



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION

(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2019/06/25
(87) Date publication PCT/PCT Publication Date: 2020/01/02
(85) Entrée phase nationale/National Entry: 2020/12/17
(86) N° demande PCT/PCT Application No.: US 2019/039030
(87) N° publication PCT/PCT Publication No.: 2020/005974
(30) Priorité/Priority: 2018/06/25 (US62/689,660)

(51) Cl.Int./Int.Cl. *C12N 15/85* (2006.01),
A61K 35/30 (2015.01), *A61K 38/00* (2006.01),
A61P 27/16 (2006.01), *C07K 14/47* (2006.01),
C12N 5/079 (2010.01), *C12N 5/0793* (2010.01)
(71) Demandeur/Applicant:
AKOUOS, INC., US
(72) Inventeurs/Inventors:
SIMONS, EMMANUEL J., US;
NG, ROBERT, US
(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : PROCEDES DE TRAITEMENT DE LA PERTE AUDITIVE ET/OU DE LA PERTE DE LA VISION ASSOCIEES
A CLRN1

(54) Title: METHODS OF TREATING CLRN1-ASSOCIATED HEARING LOSS AND/OR VISION LOSS

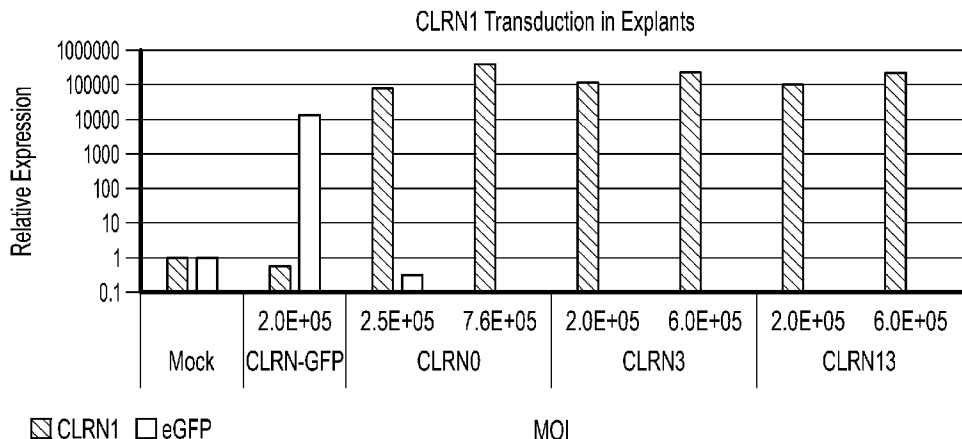


FIG. 34

(57) **Abrégé/Abstract:**

Provided herein are compositions that include a single nucleic acid vector or two different nucleic acid vectors, and the use of these compositions to treat hearing loss and/or vision loss in a subject.

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

02 January 2020 (02.01.2020)



(10) International Publication Number

WO 2020/005974 A1

(51) International Patent Classification:

C12N 15/85 (2006.01) A61K 35/30 (2006.01)

C12N 5/0793 (2010.01) A61P 27/16 (2006.01)

C12N 5/079 (2010.01) A61K 38/00 (2006.01)

C07K 14/47 (2006.01)

(21) International Application Number:

PCT/US2019/039030

(22) International Filing Date:

25 June 2019 (25.06.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/689,660 25 June 2018 (25.06.2018) US

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

(72) Inventors: SIMONS, Emmanuel J.; 7 Stetson Street, Brookline, Massachusetts 02446 (US). NG, Robert; 47 Lincoln Rd, Unit 1, Newton, Massachusetts 02458 (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).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: METHODS OF TREATING CLRN1-ASSOCIATED HEARING LOSS AND/OR VISION LOSS

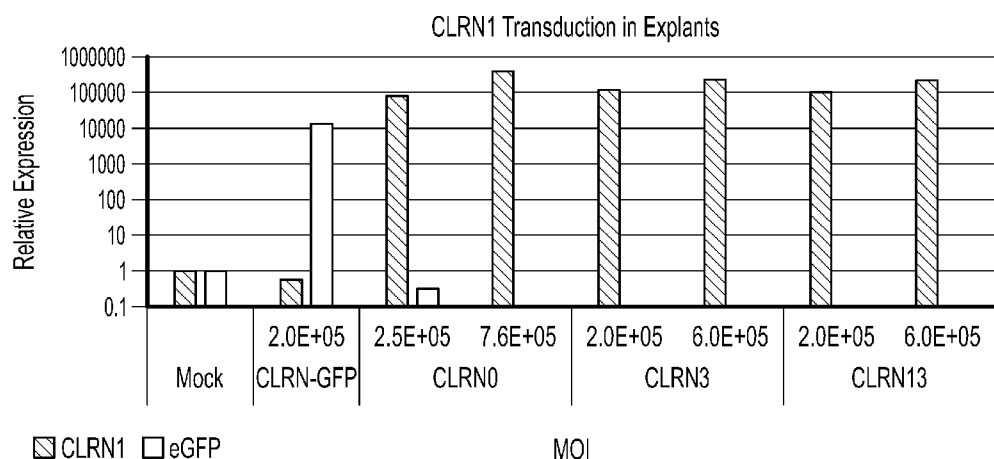


FIG. 34

(57) Abstract: Provided herein are compositions that include a single nucleic acid vector or two different nucleic acid vectors, and the use of these compositions to treat hearing loss and/or vision loss in a subject.

WO 2020/005974 A1

METHODS OF TREATING CLRN1-ASSOCIATED HEARING LOSS AND/OR VISION LOSS

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims priority to U.S. Provisional Patent Application Serial No. 62/689,660, filed June 25, 2018; the entire contents of which are herein incorporated by reference.

TECHNICAL FIELD

10 The present disclosure relates generally to the use of nucleic acids to treat hearing loss, vision loss, or both, in a human subject.

BACKGROUND OF THE INVENTION

15 Current treatments for hearing loss 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 syndromic deafness.

 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 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 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
20 may be damaged during the lifetime of an individual (e.g., as a result of noise trauma or
25 infection).

SUMMARY

 The present invention relates to a composition including at least two different nucleic acid vectors, where each of the at least two different vectors includes a coding
30 sequence that encodes a different portion of a CLRN1 protein, can be used to generate a sequence encoding an active CLRN1 protein (e.g., a full-length CLRN1 protein) in a

mammalian cell, and thereby treat CLRN1-associated hearing loss and/or vision loss in a subject in need thereof. The invention also related to compositions including a single nucleic acid vector that includes a coding sequence for a first and/or second isoform of CLRN1 protein.

5 Provided herein are compositions 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 CLRN1 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
10 different one of the encoded portions; no single vector of the at least two different vectors encodes a full-length CLRN1 protein; at least one of the coding sequences includes a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons; and when introduced into a mammalian cell, the at least two different vectors undergo homologous
15 recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein. In some embodiments of any of the compositions provided herein, each of the at least two different vectors includes a coding sequence that encodes a different portion of a first isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a
20 sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, one of the at least two different nucleic acid vectors
25 further includes a sequence that encodes a second isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the
30 compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, each of the

at least two different vectors includes a coding sequence that encodes a different portion of a second isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, one of the at least two different nucleic acid vectors further includes a sequence that encodes a first isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

In some embodiments of any of the compositions provided herein, at least one of the at least two different vectors includes a 5' untranslated region (UTR), a 3' UTR, or both. In some embodiments of any of the compositions provided herein, the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5'UTR includes a sequence that is at least 80% identical to SEQ ID NO: 12. In some embodiments of any of the compositions described herein, the 3' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions described herein, the 3' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR

includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 15.

In some embodiments of any of the compositions provided 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 provided herein, 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 provided 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 lentivirus vector, or a retrovirus vector. In some embodiments of any of the compositions provided herein, each of the at least two different vectors is an AAV vector.

In some embodiments of any of the compositions provided 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 compositions provided 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 provided herein, the overlapping amino acid sequence is between about 30 amino acid residues to about 202 amino acid residues in length. In some embodiments of any of the compositions provided 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 CLRN1 genomic DNA, and wherein the two different intron segments overlap in sequence by at least 100 nucleotides. In some embodiments of any of the compositions provided herein, the two different intron segments overlap in sequence by 100 nucleotides to about 800 nucleotides.

In some embodiments of any of the compositions provided herein, the entire nucleotide sequence of each of the at least two different vectors is between about 500

nucleotides to about 10,000 nucleotides in length. In some embodiments of any of the compositions provided herein, the entire nucleotide sequence of each of the at least two different vectors is between 500 nucleotides to 5,000 nucleotides in length. In some embodiments of any of the compositions provided herein, the number of different vectors

5 in the composition is two. In some embodiments of any of the compositions provided herein, a first of the two different vectors includes a coding sequence that encodes an N-terminal portion of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the N-terminal portion of the CLRN1 protein is between 30 amino acids to 202 amino acids in length. In some embodiments of any of the

10 compositions provided herein, the N-terminal portion of the CLRN1 protein is between 60 amino acids to 170 amino acids in length. In some embodiments of any of the compositions provided herein, the first vector further includes a 5' UTR sequence. In some embodiments of any of the compositions provided herein, the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some

15 embodiments of any of the compositions provided herein, the 5' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 80 contiguous nucleotides

20 from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5'UTR includes a sequence that is at least 80% identical to SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the first vector further includes one or both of a promoter and a Kozak sequence. In some embodiments of any of the compositions provided herein, the first vector

25 includes a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter. In some embodiments of any of the compositions provided herein, the second of the two different vectors includes a coding sequence that encodes a C-terminal portion of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the C-terminal portion of the CLRN1 protein is between 30 amino acids

30 to 202 amino acids in length. In some embodiments of any of the compositions provided herein, the C-terminal portion of the CLRN1 protein is between 60 amino acids to 170

amino acids in length. In some embodiments of any of the compositions provided herein, the second vector further includes a polyadenylation signal sequence. In some embodiments of any of the compositions provided herein, the second vector further includes a 3'UTR sequence. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 15.

Also provided herein are compositions that include a single nucleic acid vector, where the vector includes one or both of (i) a first coding sequence encoding a first isoform of CLRN1 protein, and (ii) a second coding sequence encoding a second isoform of CLRN1 protein, where one or both of the first and second coding sequences includes a nucleotide sequence spanning two consecutive exons of a CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive introns. In some embodiments of any of the compositions provided herein, the single nucleic acid vector contains the first coding sequence and not the second coding sequence. In some embodiments of any of the compositions provided herein, the single nucleic acid vector contains the second coding sequence and not the first coding sequence. In some embodiments of any of the compositions provided herein, the single nucleic acid vector contains both the first coding sequence and the second coding sequence. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a

sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

5 In some embodiments of any of the compositions provided herein, the single nucleic acid vector further includes a 5' untranslated region (UTR), a 3' UTR, or both. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some
10 embodiments of any of the compositions provided herein, the 5' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 50 contiguous
15 nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the
20 compositions provided herein, the 5' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 10 contiguous nucleotides from anywhere within
25 SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR
30 includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some
embodiments of any of the compositions provided herein, the 3' UTR includes a
sequence that is at least 80% identical to SEQ ID NO: 15.

 In some embodiments of any of the compositions provided herein, the single nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector. In some embodiments of any of the compositions provided herein, the single
30 nucleic acid vector 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 provided herein, the single

nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector. In some embodiments of any of the compositions provided herein, the single nucleic acid vector is an AAV vector. In some embodiments of any of the compositions provided herein, the single nucleic acid vector further includes one or both of a promoter and a Kozak sequence. In some embodiments of any of the compositions provided herein, the first vector includes a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter. In some embodiments of any of the compositions provided herein, the single nucleic acid vector further includes a polyadenylation signal sequence. Some embodiments of any of the compositions provided herein further include a pharmaceutically acceptable excipient.

Also provided herein are kits that include any of the compositions provided herein. Some embodiments of any of the kits provided herein further include a pre-loaded syringe including or containing 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 provided herein. In some embodiments of any of the methods provided herein, the mammal is a human. In some embodiments of any of the methods provided herein, the mammal has been previously identified as having a defective CLRN1 gene.

Also provided herein are methods of increasing expression of a full-length CLRN1 protein in a mammalian cell that include introducing any of the compositions provided herein into the mammalian cell. In some embodiments of any of the methods provided herein, the mammalian cell is a cochlear inner hair cell or a cochlear outer hair cell. In some embodiments of any of the methods provided herein, the mammalian cell is a retinal cell. In some embodiments of any of the methods provided herein, the mammalian cell is a human cell. In some embodiments of any of the methods provided herein, the mammalian cell has previously been determined to have a defective CLRN1 gene.

Also provided herein are methods of increasing expression of a full-length CLRN1 protein in an inner hair cell, an outer hair cell, or both, in a cochlea of a mammal, that include: introducing into the cochlea of the mammal a therapeutically effective

amount of any of the compositions provided herein. Also provided herein are methods of increasing expression of a full-length CLRN1 protein in an eye of a mammal that include: intraocularly administering to the eye of the mammal a therapeutically effective amount of any of the compositions provided herein. In some embodiments of any of the methods
5 provided herein, the mammal has been previously identified as having a defective CLRN1 gene. In some embodiments of any of the methods provided herein, the mammal is a human.

Also provided herein are methods of treating hearing loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective
10 amount of any of the compositions provided herein into the cochlea of the subject. Also provided herein are methods of treating vision loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions provided herein into the eye of the subject. In some
15 embodiments of any of the methods provided herein, the subject has Usher syndrome type III. In some embodiments of any of the methods provided herein, the subject is a human. Some embodiments of any of the methods provided herein further include, prior to the administering step, determining that the subject has a defective CLRN1 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
20 includes a promoter, a first coding sequence that encodes an N-terminal portion of a CLRN1 protein positioned 3' of the promoter, and a splice donor sequence positioned at the 3' end of the first coding sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a splice acceptor sequence, a second coding
25 sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal 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 sequences of the two encoded portions do not overlap with each other; where no single vector of the two different vectors encodes a full-length CLRN1 protein; and when introduced into a mammalian cell, splicing occurs
30 between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein. In some

embodiments of any of the compositions provided herein, the first coding sequence encodes an N-terminal portion of a first isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the first isoform of CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first nucleic acid vector of the second nucleic acid vector further includes a sequence that encodes a second isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

In some embodiments of any of the compositions provided herein, the first coding sequence encodes an N-terminal portion of a second isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the second isoform of CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the first nucleic acid vector or the second nucleic acid vector further includes a sequence that encodes a first isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ

ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

In some embodiments of any of the compositions provided herein, one or both of the first nucleic acid vector and the second nucleic acid vector includes a 5' untranslated region (UTR), a 3' UTR, or both. In some embodiments of any of the compositions
5 provided herein, the 5' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the
10 5' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5'UTR includes a sequence that is at least 80% identical to SEQ ID NO: 12. In some embodiments of any
15 of the compositions provided herein, the 3' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 50 contiguous nucleotides
20 from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 15.

25 In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector. In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is a human artificial chromosome (HAC), yeast artificial chromosome (YAC),
30 bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC). In some embodiments of any of the compositions provided herein, each of the first nucleic

acid vector and the second nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector. In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is an AAV vector. In some
5 embodiments of any of the compositions provided herein, at least one of the coding sequences includes a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons.

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
10 includes a promoter, a first coding sequence that encodes an N-terminal portion of a CLRN1 protein positioned 3' of the promoter, a splice donor sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splice donor sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a second detectable marker gene, a splice acceptor sequence positioned
15 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal 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 sequences of the encoded portions do not overlap with each
20 other; where no single vector of the two different vectors encodes a full-length CLRN1 protein; and when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein. In some embodiments of any of the compositions provided herein, the first coding sequence encodes an N-terminal portion of
25 a first isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the first isoform of CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ
30 ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3. In some embodiments of any

of the compositions provided herein, the first nucleic acid vector of the second nucleic acid vector further includes a sequence that encodes a second isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the first coding sequence encodes an N-terminal portion of a second isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the second isoform of CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the first nucleic acid vector or the second nucleic acid vector further includes a sequence that encodes a first isoform of the CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

In some embodiments of any of the compositions provided herein, one or both of the first nucleic acid vector and the second nucleic acid vector includes a 5' untranslated region (UTR), a 3' UTR, or both. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the

5' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

In some embodiments of any of the compositions provided herein, the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 15.

In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector. In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector 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 provided herein, each of the first nucleic acid vector and the second nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector. In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is an AAV vector. In some embodiments of any of the compositions provided herein, at least one of the coding sequences comprises a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons. In

some embodiments of any of the compositions provided herein, the first or second detectable marker gene is alkaline phosphatase.

