DNA. In some embodiments, the therapeutic compositions are formulated to comprise a CELiD

DNA. In some embodiments, the therapeutic compositions are formulated to comprise a synthetic perilymph solution. An exemplary synthetic perilymph solution includes 20-200 mM NaCl; 1-5 mM KCl; 0.1-10 mM CaCl₂; 1-10 mM glucose; 2-50 mM HEPES, having a pH of between about 6 and about 9.

5 Also provided are kits including any of the compositions described herein. In some embodiments, a kit can include a solid composition (e.g., a lyophilized composition including the at least two different vectors described herein) and a liquid for solubilizing the lyophilized composition. In some embodiments, a kit can include a pre-loaded syringe including any of the compositions described herein.

10 In some embodiments, the kit includes a vial comprising any of the compositions described herein (e.g., formulated as an aqueous composition, e.g., an aqueous pharmaceutical composition).

In some embodiments, the kits can include instructions for performing any of the methods described herein.

15

Routes of Administration

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. Dispersions may also be prepared in
20 glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms. In many cases the form is sterile and fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of
25 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
30 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.

For administration of an injectable aqueous solution, for example, the solution
5 may be suitably buffered, if necessary, and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art. For example, one dosage may be dissolved
10 in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the host. The person responsible for administration will, in any event, determine the appropriate
15 dose for the individual host.

Sterile injectable solutions are prepared by incorporating the active AAV in the required amount in the appropriate solvent with various of the other ingredients enumerated herein, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into
20 a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered
25 solution thereof.

Any of the compositions disclosed herein may also be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as
30 acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation

and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a host.

Delivery vehicles such as liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, may be used for the introduction of the compositions of the present invention into suitable host cells. In particular, the vector delivered transgenes may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically acceptable formulations of the nucleic acids or any of the vectors disclosed herein. The formation and use of liposomes is generally known to those of skill in the art. Recently, liposomes were developed with improved serum stability and circulation half-times (U.S. Pat. No. 5,741,516). Further, various methods of liposome and liposome like preparations as potential drug carriers have been described (U.S. Pat. Nos. 5,567,434; 5,552,157; 5,565,213; 5,738,868 and 5,795,587).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures. In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs, radiotherapeutic agents, viruses, transcription factors and allosteric effectors into a variety of cultured cell lines and animals. In addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have been completed.

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)). MLVs generally have diameters of from 25 nm to 4 μ m. Sonication of MLVs results in the formation of small unilamellar

vesicles (SUVs) with diameters in the range of 200 to 500 .ANG., containing an aqueous solution in the core.

Alternatively, nanocapsule formulations of any of the vectors described herein may be used. Nanocapsules can generally entrap substances in a stable and
5 reproducible way. To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded in vivo. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use.

In addition to the methods of delivery described above, the following
10 techniques are also contemplated as alternative methods of delivering any of the compositions described herein to a host. Sonophoresis (ie., ultrasound) has been used and described in U.S. Pat. No. 5,656,016 as a device for enhancing the rate and efficacy of drug permeation into and through the circulatory system. Other drug delivery alternatives contemplated are intraosseous injection (U.S. Pat. No.
15 5,779,708), microchip devices (U.S. Pat. No. 5,797,898), ophthalmic formulations (Bourlais et al., 1998), transdermal matrices (U.S. Pat. Nos. 5,770,219 and 5,783,208) and feedback-controlled delivery (U.S. Pat. No. 5,697,899).

The administration of the subject compositions may be carried out in any convenient manner, including by aerosol inhalation, injection, ingestion, transfusion,
20 implantation or transplantation. The compositions described herein may be administered to a patient trans arterially, subcutaneously, intradermally, intranodally, intramedullary, intramuscularly, by intravenous (i.v.) injection, or intraperitoneally. In one aspect, the nucleic acid compositions of the present invention are administered to a patient by intradermal or subcutaneous injection. In one aspect, the nucleic
25 compositions of the present invention are administered by i.v. injection.

Devices and Surgical Methods

Provided herein are therapeutic delivery systems for treating non-symptomatic sensorineural hearing loss. In one aspect, the therapeutic delivery systems include i) a
30 medical device capable of creating one or a plurality of incisions in a round window membrane of an inner ear of a human subject in need thereof, and ii) an effective dose of a composition (e.g., any of the compositions described herein). In some embodiments, the medical device includes a plurality of micro-needles.

Also provided herein are surgical methods for treatment of hearing loss (e.g., non-symptomatic sensorineural hearing loss). In some embodiments, the methods include the steps of: introducing into a cochlea of a human subject a first incision at a first incision point; and administering intra-cochlearly a therapeutically effective
5 amount of any of the compositions provided herein. In some embodiments, the composition is administered to the subject at the first incision point. In some embodiments, the composition is administered to the subject into or through the first incision.

In some embodiments of any of the methods described herein, any of the
10 compositions described herein is administered to the subject into or through the cochlea oval window membrane. In some embodiments of any of the methods described herein, any of the compositions described herein is administered to the subject into or through the cochlea round window membrane. In some embodiments of any of the methods described herein, the composition is administered using a
15 medical device capable of creating a plurality of incisions in the round window membrane. In some embodiments, the medical device includes a plurality of micro-needles. In some embodiments, the medical device includes a plurality of micro-needles including a generally circular first aspect, where each micro-needle has a diameter of at least about 10 microns. In some embodiments, the medical device
20 includes a base and/or a reservoir capable of holding the composition. In some embodiments, the medical device includes a plurality of hollow micro-needles individually including a lumen capable of transferring the composition. In some embodiments, the medical device includes a means for generating at least a partial vacuum.

25 The invention is further described in detail by reference to the following experimental examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified. Thus, the invention should in no way be construed as being limited to the following examples, but rather should be construed to encompass any and all variations that become evident as a
30 result of the teaching provided herein.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The following working examples specifically point out various aspects of the present

invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXAMPLES

5 **Example 1: Construction of Viral Vectors**

Recombinant AAV is generated by transfection with an adenovirus-free method as used by Xiao et al. *J. Virol.* 73(5):3994-4003, 1999. The cis plasmids with AAV ITRs, the trans plasmid with AAV Rep and Cap genes, and a helper plasmid with an essential region from an adenovirus genome are co-transfected in 293 cells in
10 a ratio of 1:1:2. The AAV vectors used here express human stereocilin or mouse stereocilin under multiple dual vector strategies using the constructs described below. AAV serotypes 1, 2, 3, 4, 5, 6, 7, 8, 9, rh8, rh10, rh39, rh43, and Anc80 are each prepared to encapsulate three sets of stereocilin constructs to test (i) a concatemerization-transplicing strategy, (ii) a hybrid intronic-homologous
15 recombination-transplicing strategy, and (iii) an exonic homologous recombination strategy, as summarized by Pryadkina et al., *Meth. Clin. Devel.* 2:15009, 2015.

Example 2: Generating and Purifying Viral Particles

Recombinant AAV-1 is produced using a triple transfection protocol and
20 purified by two sequential cesium chloride (CsCl) density gradients, as described by Pryadkina et al., *Mol. Ther.* 2:15009, 2015. At the end of second centrifugation, 11 fractions of 500 µl are recovered from the CsCl density gradient tube and purified through dialysis in 1× PBS. The fractions are analyzed by dot blot to determine those containing rAAV genomes. The viral genome number (vg) of each preparation is
25 determined by a quantitative real-time PCR-based titration method using primers and probe corresponding to the ITR region of the AAV vector genome (Bartoli et al. *Gene. Ther.* 13:20-28, 2006).

Example 3: Formulation of Viral Particles

30 AAV produced at a titer of 1e14 vg / mL is prepared at dilutions of 3.2e13, 1.0e13, 3.2e12, 1.0e12 vg/mL in artificial perilymph. Artificial perilymph is prepared by combining the following reagents: NaCl, 120 mM; KCl, 3.5 mM; CaCl₂, 1.5 mM; glucose, 5.5 mM; HEPES, 20 mM. The artificial perilymph is titrated with NaOH to

adjust its pH to 7.5 (total Na^+ concentration of 130 mM) (Chen et al., *J. Controlled Rel.* 110:1-19, 2005).

Example 4: Device Description

5 The AAV- STRC formulation is delivered to the cochlea using a specialized microcatheter designed for consistent and safe penetration of the round window membrane (RWM). The microcatheter is shaped such that the surgeon performing the delivery procedure can enter the middle ear cavity via the external auditory canal and contact the end of the microcatheter with the RWM. The distal end of the
10 microcatheter is comprised of at least one microneedle with diameter between 10 and 1,000 microns, which produces perforations in the RWM that are sufficient to allow AAV-STRC to enter the cochlear perilymph of the scala tympani at a rate of approximately 1 $\mu\text{L}/\text{min}$, but small enough to heal without surgical repair. The remaining portion of the microcatheter, proximal to the microneedle(s), is loaded with
15 the AAV- STRC /artificial perilymph formulation at a titer of approximately 1×10^{13} vg/mL. The proximal end of the microcatheter is connected to a micromanipulator that allows for precise, low volume infusions of approximately 1 μL / min.

Example 5: Animal Model 1A: Surgical Method in Aged Mice

20 AAV- STRC prepared in artificial perilymph is administered to the scala tympani in mice as described by Shu et al. (*Human Gene Therapy*, doi:10.1089/hum.2016.053, June 2016,). Six-week-old male mice are anesthetized using an intraperitoneal injection of xylazine (20 mg/kg) and ketamine (100 mg/kg). Body temperature is maintained at 37 °C using an electric heating pad. An incision is
25 made from the right post-auricular region and the tympanic bulla is exposed. The bulla is perforated with a surgical needle and the small hole is expanded to provide access to the cochlea. The bone of the cochlear lateral wall of the scala tympani is thinned with a dental drill so that the membranous lateral wall is left intact. A Nanoliter Microinjection System in conjunction with glass micropipette is used to
30 deliver a total of approximately 300 nL of AAV- STRC in artificial perilymph to the scala tympani at a rate of 2 nL/second. The glass micropipette is left in place for 5 minutes post-injection. Following cochleostomy and injection, the opening in the tympanic bulla is sealed with dental cement, and the muscle and skin are sutured. The

mice are allowed to awaken from anesthesia and their pain is controlled with 0.15 mg/kg buprenorphine hydrochloride for 3 days.

Example 6: Animal Model 2: Reciprocating Micropump in Guinea Pig

5 *Surgical Procedure*

AAV- STRC prepared in artificial perilymph is administered to guinea pigs to assess distribution and toxicity following intracochlear delivery with a reciprocating micropump as described by Tandon et al., *Lab Chip*, DOI: 10.1039/c5lc01396h, 2015. Male guinea pigs weighing approximately 350 g each (n=16) are anesthetized with a
 10 combination of pentobarbital sodium (Nembutal; 25 mg kg⁻¹, injected intraperitoneally), fentanyl (0.2 mg kg⁻¹, intramuscularly), and haloperidol (10 mg kg⁻¹, intramuscularly). Lidocaine with epinephrine is given subcutaneously at the incision site as a topical anesthetic. Using a dorsal approach, a 5 mm diameter hole is made in the bulla and a cochleostomy is created approximately 0.5 mm distal to the
 15 round window membrane. The cannula of the micropump (described below) is inserted into the cochleostomy, threaded into the cochlea 3 mm apically, and glued to the bulla with a common cyanoacrylate glue. For compound action potential (CAP) measurements, a perfluoroalkoxy-alkane-insulated silver wire electrode (203 μm uncoated diameter) is inserted near the round window niche and glued to the bulla.

20 Procedures for measurement of distortion product otoacoustic emissions (DPOAEs) and CAPs are performed as previously described in Tandon et al. *Biomed Microdevices* 17:3-21, 2015. DPOAEs are measured before and after the cochleostomy procedure at the characteristic frequencies: 32, 24, 16, 12, 8, 5.6, 4, and 2.78 kHz in order to monitor any damage that occurs as a result of the surgery.

25 AAV- STRC at a maximum titer of 1e14 vg/mL is administered to the guinea pig using a micropump as described by Tandon et al. *Lab Chip*, DOI: 10.1039/c5lc01396h, 2015. The micropump system has 4 selectable ports. These ports are connected to: (i) a large fluidic capacitor used for artificial perilymph storage; (ii) an outlet that connects to the cochlea; (iii) the outlet from an integrated
 30 AAV- STRC reservoir; (iv) the inlet to the integrated AAV- STRC reservoir. Each port is fluidically connected to a central pump chamber, and each is individually addressed with a valve. The sequence of events for reciprocating AAV- STRC delivery is as follows: (i) an internal AAV- STRC -refresh loop is run, transferring

AAV- STRC from the AAV- STRC reservoir into the main infuse–withdraw line; (ii) AAV- STRC is infused into the cochlea and some artificial perilymph is drained from the artificial perilymph storage capacitor; (iii) the first two steps can be repeated several times for additional doses; (iv) after the AAV- STRC has been allowed to
 5 diffuse for some time, a volume of perilymph is withdrawn from the cochlea that is equal to the volume infused in steps (i)–(iii), refilling the artificial perilymph storage capacitor. This process results in net delivery of drug with zero net fluid volume added to the cochlea.

The fluidic capacitors in the micropump are cylindrical chambers whose
 10 ceilings are a thin (25.4 μm), flexible, polyimide membrane. The pump chamber has a diameter of 3.5 mm, the fluidic storage capacitor has a diameter of 14 mm, and all of the remaining capacitors have diameters of 4 mm. The same membrane is deflected to block flow at each of the valves. The valve chambers have diameters of 3.1 mm. The serpentine channel that comprises the drug reservoir has a square cross section of
 15 width 762 μm and a length of 410 mm for a total volume of 238 μL . All of the other microchannels in the pump have a width of 400 μm and a height of 254 μm . *Acute drug delivery in guinea pigs*

The micropump is loaded with AAV- STRC and artificial perilymph, and the cannula inserted into a cochleostomy made in the region of the cochlea between the
 20 locations with characteristic frequency sensitivity of 24 and 32 kHz, and threaded apically 3 mm, terminating in the 12–16 kHz region. Baseline DPOAE and CAP hearing tests are performed prior to the start of AAV- STRC /artificial perilymph infusion. The pump is then activated and approximately 1 μL of artificial perilymph is infused every 5 min until a total of approximately 10 μL of artificial perilymph is
 25 delivered to the cochlea. After a 20 min wait time, approximately 10 μL of perilymph is withdrawn from the cochlea. AAV- STRC delivery is then initiated at a rate of approximately 1 μL every 5 min until a total of approximately 10 μL of fluid delivered.

Animals are sacrificed at 1 week, 1 month, 3 months, and 6 months post-
 30 treatment (n=4 per group) and their cochleae extracted. Extent of AAV transduction and STRC expression along the organ of Corti is assessed via immunostaining with anti- STRC antibodies. Antibodies against markers for hair cells (Myo7a) and supporting cells (Sox2) are used to quantify IHCs, OHCs, supporting cells and

stereocilia morphology. Annexin V staining is used to assess evidence of apoptosis in cells along the cochlear sensory epithelium.

Example 7: Animal Model 3: Sheep

5 AAV- STRC prepared in artificial perilymph is administered to juvenile sheep to assess distribution and toxicity following delivery to the cochlea via trans-RWM infusion. Baseline auditory brainstem response (ABR) and distortion product optoacoustic emissions (DPOAEs) are measured in female sheep at 3 months of age (n=40), bilaterally, to assess pre-treatment inner hair cell (IHC) and outer hair cell
10 (OHC) function. Following baseline ABR and DPOAE measurements, 20 uL of AAV1- STRC at titers of 1.0e14, 3.2e13, 1.0e13 and 3.2e12 vg/mL is injected into the left scala tympani of the sheep (n=10 per group). Each animal's right ear is left as an untreated control. ABR and DPOAE measurements are taken again bilaterally 1, 5 and 10 days following the surgical procedure. At 6 months post-procedure, additional
15 bilateral ABR and DPOAE measurements are taken from all animals, and the animals are subsequently sacrificed and their cochleae removed.

In half of the sacrificed animals (n=5 from each of the dose cohorts), immunostaining is performed to identify hair cell structures and to assess STRC protein expression along the cochlear sensory epithelium. Antibodies against markers
20 for hair cells (Myo7a), supporting cells (Sox2) and stereocilin are used as described previously (Duncker et al. 2013, *J Neurosci* 33(22):9508-9519). At the basal, middle and apical turns of the organ of corti, total numbers of hair cells and those hair cells expressing STRC are counted within 200 um regions.

In the remaining half of the sacrificed animals (remaining 5 animals from each
25 dose cohort), cochlear tissue samples are collected from the same basal, middle and apical regions as described above, and assayed for stereocilin mRNA transcript.

Example 8: Animal Model 3A: CRISPR Generated Transgenic Large Animal Model (Sheep)

30

Generation of plasmid co-expressing Cas9 and sgRNA

The pX330-U6-Chimeric_BB-CBh-hSpCas9 plasmid (Addgene plasmid #42230) is digested with BsbI, dephosphorylated using Antarctic Phosphatase, and the

linearized vector is gel purified. To generate the bicistronic vector (pX330-cas9-STRC) expressing Cas9 and sgRNA against STRC, a pair of oligos for targeting stereocilin exon 1 is annealed, phosphorylated and ligated to a linearized vector (Cong et al., *Science* 339(6121):819-23, 2013).

5

Genome editing assay in cells

The A15 astroglial sheep cell line (Vilette et al., 2000 *In Vitro Cell Dev Biol Anim* 36(1):45-9) is maintained in DMEM in 10% Fetal Bovine Serum, 2mM glutamine, 1% sodium pyruvate and 1% penicillin/streptomycin. Cells are transfected in 24-well plates with 2 µg of pX330-cas9- STRC co-expressing Cas9 and sgRNA against stereocilin using lipofectamine LTX reagent. Three days later, genomic DNA from transfected cells is extracted and quantified using a NanoDrop2000 spectrophotometer, measuring A260/A280 and A260/A230 ratios to account for sample purity.

Gene mutation activity of sgRNA sequence at the target locus of STRC exon 1 is quantified using the T7EI mismatch detection assay. DNA sequence of interest is PCR-amplified with a high-fidelity polymerase (Herculase II fusion polymerase) using specific primers. The resultant PCR product is then denatured and slowly re-annealed (95°C, 2 min; 95°C to 85°C, -2°C/sec; 85°C to 25°C, -1°C/sec) to produce a homoduplex/ heteroduplex mix. This is then digested by 5U of T7EI restriction enzyme at 37°C for 30 minutes. Digestion products are separated by 2% agarose gel electrophoresis. The ratio of cleaved to uncleaved products is used to calculate NHEJ frequency as previously described using Image J software (Menoret et al. 2011 *Advanced protocols for Animal Transgenesis. An ISTT Manual*. Heidelberg: Springer. p117-36). NHEJ frequency is calculated as % gene modification = $100 \times (1 - (1 - \text{fraction cleaved})^{1/2})$.

25

Production of sgRNA and Cas9 mRNA

As described previously (Bellec et al., *Current Gene Ther.*, 2015), T7 promoter is added to sgRNA template by PCR amplification of the pX330-cas9-STRC plasmid. The PCR product is purified using NucleoSpin® Gel and PCR Clean-up. It is used as the template for *in vitro* transcription using MEGAshortscript™ T7 kit according to the manufacturer's manual. Following completion of transcription, DNase I treatment is performed.

30

The Cas9 mRNA is transcribed using a PmeI-digested Cas9 expression JDS246 plasmid (Addgene plasmid # 43861) and the mMESSAGE mMACHINE® T7 ULTRA Transcription Kit according to the manufacturer's manual. Following completion of transcription, the poly(A) tailing reaction and DNase I treatment are performed. Both the Cas9 mRNA and the sgRNAs are purified using a MEGAclean™ kit and eluted in elution buffer.

In vitro production of embryos

The sheep embryos are produced by *in vitro* fertilization according to routine procedure as described previously (Crispo et al. 2014 *Transgenic Res*, 24(1):31-41). Briefly, sheep ovaries from a slaughterhouse are transported to the laboratory and cumulus oocyte complexes (COCs) are aspirated in recovery medium. The selected COCs are placed in maturation medium for 24 h in 5% CO₂ in humidified air atmosphere at 39°C. Then, expanded COCs are inseminated in 100 µl drops with 1 x 10⁶ dose of frozen-thawed semen selected by ascendant migration on a swim up method. Fertilization is carried out in 5% CO₂ with humidified atmosphere at 39°C for 22 h.

Microinjection into zygotes

Soon after fertilization, 600 presumptive zygotes are randomly assigned to three experimental groups to be microinjected (CRISPR group, n = 200; and Buffer group, n = 200) or not (Control group, n = 200). Microinjection of the CRISPR group is performed by microinjection into the cytoplasm with 5 ng/µl of sgRNA and 20ng/µl of Cas9 mRNA diluted in injection buffer (10mM Tris pH 7.5, 0.1mM EDTA), while the Buffer group is injected with the same procedure but with buffer alone. Lastly, injected and non-injected embryos are transferred to culture medium under mineral oil, in 5% CO₂, 5% O₂ and 90% N₂ in humidified atmosphere at 39°C. Cleavage rate on Day 2 (cleaved zygotes per total oocytes) and development rate on Day 6 (morulae and blastocysts per total oocytes) are recorded for all experimental groups. After Day 6, DNA from 20 CRISPR group embryos are analyzed by Sanger sequencing to detect the mutation at the STRC gene level.

To determine the *in vivo* efficiency of the system, blastocysts produced by CRISPR/Cas9 zygote microinjection are transferred to recipient females. Only early blastocysts, blastocysts and expanded blastocysts classified as excellent or good (i.e.

Grade 1 as defined in Stringfellow et al. 2010, *Manual of the International Embryo Transfer Society*) are transferred on Day 6 after fertilization. Embryo transfer is performed by minimally invasive surgery assisted by laparoscopy to place the embryos into the cranial side of the ipsilateral uterine horn to the corpus luteum.

- 5 Recipient ewes are previously synchronized to be on Day 6 of the estrous cycle using a standard protocol to control ovulation described previously, as described by Menchaca et al. 2004, *Reprod Fertil Dev.* 16(4):403-413.