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
5 includes a promoter, a first coding sequence that encodes an N-terminal portion of a CLRN1 protein positioned 3' to the promoter, a splice donor sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splice donor sequence; and a second nucleic acid vector of the two different nucleic acid vectors includes a F1 phage recombinogenic region, a splice acceptor sequence
10 positioned 3' of the F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal sequence positioned at the 3' end of the second coding sequence; where each of the two encoded portions is at least 30 amino acid residues in length, where the amino acid sequences of the two encoded portions do
15 not overlap with each other; where no single vector of the two different vectors encodes a full-length CLRN1 protein; and when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein. In some embodiments of any of the compositions provided herein, the first coding sequence
20 encodes an N-terminal portion of a first isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the first isoform of CLRN1 protein. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the
25 CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first nucleic acid vector of the second nucleic acid vector further includes a sequence that encodes a second isoform of the CLRN1 protein. In some embodiments of any of the
30 compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of

the compositions provided herein, the second isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

5 In some embodiments of any of the compositions provided herein, the first coding sequence encodes an N-terminal portion of a second isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the second isoform of CLRN1 protein. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second
10 isoform of the CLRN1 protein includes SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the second isoform of the CLRN1 protein consists of SEQ ID NO: 5. In some embodiments of any of the compositions provided herein, the first nucleic acid vector or the second nucleic acid vector further includes a sequence that encodes a first isoform of the CLRN1 protein. In some embodiments of any of the
15 compositions provided herein, the first isoform of the CLRN1 protein includes a sequence that is at least 95% identical to SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein includes SEQ ID NO: 3. In some embodiments of any of the compositions provided herein, the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

20 In some embodiments of any of the compositions provided herein, one or both of the first nucleic acid vector and the second nucleic acid vector comprises a 5' untranslated region (UTR), a 3' UTR, or both. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the
25 compositions provided herein, the 5' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes at least 80 contiguous nucleotides
30 from anywhere within SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 5' UTR includes a sequence that is at least 80%

identical to SEQ ID NO: 12. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15. In some embodiments of any of the compositions provided herein, the 3' UTR includes a sequence that is at least 80% identical to SEQ ID NO: 15.

In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector. In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector 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 provided herein, each of the first nucleic acid vector and the second nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector. In some embodiments of any of the compositions provided herein, each of the first nucleic acid vector and the second nucleic acid vector is an AAV vector. In some embodiments of any of the compositions provided herein, at least one of the coding sequences includes a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons.

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

In some embodiments of any of the methods provided herein, the mammal is a human. In

some embodiments of any of the methods provided herein, the mammal has been previously identified as having a defective CLRN1 gene.

Also provided herein are methods of increasing expression of a full-length CLRN1 protein in a mammalian cell that include introducing any of the compositions
5 provided herein into the mammalian cell. In some embodiments of any of the methods provided herein, the mammalian cell is a cochlear inner hair cell or a cochlear outer hair cell. In some embodiments of any of the methods provided herein, the mammalian cell is a retinal cell. In some embodiments of any of the methods provided herein, the
10 mammalian cell is a human cell. In some embodiments of any of the methods provided herein, the mammalian cell has previously been determined to have a defective CLRN1 gene.

Also provided herein are methods of increasing expression of a full-length CLRN1 protein in an inner hair cell, an outer hair cell, or both, in a cochlea of a mammal, that include: introducing into the cochlea of the mammal a therapeutically effective
15 amount of any of the compositions described herein. Also provided herein are methods of increasing expression of a full-length CLRN1 protein in an eye of a mammal that include: intraocularly administering to the eye of the mammal a therapeutically effective amount of any of the compositions provided herein. In some embodiments of any of the methods provided herein, the mammal has been previously identified as having a
20 defective CLRN1 gene. In some embodiments of any of the methods provided herein, the mammal is a human.

Also provided herein are methods of treating hearing loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions provided herein into the cochlea of the subject. Also
25 provided herein are methods of treating vision loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions provided herein into the eye of the subject. In some embodiments of any of the methods provided herein, the subject has Usher syndrome type III. In some embodiments of any of the methods provided herein, the subject is a
30 human. Some embodiments of any of the methods provided herein further include, prior to the administering step, determining that the subject has a defective CLRN1 gene.

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 CLRN1 gene” refers to a modification in a wildtype
5 CLRN1 gene that results in the production of a CLRN1 protein having 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 CLRN1 protein, and/or results in a decrease in the expressed level of the encoded CLRN1 protein in a mammalian cell as compared to the expressed level of the encoded CLRN1 protein in a mammalian cell not
10 having a mutation. In some embodiments, a mutation can result in the production of a CLRN1 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 CLRN1 gene. The term “frameshift” is known in the art to encompass any mutation in a coding sequence that results in a shift in the
15 reading 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
20 embodiments, the mutation can result in the loss (or a decrease in the level) of expression of CLRN1 mRNA or CLRN1 protein or both the mRNA and protein. In some embodiments, the mutation can result in the production of an altered CLRN1 protein having a loss or decrease in one or more biological activities (functions) as compared to a wildtype CLRN1 protein.

25 In some embodiments, the mutation is an insertion of one or more nucleotides into a CLRN1 gene. In some embodiments, the mutation is in a regulatory sequence of the CLRN1 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 CLRN1 gene. The term
30 “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
5 that is transiently expressed in a cell (e.g., a mammalian cell) is lost upon multiple rounds of cell division.

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 sheep, a goat, a pig, a dog, a cat, a non-human primate, or a human. In some embodiments, the subject has or is at risk of
10 hearing loss and/or vision loss. In some embodiments, the subject has been previously identified as having a mutation in a CLRN1 gene. In some embodiments, the subject has been identified as having a mutation in a CLRN1 gene and has been diagnosed with hearing loss and/or vision loss. In some embodiments, the subject has been identified as having hearing loss and/or vision loss.

15 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 state (e.g., hearing loss or vision 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 CLRN1 protein (e.g., a wildtype, full-length CLRN1 protein
20 or a variant of a CLRN1 protein that has the desired activity) (e.g., as compared to the expression level prior to treatment with the composition). In some embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of an active CLRN1 protein (e.g., a wildtype, full-length CLRN1 protein or an active variant) in a target cell (e.g., a cochlear inner hair cell). In some
25 embodiments, a therapeutically effective amount of a composition can result in an increase in the expression level of an active CLRN1 protein (e.g., a wildtype, full-length CLRN1 protein or active variant), and/or an increase in one or more activities of a CLRN1 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 CLRN1
30 gene, or the level(s) in a subject or a population of subjects having hearing loss and/or vision 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 nucleic acids described herein, the nucleic acid is RNA.

The term “active CLRN1 protein” means a protein encoded by DNA that, if substituted for both wildtype alleles encoding full-length CLRN1 protein in auditory hair cells, or ocular cells, of what is otherwise a wildtype mammal, and if expressed in the auditory hair cells, or ocular cells, of that mammal, results in that mammal’s having a level of hearing, or vision, approximating the normal level of hearing, or vision, of a similar mammal that is entirely wildtype. Non-limiting examples of active CLRN1 proteins are full-length CLRN1 proteins (e.g., any of the full-length CLRN1 proteins described herein).

For example, an active CLRN1 protein can include a sequence of a wildtype, full-length CLRN1 protein (e.g., a wildtype, human, full-length CLRN1 protein) including 1 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 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 acid substitution to about 35 amino acid substitutions, 1 amino acid substitution to about 30 amino acid substitutions, 1 amino acid substitution to about 25 amino acid substitutions, 1 amino

23

24

25

26

27

28

29

30

acid substitutions to about 100 amino acid substitutions, about 90 amino acid substitutions to about 95 amino acid substitutions, or about 95 amino acids to about 100 amino acids.

One skilled in the art would appreciate that amino acids that are not conserved
5 between wildtype CLRN1 proteins from different species can be mutated without losing activity, while those amino acids that are conserved between wildtype CLRN1 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 CLRN1 protein can include, e.g., a sequence of a wildtype, full-length
10 CLRN1 protein (e.g., a wildtype, human, full-length CLRN1 protein) that has about 1 amino acid to about 100 amino acids, about 1 amino acid to about 95 amino acids, about 1 amino acid to about 90 amino acids, about 1 amino acid to about 85 amino acids, about 1 amino acid to about 80 amino acids, about 1 amino acid to about 75 amino acids, about 1 amino acid to about 70 amino acids, about 1 amino acid to about 65 amino acids, about
15 1 amino acid to about 60 amino acids, about 1 amino acid to about 55 amino acids, about 1 amino acid to about 50 amino acids, about 1 amino acid to about 45 amino acids, about 1 amino acid to about 40 amino acids, about 1 amino acid to about 35 amino acids, about 1 amino acid to about 30 amino acids, about 1 amino acid to about 25 amino acids, about 1 amino acid to about 20 amino acids, about 1 amino acid to about 15 amino acids, about
20 1 amino acid to about 10 amino acids, about 1 amino acid to about 9 amino acids, about 1 amino acid to about 8 amino acids, about 1 amino acid to about 7 amino acids, about 1 amino acid to about 6 amino acids, about 1 amino acid to about 5 amino acids, about 1 amino acid to about 4 amino acids, about 1 amino acid to about 3 amino acids, about 2 amino acids to about 100 amino acids, about 2 amino acids to about 95 amino acids,
25 about 2 amino acids to about 90 amino acids, about 2 amino acids to about 85 amino acids, about 2 amino acids to about 80 amino acids, about 2 amino acids to about 75 amino acids, about 2 amino acids to about 70 amino acids, about 2 amino acids to about 65 amino acids, about 2 amino acids to about 60 amino acids, about 2 amino acids to about 55 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,
5 about 2 amino acids to about 5 amino acids, about 2 amino acids to about 4 amino acids,
about 3 amino acids to about 100 amino acids, about 3 amino acids to about 95 amino
acids, about 3 amino acids to about 90 amino acids, about 3 amino acids to about 85
amino acids, about 3 amino acids to about 80 amino acids, about 3 amino acids to about
75 amino acids, about 3 amino acids to about 70 amino acids, about 3 amino acids to
10 about 65 amino acids, about 3 amino acids to about 60 amino acids, about 3 amino acids
to about 55 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,
15 about 3 amino acids to about 15 amino acids, about 3 amino acids to about 10 amino
acids, about 3 amino acids to about 9 amino acids, about 3 amino acids to about 8 amino
acids, about 3 amino acids to about 7 amino acids, about 3 amino acids to about 6 amino
acids, about 3 amino acids to about 5 amino acids, about 4 amino acids to about 100
amino acids, about 4 amino acids to about 95 amino acids, about 4 amino acids to about
20 90 amino acids, about 4 amino acids to about 85 amino acids, about 4 amino acids to
about 80 amino acids, about 4 amino acids to about 75 amino acids, about 4 amino acids
to about 70 amino acids, about 4 amino acids to about 65 amino acids, about 4 amino
acids to about 60 amino acids, about 4 amino acids to about 55 amino acids, about 4
amino acids to about 50 amino acids, about 4 amino acids to about 45 amino acids, about
25 4 amino acids to about 40 amino acids, about 4 amino acids to about 35 amino acids,
about 4 amino acids to about 30 amino acids, about 4 amino acids to about 25 amino
acids, about 4 amino acids to about 20 amino acids, about 4 amino acids to about 15
amino acids, about 4 amino acids to about 10 amino acids, about 4 amino acids to about 9
amino acids, about 4 amino acids to about 8 amino acids, about 4 amino acids to about 7
30 amino acids, about 4 amino acids to about 6 amino acids, about 5 amino acids to about
100 amino acids, about 5 amino acids to about 95 amino acids, about 5 amino acids to

about 90 amino acids, about 5 amino acids to about 85 amino acids, about 5 amino acids to about 80 amino acids, about 5 amino acids to about 75 amino acids, about 5 amino acids to about 70 amino acids, about 5 amino acids to about 65 amino acids, about 5 amino acids to about 60 amino acids, about 5 amino acids to about 55 amino acids, about 5 amino acids to about 50 amino acids, about 5 amino acids to about 45 amino acids, about 5 amino acids to about 40 amino acids, about 5 amino acids to about 35 amino acids, about 5 amino acids to about 30 amino acids, about 5 amino acids to about 25 amino acids, about 5 amino acids to about 20 amino acids, about 5 amino acids to about 15 amino acids, about 5 amino acids to about 10 amino acids, about 5 amino acids to about 9 amino acids, about 5 amino acids to about 8 amino acids, about 5 amino acids to about 7 amino acids, about 6 amino acids to about 100 amino acids, about 6 amino acids to about 95 amino acids, about 6 amino acids to about 90 amino acids, about 6 amino acids to about 85 amino acids, about 6 amino acids to about 80 amino acids, about 6 amino acids to about 75 amino acids, about 6 amino acids to about 70 amino acids, about 6 amino acids to about 65 amino acids, about 6 amino acids to about 60 amino acids, about 6 amino acids to about 55 amino acids, about 6 amino acids to about 50 amino acids, about 6 amino acids to about 45 amino acids, about 6 amino acids to about 40 amino acids, about 6 amino acids to about 35 amino acids, about 6 amino acids to about 30 amino acids, about 6 amino acids to about 25 amino acids, about 6 amino acids to about 20 amino acids, about 6 amino acids to about 15 amino acids, about 6 amino acids to about 10 amino acids, about 6 amino acids to about 9 amino acids, about 6 amino acids to about 8 amino acids, about 7 amino acids to about 100 amino acids, about 7 amino acids to about 95 amino acids, about 7 amino acids to about 90 amino acids, about 7 amino acids to about 85 amino acids, about 7 amino acids to about 80 amino acids, about 7 amino acids to about 75 amino acids, about 7 amino acids to about 70 amino acids, about 7 amino acids to about 65 amino acids, about 7 amino acids to about 60 amino acids, about 7 amino acids to about 55 amino acids, about 7 amino acids to about 50 amino acids, about 7 amino acids to about 45 amino acids, about 7 amino acids to about 40 amino acids, about 7 amino acids to about 35 amino acids, about 7 amino acids to about 30 amino acids, about 7 amino acids to about 25 amino acids, about 7 amino acids to about 20 amino acids, about 7 amino acids to about 15 amino acids, about 7 amino

acids to about 10 amino acids, about 7 amino acids to about 9 amino acids, about 8 amino acids to about 100 amino acids, about 8 amino acids to about 95 amino acids, about 8 amino acids to about 90 amino acids, about 8 amino acids to about 85 amino acids, about 8 amino acids to about 80 amino acids, about 8 amino acids to about 75 amino acids, about 8 amino acids to about 70 amino acids, about 8 amino acids to about 65 amino acids, about 8 amino acids to about 60 amino acids, about 8 amino acids to about 55 amino acids, about 8 amino acids to about 50 amino acids, about 8 amino acids to about 45 amino acids, about 8 amino acids to about 40 amino acids, about 8 amino acids to about 35 amino acids, about 8 amino acids to about 30 amino acids, about 8 amino acids to about 25 amino acids, about 8 amino acids to about 20 amino acids, about 8 amino acids to about 15 amino acids, about 8 amino acids to about 10 amino acids, about 10 amino acids to about 100 amino acids, about 10 amino acids to about 95 amino acids, about 10 amino acids to about 90 amino acids, about 10 amino acids to about 85 amino acids, about 10 amino acids to about 80 amino acids, about 10 amino acids to about 75 amino acids, about 10 amino acids to about 70 amino acids, about 10 amino acids to about 65 amino acids, about 10 amino acids to about 60 amino acids, about 10 amino acids to about 55 amino acids, about 10 amino acids to about 50 amino acids, about 10 amino acids to about 45 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 100 amino acids, about 15 amino acids to about 95 amino acids, about 15 amino acids to about 90 amino acids, about 15 amino acids to about 85 amino acids, about 15 amino acids to about 80 amino acids, about 15 amino acids to about 75 amino acids, about 15 amino acids to about 70 amino acids, about 15 amino acids to about 65 amino acids, about 15 amino acids to about 60 amino acids, about 15 amino acids to about 55 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 100 amino acids, about 20 amino acids to about 95 amino

acids, about 20 amino acids to about 90 amino acids, about 20 amino acids to about 85 amino acids, about 20 amino acids to about 80 amino acids, about 20 amino acids to about 75 amino acids, about 20 amino acids to about 70 amino acids, about 20 amino acids to about 65 amino acids, about 20 amino acids to about 60 amino acids, about 20 amino acids to about 55 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 100 amino acids, about 25 amino acids to about 95 amino acids, about 25 amino acids to about 90 amino acids, about 25 amino acids to about 85 amino acids, about 25 amino acids to about 80 amino acids, about 25 amino acids to about 75 amino acids, about 25 amino acids to about 70 amino acids, about 25 amino acids to about 65 amino acids, about 25 amino acids to about 60 amino acids, about 25 amino acids to about 55 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 100 amino acids, about 30 amino acids to about 95 amino acids, about 30 amino acids to about 90 amino acids, about 30 amino acids to about 85 amino acids, about 30 amino acids to about 80 amino acids, about 30 amino acids to about 75 amino acids, about 30 amino acids to about 70 amino acids, about 30 amino acids to about 65 amino acids, about 30 amino acids to about 60 amino acids, about 30 amino acids to about 55 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 100 amino acids, about 40 amino acids to about 95 amino acids, about 40 amino acids to about 90 amino acids, about 40 amino acids to about 85 amino acids, about 40 amino acids to about 80 amino acids, about 40 amino acids to about 75 amino acids, about 40 amino acids to about 70 amino acids, about 40 amino acids to about 65 amino acids, about 40 amino acids to about 60 amino acids, about 40 amino acids to about 55 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, about 50 amino acids to about 100 amino acids, about 50 amino acids to about 95 amino acids, about 50 amino acids to about 90 amino acids, about 50 amino acids to about 85 amino acids, about 50 amino acids to about 80 amino acids, about 50 amino acids to about 75 amino acids, about 50 amino acids to about 70 amino acids, about 50 amino acids to about 65 amino acids, about 50 amino acids to about 60 amino acids, or about 50 amino acids to about 55 amino acids, deleted. In some embodiments where two or more amino acids are deleted from the sequence of a wildtype, full-length CLRN1 protein, at least two of 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 CLRN1 protein, some or all of 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 CLRN1 proteins from different species can be deleted without losing activity, while those amino acids that are conserved between wildtype, full-length CLRN1 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 CLRN1 protein can, e.g., include a sequence of a wildtype, full-length CLRN1 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 acids, 1 amino acid to about 85 amino acids, 1 amino acid to about 80 amino acids, 1 amino acid to about 75 amino acids, 1 amino acid to about 70 amino acids, 1 amino acid to about 65 amino acids, 1 amino acid to about 60 amino acids, 1 amino acid to about 55 amino acids, 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 100 amino acids, about 2 amino acid to about 95 amino acids, about 2 amino acids to about 90 amino acids, about 2 amino acids to about 85 amino acids, about 2 amino acids to about 80 amino acids, about 2 amino acids to about 75 amino acids, about 2 amino acids to about 70 amino acids, about 2 amino acids to about 65 amino acids, about 2 amino acids to about 60 amino acids, about 2 amino acids to about 55 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 100 amino acids, about 3 amino acid to about 95 amino acids, about 3 amino acids to about 90 amino acids, about 3 amino acids to about 85 amino acids, about 3 amino acids to about 80 amino acids, about 3 amino acids to about 75 amino acids, about 3 amino acids to about 70 amino acids, about 3 amino acids to about 65 amino acids, about 3 amino acids to about 60 amino acids, about 3 amino acids to about 55 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 about 10 amino acids, about 3 amino acids to about 9 amino acids, about 3 amino acids to about 8 amino acids, about 3 amino acids to about 7 amino acids, about 3 amino acids to about 6 amino acids, about 3 amino acids to about 5 amino acids, about 4 amino acids to about 100 amino acids, about 4 amino acid to about 95 amino acids, about 4 amino acids to about 90 amino acids, about 4 amino acids to about 85 amino acids, about 4 amino acids to about 80 amino acids, about 4 amino acids to about 75 amino acids, about 4 amino acids to about 70 amino acids, about 4 amino acids to about 65 amino acids, about 4 amino acids to about 60 amino acids, about 4 amino acids to about 55 amino acids, about

4 amino acids to about 50 amino acids, about 4 amino acids to about 45 amino acids,
about 4 amino acids to about 40 amino acids, about 4 amino acids to about 35 amino
acids, about 4 amino acids to about 30 amino acids, about 4 amino acids to about 25
amino acids, about 4 amino acids to about 20 amino acids, about 4 amino acids to about
5 15 amino acids, about 4 amino acids to about 10 amino acids, about 4 amino acids to
about 9 amino acids, about 4 amino acids to about 8 amino acids, about 4 amino acids to
about 7 amino acids, about 4 amino acids to about 6 amino acids, about 5 amino acids to
about 100 amino acids, about 5 amino acid to about 95 amino acids, about 5 amino acids
to about 90 amino acids, about 5 amino acids to about 85 amino acids, about 5 amino
10 acids to about 80 amino acids, about 5 amino acids to about 75 amino acids, about 5
amino acids to about 70 amino acids, about 5 amino acids to about 65 amino acids, about
5 amino acids to about 60 amino acids, about 5 amino acids to about 55 amino acids,
about 5 amino acids to about 50 amino acids, about 5 amino acids to about 45 amino
acids, about 5 amino acids to about 40 amino acids, about 5 amino acids to about 35
15 amino acids, about 5 amino acids to about 30 amino acids, about 5 amino acids to about
25 amino acids, about 5 amino acids to about 20 amino acids, about 5 amino acids to
about 15 amino acids, about 5 amino acids to about 10 amino acids, about 5 amino acids
to about 9 amino acids, about 5 amino acids to about 8 amino acids, about 5 amino acids
to about 7 amino acids, about 6 amino acids to about 100 amino acids, about 6 amino
20 acid to about 95 amino acids, about 6 amino acids to about 90 amino acids, about 6
amino acids to about 85 amino acids, about 6 amino acids to about 80 amino acids, about
6 amino acids to about 75 amino acids, about 6 amino acids to about 70 amino acids,
about 6 amino acids to about 65 amino acids, about 6 amino acids to about 60 amino
acids, about 6 amino acids to about 55 amino acids, about 6 amino acids to about 50
25 amino acids, about 6 amino acids to about 45 amino acids, about 6 amino acids to about
40 amino acids, about 6 amino acids to about 35 amino acids, about 6 amino acids to
about 30 amino acids, about 6 amino acids to about 25 amino acids, about 6 amino acids
to about 20 amino acids, about 6 amino acids to about 15 amino acids, about 6 amino
acids to about 10 amino acids, about 6 amino acids to about 9 amino acids, about 6 amino
30 acids to about 8 amino acids, about 7 amino acids to about 100 amino acids, about 7
amino acid to about 95 amino acids, about 7 amino acids to about 90 amino acids, about

7 amino acids to about 85 amino acids, about 7 amino acids to about 80 amino acids,
about 7 amino acids to about 75 amino acids, about 7 amino acids to about 70 amino
acids, about 7 amino acids to about 65 amino acids, about 7 amino acids to about 60
amino acids, about 7 amino acids to about 55 amino acids, about 7 amino acids to about
5 50 amino acids, about 7 amino acids to about 45 amino acids, about 7 amino acids to
about 40 amino acids, about 7 amino acids to about 35 amino acids, about 7 amino acids
to about 30 amino acids, about 7 amino acids to about 25 amino acids, about 7 amino
acids to about 20 amino acids, about 7 amino acids to about 15 amino acids, about 7
amino acids to about 10 amino acids, about 7 amino acids to about 9 amino acids, about 8
10 amino acids to about 100 amino acids, about 8 amino acid to about 95 amino acids, about
8 amino acids to about 90 amino acids, about 8 amino acids to about 85 amino acids,
about 8 amino acids to about 80 amino acids, about 8 amino acids to about 75 amino
acids, about 8 amino acids to about 70 amino acids, about 8 amino acids to about 65
amino acids, about 8 amino acids to about 60 amino acids, about 8 amino acids to about
15 55 amino acids, about 8 amino acids to about 50 amino acids, about 8 amino acids to
about 45 amino acids, about 8 amino acids to about 40 amino acids, about 8 amino acids
to about 35 amino acids, about 8 amino acids to about 30 amino acids, about 8 amino
acids to about 25 amino acids, about 8 amino acids to about 20 amino acids, about 8
amino acids to about 15 amino acids, about 8 amino acids to about 10 amino acids, about
20 10 amino acids to about 100 amino acids, about 10 amino acid to about 95 amino acids,
about 10 amino acids to about 90 amino acids, about 10 amino acids to about 85 amino
acids, about 10 amino acids to about 80 amino acids, about 10 amino acids to about 75
amino acids, about 10 amino acids to about 70 amino acids, about 10 amino acids to
about 65 amino acids, about 10 amino acids to about 60 amino acids, about 10 amino
25 acids to about 55 amino acids, about 10 amino acids to about 50 amino acids, about 10
amino acids to about 45 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 20 amino acids to
30 about 100 amino acids, about 20 amino acid to about 95 amino acids, about 20 amino
acids to about 90 amino acids, about 20 amino acids to about 85 amino acids, about 20

amino acids to about 80 amino acids, about 20 amino acids to about 75 amino acids,
about 20 amino acids to about 70 amino acids, about 20 amino acids to about 65 amino
acids, about 20 amino acids to about 60 amino acids, about 20 amino acids to about 55
amino acids, about 20 amino acids to about 50 amino acids, about 20 amino acids to
5 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 30 amino acids to about 100 amino acids,
about 30 amino acid to about 95 amino acids, about 30 amino acids to about 90 amino
acids, about 30 amino acids to about 85 amino acids, about 30 amino acids to about 80
10 amino acids, about 30 amino acids to about 75 amino acids, about 30 amino acids to
about 70 amino acids, about 30 amino acids to about 65 amino acids, about 30 amino
acids to about 60 amino acids, about 30 amino acids to about 55 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
15 acids, about 40 amino acids to about 100 amino acids, about 40 amino acid to about 95
amino acids, about 40 amino acids to about 90 amino acids, about 40 amino acids to
about 85 amino acids, about 40 amino acids to about 80 amino acids, about 40 amino
acids to about 75 amino acids, about 40 amino acids to about 70 amino acids, about 40
amino acids to about 65 amino acids, about 40 amino acids to about 60 amino acids,
20 about 40 amino acids to about 55 amino acids, about 40 amino acids to about 50 amino
acids, about 40 amino acids to about 45 amino acids, about 50 amino acids to about 100
amino acids, about 50 amino acid to about 95 amino acids, about 50 amino acids to about
90 amino acids, about 50 amino acids to about 85 amino acids, about 50 amino acids to
about 80 amino acids, about 50 amino acids to about 75 amino acids, about 50 amino
25 acids to about 70 amino acids, about 50 amino acids to about 65 amino acids, about 50
amino acids to about 60 amino acids, about 50 amino acids to about 55 amino acids,
about 60 amino acids to about 100 amino acids, about 60 amino acid to about 95 amino
acids, about 60 amino acids to about 90 amino acids, about 60 amino acids to about 85
amino acids, about 60 amino acids to about 80 amino acids, about 60 amino acids to
30 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 CLRN1 protein can, e.g., include the sequence of a wildtype, full-length CLRN1 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 8 amino acids, about 4 amino acids to 7 amino acids, about 4 amino acids to 6
5 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
acids to 10 amino acids, about 5 amino acids to 9 amino acids, about 5 amino acids to 8
10 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 amino
acids to 15 amino acids, about 6 amino acids to 10 amino acids, about 6 amino acids to 9
15 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, about 7 amino
acids to 15 amino acids, about 7 amino acids to 10 amino acids, about 7 amino acids to 9
20 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 acids, about 8 amino
acids to 10 amino acids, about 10 amino acids to 50 amino acids, about 10 amino acids to
25 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, 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
30 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 inserted in multiple, non-contiguous places in 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.

Other features and advantages of the invention will be apparent from the following detailed description and figures, and from the claims

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1 is an exemplary schematic representation of a genetic map of a CLRN-1 vector (SEQ ID NO: 40; 3397 basepairs (bp)) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID

NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 6), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 2 is an exemplary schematic representation of a genetic map of a CLRN-2GFPvector (SEQ ID NO: 41; 4177 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 6), T2A sequence (SEQ ID NO: 31), an eGFP (SEQ ID NO: 32), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 3 is an exemplary schematic representation of a genetic map of a CLRN-3 vector (SEQ ID NO: 42; 4607 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), 5'UTR-291, CLRN1 isoform 1 (SEQ ID NO: 1), 3'UTR-1357 (SEQ ID NO: 36), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 4 is an exemplary schematic representation of a genetic map of a CLRN-4 vector (SEQ ID NO: 43; 4796 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 6), 3'UTR-1406 (SEQ ID NO: 37), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 5 is an exemplary schematic representation of a genetic map of a pITR-CBA-5'UTR-tGFP-3'UTR vector (5026 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), 5'UTR-291, tGFP (SEQ ID NO: 19), 3'UTR-1595, bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 6 is an exemplary schematic representation of a genetic map of a CLRN-6eGFP vector (SEQ ID NO: 44; 4756 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17),

chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), eGFP (SEQ ID NO: 32), 3'UTR-1406 (SEQ ID NO: 37), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 7 is an exemplary schematic representation of a genetic map of a pITR-CBA-3'UTR-600A vector (3982 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 4), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 6), 3'UTR-600 (SEQ ID NO: 27), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 8 is an exemplary schematic representation of a genetic map of a CLRN-8 vector (SEQ ID NO: 46; 3982 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 6), 3'UTR-600B (SEQ ID NO: 28), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 9 is an exemplary schematic representation of a genetic map of a CLRN-9 vector (SEQ ID NO: 47; 3982 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 6), 3'UTR-600C (SEQ ID NO: 29), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 10 is an exemplary schematic representation of a genetic map of a CLRN-0 vector (SEQ ID NO: 39; 4732 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, a chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), a 3'UTR 1773 (SEQ ID NO: 15), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 11 is an exemplary schematic representation of a genetic map of a CLRN-7eGFP vector (SEQ ID NO: 45; 3580 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17),

chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), an eGFP sequence (SEQ ID NO: 32), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 12 is an exemplary schematic representation of a genetic map of a CLRN-10 (SEQ ID NO: 48; 3511 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 4), a HA sequence (SEQ ID NO: 34), a FP sequence (SEQ ID NO: 30), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 5), 3x FLAG tag sequence (SEQ ID NO: 35), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 13 is an exemplary schematic representation of a genetic map of a CLRN-10myc vector (SEQ ID NO: 49; 3574 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), a myc sequence (SEQ ID NO: 33), a FP sequence (SEQ ID NO: 30), T2A sequence (SEQ ID NO: 31), 3x FLAG tag sequence (SEQ ID NO: 35), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 14 is an exemplary schematic representation of a genetic map of a CLRN-10NF vector (SEQ ID NO: 50; 3499 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), T2A sequence (SEQ ID NO: 31), an HA sequence (SEQ ID NO: 34), CLRN-1 isoform 2 (SEQ ID NO: 5), 3x FLAG tag sequence (SEQ ID NO: 35), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 15 is an exemplary schematic representation of a genetic map of a CLRN-11 vector (SEQ ID NO: 51; 4908 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), a HA sequence (SEQ ID NO: 34), a FP sequence (SEQ ID NO: 30), T2A sequence (SEQ ID NO: 31), CLRN-1 isoform 2 (SEQ ID NO: 5), 3x FLAG tag sequence (SEQ ID NO: 35), 3'UTR-1406 (SEQ ID NO: 37), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 16 is an exemplary schematic representation of a genetic map of a CLRN-11myc vector (SEQ ID NO: 52; 4971 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), a myc sequence (SEQ ID NO: 33), a FP sequence (SEQ ID NO: 30), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 5), 3x FLAG tag sequence (SEQ ID NO: 35), 3'UTR-1406 (SEQ ID NO: 37), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 17 is an exemplary schematic representation of a genetic map of a CLRN-11NF vector (SEQ ID NO: 53; 4896 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 1 (SEQ ID NO: 1), an HA sequence (SEQ ID NO: 34), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 2 (SEQ ID NO: 5), 3x FLAG sequence (SEQ ID NO: 33), 3'UTR-1406 (SEQ ID NO: 37), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 18 is an exemplary schematic representation of a genetic map of a CLRN-12 vector (SEQ ID NO: 54; 4640 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), a 5'UTR-291 sequence (SEQ ID NO: 12), CLRN1 isoform 1 (SEQ ID NO: 1), an HA sequence (SEQ ID NO: 34), 3'UTR-1357 (SEQ ID NO: 36), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 19 is an exemplary schematic representation of a genetic map of a CLRN-13 vector (SEQ ID NO: 55; 4291 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 4 (SEQ ID NO: 7), 3x FLAG tag sequence (SEQ ID NO: 35), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 1 (SEQ ID NO: 1), an HA sequence (SEQ ID NO: 34), 3'UTR-600 (SEQ ID NO: 27), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 20 is an exemplary schematic representation of a genetic map of a CLRN-14 vector (SEQ ID NO: 56; 4192 bp) that can be used in any of the present methods

described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 4 (SEQ ID NO: 7), T2A sequence (SEQ ID NO: 31), CLRN1 isoform 1 (SEQ ID NO: 1), 3'UTR-600 (SEQ ID NO: 27), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 21 is an exemplary schematic representation of a genetic map of a CLRN-15 vector (SEQ ID NO: 57; 3505 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 4 (SEQ ID NO: 7), 3x FLAG tag sequence (SEQ ID NO: 35), 3'UTR-600 (SEQ ID NO: 27), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 22 is an exemplary schematic representation of a genetic map of a CLRN-16 vector (SEQ ID NO: 58; 3439 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, chimeric intron (SEQ ID NO: 16), CLRN1 isoform 4 (SEQ ID NO: 7), 3'UTR-600 (SEQ ID NO: 27), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 23 is an exemplary schematic representation of a genetic map of a CLRN-17 vector (SEQ ID NO: 59; 130 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, a sh-chimeric intron (SEQ ID NO: 26), 5'UTR-291 (SEQ ID NO: 12), CLRN1 isoform 1 (SEQ ID NO: 1), an HA sequence (SEQ ID NO: 34), 3'UTR-1773 (SEQ ID NO: 15), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 24 is an exemplary schematic representation of a genetic map of a CLRN-18 vector (SEQ ID NO: 60; 4277 bp) that can be used in any of the present methods described herein. The vector includes an AAV2 ITR, CMV enhancer (SEQ ID NO: 17), chicken β -actin promoter, a sh-chimeric intron (SEQ ID NO: 26), CLRN1 isoform 1 (SEQ ID NO: 1), 3'UTR-1773 (SEQ ID NO: 15), bGH poly(A) signal (SEQ ID NO: 20) and an AAV2 ITR.