Monitoring of fetuses and lambs

- 10 Pregnancy diagnosis and fetal development are performed on Day 30 and 105, respectively, by using B-mode ultrasonography equipped with a 5 and 3.5 MHz probe. Day 0 of the experiment is defined as the moment of embryo fertilization. Several parameters are measured to study the development of fetuses at Day 105 of gestation: thoracic diameter, biparietal diameter, occipitonasal length and heart rate.
- 15 At delivery, length of gestation, gender, rectal temperature, heart and respiratory rates, body weight, thoracic perimeter, biparietal diameter, crown-rump and occipitonasal length, height at withers, height at hips, width at hips and width at chest were recorded. Body weight and morphometric variables are determined at birth, and 15, 30 and 60 days later.

20

Identification and genotyping of transgenic animals

- Samples from skin and limb muscle of the lambs are taken seven days after birth and T7EI assay, western blot test and histology examinations are performed in order to identify and characterize KO founders and off-target sites. Total DNA is
- 25 isolated from skin biopsies for all animals and from muscle for some animals. Samples are analyzed using capillary electrophoresis. Genotyping of STRC exon 1 is performed by direct sequencing of PCR amplicons and in muscle biopsies by additional sequencing of isolated bacterial clones with individual amplicon sequences.

30 *Analysis of stereocilin expression*

Western blotting is performed to determine the presence of myostatin in the muscle fiber. Equal amounts of total proteins are run on 12% (v/v) gel electrophoresis and electrophoretically transferred to a PVDF membrane. Monoclonal mouse anti-stereocilin antibody is used in the western blotting. The washed membranes are

incubated with 1:50000 dilution of secondary antibody linked to horseradish peroxidase (HPR). HPR activity is detected using western blot chemiluminescence.

AAV- STRC rescue therapy in transgenic sheep model

5 AAV- STRC prepared in artificial perilymph is administered to STRC knockout transgenic sheep to assess the ability to restore normal hearing function following delivery to the cochlea via trans-RWM infusion. Baseline auditory brainstem response (ABR) and distortion product otoacoustic emissions (DPOAEs) are measured in female sheep at 3 months of age (n=30), bilaterally, to assess pre-
10 treatment inner hair cell (IHC) and outer hair cell (OHC) function. Following baseline ABR and DPOAE measurements, 20 uL of AAV1- STRC at titers of 1.0e14, 3.2e13, and 1.0e13 vg/mL is injected into the left scala tympani of the sheep (n=10 per group). Each animal's right ear is left as an untreated control. ABR and DPOAE measurements are taken again bilaterally 1, 5 and 10 days following the surgical
15 procedure. At 6 months post-procedure, additional bilateral ABR and DPOAE measurements are taken from all animals, and the animals are subsequently sacrificed and their cochleae removed.

 In half of the sacrificed animals (n=5 from each of the dose cohorts), immunostaining is performed to identify hair cell structures and to assess STRC
20 protein expression along the cochlear sensory epithelium. Antibodies against markers for hair cells (Myo7a), supporting cells (Sox2) and stereocilin are used as described previously (Duncker et al., *J Neurosci* 33(22):9508-9519, 2013). At the basal, middle and apical turns of the organ of corti, total numbers of hair cells and those hair cells expressing STRC are counted within 200 um regions.

25 In the remaining half of the sacrificed animals (remaining 5 animals from each dose cohort), cochlear tissue samples are collected from the same basal, middle and apical regions as described above, and assayed for stereocilin mRNA transcript as described previously (Duncker et al. 2013, *J Neurosci* 33(22):9508-9519, 2013; Heidrych et al., *Hum. Mol. Genet.* 17:3814-3821, 2008; Heidrych et al., *Hum. Mol.*
30 *Genet.* 18:2779-2790, 2009).

Example 9: Human Clinical Example (Pediatric Treatment)

The patient is put under general anesthesia. The surgeon approaches the tympanic membrane from external auditory canal, makes a small incision at the inferior edge of the external auditory canal where it meets the tympani membrane, and

5 lifts the tympanic membrane as a flap to expose the middle ear space. A surgical laser is used to make a small opening (approximately 2 mm) in the stapes footplate. The surgeon then penetrates the round window membrane with a microcatheter loaded with a solution of AAV- STRC prepared in artificial perilymph at a titer of 1×10^{13} vg/mL. The microcatheter is connected to a micromanipulator that infuses

10 approximately 20 uL of the AAV- STRC solution at a rate of approximately 1 uL / min. At the conclusion of the AAV- STRC infusion, the surgeon withdraws the microcatheter and patches the holes in the stapes foot plate and RWM with a gel foam patch. The procedure concludes with replacement of the tympanic membrane flap.

15 **Example 10: Non-Invasive Prenatal Testing of Maternal Blood to Detect STRC Mutation**

Maternal blood samples (20-40 mL) are collected into Cell-free DNA tubes. At least 7 mL of plasma is isolated from each sample via a double centrifugation protocol of 2,000 g for 20 minutes, followed by 3,220 g for 30 minutes, with

20 supernatant transfer following the first spin. cfDNA is isolated from 7-20 mL plasma using a QIAGEN QIAmp Circulating Nuclei Acid kit and eluted in 45 μ L TE buffer. Pure maternal genomic DNA is isolated from the buffy coat obtained following the first centrifugation.

By combining thermodynamic modeling of the assays to select probes with

25 minimized likelihood of probe-probe interaction with amplification approaches described previously (Stiller et al., *Genome Res.* 19(10):1843-1848, 2009), multiplexing of 11,000 assays can be achieved. Maternal cfDNA and maternal genomic DNA samples are pre-amplified for 15 cycles using 11,000 target-specific assays and an aliquot is transferred to a second PCR reaction of 15 cycles using

30 nested primers. Samples are prepared for sequencing by adding barcoded tags in a third 12-cycle round of PCR. The targets include a genomic sequence in chromosome 15 known to be the site of a stereocilin loss-of-function mutation (Mandelker et al., *J. Mol. Diagnostics* 16(6):639-647, 2014). The amplicons are then sequenced using an

Illumina HiSeq sequencer. Genome sequence alignment is performed using commercially available software.

Example 11: Adenovirus (AAV) Trans-Splicing Strategy

At least two different nucleic acid vectors (e.g., AAV vectors) can be to reconstitute an active stereocilin gene (e.g., a full-length stereocilin gene) within a cell following intermolecular concatamerization and trans-splicing. See, e.g., Yan et al., *Proc. Natl. Acad. Sci. U.S.A.* 97:12; 6716-6721, 2000, incorporated in its entirety herein.

In some examples, two different nucleic acid vectors will be used. A first nucleic acid vector can include a promoter (e.g., any of the promoters described herein), a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a stereocilin protein described herein and/or any of the N-terminal portions of a stereocilin protein described herein), and a splicing donor signal sequence positioned at the 3' end of the first coding sequence. A second nucleic acid vector can include a splicing acceptor signal sequence, a second coding sequence that encodes a C-terminal portion of a stereocilin protein (i.e., the entire portion of the stereocilin protein that is not included in the N-terminal portion) positioned at the 3' end of the splicing acceptor signal sequence (e.g., any of the sizes of a portion of a stereocilin protein described herein and/or any of the C-terminal portions of a stereocilin protein described herein), and a polyadenylation sequence at the 3' end of the second coding sequence (e.g., any of the polyadenylation sequences described herein). In some embodiments, each of the encoded portions is at least 30 amino acid residues in length (e.g., at least 50 amino acids, at least 75 amino acids, or at least 100 amino acids in length), the amino acid sequence of each of the encoded portions does not overlap with the sequence of the other encoded portion, and no single vector of the two different vectors encodes an active stereocilin protein (e.g., a full-length stereocilin protein). When introduced into a mammalian cell (e.g., any of the mammalian cells described herein) splicing occurs between the splicing donor signal sequence and the splicing acceptor signal sequence, thereby forming a recombined nucleic acid that encodes an active stereocilin protein (e.g., a full-length stereocilin protein).

In another example, three different nucleic acid vectors can be used. A first nucleic acid vector can include a portion of a promoter sequence (e.g., any of the promoter sequences described herein), a first coding sequence of a stereocilin gene that encodes a first portion of a stereocilin protein (e.g., any of the stereocilin coding sequences described herein) positioned 3' of the promoter, and a first splicing donor signal sequence positioned at the 3' end of the first coding sequence. A second nucleic acid vector can include a first splicing acceptor signal sequence, a second coding sequence of a stereocilin gene that encodes a second portion of a stereocilin protein positioned at the 3' end of the first splicing acceptor signal sequence, and a second splicing donor signal sequence positioned at the 3' end of the second coding sequence (e.g., any of the splicing donor signals described herein). A feature of the second nucleic acid vector will be that self-splicing cannot occur (i.e., splicing will not occur between the second splicing donor signal sequence and the first splicing acceptor signal sequence of the second nucleic acid vector). In some embodiments, the splicing donor signal sequence of the first nucleic acid vector and the second splicing donor signal of the second nucleic acid vector are the same (e.g., any of the splicing donor signals described herein or known in the art). In some embodiments, the first splicing donor signal sequence of the first nucleic acid vector and the second splicing donor signal sequence of the second nucleic acid vector are different (e.g., any of the splicing donor signal sequences described herein or known in the art). A third nucleic acid vector will include a second splicing acceptor signal sequence, a third coding sequence of a stereocilin gene that encodes a third portion of a stereocilin protein positioned at the 3' end of the second splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the third coding sequence (e.g., any of the polyadenylation sequences described herein). In such methods where three nucleic acid vectors are used, the first splicing donor sequence and the first splicing acceptor sequence can assemble together (recombine) and the second splicing donor sequence and the second splicing acceptor sequence can assemble together (recombine), and the portion of stereocilin protein encoded by the first, second, and third coding sequences do not overlap, and when introduced into a mammalian cell (e.g., any of the mammalian cells described herein), splicing occurs between the first splicing donor sequence and the first splicing acceptor sequence, and between the second splicing donor sequence and the second splicing acceptor sequence, to form a recombined nucleic acid that encodes an active stereocilin protein (e.g., a full-length

stereocilin protein). Based on the strategies provided above, one skilled in the art would understand how to develop a strategy using four, five, or six different nucleic acid vectors.

5 In any of the examples of these methods, the amino acid sequence of each of the encoded portions do not overlap, and no single vector encodes an active stereocilin protein (e.g., a full-length stereocilin protein).

Each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions can be at least 30 amino acids (e.g., between about 30 amino acids to about 1600 amino
10 acids, or any of the other subranges of this range described herein).

In some embodiments, each of the coding sequences can include at least one exon and at least one intron of SEQ ID NO: 4 (e.g., at least two exons and at least one intron, at least two exons and at least two introns, at least three exons and at least one intron, at least three exons and at least two introns, or at least three exons and at least
15 three introns). In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions can encode up to 80% of the amino acid sequence of SEQ ID NO: 1 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%, up to 60%, or up to 70% of SEQ ID NO: 1) such that each of the encoded portions is non-
20 overlapping. In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding up to 80% of the amino acid sequence of SEQ ID NO: 1 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%, up to 60%, or up to 70% of SEQ ID NO: 1), such that each of the encoded portions is non-overlapping.

25 Each of the at least two nucleic acid vectors may further include an inverted terminal repeat (ITR) to allow head-to-tail recombination. The ITR will be subsequently removed via splicing. For example, the ITR could be a palindromic double-D ITR as described in Yan et al., *Proc. Natl. Acad. Sci. U.S.A.* 97(12):6716-6721, 2000, incorporated in its entirety herein. For example, the ITR could be a AAV serotype-2 ITR as described in Gosh et al., *Mol. Ther.* 16:124-130, 2008, and Gosh et al., *Human Gene Ther.* 22: 77-83, 2011. In some embodiments of any of the vectors described herein, the vector can include a 5' ITR and/or a 3' ITR. In some
30 embodiments, the 5' ITR includes or consists of SEQ ID NO: 18. In some embodiments, the 3' ITR includes or consists of SEQ ID NO: 36.

Non-limiting examples of splicing acceptor and/or donor signal sequences are known in the art. See, e.g., Reich et al., *Human Gene Ther.* 14(1):37-44, 2003, and Lai et al. (2005) *Nat. Biotechnol.* 23(11):1435-1439, 2005, 2005. The splicing donor and acceptor signal sequences can be any endogenous intron splicing signal of a gene (e.g., a stereocilin gene). For example, the splicing donor signal sequence can be: 5'-GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGA AACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCT-3' (SEQ ID NO: 6) and the splicing acceptor signal can be 5'-GATAGGCACCTATTGGTCTTACTG ACATCCACTTTGCCTTTCTCTCCACAG-3' (SEQ ID NO: 7) (see, e.g., Trapani et al., *EMBO Mol. Med.* 6(2):194-211, 2014). Methods of evaluating splicing and splicing efficiency are known in the art (see, e.g., Lai et al., *Nat. Biotechnol.* 23(11):1435-1439, 2005).

Example 12: Hybrid Vector Trans-Splicing Strategy using an alkaline phosphatase (AP) highly recombinogenic exogenous gene region

At least two (e.g., two, three, four, five, or six) different nucleic acid vectors (e.g., AAV vectors) can also be used in any of the methods described herein to reconstitute an active stereocilin gene (e.g., a full-length stereocilin gene) within a cell following intermolecular concatamerization, marker gene-mediated recombination, and trans-splicing. This strategy is a hybrid strategy as it will include homologous recombination and/or trans-splicing. See, e.g., Gosh et al., *Mol. Ther.* 16: 124-130, 2008; Gosh et al., *Human Gene Ther.* 22: 77-83, 2011; and Duan et al., *Mol. Ther.* 4: 383-391, 2001, each incorporated in its entirety herein. As used herein, a detectable marker gene can be a highly recombinogenic DNA sequence that will allow for coding sequence-independent recombination. An non-limiting example of a detectable marker gene is an alkaline phosphatase (AP) gene. For example, the detectable marker gene can be the middle one-third of the human placental AP complementary DNA, which is 872 bp in length (see, e.g., Gosh et al., 2008). At least two different nucleic acid vectors will contain a detectable marker gene (e.g., any of the detectable marker genes described herein). Since the hybrid vector will be constructed based on a trans-splicing vector as described in Example 10, an active stereocilin gene (e.g., a full-length stereocilin gene) may be reconstituted using either ITR-mediated recombination and trans-splicing or detectable marker gene-mediated

(e.g., AP-gene mediated) recombination and trans-splicing. After trans-splicing, an active stereocilin gene (e.g., a full-length stereocilin gene) will be reconstituted in the genomic DNA of a mammalian cell (e.g., any mammalian cell described herein).

In one example, two different nucleic acid vectors will be used. A first nucleic acid vector can include a promoter (e.g., any of the promoters described herein), a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a stereocilin protein described herein and/or any of the N-terminal portions of a stereocilin protein described herein), a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splicing donor signal sequence. A second nucleic acid vector can include a second detectable marker gene, a splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence (e.g., any of the sizes of a portion of a stereocilin protein described herein and/or any of the C-terminal portions of a stereocilin protein described herein), and a polyadenylation sequence at the 3' end of the second coding sequence (e.g., any of the polyadenylation sequences described herein). In some embodiments, each of the encoded portions is at least 30 amino acid residues in length (e.g., at least 50 amino acids, at least 75 amino acids, or at least 100 amino acids in length), the amino acid sequence of each of the encoded portions do not overlap, and no single vector of the two different vectors encodes an active stereocilin protein (e.g., a full-length stereocilin protein). When introduced into a mammalian cell (e.g., any of the mammalian cells described herein) splicing occurs between the splicing donor signal sequence and the splicing acceptor signal sequence, thereby forming a recombined nucleic acid that encodes an active stereocilin protein (e.g., a full-length stereocilin protein).

In another example, three different nucleic acid vectors can be used. A first nucleic acid vector can include a portion of promoter sequence (e.g., any of the promoter sequences described herein), a first coding sequence of a stereocilin gene that encodes a first portion of a stereocilin protein (e.g., any of the stereocilin coding sequences described herein) positioned 3' of the promoter, a first splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene. A second nucleic acid vector can include a second detectable

marker gene, a first splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence of a stereocilin gene that encodes a second portion of a stereocilin protein positioned at the 3' end of the first splicing acceptor signal sequence, a second splicing donor signal sequence positioned at the 3' end of the second coding sequence (e.g., any of the splicing donor signals described herein), and a third detectable marker gene. A feature of the second nucleic acid vector will be that self-splicing cannot occur (i.e., splicing will not occur between the second splicing donor signal sequence and the first splicing acceptor signal sequence of the second nucleic acid vector). In some embodiments, the splicing donor signal sequence of the first nucleic acid vector and the second splicing donor signal of the second nucleic acid vector are the same (e.g., any of the splicing donor signals described herein or known in the art). In some embodiments, the first splicing donor signal sequence of the first nucleic acid vector and the second splicing donor signal sequence of the second nucleic acid vector are different (e.g., any of the splicing donor signal sequences described herein or known in the art). A third nucleic acid vector can include a fourth detectable marker gene, a second splicing acceptor signal sequence positioned 3' of the fourth detectable marker gene, a third coding sequence of a stereocilin gene that encodes a third portion of a stereocilin protein positioned at the 3' end of the second splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the third coding sequence (e.g., any of the polyadenylation sequences described herein). In such methods where three nucleic acid vectors are used, the first splicing donor sequence and the first splicing acceptor sequence can assemble together (recombine) and the second splicing donor sequence and the second splicing acceptor sequence can assemble together (recombine), and the portion of stereocilin protein encoded by the first, second, and third coding sequences do not overlap, and when introduced into a mammalian cell (e.g., any of the mammalian cells described herein), splicing occurs between the first splicing donor sequence and the first splicing acceptor sequence, and between the second splicing donor sequence and the second splicing acceptor sequence, to form a recombined nucleic acid that encodes an active stereocilin protein (e.g., a full-length stereocilin protein). As can be appreciated in the art when three nucleic acid vectors are used, two of the at least two different nucleic acid vectors can include a detectable marker gene (e.g., an AP marker gene) and one of the at least two different nucleic acid vectors may include a splicing acceptor signal sequence that is complementary to a

splicing donor signal sequence in a nucleic acid vector that includes a detectable marker gene. For example, in some embodiments, the first and second nucleic acid vectors can include a detectable marker gene (e.g., an AP marker gene), and the third nucleic acid vector will include a splicing acceptor signal sequence that is

5 complementary to the splicing donor signal sequence in the second nucleic acid vector, and the third nucleic acid vector will not include a detectable marker gene (e.g., an AP marker gene). In other examples, the second and third nucleic acid vector can include a detectable marker gene (e.g., an AP marker gene), and the first nucleic acid vector will include a splicing donor signal sequence that is complementary to the
10 splicing acceptor signal sequence in the second nucleic acid vector and the first nucleic acid vector will not include a detectable marker gene (e.g., an AP marker gene).

Based on the strategies provided above, one skilled in the art would understand how to develop a strategy using four, five, or six vectors.

15 The coding sequences provided in the at least two nucleic acid vectors (e.g., two, three, four, five or six) will not be overlapping. Each of the at least two different vectors can include a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions being, e.g., at least 30 amino acids (e.g., about 30 amino acids to about 1600 amino acids, or any of the other subranges of this range
20 described herein).

In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding at least one exon and at least one intron of SEQ ID NO: 4 (e.g., at least two exons and at least one intron, at least two exons and at least two
25 introns, at least three exons at least one intron, at least three exons and at least two introns, or at least three exons and at least three introns). In some embodiments, each of the at least two different vectors include a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 1 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%, up to
30 60%, up to 70% of SEQ ID NO:1) such that each of the encoded portions is non-overlapping. In some embodiments, each of the at least two different vectors include a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 1 (e.g., up to 10%, up to 20%,

up to 30%, up to 40%, up to 50%, up to 60%, or up to 70% of SEQ ID NO:1), such that each of the encoded portions is non-overlapping.

As described in Example 10, each of the at least two nucleic acid vectors may further include an inverted terminal repeat (ITR) to allow head-to-tail recombination.

- 5 The ITR will be subsequently removed via splicing. Examples of ITRs and splicing acceptor and/or donor signal sequences are known in the art and have been described in Example 10.

Example 13: Hybrid Vector Trans-Splicing Strategy using a F1 phage highly recombino-
10 **genic exogenous gene region (AK)**

- At least two (e.g., two, three, four, five, or six) different nucleic acid vectors (e.g., AAV vectors) can also be used in any of the methods described herein to reconstitute an active stereocilin gene (e.g., a full-length stereocilin gene) within a cell following intermolecular concatamerization, marker gene-mediated recombination, and trans-splicing. This strategy is a hybrid strategy as it will include homologous recombination and/or trans-splicing. See, e.g., Trapani et al., *EMBO Mol. Med.* 6(2):194-211, 2014, incorporated in its entirety herein. As used herein, an F1 phage recombino-
- 15 genic region (AK) will be used to allow coding sequence-independent recombination. The F1 phage recombino-
- 20 genic region may be a 77 bp recombino-
- genic region from the F1 phage genome as described in Trapani et al. (2014). At least two different nucleic acid vectors will contain an F1 phage recombino-
- genic region. Since the hybrid vector will be constructed based on a trans-splicing vector as described in Example 10, a nucleic acid encoding an active stereocilin protein (e.g., a full-length stereocilin protein) may be generated using either ITR-mediated recombination and trans-splicing or F1 phage recombino-
- 25 genic region-induced recombination and trans-splicing. After trans-splicing, a nucleic acid encoding an active stereocilin protein (e.g., a full-length stereocilin protein) will be generated in a mammalian cell (e.g., any of the mammalian cells described herein).

- 30 In one example, two different nucleic acid vectors will be used. A first nucleic acid vector can include a promoter (e.g., any of the promoters described herein), a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a stereocilin protein described herein and/or any of the N-terminal portions of a stereocilin protein

described herein), a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and an F1 phage recombinogenic region positioned 3' of the splicing donor signal sequence. A second nucleic acid vector can include an F1 phage recombinogenic region, a splicing acceptor signal sequence positioned 3' of the F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence (e.g., any of the sizes of a portion of a stereocilin protein described herein and/or any of the C-terminal portions of a stereocilin protein described herein), and a polyadenylation sequence at the 3' end of the second coding sequence (e.g., any of the polyadenylation sequences described herein). In some embodiments, each of the encoded portions is at least 30 amino acid residues in length (e.g., at least 50 amino acids, at least 75 amino acids, or at least 100 amino acids in length), the amino acid sequence of each of the encoded portions do not overlap, and no single vector of the two different vectors encodes an active stereocilin protein (e.g., a full-length stereocilin protein). When introduced into a mammalian cell (e.g., any of the mammalian cells described herein) splicing occurs between the splicing donor signal sequence and the splicing acceptor signal sequence, thereby forming a recombined nucleic acid that encodes an active stereocilin protein (e.g., a full-length stereocilin protein).