FIG. 25 is an image of an immunoblot of CLRN1 protein levels from transfected HEK293FT cells 48-hours post-transfection using anti-HA and anti-FLAG antibodies. Lane 1 – PageRuler Plus Prestained; Lane 2 – CLRN-10 (37 °C denaturing); Lane 3 –

CLRN-11 (37 °C denaturing); Lane 4 – CLRN-12 (37 °C denaturing); Lane 5 – negative control; Lane 6 – CLRN-10 (56 °C denaturing); Lane 7 – CLRN-11 (56 °C denaturing); Lane 8 – CLRN-12 (56 °C denaturing); Lane 9 – negative control. CLRN-1 isoform protein is a glycosylated protein and often migrates as smear bands.

5 **FIG. 26** is an image of an immunoblot of CLRN1 protein levels from transfected HEK293FT cells 48-hours post-transfection using an anti- FLAG antibody. Lane 1 – PageRuler Plus Prestained; Lane 2 – CLRN-10; Lane 3 – CLRN-11; Lane 4 – CLRN-12; Lane 5 – CLRN-13; Lane 6 – CLRN-15; Lane 7 – negative control

10 **FIG. 27** is an image of an immunoblot of CLRN1 protein levels from transfected HEK293FT cells 48-hours post-transfection using an anti- FLAG antibody. Lane 1 – PageRuler Plus Prestained; Lane 2 – CLRN-10; Lane 3 – CLRN-10NF; Lane 4 – CLRN-10myc; Lane 5 – CLRN-11NF; Lane 6 – CLRN-11myc; Lane 7 – negative control. All samples were kept at room temperature.

15 **FIG. 28** is an image of an immunoblot of CLRN1 protein levels from transfected HEK293FT cells 48-hours post-transfection using an anti-myc antibody. Lane 1 – PageRuler Plus Prestained; Lane 2 – CLRN-10; Lane 3 – CLRN-10NF; Lane 4 – CLRN-10myc; Lane 5 – CLRN-11NF; Lane 6 – CLRN-11myc; Lane 7 – negative control.

20 **FIG. 29** is an image of an immunoblot of CLRN1 protein levels harvested from transfected HEK293FT cells 48 hours post-transfection using anti-HA and anti-FLAG antibodies. Lane 1 – PageRuler Plus Prestained; Lane 2 – CLRN-13; Lane 3 – CLRN-10; Lane 4 – CLRN-10NF; Lane 5 – CLRN-10myc; Lane 6 – CLRN-11NF; Lane 7 – CLRN-11myc; Lane 8 – negative control.

FIG. 30 is an image of an immunoblot using anti-CLRN (EKIANYKEGTYVYKTQSEKY; SEQ ID NO: 38) rabbit polyclonal antibody of
25 CLRN1 isoform 1 protein levels harvested from HEK293FT cells transfected with plasmids described herein 48 hours post-transfection. Lane 1 – PageRuler Plus Prestained; Lane 2 – CLRN-10; Lane 3 – CLRN-1; Lane 4- CLRN-2; Lane 5 – CLRN-3; Lane 6 – CLRN-4; Lane 7 – CLRN-8; Lane 8 – CLRN-9; Lane 9 – CLRN-10; Lane 10 – CLRN-11; Lane 11 – CLRN-12; Lane 12 – CLRN-13, Lane 13 – CLRN-14; Lane 14 –
30 CLRN-15; Lane 15 – CLRN-16; Lane 16 – negative control; Lane 17 – negative control.

FIG. 31 is an image of an immunoblot CLRN1 protein levels harvested from transfected HEK293FT cells 48 hours post-transfection using an anti-CLRN rabbit polyclonal antibody. Lane 1 – PageRuler Plus Prestained; Lane 2 – Anc80-CLRN-0; Lane 3 – Anc80-CLRN-0 +PNGase F; Lane 4- Anc80-CLRN-3; Lane 5 – Anc80-CLRN-3 +PNGase F; Lane 6 – Anc80-CLRN-6eGFP; Lane 7 – Anc80-CLRN-6eGFP +PNGase F; Lane 8 – Anc80-CLRN-13; Lane 9 – Anc80-CLRN-13 +PNGase F; Lane 10 – no vector; Lane 11 – no vector +PNGase F. CLRN-1 isoform protein is a glycosylated protein and often migrates as smear bands (lanes 2, 4 and 8). After treatment with PNGase F, the smeared bands disappeared and shifted to distinct bands (lanes 3, 5 and 8).

FIG. 32 is a set of immunofluorescent images of AAVanc80-CLRN6eGFP transduced HEK293FT cells taken 24 hours and 48 hours post-transfection at MOI 8.41E+04 and MOI 2.53E+05, respectively.

FIG. 33 is a bar graph showing the relative CLRN1 and GFP expression in HEK293FT cells transduced with AAVanc80-CLRN-6eGFP (at MOI 1.05E+05 and MOI 3.15E+05), AAVanc80-CLRN-0 (at MOI 8.23E+04 and MOI 2.47E+05), AAVanc80-CLRN-3 (at MOI 8.41E+04 and MOI 2.53E+04), and AAVanc80-CLRN-13 (at MOI 8.33E+04 and MOI 2.50E+05), respectively.

FIG. 34 is a bar graph showing the relative CLRN1 and GFP expression in P2 cochlear explants from WT mice infected 16-hours with AAVanc80-CLRN-6eGFP (at MOI 2.0E+05), AAVanc80-CLRN-0 (at MOI 2.5E+05 and MOI 7.6E+05), AAVanc80-CLRN-3 (at MOI 2.0E+05 and 6.03E+05) and AAVanc80-CLRN-13 (at MOI 2.0E+05 and MOI 6.0E+05), respectively.

FIG. 35 is a set of fluorescent images of P2 cochlear explants from WT mice infected 72-hours with 1.3E10 AAVanc80-CLRN-0 VG/cochlea, 9.9E9 AAVanc80-CLRN-3 VG/cochlea and 1.0E10 AAVanc80-CLRN-13 VG/cochlea showing Myo7a and DAPI staining.

FIG. 36 is a set of fluorescent images of P2 cochlear explants from WT mice infected 72-hours with 1E09 VG/cochlea AAV Anc80.CAG.eGFP.3'UTR showing eGFP, Myo7a and DAPI staining.

DETAILED DESCRIPTION

Deficiency or mutations in “clarin 1,” the protein encoded by the CLRN1 gene, causes hearing loss and vision loss. For example, mutations in CLRN1 lead to Usher syndrome type III and retinitis pigmentosa.

5 Provided herein are compositions that include 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 CLRN1 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
10 different one of the encoded portions; no single vector of the at least two different vectors encodes a full-length CLRN1 protein; at least one of the coding sequences comprises a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons; and when introduced into a mammalian cell, the at least two different vectors undergo homologous
15 recombination with each other, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

 Provided herein are compositions that include a single nucleic acid vector, wherein the vector comprises one or both of (i) a first coding sequence encoding a first isoform of CLRN1 protein; and (ii) a second coding sequence encoding a second isoform
20 of CLRN1 protein, where one or both of the first and second coding sequences comprises a nucleotide sequence spanning two consecutive exons of a CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive introns.

 Provided herein are compositions that include two different nucleic acid vectors, wherein: a first nucleic acid vector of the two different nucleic acid vectors comprises a
25 promoter, a first coding sequence that encodes an N-terminal portion of a CLRN1 protein positioned 3' of the promoter, and a splice donor 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 splice acceptor sequence, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor
30 sequence, and a polyadenylation signal 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 sequences of the two encoded portions do not overlap with each other; wherein no single vector of the two different vectors encodes a full-length CLRN1 protein; and when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

Provided herein are compositions that include 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 CLRN1 protein positioned 3' of the promoter, a splice donor sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splice donor sequence; and a second nucleic acid vector of the two different nucleic acid vectors comprises a second detectable marker gene, a splice acceptor sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal 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 sequences of the encoded portions do not overlap with each other; wherein no single vector of the two different vectors encodes a full-length CLRN1 protein; and when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

Also provided herein are compositions that include 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 CLRN1 protein positioned 3' to the promoter, a splice donor sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splice donor sequence; and a second nucleic acid vector of the two different nucleic acid vectors comprises a F1 phage recombinogenic region, a splice acceptor sequence positioned 3' of the F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal sequence positioned at the 3' end of the

second coding sequence; wherein each of the two encoded portions is at least 30 amino acid residues in length, wherein the amino acid sequences of the two encoded portions do not overlap with each other; wherein no single vector of the two different vectors encodes a full-length CLRN1 protein; and when introduced into a mammalian cell, splicing
5 occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

Provided herein are methods that include: introducing into a cochlea of a mammal a therapeutically effective amount of any of the compositions described herein.

Provided herein are methods of increasing expression of a full-length CLRN1
10 protein in a mammalian cell, that include: introducing any of the compositions described herein into the mammalian cell.

Provided herein are methods of increasing expression of a full-length CLRN1 protein in an inner hair cell, an outer hair cell, or both, in a cochlea of a mammal that include: introducing into the cochlea of the mammal a therapeutically effective amount of
15 any of the compositions described herein.

Provided herein are methods of increasing expression of a full-length CLRN1 protein in an eye of a mammal that include: intraocularly administering to the eye of the mammal a therapeutically effective amount of any of the compositions described herein.

Provided herein are methods of treating hearing loss in a subject identified as
20 having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions described herein into the cochlea of the subject.

Also provided herein are methods of treating vision loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions described herein into the eye of the subject.

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

CLRN1

The CLRN1 gene encodes “clarin 1” (CLRN1), a protein that is expressed in hair
30 cells of the inner ear (e.g., inner ear hair cells, outer ear hair cells) and in the retina.

The human CLRN1 gene is located on chromosome 3q25.1. It contains 7 exons encompassing ~ 47 kilobases (kb) (Vastinsalo et al. (2011) Eur J Hum Genet 19(1): 30-35; NCBI Accession No. NG_009168.1).

Various mutations in the CLRN1 genes have been associated with Usher syndrome type III (e.g., Usher syndrome type IIIA (MIM#606397) (see, e.g., Fields et al. (2002) Am J Hum Genet 71: 607-617, and Joensuu et al. (2001) Am J Hum Genet 69: 673-684) and retinitis pigmentosa (see, e.g., Khan et al. (2011) Ophthalmology 118: 1444-1448). Usher syndrome type III-causing mutations have been predominantly found in exon 3 of CLRN1. Usher syndrome type III-deafness can be modeled by generating CLRN1- deficient mice (see, e.g., Geng et al. (2017) Sci Rep 7(1): 13480). Exemplary mutations CLRN1-associated with Usher syndrome type III include: T528G, M120K, M44K, N48K, and C40G.

Exemplary mutations CLRN1-associated with retinitis pigmentosa include L154W and P31L (see, e.g., Khan et al. (2011) Ophthalmology 118: 1444-1448).

Additional exemplary mutations in a CLRN1 gene that have been detected in subjects having hearing loss and methods of sequencing a nucleic acid encoding CLRN1 are described in, e.g., Fields et al. (2002) Am J Hum Genet 71: 607-617, Joensuu et al. (2001) Am J Hum Genet 69: 673-684, Adato et al. (2002) Europ J Hum Genet 10: 339-350, Aller et al. (2004), Clin Genet 66: 525-529. 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 CLRN1 protein is or includes the sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 and SEQ ID NO: 7. Non-limiting examples of nucleotide sequences encoding a wildtype CLRN1 protein are or include SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 and SEQ ID NO: 8.

In some embodiments of any of the compositions described herein, the CLRN1 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%) identical to SEQ ID NO: 1.

In some embodiments of any of the compositions described herein, the CLRN1 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%) identical to SEQ ID NO: 3.

5 In some embodiments of any of the compositions described herein, the CLRN1 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%) identical to SEQ ID NO: 5.

In some embodiments of any of the compositions described herein, the CLRN1 protein
10 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%) identical to SEQ ID NO: 7.

Human Full-length Wildtype CLRN1 Protein Isoform D (SEQ ID NO: 1)

15 MPSQQKKIIFCMAGVFSFACALGVVTALGTPLWIKATVLCKTGALLVNASGQELDKFMGEMQYGLFHGEGV
RQCGLGARPFRESFFPDLLKAIPVSIHVNVLFSAILIVLTMVGTAFFMYNAFGKPFETLHGPLGLYLLSF
ISVALWLPATRHQAQSGCGLVMILFASEVKIHHLSEKIANANYKEGTYYVYKTQSEKYTTSEFWVIFFCFFVHF
LNGLLIRLAGFQFPFAKSKDAETTNVAADLMY

20 Human Wildtype CLRN1 Isoform D cDNA (SEQ ID NO: 2)

atgccaaagc caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg
tgccctcgga gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg
caaaacggga gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga
aatgcagtac gggcttttcc acggagaggg tgtgaggcag tgtgggttgg gagcaaggcc
25 ctttcggttc tcattttttc cagatttgct caaagcaatc ccagttagca tccacgtcaa
tgtcattctc ttctctgcca tccttattgt gttaaccatg gtggggacag ccttcttcat
gtacaatgct tttggaaaac cttttgaaac tctgcatggt cccctagggc tgtacctttt
gagcttcatt tcagttgccc tttggtgccc agctaccagg caccaggctc aaggctcctg
tggctgtcct gtcatgatat tgtttgcctc tgaagtgaat atccatcacc tctcagaaaa
30 aattgcaaat tataaagaag ggacttatgt ctacaaaacg caaagtgaat aatataccac
ctcattctgg gtcattttct tttgcttttt tgttcatttt ctgaatgggc tctaataacg
acttgctgga tttcagttcc cttttgcaaa atctaaagac gcagaaacaa ctaatgtagc
tgcagatcta atgtactga

35

Human Full-length Wildtype CLRN1 Protein Isoform A (SEQ ID NO: 3)

MPSQQKKIIFCMAGVFSFACALGVVTALGTPLWIKATVLCKTGALLVNASGQELDKFMGEMQYGLFHGEGV
 RQCGLGARPFREFSFFPDLLKAIPVSIHVNILFSAILIVLTMVGTAFFMYNAFGKPFETLHGPLGLYLLSF
 ISGSCGCLVMILFASEVKIHHLSEKIANEYKEGTYVYKTQSEKYTTSEFWVIFFCFFVHFLNGLLIRLAGFQF
 5 PFAKS KDAETTNVAADLMY

Human Wildtype CLRN1 Isoform A cDNA (SEQ ID NO: 4)

atgccaaagc caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg
 tgccctcgga gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg
 10 caaaacggga gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga
 aatgcagtac gggcttttcc acggagaggg tgtgagggcag tgtgggttg gagcaaggcc
 ctttcgggtc tcatTTTTTt cagatttgct caaagcaatc ccagttagca tccacgtcaa
 tgtcattctc ttctctgcca tccttattgt gttaccatg gtggggacag cttcttcat
 gtacaatgct tttggaaaac cttttgaaac tctgcatggt cccctagggc tgtacctttt
 15 gagcttcatt tcaggctcct gtggctgtct tgtcatgata ttgtttgcct ctgaagtga
 aatccatcac ctctcagaaa aaattgcaaa ttataaagaa gggacttatg tctacaaaac
 gcaaagtga aaatatacca cctcattctg ggtcattttc ttttgctttt ttgttcattt
 tctgaatggg ctccataatac gacttgctgg atttcagttc ccttttgcaa aatctaaaga
 cgcagaaaca actaatgtag ctgcagatct aatgtactga
 20

Human Full-length Wildtype CLRN1 Protein Isoform C (SEQ ID NO: 5)