In another example, three different nucleic acid vectors will be used. A first nucleic acid vector can include a portion of promoter sequence (e.g., any of the promoter sequences described herein), a first coding sequence of a stereocilin gene that encodes a first portion of a stereocilin protein (e.g., any of the stereocilin coding sequences described herein) positioned 5' of the promoter, a first splicing donor signal sequence positioned at the 3' end of the first coding sequence, and an F1 phage recombinogenic region. A second nucleic acid vector can include an F1 phage recombinogenic region, a first splicing acceptor signal sequence positioned 3' of the F1 phage recombinogenic region, a second coding sequence of a stereocilin gene that encodes a second portion of a stereocilin protein positioned at the 3' end of the first splicing acceptor signal sequence, a second splicing donor signal sequence positioned at the 3' end of the second coding sequence (e.g., any of the splicing donor signals described herein), and an F1 phage recombinogenic region. A feature of the second nucleic acid vector will be that self-splicing cannot occur (i.e., splicing will not occur between the second splicing donor signal sequence and the first splicing acceptor

signal sequence of the second nucleic acid vector). In some embodiments, the splicing donor signal sequence of the first nucleic acid vector and the second splicing donor signal of the second nucleic acid vector are the same (e.g., any of the splicing donor signals described herein or known in the art). In some embodiments, the first
5 splicing donor signal sequence of the first nucleic acid vector and the second splicing donor signal sequence of the second nucleic acid vector are different (e.g., any of the splicing donor signal sequences described herein or known in the art). A third nucleic acid vector can include an F1 phage recombinogenic region, a second splicing acceptor signal sequence positioned 3' of the F1 phage recombinogenic region, a third
10 coding sequence of a stereocilin gene that encodes a third portion of a stereocilin protein positioned at the 3' end of the second splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the third coding sequence (e.g., any of the polyadenylation sequences described herein). In such methods where three nucleic acid vectors are used, the first splicing donor sequence and the first splicing
15 acceptor sequence can assemble together (recombine) and the second splicing donor sequence and the second splicing acceptor sequence can assemble together (recombine), and the portion of stereocilin protein encoded by the first, second, and third coding sequences do not overlap, and when introduced into a mammalian cell (e.g., any of the mammalian cells described herein), splicing occurs between the first
20 splicing donor sequence and the first splicing acceptor sequence, and between the second splicing donor sequence and the second splicing acceptor sequence, to form a recombined nucleic acid that encodes an active stereocilin protein (e.g., a full-length stereocilin protein). As can be appreciated in the art when three nucleic acid vectors are used, two of the different nucleic acid vectors can include an F1 phage
25 recombinogenic region and one of the different nucleic acid vectors may include a splicing acceptor signal sequence that is complementary to a splicing donor signal sequence in a nucleic acid vector that includes an F1 phage recombinogenic region. For example, in some embodiments, the first and second nucleic acid vectors can include an F1 phage recombinogenic region, and the third nucleic acid vector will
30 include a splicing acceptor signal that is complementary to the splicing donor signal sequence in the second nucleic acid vector, and the third nucleic acid vector will not include an F1 phage recombinogenic region (e.g., an AP marker gene). In other examples, the second and third nucleic acid vector can include an F1 phage recombinogenic region and the first nucleic acid vector will include a splicing donor

signal sequence that is complementary to the splicing acceptor signal sequence in the second nucleic acid vector and the first nucleic acid vector will not include an F1 phage recombinogenic region. Based on the strategies provided above, one skilled in the art would understand how to develop a strategy using four, five, or six vectors.

5 The coding sequences provided in each of the at least two nucleic acid vectors (e.g., two, three, four, five or six) will not be overlapping. Each of the at least two different vectors include a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions being at least 30 amino acids (e.g., about 30 amino acids to about 1600 amino acids, or any of the subranges of this range
10 described herein).

 In some embodiments, each of the at least two different vectors include a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding at least one exon and at least one intron of SEQ ID NO: 4 (e.g., at least two exons and at least one intron, at least two exons and at least two
15 introns, at least three exons at least one intron, at least three exons and at least two introns, or at least three exons and at least three introns). In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 1 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%,
20 up to 60%, or up to 70% of SEQ ID NO: 1) such that each of the encoded portions is non-overlapping. In some embodiments, each of the at least two different vectors include a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 1 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%, up to 60%, or up to 70% of SEQ ID
25 NO: 1), such that each of the encoded portions is non-overlapping.

 As described in Example 10, each of the at least two nucleic acid vectors may further include an inverted terminal repeat (ITR) to allow head-to-tail recombination. The ITR will be subsequently removed via splicing. Examples of ITRs and splicing acceptor and/or donor signals are known in the art and have been described in
30 Example 10.

Example 14: CRISPR-Cas9-Mediated Knockin *STRC* Pseudogene Lacking Undesired Residues

The genomic sequences of the functional *STRC* gene and the *STRC* pseudogene are 98.9% identical and the coding regions are 99.6% homologous (Francey et al., *Am. J. Med. Genet. A.* 158A(2):298-308, 2012). Francey et al. determined that 56 bp between exons 16 and 28 were divergent. Verpy et al. determined that a nonsense mutation in exon 20 of the *STRC* pseudogene rendered the protein expressed from the pseudogene inactive (Verpy et al., *Nat. Genet.* 29:345-349, 2001). Expression of a functional *STRC* gene may be obtained using a composition including a CRISPR-Cas9 vector to correct (e.g., modify, replace) the nonsense mutation in exon 20 of the *STRC* pseudogene (SEQ ID NO: 5) to generate a sequence encoding an active stereocilin protein (e.g., a full-length stereocilin protein) in a mammalian cell, and thereby treat non-syndromic sensorineural hearing loss in a subject in need thereof. Various CRISPR-Cas9 techniques that can be used to introduce a mutation into an endogenous gene are also known in the art. See, e.g., Merkle et al., *Cell Rep.* 11(6):875-883, 2015; He et al., *Nucleic Acids Res.* 44(9):e85, 2016; Wang et al., *Mol. Ther. Nucleic Acids* 5: e396, 2016, and Verma et al., *Methods Mol. Biol.* 1513:119-140, 2017. For example, a composition including a Cas9 nuclease (e.g., *Streptococcus pyogenes*) and a guide RNA that includes at the 5' end of the guide RNA a complementary region consisting of about 20 nucleotides (e.g., 20, 21 or 22) that is complementary to about 20 (e.g., 20, 21 or 22) consecutive nucleotides within positions 13955-14151 of SEQ ID NO:5 can be used in CRISPR/Cas9 RNA-guided genome editing to remove a nonsense mutation at a predetermined site in an endogenous stereocilin pseudogene sequence of SEQ ID NO: 5; and when introduced into a mammalian cell, a nucleic acid encoding an active stereocilin protein (e.g., a full-length stereocilin protein) is reconstituted at the locus of the stereocilin pseudogene.

Example 15: AAV-mediated Antisense Oligonucleotide Delivery

An additional approach to generate a functional *STRC* transcript from the *STRC* pseudogene will include using AAV-mediated antisense oligonucleotide delivery via exon skipping. See, e.g., Chamorro et al., *Mol. Ther. Nucleic Acids* 5: e307, 2016; Kawecka et al., *Curr. Gene Ther.* 15(4):395-415, 2015; Bremer et al.,

Mol. Ther. 5:e3379, 2016; Tei et al., *Biochem. Biophys. Res. Commun.* 461(3):481-486, 2015; and Jirka et al., *Nucleic Acid Ther.* 24(1):25-36, 2014. As shown in Bremer et al., antisense oligonucleotides can be designed to bind to a particular exon of interest and lead to in-frame exon skipping at the RNA level, thereby restoring

5 gene function.

In the case of a stereocilin gene, a study can be conducted to determine the eligibility of exon 20 of the *STRC* pseudogene as a target for exon skipping, using techniques such as those used in Bremer et al. (2016). Upon determination of the eligibility of exon 20 of the *STRC* pseudogene, antisense oligonucleotides with

10 sequences corresponding to 17-23 base pairs in exon 20 can be generated and analyzed for their melting temperature, GC-content, off-target binding and binding energy. In some embodiments, the oligonucleotides will be synthesized as 2'-O-methyl phosphorothioate antisense oligonucleotides and analyzed for exon skipping by RT-PCR. To determine *de novo* expression of the corrected *STRC* pseudogene in a

15 cell, cells (e.g., a mammalian cell, a human cell, or any cell described herein) will be transfected within a nucleic acid vector (e.g., an AAV vector) including an antisense oligonucleotide of a stereocilin gene (e.g., an antisense oligonucleotide corresponding to part of exon 20 of a *STRC* pseudogene) using known transfection methods in the art, and expression of the corrected *STRC* pseudogene will be determined using

20 methods such as RT-PCR, immunofluorescence, or Western blot, as well as functional biochemical assays for a stereocilin protein. For example, in some embodiments, a composition including a nucleic acid vector that includes an antisense oligonucleotide that is complementary to at least a portion (e.g., at least 5 contiguous nucleotides, at least 10 contiguous nucleotides, at least 15 contiguous nucleotides, at least 20

25 contiguous nucleotides, at least 25 contiguous nucleotides, at least 30 contiguous nucleotides, at least 35 contiguous nucleotides, at least 40 contiguous nucleotides, at least 45 contiguous nucleotides, at least 50 contiguous nucleotides, at least 55 contiguous nucleotides, at least 60 contiguous nucleotides, at least 65 contiguous nucleotides, at least 70 contiguous nucleotides, at least 75 contiguous nucleotides, at least 80 contiguous nucleotides, at least 85 contiguous nucleotides, at least 90

30 contiguous nucleotides, at least 95 contiguous nucleotides, at least 100 contiguous nucleotides, at least 105 contiguous nucleotides, at least 110 contiguous nucleotides, at least 115 contiguous nucleotides, at least 120 contiguous nucleotides, at least 125 contiguous nucleotides, at least 130 contiguous nucleotides, at least 135 contiguous

nucleotides, at least 140 contiguous nucleotides, at least 145 contiguous nucleotides, at least 150 contiguous nucleotides, at least 155 contiguous nucleotides, at least 160 contiguous nucleotides, at least 165 contiguous nucleotides, at least 170 contiguous nucleotides, at least 175 contiguous nucleotides, at least 180 contiguous nucleotides, at least 185 contiguous nucleotides, or at least 190 contiguous nucleotides) of the nucleotide sequence of SEQ ID NO: 5 within positions 13955-14151 will be introduced into a mammalian cell and a functional stereocilin protein will be generated. One skilled in the art will appreciate that an exon 20 antisense oligonucleotide (e.g., an antisense oligonucleotide of SEQ ID NO: 5 complementary to sequence within the region defined by positions 13955-14141) will be 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides in length, or any range of numbers having one of those numbers at the low end and a higher number at the high end. In some embodiments, the antisense oligonucleotides of exon 20 will be 20 nucleotides in length. In some embodiments, the antisense oligonucleotides of exon 20 (e.g., an antisense oligonucleotide of SEQ ID NO: 5 complementary to sequence comprising of positions 13955-14141) will be 15 to 30 nucleotides in length. In some embodiments, the antisense oligonucleotides of exon 20 (e.g., an antisense oligonucleotide of SEQ ID NO: 5) will be at least 80% identical (e.g., at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, at least 99%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%) to the complement of a portion of the region defined by positions 13955-14151 of SEQ ID NO: 5.

25 **Example 16. STRC Expression in HEK293FT Cells**

HEK293FT cells were seeded overnight at 7E4 cells/well in a 24-well format and were transfected with exemplary single or dual STRC vectors using jetprime reagent. 72 hours post-transfection, HEK293FT cells were harvested with 200 μ L RIPA buffer/well. Thirty microliters of samples were loaded into each well onto a 4- 12% Bis-Tris protein gel. STRC protein expression was determined using a STRC polyclonal antibody. Beta-actin (ACTB) was used as a loading control.

As shown in Figure 7, the dual STRC vectors using CBA/CAG promoter (pITR-205 plus pITR-202 and pITR-205 plus pITR-202GFP) resulted in more intense

full-length STRC protein in HEK293FT cells. This result confirmed that STRC protein can be expressed by exemplary STRC vectors described herein.

Next, HEK293FT cells were transduced with dual STRC vectors at different MOIs and were seeded overnight at 1.2E5 cells/well in a 24-well format with etoposide reagent. HEK293FT cells were harvested 72 hours post-infection with 200 μ L RIPA buffer/well. As shown in Figure 8, HEK293FT cells transduced with AAV2 vectors showed higher expression of full-length STRC protein as compared to HEK293FT cells transduced with AAV.Anc80 vector.

Figure 9 showed high levels of STRC RNA expression relative to GAPDH in HEK293FT cells infected with either Anc80.STRC (at MOI 8.6E4 and MOI 2.6E5) or AAV2.STRC (at MOI 8.6E4 and MOI 2.6E5).

Example 17. STRC Expression in P1-3 Cochlear Mouse Explants

P2 cochlear from WT mice were infected after plating and were harvested for RNA and immunofluorescence 72 hours after infection. As shown in Figure 10, full-length STRC was efficiently expressed in cochlear explants infected with Anc80.STRC at MOI 3.9E10 and AAV2.STRC at MOI 1.3E10, respectively.

As shown in Figures 11A-E and 12A-C, outer hair cells and inner cells of P1-3 cochlear explants expressed Myo7a and phalloidin when transduced with AAV.Anc80.dSTRC (7.8E VG/cochlea, 2.6E8 VG/cochlea) and AAV2.dSTRC (7.8E8 VG/cochlea and 2.6E8 VG/cochlea). Figures 11A and 11B showed the natural cochlear structure, in which Myo7a is expressed in both outer and inner hair cells, while phalloidin is expressed exclusively in the inner hair cells. A similar staining pattern was observed in P1-3 cochlear explants transduced with AAV.Anc80.STRC (Figures 11C-E), and in P1-3 cochlear explants transduced with AAV2.STRC (Figures 12A-C).

Transfection with either vector did not disrupt the structural integrity of outer hair cells or inner hair cells of the cochlea. Comparison between non-infected with AAV2.dSTRC and AAV.Anc80.dSTRC infected explants showed robust numbers of living hair cells suggesting the relative safety of the vectors. Thus, Figures 11A-E and 12A-C showed lack of toxicity of STRC vectors with viable and organized outer hair cells, inner hair cells, and stereociliary bundles.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of
5 the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, section headings, the
10 materials, methods, and examples are illustrative only and not intended to be limiting.

What is claimed is:

1. A composition comprising at least two different nucleic acid vectors, wherein:
 - each of the at least two different vectors comprises a coding sequence that encodes a different portion of a stereocilin protein, each of the encoded portions being at least 30 amino acid residues in length, wherein the amino acid sequence of each of the encoded portions may optionally partially overlap with the amino acid sequence of a different one of the encoded portions;
 - no single vector of the at least two different vectors encodes a full-length stereocilin protein;
 - at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons; and
 - when introduced into a mammalian cell the at least two different vectors undergo homologous recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein.
2. The composition of claim 1, wherein each of the at least two different vectors is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector.
3. The composition of claim 1, wherein each of the at least two different vectors is a human artificial chromosome (HAC), yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC).
4. The composition of claim 1, wherein each of the at least two different vectors is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector.
5. The composition of claim 4, wherein the viral vector is an AAV vector.

6. The composition of any one of claims 1-5, wherein the amino acid sequence of none of the encoded portions overlaps with the amino acid sequence of a different one of the encoded portions.

7. The composition of claim 7, wherein the amino acid sequence of each of the encoded portions partially overlaps with the amino acid sequence of a different one of the encoded portions.

8. The composition of claim 7, wherein the overlapping amino acid sequence is between about 30 amino acid residues to about 1000 amino acid residues in length.

9. The composition of any one of claims 1-5, wherein the vectors include two different vectors, each of which comprises a different segment of an intron, wherein the intron comprises the nucleotide sequence of an intron that is present in stereocilin genomic DNA, and wherein the two different segments overlap in sequence by at least 100 nucleotides.

10. The composition of claim 9, wherein the two different segments overlap in sequence by about 100 nucleotides to about 800 nucleotides.

11. The composition of any one of claims 1-10, wherein the nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 10,000 nucleotides in length.

12. The composition of claim 11, wherein the nucleotide sequence of each of the at least two different vectors is between 500 nucleotides to 5,000 nucleotides in length.

13. The composition of any one of claims 1-12, wherein the number of different vectors in the composition is two.

14. The composition of claim 13, wherein a first of the two different vectors comprises a coding sequence that encodes an N-terminal portion of the stereocilin protein.

15. The composition of claim 14, wherein the N-terminal portion of the stereocilin protein is between 30 amino acids to 1600 amino acids in length.
16. The composition of claim 15, wherein the N-terminal portion of the stereocilin protein is between 200 amino acids to 1500 amino acids in length.
17. The composition of any one of claims 14-16, wherein the first vector further comprises one or both of a promoter and a Kozak sequence.
18. The composition of claim 17, wherein the first vector comprises a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter.
19. The composition of any one of claims 14-18, wherein the second of the two different vectors comprises a coding sequence that encodes a C-terminal portion of the stereocilin protein.
20. The composition of claim 19, wherein the C-terminal portion of the stereocilin protein is between 30 amino acids to 1600 amino acids in length.
21. The composition of claim 20, wherein the C-terminal portion of the stereocilin protein is between 200 amino acids to 1500 amino acids in length.
22. The composition of any one of claims 19-21, wherein the second vector further comprises a poly(dA) sequence.
23. The composition of any one of claims 1-22, further comprising a pharmaceutically acceptable excipient.
24. A kit comprising a composition of any one of claims 1-23.
25. A kit of claim 24, further comprising a pre-loaded syringe comprising the composition.

26. A method comprising:
introducing into a cochlea of a mammal a therapeutically effective amount of the composition of any one of claims 1-23.

27. The method of claim 26, wherein the mammal is a human.

28. The method of claim 26 or 27, wherein the mammal has been previously identified as having a defective stereocilin gene.

29. A method of increasing expression of a full-length stereocilin protein in a mammalian cell, the method comprising introducing the composition of any one of claims 1-23 into the mammalian cell.

30. The method of claim 29, wherein the mammalian cell is a cochlear outer hair cell.

31. The method of claim 29 or 30, wherein the mammalian cell is a human cell.

32. The method of any one of claims 29-31, wherein the mammalian cell has previously been determined to have a defective stereocilin gene.

33. A method of increasing expression of a full-length stereocilin protein in an outer hair cell in a cochlea of a mammal, the method comprising:

introducing into the cochlea of the mammal a therapeutically effective amount of the composition of any one of claims 1-23.

34. The method of claim 33, wherein the mammal has been previously identified as having a defective stereocilin gene.

35. The method of claim 33 or 34, wherein the mammal is a human.

36. A method of treating non-symptomatic sensorineural hearing loss in a subject identified as having a defective stereocilin gene, the method comprising:

administering a therapeutically effective amount of a composition of any one of claims 1-23 into the cochlea of the subject.

37. The method of claim 36, wherein the subject is a human.

38. The method of claim 36 or 37, further comprising, prior to the administering step, determining that the subject has a defective stereocilin gene.

39. A composition comprising two different nucleic acid vectors, wherein:

a first nucleic acid vector of the two different nucleic acid vectors comprises a promoter, a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter, and a splicing donor signal sequence positioned at the 3' end of the first coding sequence; and

a second nucleic acid vector of the two different nucleic acid vectors comprises a splicing acceptor signal sequence, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence at the 3' end of the second coding sequence;

wherein each of the encoded portions is at least 30 amino acid residues in length, wherein the amino acid sequence of each of the encoded portions do not overlap; wherein no single vector of the two different vectors encodes a full-length stereocilin protein; and

when introduced into a mammalian cell, splicing occurs between the splicing donor signal sequence and the splicing acceptor signal sequence, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein.

40. The composition of claim 39, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons.

41. A composition comprising two different nucleic acid vectors, wherein:

a first nucleic acid vector of the two different nucleic acid vectors comprises a promoter, a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' of the promoter, a splicing donor signal sequence positioned at

the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splicing donor signal sequence; and

a second nucleic acid vector of the two different nucleic acid vectors comprises a second detectable marker gene, a splicing acceptor signal sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence;

wherein each of the encoded portions is at least 30 amino acid residues in length, wherein the amino acid sequence of each of the encoded portions do not overlap; wherein no single vector of the two different vectors encodes a full-length stereocilin protein; and

when introduced into a mammalian cell, splicing occurs between the splicing donor signal and the splicing acceptor signal, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein.

42. The composition of claim 41, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons.

43. The composition of claim 42, wherein the first or second detectable marker gene is alkaline phosphatase.

44. The composition of claim 41 or 43, wherein the first and second detectable marker genes are the same.

45. A composition comprising two different nucleic acid vectors, wherein:

a first nucleic acid vector of the two different nucleic acid vectors comprises a promoter, a first coding sequence that encodes an N-terminal portion of a stereocilin protein positioned 3' to the promoter, a splicing donor signal sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splicing donor signal sequence; and

a second nucleic acid vector of the two different nucleic acid vectors comprises a F1 phage recombinogenic region, a splicing acceptor signal sequence

positioned 3' of the F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a stereocilin protein positioned at the 3' end of the splicing acceptor signal sequence, and a polyadenylation sequence positioned at the 3' end of the second coding sequence;

wherein each of the encoded portions is at least 30 amino acid residues in length, wherein the amino acid sequence of each of the encoded portions do not overlap; wherein no single vector of the two different vectors encodes a full-length stereocilin protein; and

when introduced into a mammalian cell, splicing occurs between the splicing donor signal and the splicing acceptor signal, thereby forming a recombined nucleic acid that encodes a full-length stereocilin protein.

46. The composition of claim 45, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two neighboring exons of stereocilin genomic DNA, and lacks an intronic sequence between the two neighboring exons.