MQALQQQPVFPDLLKAIPVSIHVNILFSAILIVLTMVGTAFFMYNAFGKPFETLHGPLGLYLLSFISGSC
 GCLVMILFASEVKIHHLSEKIANEYKEGTYVY KTQSEKYTTSEFWLTKGHS

Human Wildtype CLRN1 Isoform C cDNA (SEQ ID NO: 6)

atgcaggc cctgcagcag caaccagttt ttccagattt gctcaaagca atcccagtga
 gcatccacgt caatgtcatt ctcttctctg ccatccttat tgtgttaacc atgggtggga
 cagccttctt catgtacaat gcttttgga aaccttttga aactctgcat ggtcccctag
 ggctgtacct tttgagcttc atttcaggct cctgtggctg tcttgctcatg atattgtttg
 30 cctctgaagt gaaaatccat cacctctcag aaaaaattgc aaattataaa gaagggactt
 atgtctacaa aacgcaaagt gaaaaatata ccacctcatt ctggctgact aaaggccaca gctga

Human Full-length Wildtype CLRN1 Protein Isoform E (SEQ ID NO: 7)

MPSQQKKIIFCMAGVFSFACALGVVTALGTPLWIKATVLCKTGALLVNASGQELDKFMGEMQYGLFHGEGV
 35 RQCGLGARPFREFSCYFLDPFMGLPTGVPHLLSLPCSTSCRREHTSERVQEPAGCFSAVRSKLHAGPAAATS
 FSRFAQSNPSEHPRQCHS LLCHPYCVNHGGDSLLHVQCFWKTF

Human Wildtype CLRN1 Isoform E cDNA (SEQ ID NO: 8)

atgccaagc caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg
 tgccctcgga gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg
 caaaacggga gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga
 5 aatgcagtac gggcttttcc acggagaggg tgtgaggcag tgtgggttg gagcaaggcc
 ctttcggttc tcatgctatt ttcttgacct cttcatggga ctccaacag gggtaaccca
 ttactcagc ctgccctgct caacctcttg caggagggag cacacgagt aacgagtgca
 ggaaccagct ggctgcttta gtgctgtgag gtagaaact catgcaggcc ctgcagcagc
 aaccagtttt tccagatttg ctcaaagcaa tcccagtgag catccacgtc aatgtcattc
 10 tcttctctgc catccttatt gtgttaacca tggtagggac agccttcttc atgtacaatg
 cttttgaaa accttttga

Human Wildtype CLRN1 Genomic sequence (SEQ ID NO: 9)

aggagatact tgaaggcagt ttgaaagact tgttttacag attcttagtc caaagatttc 61
 15 caattagga gaagaagcag cagaaaagga gaaaagccaa gtatgagtga tgatgaggcc 121
 ttcactact gacatttaac ctggcgagaa ccgtcgatgg tgaagttgcc ttttcagctg 181
 ggagctgtcc gttcagcttc cgtaataaat gcagtcaaag aggcagtccc ttcccattgc 241
 tcacaaaggt cttgtttttg aacctgccc tcacagaagc cgtttctcat catgccaaagc 301
 caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg tgccctcgga 361
 20 gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg caaaacggga 421
 gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga aatgcagtac 481
 gggcttttcc acggagaggg tgtgaggcag tgtgggttg gagcaaggcc ctttcggttc 541
 tcatgtaagt agcaattgca tttagattat ttaatgcttt aggcagactc ttcccagtg 601
 tgcgaggaat tatatttgag aattttgccg tgtttactgc aggacttttt aaatcggtgt 661
 25 gaaccatag aaaaacctat gactctgagc aatttcttct tcctagtttt tattatttta 721
 tacttgcttt ttattataat atagagttaa ttcattgtta cataattaag gtttttgga 781
 atattggcaa ttaagatgct taagtattaa tatttatgta aaaaattatg gagtcttttt 841
 aaaaaagtaa acttggggaa ataggaaagc tgtaaagaat gatctttatg ctttttggtc 901
 ttataaaaa gaaccaaggt catgggctcc gtatttaacc aggttgccac ctttctcatg 961
 30 attttgtttc ctgctcccca ctccctccca ttattcctgc taagaccttt cctgctgcta 1021
 aatattcagt tttcattttt aactaatttg gaatcatttg gctatagaaa tttaaaatga 1081
 tctgctgtgc taactgggaa agaaatggat gcctatttag tatagaacat tttaaactga 1141
 ttgacctgca aatcatgtag agaatatgag agagattttc ttgttgatg ttttgatgaaa 1201
 tggaagtgt atccacagta ttataacct gtttatctta agaagagaat ttttaaaaat 1261
 35 taccatgtga ataggcaact cattaaatga aaattaatag gaagtcattt gttatatctc 1321
 ttacaacaca cattcagaag ttattattat ttcagaaggg ctggtttgga acaaccttat 1381
 gaagacacag tcagtaaatt actgcataaa tcaactctta ggaaaggagg ttaccaactg 1441
 aagcatttaa aatgaattat tattttgccc aggttttttt tttctttcta gtataggtag 1501

aaggctaaat taattgaatt tattattaac atatgcagtg cctaattaaa tttcagtgtc 1561
ggtctattta tattttctgca acattcctta tatcttctta gcagtcattg gacaccaacc 1621
ttcagctcac ataggttact aagtgatatg aattttcata gggctccaga aaattttcaa 1681
gaattggttg ttagctttttt aattgatgaa gtggatacca gttcttttca ctgaatggct 1741
5 tttattcatt aaggtaatgg ggctgttaga gttgcttagt tttcctgggg aagggggaagg 1801
aagaaaacaa agcagaatgt catgtgatat gcaactgtat taaaaaacccg aaaaggaaaa 1861
aagttgagag agatgattta accgtgagtc accggcagcc aaagcgtgag taaagcttct 1921
cacagatgaa tttagacaaa agcggagaag gtactggtga attttctgga gcctttacat 1981
tttctacagt gaaatggaga taaactttac tcatgccata ggacatgttt caaaacaata 2041
10 ataagatgtt ttctgaacac ttactacata ctaagcactt tatatgcttt gtctcattta 2101
atccttacac agccacattc ttctgggggt tagcgaatga tttttgtggt tgtgtcctat 2161
gcttgtcctg tctaaggatg aagttgttct aattgggtgc ccctcctttt gctttctgtg 2221
aggacttgca gaactggtgg ggtttaacaa gtaccctcac ttatctcaca gaatttcatt 2281
agctcccaga taccctgac attctcccc tagcctagtg aagaaaatct tccatttact 2341
15 tgttcattct gcagtgacag ctccatcaat atacaataga ctatacatat taagtgtact 2401
gtatatacta tacatgttaa aaatctcatt cattttggtg aggcccagct aagaatactt 2461
acagtagagc tttttttttt ttcctaagca taaaagtatc tttttcaatg cagcatgaga 2521
cagagttggg aaaacccaaa taaatagatc caatggactc cccaaagagg ataataattca 2581
tttaaataaa caccctctc agtggttaaaa ctttctaate aacatgcctt tgggacacat 2641
20 tgcaccctca aagtttacac tcccattgca acgcagcttt gtggttcacg ttttttccat 2701
tcagaatgtc attaccctgt caatgatgtt tcatcaacgt ttgcttggat gagaatcctc 2761
tgatattctt cctgatagaa atgtataagc cctgttcata taaatgaata aaagatctaa 2821
ccttactttc tcagtagtgg cttccttgga gcaaaaagca gggacctcca gagagctcag 2881
gtggatgact cttttctgtt tcttccagag ctcaacttac aattagtga caattcattt 2941
25 ccagaatgt cttctttctt attgtgcctt tagaaagtta ttaagcaaac atttgaattc 3001
acagaatctt accagtgtaa gaggaatgga aaaggtaact tatcaaggta acaatcactt 3061
cgtggccagt tttttcggct cactgcaact accctcctg ggttcaagcg attctcctgc 3121
ctcagcctcc caagtagctg ggactacagg cgtgcaccac catgcccagc taattttttt 3181
tttgtatttt tagtaaagac agggtttcac catgtgggccc aggctggtct catggcaagt 3241
30 tttctttgtg ttgtcatgtt attatcaatt aataggaatt tatatttcag ttctgttagg 3301
tgataaaca ctattttgca tacctaaatg tttcatttat atcagcactg gccaataaaa 3361
atatactata agcaggccgg gtgcagtggc tcacgcctgt aatcccaact tttgggaggg 3421
caacactttg ggaggacaca gggtcaggag atcgagacca tcctggctaa cctggtgaaa 3481
tcccgtctct actaaaaata caaaaaatta gccgggctg gcgggggccc ctgtagtccc 3541
35 agctacttgg gaggtgagg caggagaatg gcacgaaccc gggaggtgga gcttgcagtg 3601
agccgagatc tcgccactgc actccagcct tggtgacaga gcaagactct gtctcaaaaa 3661
aaaaaaaaa aaaaaaaaaa aaaaaatata tatatatata tatatattac aagccacaag 3721
ccacatatgt acttttaaat gtctagtag ccatattaga aaacaaaagt aaaaaggaat 3781

agatgaaatt aatttttaattg attttttttaa acccaagata tccaaaatat tatcattttta 3841
 acatataaatt aaaaattttat tgagatgttt tacattgttt ttttttttct tgacgctgtc 3901
 ttggaaatct agagtgtatt ttacattttac attacatcta attcagtcta gccacatttc 3961
 acatgctcat ttttttgtgt gtgggttttat ggagcagaga gtttaatagg caagaaagaa 4021
 5 aagagaaggc agaagaaaat ggctccctg tacagagacg cgggggtggg gcgctccaaa 4081
 gccaaaagag gaggtcccta agtatggtag acaccagcca ggaatatatg cagtgtctgg 4141
 aggaggggat gtctgatttg cataggggtca catgctcatt tttatggcta ctgtattggg 4201
 cagtacagat ttagatgggg atttgctctg aacaagttgg tgattgatgg tgttcataatt 4261
 ttaattgaat ttttcctggg ttctgctgta cacttgtagt tgtgttagtt tcatgtgcaa 4321
 10 tgcttgtgtc acttttataaa cccagatata tggcataaca tgagaatgaa aaatggacca 4381
 gaaaaatagt ttggcaatgt agtcatgttt gttcctatta aatgttcctt attgaccact 4441
 ctatctcttt taattataac aagaatctgc cctgccagca tgcccagtta cgctgggaaa 4501
 acttctgcct catttactct ggctgattct catccactta tgggtcagtg gttcattttc 4561
 tagaggcac cagcattcat acctagcata caatttcatt cattataatg aaggaatgtt 4621
 15 ttcccttcaa agagacacaa ctagtgggct taatttttct tgatatgtca cctgtaaaat 4681
 tttaatgatg atgttttaaac tctaaatgta gccatcaaga caaaaactgc aaattttgag 4741
 cctcagtggtg tgtgggtggg tttctgtttc ggtaatttga aacattgcag aatatcatca 4801
 aaatatgata cccaagaatc atatggtatc aatcattcct agatatactg atctattcat 4861
 tgccaagata gttcaatgag ctggcaaaaa catatggaat tattttctta aaatgtgaaa 4921
 20 aataaaattt aaccaatcat gtatcacagc ttgcaacttt agtcatactt tgaaaagcat 4981
 tttaatgttg acctcatgat tgaaaattta taaaagctg aacagaaatt agttcacttc 5041
 atattttaga aaagcagagt ttctttacta aatgaggcat ttgacccaaa ttggagagaa 5101
 aatgttgaaa ctacttctgt gagcaagcag gtggcttctc aaacacatgt tgggatgaaa 5161
 tgggtgggcc tcaggggtctc agtgctctgc actgagagtt ggcactctct atctccatgg 5221
 25 tctcctccaa gtgtgactct tgtctcttgc tgacctgacc tgccccaagt gactcactgg 5281
 tcatgaccct gcacaccttg cgtctctcct atcaccctgc cgatggcaga gctacaaaagg 5341
 tctttgatgt agctctgtct gatatcctgt gtttccccct atgggtctgtg tggaagcagg 5401
 tatgggggtg tgtaaagggg aagcctatga agttcatctg caaagactac ctggttaggg 5461
 gaggagagga agaagctata tgcaccattt caccagcaag catgggctct tctgcctttt 5521
 30 agcttagggg tcctgttggt tagtctcact cacctattaa aacagtccag caaatagagg 5581
 ttttgtttac ctccatttaa aaaaaagca attaatggaa tagaagataa taatgtatga 5641
 gaagcactat tgtgaaagaa aaaccttcaa cttctctcag ccaaataatt gttcctcct 5701
 ctgtccttcc cagaccttga tgtttgctct attatttcaa aacaactata ttataaatat 5761
 ttgagaatgt gtatttccct gcaaggagat cttaatcccc aagaaggcag gagctgtgta 5821
 35 ttattcatcc cagtgtcctt tcaaaccagg gcctcagaca gtgcatggcc caaagaggta 5881
 cttataaac gtttggtaaa ctgacactat tgaaattaag caacctggat ttgaggtggg 5941
 tctctgccac tcacaagaga ttacgctttg agaaaattcc atcacttcat tgattttcag 6001
 ttcttgcatc tgtatatggg agacgatact aggtgatttc tgacatctct caccgtttta 6061

atgctctgtg atctatacaa cgaggggctc gctgttctag acaagttcct tccagcttta 6121
 cagttgcata acccttctaa tcttagtcac atgatgactc cactgacaga tttttggcca 6181
 ccatcattag acatgctgag ttacgtgtgc ctttgctctg atcctcaaaa ctcatgattt 6241
 ttaaagtttt ctgaaatata taccatttat caggatccag atggatttca tgaccaaagt 6301
 5 ggatgtttct tttctctccc attacaatct tttacttttt gtgtgggaaa ttgcatgtta 6361
 aagaaaggga aattgaagaa tgggatgctt tggaattctg gcaagatgga ttagtgggtt 6421
 ccagaaagta ggggcagcca caaataccga aataaatgag atcgtattat tgagaaagca 6481
 caaatggaag aaggtcaaaa gcaagaagaa gctgacatcc tgcttcctcc aatttttgct 6541
 ttctctgttt ttccaagaaa ctctcttcc aagccttgct gaaaaactcc actttcctaa 6601
 10 atctaacttc ttaaaactgat aatggcaaga agtttagaat gaccaatatg tttaacactc 6661
 caagagtatt tgttttgttt tgttttgaga ctgggtctca ctctgtcgtc cagcctggag 6721
 tgcagtgggtg tgatcacagc tcaactgcagc cttgacctct cctgccaag gaatcctccc 6781
 acctcagcct cctgagtaga agggaccaca ggcattgtgc actacacctg gctaactttt 6841
 aaaaaatttt tttgtagaga tggggttgcc caggcaggtc tcaaactgct gggctcaagc 6901
 15 aatcctcctg cattggcctc ccaaagtgtc ggaattatgg gcataagcca ctacgcctga 6961
 cctctccaag ggcattcttt acccagaaga ggaacttggc agaacttatc ctccaattgg 7021
 tgaggaatat ggagaaaatg actttaagca aaggaacttc tggttctgcc tacctaattc 7081
 agaaaaagaa gttttatttc tcccttcccc tagtaactat cttcccatat tcacataaaa 7141
 aagtacagaa tcaacattgt tcaagaatta taattttact tgtaagcaca tgtgcacagc 7201
 20 cacaccata taccttcctt ccctttaaat catccacac cctaatagta gtaaaatcat 7261
 tgacccgagc atacctggga gaggaagagg agtctgacag gggcagggtc taagtggcac 7321
 tcctggaact taaccctggt gtatatgaac ttacctatt gaaggatgac tcctcaactg 7381
 ttctcacaat ttgctgctct gctttctttt ctaatttctg aaggtgactc atcttcccca 7441
 aggactttca gacttctcag aagaaaaaaa tattgggtgg gtctctgcca ctggcaaaag 7501
 25 attagacttt gagaatcata aaagtatata agtatatact cattaatatt gaattactat 7561
 aattaatatt atgatattga tataatgata gaatgatatt gataaaagca atattcaata 7621
 atgaatatta ttctagctgc ccacttattg ggtgcctcat aggtgccagg cattttgtat 7681
 gtattatcta caacccttac atgggacata ttatgatgct gtttctcttg aagaaatatg 7741
 gaaactggaa acagagagggt caccacaatt ttccaaagtc acatagctaa taggtagcag 7801
 30 acttgggatt caaattcata tgcatatggt aaatcatgct ctctctctgc tacattttgc 7861
 ccccttagaa tatgaaaaag ggatacaaag agatgaagaa aatatgtaag attatccttc 7921
 aatttcacta tcttttttaa agtttttttt attatacttt aaattctgaa atacatgtgc 7981
 agaacatgca ggtttggttac ataggtatac acgtgccatg atggtttggt gcacccatca 8041
 acctgtcatc tacattagggt atttttccta atgctatccc tcccctagtc cccacccac 8101
 35 cgacaagccc cggtgtgtga tattccctc cctgtgtcca tgtgttctca ttgttcaact 8161
 cccacttatg agtgagaaca tgcgggtgtt ggttttctgt tcctgctca gtttctgag 8221
 aatgatgggt tccagcttca ttcattgtcc tgcaaaaggc atgaactcat tttatggctg 8281
 cataccattc tatggtatac atgtgccaca tttctttaat ctagtctatc attgatgggc 8341

atttggttta gttccaagtc cttgctattg tgaacagtgc tgcaataaac atatgtgggc 8401
 atatgtcttt atagtagaat gattttataat cttttgggta tagaccagc aatgggattg 8461
 ctgggtcaaa tggcatttct ggttctaggt ctttcaggaa ttgccacact gtcttccaca 8521
 atagctgaac tagtttacac tcccaccaac agtgtaaaag cgttcctatt tctccacatc 8581
 5 ctctccaaca tctgttgctt cctgactttt taatgattgc cattctaatt ggagtggat 8641
 ggtatctcag tgtggttttg attagcattt ctttaatgac aagtgatgat gagctttttt 8701
 tcatgtttgt tggccgtata aatgtcttct tttgagaagt gtccgttcat atcctttgcc 8761
 cactttttta cgggggtttt tcttgtaaat ttgtttaact tccttgtaga ttctggatat 8821
 tagtcctttg tcagatgggt agattgcaaa aattttcttc cattctgcag gttgccgtt 8881
 10 cactctgata atagtttctt ttgctgcgca gaagtttttt tagtttaatt agatcccat 8941
 tgtcaatttt ggcttttggt gccattgctt ttgggtgtta gtcataagt ctttgccac 9001
 gcctatgtcc tgaatggtaa tgcctagggt ttcttctagg atttttatgg ttttaggtct 9061
 tatgtttaaa tctttaatcc atcttgaggt aatttttgta taaggataa ggaagggtc 9121
 cagtttagtt ttctgcataat ggctagccag ttttcccaac accatttatt aaataggga 9181
 15 tcctttcccc gttgcttggt tttgtcaggt ttgtcaaaga gcagatgggt gtagatgtgt 9241
 ggcattatct ctgaggcctc tgttctgttt cattgggtctc tgtatctgtt ttgatacaag 9301
 taccatgctg ttttggttac tgtagacttg tagtataatt tgaagtcagg tagcgtata 9361
 cctccagttt tgttcttttt gcttaggatt gtcttgacta ttcaggctct tttttggtc 9421
 catatgaaat ttaaagtagt tttttctaatt tctgtgaaga aagtcaatgg tagcttgatg 9481
 20 ggaaaagcat tgaatctata agttactttg ggcagtatgg ccattttcat gatattaatt 9541
 cttcctatcc gtgagcatgg aatgtttttc catttgtttg tgtcctttct tatttccttg 9601
 agcagtagtt tgtagtcttc cttgaagagg tccatcacat cccttgtaag ttgtattcct 9661
 aggtattttg ttctctttgt agcaattgtg aatggaagtt cactcataag tttgctctct 9721
 gtttgtctgt tattgggtgta taggaatgct tgtgattttt gcacattgat tttgtatcct 9781
 25 gagactttgc cgaagttgct tatcagctta aggagatttt gggctgagac gacagggtt 9841
 tctaaatata caatcatctc atctgcaaac agagacaatt tgacttcctc tcttcctatt 9901
 tgaatacgtt ttattttctt ctcttgcttg tttgccttg ccagaacttc caatactgtg 9961
 ttgaatagga gtgttcacca cctattttaa gaatagtatt gaagcctcac aaaagctggt 10021
 tctcatgtaa ccatctgaga atatttggtc ttatgacttg aattcattca ttgccttttt 10081
 30 atttcacatt ttagtgatcc tgatgtctaa atcttaattt ttgatccttg caaggtaaaa 10141
 tagccaagtc aagcctgttt aataatattg gttgaggaag tcacatgctt atgatcaatc 10201
 tttgggttat gtaattatat taccttaatg ttggcagttt aggtgtaagg cagagatatc 10261
 tgatcacatg tgtggtttag taatttaaga tcaactgcaa ctaaaatctt catggtatga 10321
 tcttcaaagt tagctacttt gaccacagca atgatttcac cacagcaatt aacaaaatgg 10381
 35 cagactcttt cctgaggtgg catgaacagt tttaaaacaa agtcaaggac caaaagaaa 10441
 gcaggcacat ggcatttgat tcatttcaaa aactagtatt gtattaagag ccaaggggat 10501
 agaattgtag catcaattaa aatcttggtt gaaaaaaaat aaaaacaaac gtccattttt 10561
 atctctcaaa tatattagggt ttttcataaa gttatagggt tatttttttaa aaaacaaaac 10621

tcatatacat tagactgaaa aatttgccctg tcatctcatc atgcagctaa atgcaattgt 10681
 ctatggcgag atacactcctt attagaggta tgatagcact aactaatagc aaactttgta 10741
 cctggtagtc taatttatgc agggttcata ttctgctccc tctcagcatg ctgtaactgt 10801
 ggcaaagcca ttctcaggag ttattacccc aacataatca taccctgtg gattaggagc 10861
 5 agttaaatgg gtcctgttat cagagacaca tatgtgccac agccgctgtc atccctaagc 10921
 caccgtgggt gattaagact cactgatggg actacctctg aataggcttg agtgagggtg 10981
 atacacttca gctgagagaa attcaggtaa ggggctgaga aaatcagatt ttgaatgggt 11041
 ttatcatacc atcagggtctc cttttaagtg ctggggctcat ggatttcatt caacctgacc 11101
 acatagcctt tggaagcttg gctcaatacc tgagtgtgag attatgggtg atattaaagg 11161
 10 agatgaatgc attgagctga tgtcagagaa tgtcttttac tggattttca taatgactgg 11221
 ctgcagatgg gctgagggga aagtcagatc aagagcattt ctgtaagaag aaagaaatct 11281
 tccctcttat tctctttcaa gaaaatgaat agctgagcac caagaggcca aatacttttt 11341
 taaaaaacac atccttttat gtagagaagg acagggttgag acgaaaaaca ggacctctga 11401
 aactgtcttt atagctttta tctaggaaga aattgcgga ccattgctga catcattatc 11461
 15 agaggctgcc ttagttctga gagcttcaca gatggccttt tctgcatttt acatctggcc 11521
 agatgaaggc aaaaagggtg atgaaaccaa aactattaga tcagtgggtc ataactctgg 11581
 ctgcacttta gaatcatctg gggagtcctt aaaactactg atcctgaggc tccatcccag 11641
 accaattgaa tcaggatctc tgggggtggg acctgagcat tggaatgctt taaaagttcc 11701
 tcagggactc taatgtgcag ccaaggctga gaatgactga ggtggatggt ggccaagaca 11761
 20 ggtgaggcca agagttagaa gcccttactg ttagggaagg cagaagccac aaggaggagg 11821
 ggagggggaa ggagcagtat ttagcatttc ttccacacag ttgggggggt ttctgatcaa 11881
 aattaggctg ggatctttcg cttccatttt catgaaggct ttatgctttg tccacagcca 11941
 cctggcctca gggcaatgag caatctgatt gatgaatttt cagtaagaaa ctgagcacac 12001
 ttggctctca gccccagtgg cttcccctgt ttccaaatct gccaccagt tacaggagcc 12061
 25 tgctcaccaa ctggttgggt taaagaagtc ggctctgtcc ttggcaggga ggccttcagc 12121
 tgtctggccc tgtctgtgac tgcgtgggtg aagctgccta atttggggaa cttgatggaa 12181
 gatctaagct atgttctcta acagttttat cagaaataaa gttaactttt gacctccatc 12241
 tgctgtctc ctgtggaagg gctctgtccc ttccaaagag actctcaggg gttctcctta 12301
 gagggtgtgt atcagtccaa agatcatttt agaccagcca ttacagagga tgtccaagaa 12361
 30 attgcacaag ggaaatagaa gtaagaatga gagcaatata tcagatagta aggaaagtat 12421
 gtgactggtt tcaaataagt aattaaccat aaatccaaag ttctgcatt agttatagca 12481
 atagtcatcc acctggctcag gaaaaaaaaa gtaaaagact gagctgcaag atagaaagtt 12541
 tgctgaatt ccacgtactt caaggactac tatggaggat tccttgggtc caatgcagag 12601
 acatgtactc tgaccaccca tgggcccatt cctcttacc caccctcagt gattcctcct 12661
 35 gggacctcac tacattgagt tcttacatc cccactcttt tcgggggaaa tataaccctt 12721
 ctgctcttct catgatgtta aaaatattta gacaattact taaaatttac aacaatcttc 12781
 cacataaatg atctcactta tgtttcttgt gagctctgtg atttatttct attatcatct 12841
 atgctgatag ttaaggaaac tgagatatcg agaccaacat gcatagtaaa tggcaaaacc 12901

agatgaatttt taaacttttcc taactccaaa agccacatcc tgcccaaccc gccatgctgc 12961
 cgctcagtta atgcctggct gtttgtctcc ccattggccc ctctacccat tgtgctttga 13021
 tggcacactg tattccaact gcctgagact cctgggttaat gccattacgt acagacttag 13081
 gttgaattta ctaggatttt taagatttgt aggataagag atatgactgt tagactggaa 13141
 5 tcagcaatag ataaaagggtt aacaaagttt cagaataaaa tataatagaa aaccccagca 13201
 gatagaagta aaatgaatgg taagacaact gaaatgaaag ccatacattt agaaatatca 13261
 aataaaacac agattaatgg cagacaataa aggaacatac ttagcagtta gtaaaaacac 13321
 tttttacctt attcttatta cctccagta accttttttt tttttttttt tgagacaaag 13381
 tctcgctgtg tcgtccaggc tggagggttag tgatgtgac tcggctcact gcaacttctg 13441
 10 cctcttgggc tgaagagatt ctctgcccc agcctcctga gtagctggga ctataggggc 13501
 ccatcactgc acctggctaa tttttgtatt tttagtagag acgggggttc accatgttgg 13561
 ccaggctggt cttgaactcc tggcctcagg tgatccaccc gcctcagcct cccaaagtgc 13621
 tgagattaca agcgtgagcc accaagcccg gctgtaacct attgaaaata acattacttg 13681
 acatgtgaga caaattattt gtaagttaaa gagtttatgt gccttcaatg ctcccaccct 13741
 15 tccctcccct aaaaagatta ttgagtgcc atgatgtgcc taagccttct ggtagactct 13801
 gagaatgtga agagagtttag aggatacttg ttcaacaacc cagaatctag cagaagtgtt 13861
 ttgaaatgac cagccacttg ggagagctat gctagcatat cattcaggat gggctgggtc 13921
 atgttttaat aacaagtaag tttgaacttt ataggtttga aacaaaaaag atgtgtttct 13981
 tgctcacagt tcataTTTTT tggagattgg ctggagcttc attccaggat actgccacca 14041
 20 tctggaatgt tgacagttac cgtaacagag aaagagttgt ggaggctgtc aactggcaa 14101
 ttaaattcgc tgacctggaa ctgacacatg ccacctccac ttatatTTTt ttggccaagg 14161
 cttgttacac agccacatct aacttcagag aggtcaagat gagcaaTcc taccacgagc 14221
 tatagatgga tgagtagcac tcattattat cacaattatt ttatttTtta taaaaactcc 14281
 aagaaggaga ggtactctat gtgaggaata ggcataggaa aatcagaggg aagtTcttaa 14341
 25 agatgagcta gatttcacta gatggcattg aagaaacatt ccaagtaaTt gaacagcata 14401
 agcaaTgca tgaagaactt attgtggTcc tagtacagac gatgcgtgag agtgtggagg 14461
 agggaaaagg taagactgga gaggaaggca ggaaccagaa caggacagac tgtgttTact 14521
 gtcaggcagt taagtctatc ttgtaggcaa catgaagcct tTtaaggaa gcaagcagct 14581
 atgtgacagg acagaagatg gattggaaga aaatcaaaag agaagtgggg accagtTata 14641
 30 aggtcctcc aatatttggg gatccaaacc aaagcactgg cagagaaata gaaaggaaagg 14701
 gaaatatttc caaaatattc aaaacagaaa ccaagagaac ttgatgacag aagatgaacc 14761
 caggtttcta ctgaatggac agttgtTaca ttctctgaga taaggaaTat agaaggaaaa 14821
 agttgagagg gaatatgaaa ttattttTaa tcatgtTtaa tttgatcact tgtggaacat 14881
 caaccagaga tgtccatcag ttacaccaat ttgtaggTtt tgagagaggc ctgaggTtaga 14941
 35 ggcagagaag tgggaatcag cagattagcg gcagtagttg gaaccataac tgaagatgag 15001
 atttcccagg gaggggagtt gaacaggaag gtagacaaag gctgagtTca gtggctcata 15061
 cctgtaatcc cagcactttg ggaggctgca gtgggtagat cacaacaag atcaggagtt 15121
 caagaccagc ctggccaaca tggTgaaacc ccgtctctac taaaaaaaca caaaaattag 15181

ccgggctgtg tggcacacac ttgtaatccc agcctcagga ggctgaggca ggagaattgc 15241
 ttgaaccccg gaggcagagg ttgcagtgag ccaagatcgt gccattgcac tccagcctgg 15301
 gcgacagagc gagactccgt cccacccgac cccctaaaaa agaaagtaga caaagatgtg 15361
 tcctagaaaa cactaatatg aatgggtaga gaggggagac ttgtcaagga gatcgagggg 15421
 5 agagtcagtg aggtgaaacc ttagaggata gaccttcacc aaggaagtaa cagtcaacag 15481
 gcctaaatgc cacagagatg tcaaatagaga gacactggaa atggttcttt gaattattaca 15541
 gccagaacat cacaggagac catttccaaa gcagtcttga tgaagtgggtg gggaaggag 15601
 gtccctgaag gagctaagga gggactggga gatcatgacc cagagataag tgttgaggt 15661
 ataaaaggga aagactgaga tcaggaaata gccacaggat ccctaaggcc aggtctcggg 15721
 10 tggggtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tataaatgg gaagacctt 15781
 agtataattc tgtactaaga agatagagca gtagagagga aaaggctgaa gacagatggg 15841
 agggtaaata gtgaattaga aaggctctgtg caaagatgag aaaggatgag tctgaaagca 15901
 cagggagaag ggccagcctt ggacaggaga aaacatttct ttcactaaga ataaaaggga 15961
 ggttggttat acaccacaaa gaagggtgtg atgggtggca atggagtctt gtcaagttga 16021
 15 gggcattcca tgatagcctc acctttctct gtgaaatgag agagttagt ttacaagata 16081
 tttctgagca cttcataagc caggttcata gactgaaaga agtgagggtg gagtgtgata 16141
 ggtccttaag aataggggaa gtttggaaata tctgacaagg gacagagagg aaaaaactag 16201
 aaaaggcttt gcagaatgtg ggcccacaga tcagaggcta aggggtcccc atttgtgcag 16261
 gagagtaagg gcagggaggc aaggctcaca gcccaaaata tagagcccct caccaaatga 16321
 20 ctgcaggagg gcagctttcc tatgagagca tccctatcac tgttttact ccgagtcatt 16381
 aacttacgac ttactcagct ctgtttcgta atagcagact cgagtaatga gggtatgaca 16441
 gcctctctct gcattgccaag gtatgcagcg tggatttctt ttttctctt ctctctctg 16501
 tggcttaggt gccttctgtt ctgctaccag gatagagaac ccagtgacta gtttcttcta 16561
 gctctctttt tctgactagg tatcttgtca gaaatttctg cttaccagac ttcattggaga 16621
 25 gggaatcaag ctttgaatca ggggtgaaaa agtagagctt aatatatata ttacaaaatg 16681
 ccaactcacgt tcttgagggt accttgatc tataaccacaa ctagcattct tttagaaagc 16741
 accattaccg aagtaatccc tttctgga attcaccaa aaagggtatt cccacttatt 16801
 ccccatctcc aaaataaaaa agaaaaatgt gtgtgcttag agatgttctt ggaagcatga 16861
 gctgtaatac tgaaccaata gaagacgaac taacagattg cagggcattc gtttgcaaaa 16921
 30 aaacttatgc agttatctaa atgatagtta tgaagacaat atgtataaca cattatatct 16981
 tgagtgaaaa gaacagaagg tgatttcaaa actgcattgg gataatagta atataggaaa 17041
 tagtgagatg aacaaagatt tcaaaggaac aaaaataaag aaaaattttg ctttttatat 17101
 tggtaggcgt atgggtgaaa ttccaatttt aattttaatt tcatagttat aaaattgttc 17161
 ataaaaaaag gtcccccata gacagttggg ctttgggaca aactaacaga aacagagtag 17221
 35 gaagaaaatt catcttcctt caatccccct tctctgttta aaacaaaaca aaaaagagct 17281
 ttgtcatgtt cagggtgtgca acgaattctt tttccaaatc tggaacttta catctgctat 17341
 taaacaggag tcagtttcca tgtaacatgt tgacaatccc ccaagtgtgt tggaataatt 17401
 ttttttaatt aggagatttg aaattccatt tcaattgcca acctgcctct ttcaacttct 17461