47. A composition comprising:

a Cas9 nuclease;

a guide RNA that includes at the 5' end of the guide RNA a complementary region consisting of 20 nucleotides that are complementary to 20 consecutive nucleotides within positions 13955-14151 of SEQ ID NO:5, that can be used in CRISPR/Cas9 RNA-guided genome editing to remove a nonsense mutation at a predetermined site in the endogenous stereocilin pseudogene sequence of SEQ ID NO: 5; and

when introduced into a mammalian cell, a nucleic acid encoding a full-length stereocilin protein is reconstituted at the locus of the stereocilin pseudogene.

48. A composition comprising at least one nucleic acid vector, wherein the at least one nucleic acid vector comprises an adeno-associated virus (AAV) vector that comprises an antisense oligonucleotide that is at least 80% complementary to a contiguous nucleotide sequence of SEQ ID NO: 5 at positions 13955-14151,

wherein the antisense oligonucleotide is between 15-30 nucleotides in length; and

when introduced into a mammalian cell, a nucleic acid encoding a full-length stereocilin protein is generated at the locus of the stereocilin pseudogene.

49. A kit comprising a composition of any one of claims 39-48.

50. A kit of claim 49, further comprising a pre-loaded syringe comprising the composition.

51. A method comprising:
introducing into a cochlea of a mammal a therapeutically effective amount of the composition of any one of claims 39-48.

52. The method of claim 51, wherein the mammal is a human.

53. The method of claim 51 or 52, wherein the mammal has been previously identified as having a defective stereocilin gene.

54. A method of increasing expression of a full-length stereocilin protein in a mammalian cell, the method comprising introducing the composition of any one of claims 39-48 into the mammalian cell.

55. The method of claim 54, wherein the mammalian cell is a cochlear outer hair cell.

56. The method of claim 54 or 55, wherein the mammalian cell is a human cell.

57. The method of any one of claims 54-56, wherein the mammalian cell has previously been determined to have a defective stereocilin gene.

58. A method of increasing expression of a full-length stereocilin protein in an outer hair cell in a cochlea of a mammal, the method comprising:

introducing into the cochlea of the mammal a therapeutically effective amount of the composition of any one of claims 39-48.

59. The method of claim 58, wherein the mammal has been previously identified as having a defective stereocilin gene.

60. The method of claim 58 or 59, wherein the mammal is a human.

61. A method of treating non-symptomatic sensorineural hearing loss in a subject identified as having a defective stereocilin gene, the method comprising:
administering a therapeutically effective amount of a composition of any one of claims 39-48 into the cochlea of the subject.

62. The method of claim 61, wherein the subject is a human.

63. The method of claim 61 or 62, further comprising, prior to the administering step, determining that the subject has a defective stereocilin gene.

64. The composition of claim 13, wherein one of the two vectors comprises SEQ ID NO: 13 and the second of the two vectors comprises SEQ ID NO: 17.

65. The composition of claim 1, wherein one of the at least two different vectors comprises a sequence encoding a stereocilin protein.

66. The composition of claim 65, wherein the sequence encoding a stereocilin protein is at least 90% identical to SEQ ID NO: 11.