aaaaacaaag taaaacaaaa caaaaacaca ctgggtccta tcacccctc ttgctactac 17521
 tattttatct ccattgccct gaattctttc caaacttctt tccaccacgc tttgatttgt 17581
 tttcagtcgg gtttattgag gcataattta cgtacagtaa aattcatcct tcttagattt 17641
 agaggtctat gcattttgac taatgcatat tggcttataa tgactaccac aacaaagata 17701
 5 tagaacacac ccatcactcc cgggttctcc ctgtcccttt tttggtccat ctctctcct 17761
 accccaacc ccttggaac cactgatctg tttctgtcc ttatcatttt gctttttcca 17821
 ggatgttgta tataaggaat catgcagcat gcagcctctc gagtctgact tcttcagtt 17881
 agcacagtta tttaagatct atccgagtta ttgtgagtag cagcatcttt tttattgctg 17941
 actagtattt catcacatgg atgggccaca acttgtttat ctgttcacct gtcaatggat 18001
 10 actaagttgt ttccagtttt tggcaaatat gaataaagta aacatttgca tacagatttt 18061
 tgtgtggaca catgttttca attctcttag gtaataactg agcaatggga ttgtcgggtt 18121
 atatgttaag tctatgttca gttttctaag ttctgaaaca attggatata tatatgcaaa 18181
 aatacaaat taaccttgac tcaaacttgt accatacaca aaaaataacc tgaaagagat 18241
 cacaggtgta aatgtaaacc tagaactaaa aacttcaagg agagaaacat agcagaaaaat 18301
 15 atttgtgacc ttgcattagg caaagacttc ttagatttga catttgaaac atgatacata 18361
 aaaaatcttg ataaattgga gttcataaaa ataagaacta ctcttcaaaa gacactgata 18421
 agagaatgaa aatacaagcc acagacagag aaaatatttg tgaatttctt atctgataaa 18481
 gggtttgat ccagcatgca taaagacttt tcaaaactca ataataaaca atccaataag 18541
 aaatgcacaa aagatacaaa cacatcatcg aaatgacct atgaaaggca aataagccca 18601
 20 caaaaaaatt ctcaacatca ttagacactt gggaaatgca aattaaaaca aactgagat 18661
 aaactacata tctattaaat ggctaccact ttaaaacct gtcaagtgcc agccagaatg 18721
 tggaacaagt aggactctct tacattgcta gtgggaagggt gaatggtaca gccactttgg 18781
 aaaacactcc gcagcttctt atagttacac ataggacctg ggggtagagg atggattgac 18841
 tgtaaagggg caaaacttcc tggtaggggt ggggaaagtt tttagaaggg tgaaattggt 18901
 25 ctatatcttg attattgtgg tggttacaca actgtctgca tttgtcaaaa ctacagaaac 18961
 tgtacactaa aaaggatat ttttatgctc tgctaatttt actttaatct taaaaatagg 19021
 aaggaaaaaa taaatcaat gccactgtgc gactttgggc aagttacttc acttctctgt 19081
 gccttggtct tttgaaatct atacattaag gataaaataa taccttcctc atactgttag 19141
 aattaatgt gctaatttat gttttatata tataaagtac ttggccgggt gtggtgactc 19201
 30 acacctgtaa tcccgacact gtgggaggcc gaggtgggca gatcacttga ggacaggagt 19261
 tcaagactag cctggtcaac atggtgaaac cctgtctcta ctaaaaatac aaaaattagc 19321
 tgggcttggt ggtgcgtgct tgtaatccca gctacttgag tggctgaggt gggaggatca 19381
 cttgaactcc agaggtggag gctgcagtgg gctccacca ctgcctgggt gacagagcta 19441
 gactccatct cgaagaaaaa ataagtactt gaaacatagc aaatgtttta taattattgg 19501
 35 cttttttttc ttattgttat tacttgcat attgctgttt gaagaagttt ggtgataatg 19561
 gagagaaaag ggcaattagg ggtctgggat ggtttaagta tgaggagacc gagacacatt 19621
 gactcagagt gaagaaatca aagataagag aatgaaagaa agggagggtta attctgaacg 19681
 cacagacaaa gttatgttac taacgtggca tcggctgtgt gttgtaataa ataactccct 19741

ttcacttgctc aatagctaata caaataacttt ctggagacca gaagtgactt gctggatcaa 19801
 tgacaactcc tccacagatc aaaatgttca aatccttttt ctgtgttgta atcctaaatc 19861
 taaaagaaca gagagaccaa gcaaactctac ctcccaacat cattaaagtg acaactctca 19921
 gtattttattt gaatgggtctg ctctcagctt caaccaagga aaagtcaaata tagtgttggt 19981
 5 cagaaaacag aaggggtgta cgagagttct ggctgggtcat tacagacttg gggatttttg 20041
 attaaaagaa gaagaagaag aagaaacctg ataaagtgtg aatatagcaa gcagggatta 20101
 gtgtcctgct ggggtcatgtt ctcacaacag tgagaatttc agagatttca taagaattaa 20161
 actgctccac atgacaattt attttacctt ctggcttttc caggaggcaa atcagtgcga 20221
 cttctttctg cttttgtttc aatttggtta caacctcaa ttttaggaca ggctaaacct 20281
 10 agccacccta tcagagatga tgaagtagcc atctttttta caggtgggga gatgaatgga 20341
 atcaggggtt gtttggttggt ttgtttaata actgctagta aaaaccaagt caatagctga 20401
 ctgagtgtta gggaggctcc agaaggcagg ttattgtagt atagatgtga ctcgacttat 20461
 gatgatgtta cttcccgata aaccctcat aagttgaaat atcgttaaat tgaaaatgct 20521
 ttttaatacac cgaatctacc gaacatcata gcttagctca gcctacgtta aatgtgctca 20581
 15 ggacgcacat tgctacagc tgagccaaat cacctggcaa cacaggacac ttagagtagt 20641
 cggttgctgg cccttggtgat gctgtgactg actgggagct gcgcttagtg cctctacca 20701
 gcattgagag tttcttatcg cttattacta gcctgggaaa agaccaaata tcaaaactca 20761
 aagtgcggtt tctaccgaat gcttataact ttcagaccat catgatgttg aaaaatcgaa 20821
 ccatcgtagg ttgggatcca tcctataaga cgagctacac tgccggaagt gtaagactgc 20881
 20 tatgctgccg gaagatgggg catagtggac aactgcaagt cctgacaaca ggaggtcagc 20941
 atctgcgacc tttaacatcc acattgacac taccacagtc ttccaaacag agctgatgat 21001
 atagtttgga tgcctgcctt gcccacatct catgtcgaat cgtaatcccc agtgttgagg 21061
 gtggggcctg gtgggagggtg attgggtcat gggggcagag ttcttatgaa tggtttagca 21121
 cgggtccccct ttgggtactgt atagttagtg agttctcatg cgatctggtt gtttaaaagt 21181
 25 gtgtggcacc tccccctctc ctctttctcc tactctggcc atgtgaagtg ttggctcccc 21241
 ctttgccctc caccatgatt gttaaattcc cagaggtctc cctagaagct aatgctgcca 21301
 cgtacagcct ggagaactgt gagccaatta aacctcatct ctttttaaat taccagctct 21361
 caggccgggc gcggtgactg acacctgtaa tctcagcact ttgggaggct caggcaggaa 21421
 gatgatttga ggtcaggagt tcgagaccag cctggccaac atggtgaaac ccatctctca 21481
 30 ctgaaaatat aaaaattagc caggcatggt ggcggtgccc tgtaatccca gctacttggg 21541
 aggctgaggc aggagaatcg cttgaacctg ggagtcagag gttgcagaga gccaaaatgg 21601
 agccactgta ctccagcctg ggcaatggag tgagaccctg tctcaaaaaa tatatatata 21661
 ttaccagac tcagggtattt ctttacctga gactatgaga gaatggacta atatagctga 21721
 agaattttat tttattttta aaaaactttt acgtttgggg gtacctgtaa aagtctgtta 21781
 35 cataggtaaa ctctgtcat gaggatttgt tgtacagatt ctttctgct cctccccctc 21841
 ctcccaccct ccatcctcaa gaagatccca gtgtctgttg tttcttctt tgtgttcgta 21901
 agttctcatc atgtagctcc cacgtataag tgagaacatg cagtatttgg ttttctgtcc 21961
 ctgtgttagt ttgctaagga tgatagcctc caactccatc tatcttctctg caaaagacat 22021

gatctcattc attttttattg ctgcatagta ttccatgggtg tatatgtacc acatttttctt 22081
 tatccagtct gtcattgatg ggcatttagg ttgattctgt gtcttcagaa ttgtgaatag 22141
 tgctgcaacg aacattcgtg tgcttggtgc ttatagtag aatgatttct attcttctgg 22201
 tagtaatggg attgctgggt caaatgggtc ttctgctttt agctctttgc agaataacca 22261
 5 tactgctttc cacagtgggt gaactaattt acactccac taacagtgtg taagtgttcc 22321
 cttttctctg caaccttgcc agcctctgtt atcttttgac tttttaataa aaaccattct 22381
 aattagtgtg atgggtatttc attgttggtt tgatttgcat ttctctaag atcagtgatg 22441
 ttgagctttt tttcatgttc gttggctgca ggtacatctt cttttgaaa gtgtctgctc 22501
 atgtcctttg cccacttttt aatgggggtt tttttctctt gtaaatttaa gttcctcata 22561
 10 gatgctgggt attagacctt tgtcagatgt atagcttgca aatattttct cccattctgt 22621
 aggttgctctg cttactcttt tgattgtttc ttttaccatg cagaagctcc taagttaa 22681
 tagatcccat ttgtcaattt ttgcttttgt tgcaattgct tttgggtgtc ttgtcatgaa 22741
 atctttgcca ggtcctatgt ccagaatgat attgcctagg ttgtcttcta gggtttttat 22801
 agttttgggt tttacattta aatctttaat ccactctgag ttgatttttg tgtttggtgt 22861
 15 aaggaagggg tccagtttca atattctgca tatggctagc cagttatccc agcattattt 22921
 attgagtaag gagtatctcc tccgttgctt gtttttccca ggttggtga agatcagatg 22981
 gttgtaggtg tgtggcctta ttttggggct ctctatctg ttcatttggt ctatgtgcct 23041
 gttttgtac cagtaccatt ctgttttggt tactgtagcc ctatagcata tttcaaagt 23101
 gggtaacatg atgcctccag ctttattctt tttgcttaga attaccttg ccatttgggc 23161
 20 tctttttggt accatatgaa gtttaaaata gtttttttct tagttatgtg aagaatgtcg 23221
 ttggtaattt gataggaata acatgtaatg atattgattc ttcctatcca tgagcatggg 23281
 atgtttttcc atttgttgt gtctctctg atttcttcaa gcagtgtttt gtaactcata 23341
 ttgtagagat tattcacctc cttgcttagc tgtattccta ggtattgtat tctttctgta 23401
 gtaattgtga atgggattgc ttttctgatt tggccctcag cttggtattg ttggtgtata 23461
 25 ggaatgctag tgattttttg taccctgaga ctttgctgaa gttatttatc agctgaagga 23521
 gcttttgggc tgagactagg gggtttttta gatatagaat catgttctct gcaaacagat 23581
 ttagtttgac ttctctctt cctacttgga tgccctttat ttctttctct tgctgattt 23641
 ccctggccag gacttccagt accatgttga ataggcgtgg tgagagaggg cattcttgct 23701
 ttgtgccagt tttcaggagg aatgcttcca cttttgccc attcagtacg atgtttgtgg 23761
 30 tggtttgctca tatatggcta ttattatttt gaggtgtgtt cttttaatac ctagtattat 23821
 gacagttttt aacatgaagc agtgtttaatt tttattaaaa gtcttttctg cctctgttga 23881
 gatagtcatg tggcttttgt ctttagttct gtttatgtga tgaatacat ttgttgattt 23941
 cttatgttg gaccaacctt gcatccagg gatgaagcct acttgattgt ggtggtttag 24001
 ctttttgata tactactgga ttcagtttgc aagtattttg ttgaggattt ttgcattgat 24061
 35 gttcatcacg gatatcgcc tgaagtttct tttttgttg tgtctctgtc aggttttggt 24121
 atcagaatga tgctggcctc ctagaatgag ttggggagga gttcctcctc ctcaattttt 24181
 ttggaatagg ttctgtagga atggtagcag ctcttcttta tacatctggg agagtgtggc 24241
 tgtgaagcca tcaggctctg ggatttttta gttggtaggg tatttattac tgattcccta 24301

aatagaccga taatgat tttt agaagtggag tcgggtttttt cctgggtccag tcttggggaag 24361
 gtgtatgtat ccaggaattt atttagctct tctaggtttt ctagtttgtg tgcatatggg 24421
 tgttcatagt agtttctgat ggttggtttt atttccgtgg gatcagtggg aacattctct 24481
 tcatcatttc tttttttttt tttttttttt ttttttgaga cggagtctcg ctctgtcgcc 24541
 5 caggctggag tgcagtggcg cgatctcggc tcaactgcaag ctccgcctcc cgggttcacg 24601
 ccattctcct gcctcagcct cccgagtagc tgggactaca ggcgcccgt accacgtccg 24661
 gctaattttt tgtattttta gtagagacgg ggtttcaccg tgtagccag gatggtctcg 24721
 atctcctgac ctctgtgatcc gccgcctcg gcctcccaa gtgctgggat tacaggcgtg 24781
 agccaccgcg cccggcccat ttctaattgt gtttatttga atcctctctc ttctcttctt 24841
 10 tattaggcta gctagtggcc tatctatctt attaat tttt tcaaaaaacc agctcctgga 24901
 tttcttgatc ttttgaatgg tttttcatgt atcaatcctt cagttcagct ctgatttttg 24961
 ttatttcttg tcttgtgcta gctttggggg tgacttggtc ttgcttctct aattctttca 25021
 gttctgatgt tagtttggtta gtttgagatc taactttttg atgtggacat ttagtgctat 25081
 aaatttaact cttaacactg ccttagctgt gtcccagaga gtctggtatg ttgtatcttt 25141
 15 gttctcatta gtttgaaaaa acttcttgat ttctgtctta atttaattat ttatccaaag 25201
 tcattcagga acatgttggt taatttccat gtaattgcat ggttttgagc gattttctta 25261
 gtcttgactt ctatttttat tgtaccgtgg tctgaggggtg tttgatatga ctttggttct 25321
 tttgcatttg ctgaggattg ttttatgtcc aattatgtgg ttgatttttag agtatgtgcc 25381
 atgtgggtgat gagaagaatg tatattctgt tgggttttggg tacagagttc tgtagaggtc 25441
 20 tattagatcc atttggtcca atgttgagtt cagatcctga atatctttgc taatttctctg 25501
 cctccatgat ctaatactgt cagtaaagca ctgaagtctc ctactactat tgtgtgggag 25561
 tctatgtctc tttataggtc tctaagaact tgctttatga atctgggtgc ttctgtgttg 25621
 gatgcatata tatttaggat agttagatct tcttggtgaa ttgaaccctt taccattatg 25681
 taacgccctt ctttgtcttt ttttttcttt gttgggttga agtcttcttt gtctgaaatt 25741
 25 aggattgcaa cccctgcttt tttctgtttt ctgtttgctt ggtagatttt cctccatccc 25801
 tttatttttg acctatgggt gtccctacat attttatctt tatctatcca tccagccatc 25861
 cagccatcca ttcacccgta tcattttttaa ccaataagga cttttaaaag cgcaaccaca 25921
 acaccattaa cataaccaat aaaatctata acaatgataa aatatcatct aatactcagt 25981
 ccatgtccaa ttttccctcg ctatctcaaa atcgtcttct tagaaatggg ctgttcaa 26041
 30 gagatcacat ggacagagga aggcgaacct cacactgtgg ggactgttgt ggggtggggg 26101
 gaggggggag ggatagcatt gggagatata cctaagtcta gatgacaagt tagtgggtgc 26161
 agcgcaccag catggcacat gtatatgtat gtaactaacc tgcacaatgt gcacatgtac 26221
 cctaaaactt aaagtataat aataaaaaat aaaaaataaaa aaataaaaag tgaaaaaaga 26281
 aaaaaaaaaa aaaaagaaat ggtctgttca aatcacaac cagattcaga aacaatagcc 26341
 35 atacattaca ttttattaat atgtctctta aatttctttt aatctattac agtctttgga 26401
 atttttatgt ctctgtttat ccttccaatt attaaaaaa aagtattttt gtattcattg 26461
 aatagacaat gcttgcgaaa aagtaaaaaa aaaaaattt agtacaaaaa ggtacatagt 26521
 gagctgttcc ttagtctccc ttcccagaag caatgttacc acttttgtac aaatagtctc 26581