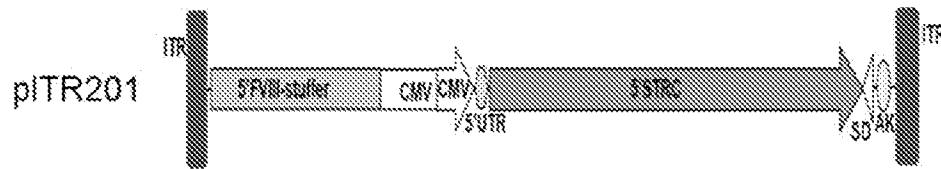


Figure 1



Figure 2



Figure 3

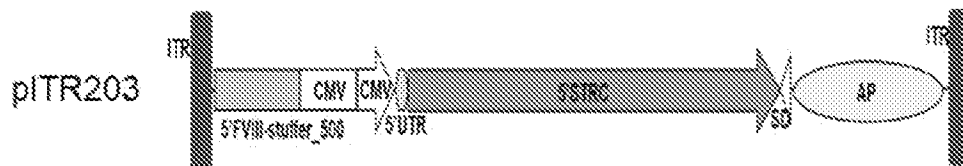


Figure 4

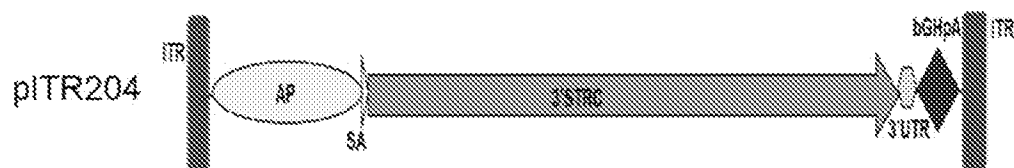


Figure 5

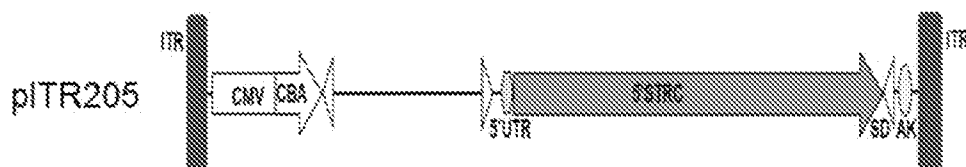


Figure 6

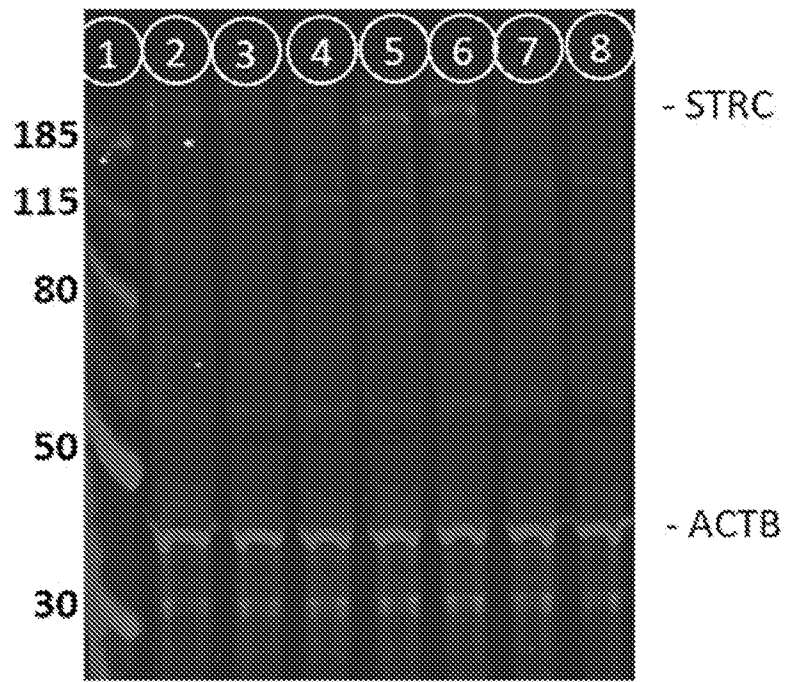


Figure 7

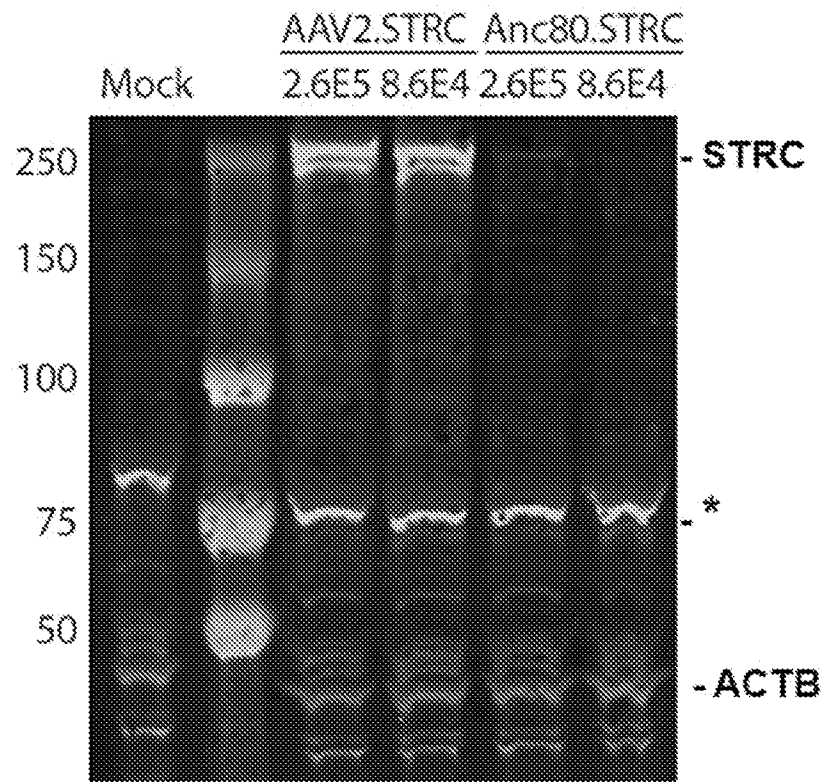


Figure 8

AAV-dSTRC Transduction in HEK293FT

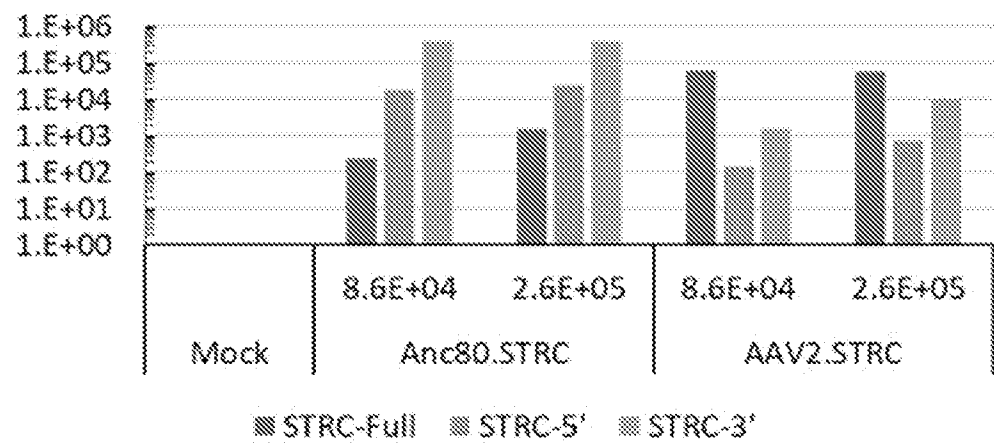


Figure 9

AAV-dSTRC Transduction in Explants

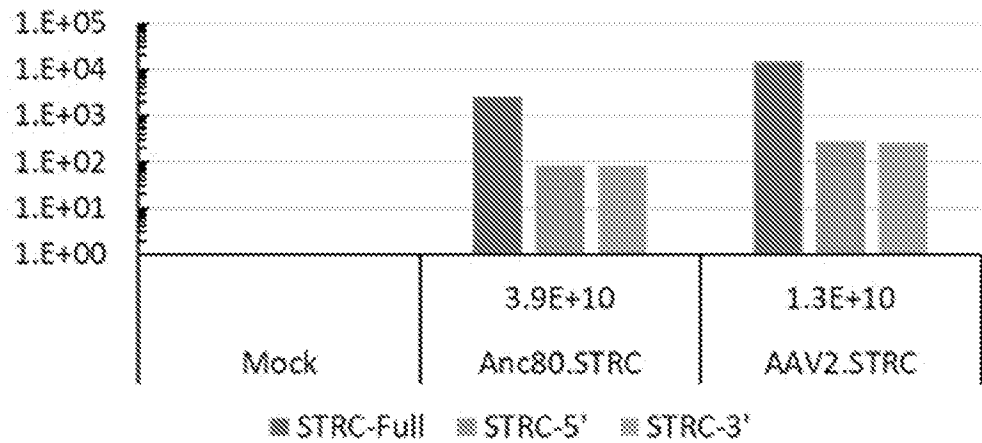


Figure 10

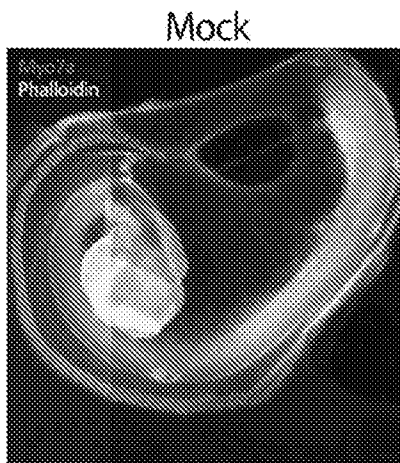


Figure 11A

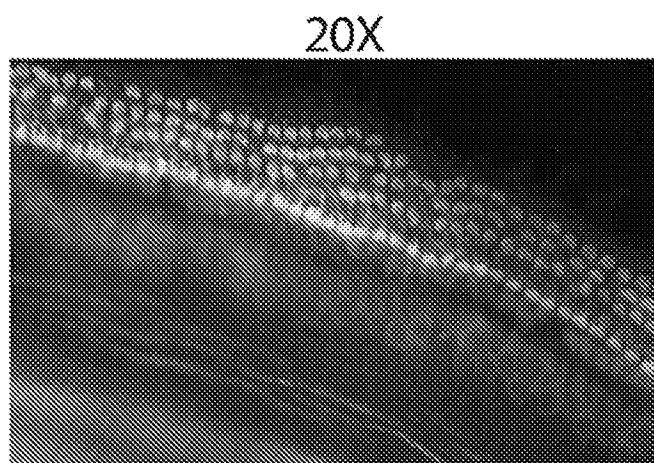


Figure 11B

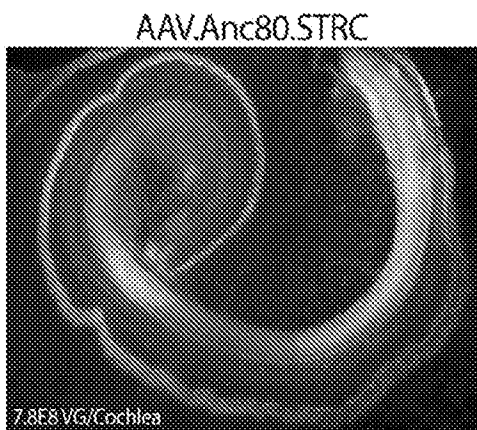


Figure 11C

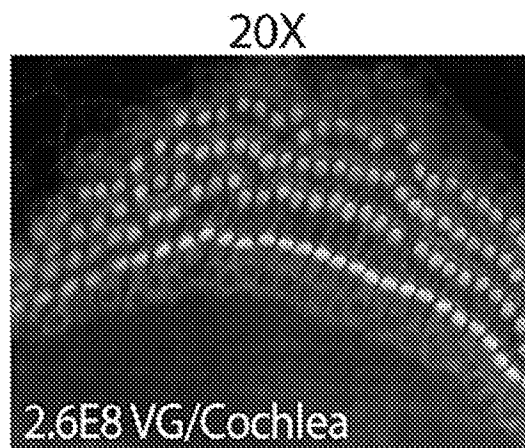


Figure 11D

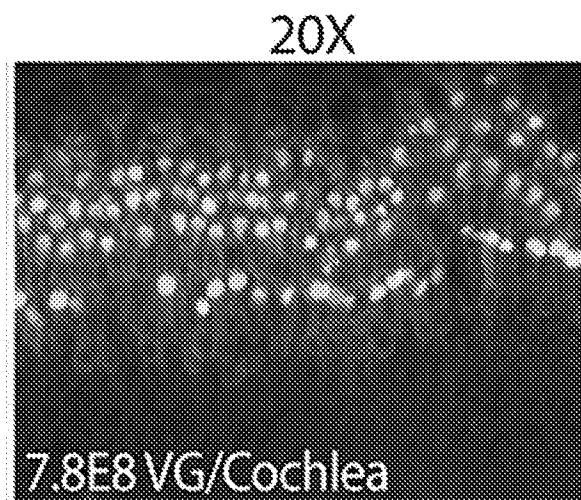


Figure 11E

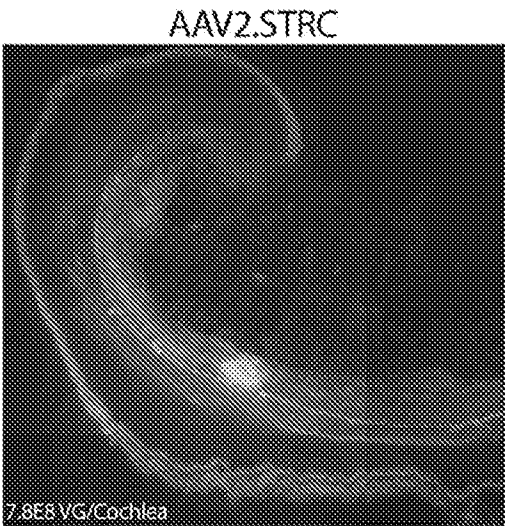


Figure 12A

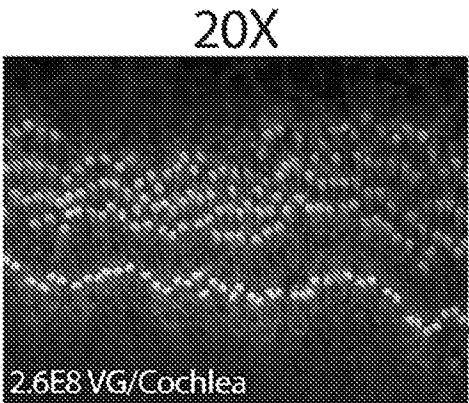


Figure 12B

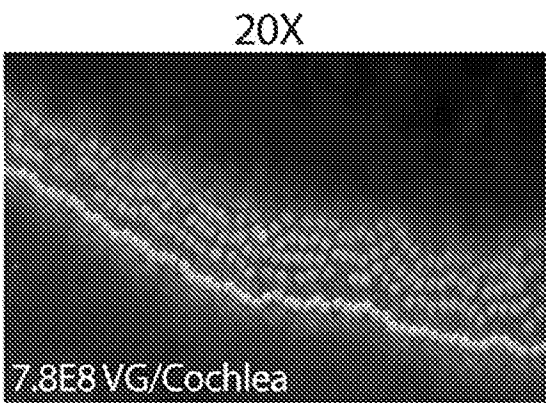


Figure 12C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2019/041625

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

a. ☒ forming part of the international application as filed:

☒ in the form of an Annex C/ST.25 text file.

☐ on paper or in the form of an image file.

b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

☐ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).

☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 26-38,51-63
because they relate to subject matter not required to be searched by this Authority, namely:
Claims 26-38, 51-63 pertain to a method for treatment of the human body by therapy or surgery, and thus relate to a subject matter which this International Searching Authority is not required, under PCT Article 17(2)(a)(i) and PCT Rule 39.1(iv), to search.
2. ☒ Claims Nos.: 12,14-16,18,20,21,25,27,30,34,37,50,52,55,59,62,64
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Each of claims 12, 14-16, 18, 20, 21, 25, 27, 30, 34, 37, 50, 52, 55, 59, 62, 64 refers to a claim which is not drafted in accordance with Rule 6.4(a).
3. ☒ Claims Nos.: 11,13,17,19,22-24,26,28-29,31-33,35-36,38,49,51,53,54,56-58,60-61,63
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of any additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/85(2006.01)i, C12N 5/0793(2010.01)i, C12N 15/65(2006.01)i, C12N 15/10(2006.01)i, C07K 14/435(2006.01)i, A61K 48/00(2006.01)i, A61P 27/16(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N 15/85; A61K 39/00; A61K 48/00; C12N 15/86; C12N 5/0793; C12N 15/65; C12N 15/10; C07K 14/435; A61P 27/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & keywords: vector, stereocilin(STRC), non-syndromic sensorineural hearing loss, CAS9, guide RNA, recombination

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2018-039375 A1 (AKOUOS, INC. et al.) 1 March 2018 See abstract; pages 4, 161, 190-191, 222; claims 1-25, 39-46.	1-10, 39-46, 65-66
Y		47-48
Y	NCBI, GenBank Accession No. NR_146078.1, 'Homo sapiens stereocilin pseudogene 1 (STRCP1), non-coding RNA', 9 March 2017 See the whole document.	47-48
A	VERPY, E. et al., 'Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus', Nature Genetics, 2001, Vol. 29, pp. 345-349 See the whole document.	1-10, 39-48, 65-66
A	WO 2017-100791 A1 (MASSACHUSETTS EYE AND EAR INFIRMARY et al.) 15 June 2017 See claims 1-13, 22-23.	1-10, 39-48, 65-66
A	US 2016-0076054 A1 (FONDAZIONE TELETHON) 17 March 2016 See the whole document.	1-10, 39-48, 65-66



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

29 October 2019 (29.10.2019)

Date of mailing of the international search report

29 October 2019 (29.10.2019)

Name and mailing address of the ISA/KR

International Application Division

Korean Intellectual Property Office

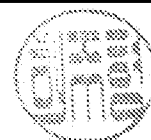
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

HEO, Joo Hyung

Telephone No. +82-42-481-8150



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2019/041625

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018-039375 A1	01/03/2018	AU 2017-315679 A1 CA 3034527 A1 EP 3510161 A1 US 2019-0185864 A1	07/03/2019 01/03/2018 17/07/2019 20/06/2019
WO 2017-100791 A1	15/06/2017	AU 2016-366846 A1 CA 3007476 A1 CN 109310745 A EP 3386537 A1 JP 2018-536420 A KR 10-2018-0097631 A SG 11201804814 A US 2018-0369414 A1	21/06/2018 15/06/2017 05/02/2019 17/10/2018 13/12/2018 31/08/2018 30/07/2018 27/12/2018
US 2016-0076054 A1	17/03/2016	AU 2014-255665 A1 AU 2014-255665 B2 BR 112015026422 A2 CA 2909733 A1 CN 105408352 A DK 2986635 T3 EA 201591992 A1 EP 2986635 A1 EP 2986635 B1 ES 2704677 T3 JP 2016-516424 A JP 6453307 B2 MX 2015014423 A WO 2014-170480 A1	12/11/2015 02/08/2018 14/11/2017 23/10/2014 16/03/2016 28/01/2019 29/04/2016 24/02/2016 03/10/2018 19/03/2019 09/06/2016 16/01/2019 28/07/2016 23/10/2014