tgccctagaca cacatgccag tccctaaggt ggctgtaaca aggtgggttaa gaggagaaac 26641
 atgaattcaa attcctatta tgccactcac taagtataaa tcttggtcat ggtacatgcc 26701
 tctgtgcctc agtttttaaat aatgggtacct acctcatagg gctgttgaga gaattaaatc 26761
 agataagtgc ttaaataact attaatatatt attattattc acattccctt ttggcttttt 26821
 5 tcccaaatag cagagtgggtg cacatatgtc ttcatttatt tggcttggtt tttcacctca 26881
 catcacattt tgatgaataa ttccatacat gttgttatag atttgcttca tttcttgtaa 26941
 tcattgacta atattccatt gtatgaatat gctactgcta aacatgtacg ttattttcaa 27001
 cctcttatta tcaagaaatg ctgcaatgaa tatccttgta atactttagt ggattcatgt 27061
 gcaaaaatat tcataggata aaatcctgaa agttaaattg ctgagctaaa gggatgtgtc 27121
 10 attttaatgc tttatagatt gcccagctgc ctcaaaggag gttataacaa tttacactcc 27181
 caagaaaaat gcacaagggg ccccatcttc ccatacccta gctaacacag gatattgcta 27241
 aatgctttca tctttgtaaa catgatgtat tgaaaatggg atctcaaagt tttaatgtgc 27301
 atttttctga ttgtgagaag ataaaggaaa tgtagtacaa ctaaacaatca gcagtcaa 27361
 gacctggcca tgactcctga gtgaggacac tggtaaacac catcaggatc caaacacctc 27421
 15 tgtatttacg aagaggatgc tccctattgg atagcactaa gcttatttca tgtatgtaca 27481
 tatgtagtta gttaattaca tccagcgggtg gcaaaagggt tgttctgacc caatgaaact 27541
 ttctctcctg gcccccttcc agcatgtggg caggagtaga gtgttggtgc catgaggcat 27601
 gcatttgtag agatgactac ttactcctcc ttgaaacatt tttttccatt tgcttccctg 27661
 ctgtctcact catgggtctg ctccctaattc acaaatcact cttttcccag tcttcttggt 27721
 20 tgggttttct cctcttctgt gcttgttagac atgggggagc cccagggtt ctctcttgaa 27781
 ctacagcttc tccctgggtt catctccttg ggatgtctgt tccaatgggt ttaaataacta 27841
 gggctaggac tagggagagg tcattgaggg acttagcaca taaagtttaa ggaacattc 27901
 tttatcaggc ccatgcaagt gcaggactgg ccctggaggg tgactaccac cttacctttt 27961
 ccaccctagg caccttggtt gccttaccct agccccagtc ctctttaaca cccagtcctt 28021
 25 ctccctggac ctccagaaac ataaatccta tttgacattt ctacttgag gttttaagggt 28081
 aactcaaacy taaaatatct aagacagaac tcttgcatca cttcccatcc tggggcccaa 28141
 gcctgtctct tctactagtc tatctcagtt aacagcatca ccattttattc agttgtctag 28201
 gacaaaaaat ttgaagtaat ccttgactct tctttttttt tttttgagac ggattctcac 28261
 tctgttgccc aggttgaggt gcagtgggtg gaccttggt cactgcaacc tctgcctcct 28321
 30 gggttcaagc aattctcctg cctcagtcct ctgacctcgt gatccactcg cctcggcctc 28381
 ccaaagtgtc gggattacag gcgtgagcca ccgcaccgg cctggattct tttttttttt 28441
 tttaaacaag tcttatcttc catctccaaa atgtatccca aatctgacaa cctctcccac 28501
 cgtaggccag ccccatctc tccctctga aaatagcctc ccttagatct ctggacattt 28561
 gttcttcccc acccccttgt gatcactatt cagcattcag aatgatcttt taatattatg 28621
 35 aaggagactg tgttctctc ctacttaaaa ttctctagt gcttcctaata aaatttagaa 28681
 taaaaagcca actcctcgcc atggtcacca ggccaggatc cagtgggtac aacaatctgt 28741
 tcccagcaca ataatcccac ttctgcctcc ccaccattca ccaggctcca ctgcaactggc 28801
 tcattcctgc ctcagttgtg cttcttttcc ctggaaagtt ctttctgtag atcttttaaag 28861

ggtgatttcc ttctcaacat tcaggtgtca gctcgtttca ctttcttgac catgcccac 28921
 cactcagaga tcaactcaaaa tcccattacc ctatttttatt tctccatcat atgtatcact 28981
 atctgaaact atcttggtgt tgatgcaggc tatttttcttg accccttcat gggactccca 29041
 acaggggtac cccattttact cagcctgccc tgctcaacct cttgcaggag ggagcacacg 29101
 5 agtgaacgag tgcaggaacc agctggctgc tttagtgtctg tgaggagtaa actccatgca 29161
 ggccctgcag cagcaaccag gtaggggtgc ctgcaacccc agggccccag aggggtgtgtt 29221
 acaatgctct cgtagctctg ccatctgtgg acagcagtgt gttgtcagct cagtgggccc 29281
 tttgcttcat catgtagggg ggtgcacctc tgcctgtgag ggcaaagggc cagggtgaca 29341
 gtcttttttg gtacccacaa tttgtgcac ctagaattctt gtttggtgcc caagaagaat 29401
 10 ggggtcacac agatgaactg aaggatgggtg aatgcagaga attagcaatg aaagtggctc 29461
 tcagcagaga gagaagctga aaagggagt ggaagggcag gtcactctcc cctgaagtca 29521
 agtcacatct ctccaatgtc cagccaccat ctctgaagtc aagtttcctc tctctgatgt 29581
 ccagccactt ctctctctta ctggctgagt ctgggggtatt tataggcaga ggataggtgg 29641
 tggggcaggc catacataat tttggaaaag gcaacattct atttggtaaa agacattatt 29701
 15 cataaagaac caattgggaa agagcgggca cacagggatg gaagttctca ctttgggctg 29761
 caggtttcag gcttttccag tcaaaagtga ggtttttcca gggacctgcc ccgtctgcct 29821
 aaaatttcta cacttctgtc attgataccg ctggataaca gctctcctat aaagttcagg 29881
 ggcttaaaac aaaaatttac taattcaccg ggacagttct tgatagggtc tctcctaact 29941
 gttgcagtta catagcaact gggttggagt catcttttct gggcttgaca tccaggacag 30001
 20 cttcttccct tgtgtgtctg gtgcctcagt gctcctccag gcagccttc tctccagaag 30061
 agtagcctgg acttcttggc aactcaagct tccaaaagaa aaaaaagcag agactgctgg 30121
 ttctcttata aaacaggcct ggaactggca caatgtactt ctgctgccat attcaggagg 30181
 tcaaagcaat cgaaggccaa tccaagttca aaggcattgg agaaaaatga gaagtgtcac 30241
 ttaaatgaga agtgacacat gcgtaaaagg gggaaaagca ttgattgtgg ccatttttgg 30301
 25 agataagcta tcacgtttat ttgtttgttt gcttttcta ggtctgtctt ctcccattag 30361
 gttataagct ctgtgagaca acaggaatct tgtccatctt gttttatggc tctacttcca 30421
 acacctagaa taatgcctgg cacatagtag gtgctcagtg aataacttaa gcacttgata 30481
 catgtttggg gaactaaaat gaacagaatt aaacttcccc agattgggtcc tgccagattt 30541
 gctgatgcca agcatgctga tgccctacca gatgaaagaa gccctaaaat gtaggggttc 30601
 30 gcttttctctg caaacaagaa aaacttgccc tgaacacaaa atctagaaat agatttggcg 30661
 tgttttctac attgaaatat ttcccgtagt accagaaatt attttccac agctttgtgc 30721
 tacattaaaa tattgaagtt gactgaaaat atctccattc tttaatcttg gtgtagacta 30781
 gaaacaattt ttttgtaaca aagtaaatat gaaaacttcc taatatttga actccccaga 30841
 tatccccaga tatctccaaa cttaaaatat cattgcaagt taagataaat ttttttaaat 30901
 35 gactaccgag aaaggtcatt aaaggcttgg ttattaaaa gtacagattt ggggtataaa 30961
 gccagaactt atttgtttaa atcattacat atgaccaagc acagaaaata aattacctca 31021
 aatctcctct ttgctaattt ttactggtaa actctataaa atgacccat ctttaaacct 31081
 ttttgtaaac cccttataag ttagtaagtg agaattgtatt catcagaagg attttagtga 31141

tgtttgaaat taaaaaagag agatttgatt tttaaaatta tacttgcaga ctactgctaa 31201
 tgaaacttct tctaacccta gttttgtctt atcttcagtt tttccagatt tgctcaaagc 31261
 aatcccagtg agcatccacg tcaatgtcat tctcttctct gccatcctta ttgtgttaac 31321
 catgggtggg acagccttct tcatgtacaa tgcttttggg aaaccttttg aaactctgca 31381
 5 tgggtccccta gggctgtacc ttttgagctt catttcaggt aagtacaaaa ttctacctct 31441
 gaagacaaat gtgcttttca atatgtcaaa aagaccgtct acctaaatat aaagttataa 31501
 tcttaacata tatacatgga tgcacactgt agtattatac ataataacaa aaagtgtgga 31561
 aataccacac ttgctcaaca gtagggaatt caataaatac tttatggaca tctatatgaa 31621
 taactgtgat gctgacatta aattatattt ttgaagatgt aatcaagagg ataaacgctt 31681
 10 gtctcaaaaa gttacatggg aaaagcagta tgtaaactta tataaacatt gtgaacctaa 31741
 tttgattata tatataaaat atagggaata tacatataaa atacatatat gtatatatgg 31801
 gggctatata tatatatata tgaaagagat agatagatag atagatagat agatagatag 31861
 acagacagac agacaatata gactccaatc tgttggctgt ggttgtctct gacctatgac 31921
 actatggggg acttttattt ttgctcatac ttttcaatat ttcttagtgt tcaataatgt 31981
 15 gctattattt atacataata ataaaaataa ataaatggca tcaaaaaaga gtaaagggcc 32041
 agtgttccgc ccacatatga gcagccatat tcaagcctgt agacactttg ttagcctaa 32101
 tgctaggtgt atctgggcaa ggataaactc taaagccaga aattagttca tcaataaaca 32161
 tgtgtactc aatagctagg gctgatggaa aagaatataa aaccagctct gtgccaaatg 32221
 gtgcttacta tctgcaagtg ggagaaggag aaagacgaga aaatgaaaaa tgtgtgtata 32281
 20 atttatatgt agctgttctg taggagatct ctgacttcac cccattctaa ctttgcaaaa 32341
 agatccaaca ctttgtcaga ttccctgggag gcaagtaatt ttattgatgg tttcatggag 32401
 ggatacagaa cgataacaac tcacacaaag caaacaatgt aatgaaaatc tctattcgac 32461
 tgtttctttt tctcctgaag ttgccctttg gctgccagct accaggcacc aggctcaagg 32521
 tactttcttg ctcttgacac tactcccttc tctcatacaa ttcaaccca accacaaacg 32581
 25 tgtatagatc tctctctcta taaaacaaag gcctgtagtt aacaggaggt cacttgcagt 32641
 gtagcctctg ttcatgttta cttgtgcaca ctgcttaggg tctcaccca tccacattct 32701
 gctaatacaca ttattcacc atccaatgta gatctctcca gtggagattc tgctaataatt 32761
 ttcttttagat ttgtcacaag tatataatac agttttaaat tgtacagatg attatcctat 32821
 aacagaagag tctaagcctt ttacatcttt gtatctctaa cgaaagcatt cagcatgaag 32881
 30 ctctgtacac agcagacaat tcaatatgaa tttgctgact tgaaacagca agcctagaaa 32941
 ggagatgtta acttggtcac ttagacagaa cagggtttcag caatcagaat tcagatgaca 33001
 tggaactggt agaacaggcg ctttgaagca ataggacatg agccagttag gagaggatg 33061
 gaatatcata aacaaaggcc aagggttttg caatcagagc tgaagagcca agagcacagg 33121
 ctcagggtgt gggcagactg aatgagaaa tgattcaatc acatgtgaaa gtccagatga 33181
 35 gaagagagag ttgggattac ttctgtcac caaacatcca aaaccaaaca ggtggatccg 33241
 ggtggtgtgc tgttttactg atgaccatta cacagaattt taacagaagg aatgtaaagc 33301
 agtggttctc aaactggagt tcccagatga gcagcgtctg catcatctgg aaacctgata 33361
 aagcagcaaa ttctcaggcc ctacccaga cccactgaat cagaaacttg ggggtctggg 33421

ggaagatggc catctgtatt ttaacaatct cccttcagga gattctgagg ctggctcaag 33481
 tttgaactac aggtagttgg ttcaacaggt gttggtggac tgacaaacaa aaagagactc 33541
 cgaggttaact ccaagatggg aatgtcagaa agcagctacc acccctaggg cttgggggaa 33601
 ccaacaaaag agtttggcat tgccagaacc tagaatcttg aggagaggcc ccagagcatt 33661
 5 gtgtctcaga ccttgaggac ttggcactgg gacaccatga ggggtttctg ggtgtgggaa 33721
 agggctggaa actccccagt tgctgccacc agggagaact acaagtgaag tggaaggtgt 33781
 gggcctttct cccttttctt ttctctgtct ctctctcccc ctggtgctca tgtttgacag 33841
 aaaacagctg aaaaggcaga actagtttgg ggagtcttga cctggcatca taaagcagag 33901
 aaaagcaaag ctggagtga ggtgagacac aacagctcat tagcagcaac agccgtctag 33961
 10 cgctccagct tctgaatgaa attctgaaga acagcgcact tggaagacaa attatttgac 34021
 agttctgaca gacgacaaa ctaacagcat ttgaaaagca agatgactca gagaatacag 34081
 aatttaatcc aaatcccaga tcctatctct gcctttggcc aggcctaata caaggagaac 34141
 ctgagcacac aaatatatgc aggaatgatc aggacctgtg cctgcattct attctgtctc 34201
 acccaccttc aaatttggtg taaaaacatg ggctcaataa aggtttgtga atcagggaag 34261
 15 gaagagaagg ggagaaagga agggaaggag ccagctccag atctgtgtct tgagaggat 34321
 aaaggccagt gtttttagat caccagtggt ttttctaagc cccaataact tattttgaaa 34381
 tatcaaatg ttcaataact aaaaaaaaaa ccgttacaac aataaaacat gtttgagggt 34441
 cagatggact ggcagtttgt gacctctggg gataaacagg tcactttgga atcacagact 34501
 tcctcattcc ccttaaatct catatggtac ccagaagccc ttggaacttt ggaaggtgtt 34561
 20 tattcacagt tgtaatgtcc atgcagacct tggctctaag acccaattgt gtaagggtag 34621
 gttttagacc cttatcccaa acattctaag tgtgagccaa tgcgtcacac actcagaggc 34681
 cagagactgt attgggtgcc tttatttcac gtacgagtca cattccatta agagacccca 34741
 gaagtcagct ctcttccact gactggttct ctcccttgt ttctcttgcc aatgtgtgct 34801
 gccaggtgg cagtgtctac tgtcagcaga gaagaaaaat gctttcctcc ttggacctct 34861
 25 tttctctttt tctcctccct actcacattc aggttcctta agcttccccg ctcttgtgct 34921
 tgaagtcatt ctatggtcat ttcttcaact gtctacttcc ctgctggatg ggcacccaag 34981
 acttggcatc ctggggcatg tagaaagggg aaaggggaagg gaagggaaaa gaagtcctcc 35041
 caattgtcta tctggacctt tccacactgc ccagagtact gctatgggca tctccttatg 35101
 tctcccgatg tgggtgcatgc cagacctgc aggtagaaaa ggaaagaaag caacccattg 35161
 30 gaccaggcca gcaaaggctt cagtcacaca gctgggtcat acttatggct tcatattctg 35221
 ttgcctcttg aaccagacat ttcttccact ctcataacct ccagtttagc tcgtattcct 35281
 cagcattctc catgtaatat tgttgcatga aacctatgca agtcagccaa tttgctcttt 35341
 ctcatcttgt catttataaa ttgatgctga gaagtctttt tccaagggtt ttagtaata 35401
 cttcatcat ccccatagt tcattttggg cagtgattcg cttctttgac tgtacattag 35461
 35 aatcatctga agaactttct aaaactactg atctcaggtc tcacacaaca ctaattaaat 35521
 tatagtctct ggtggggtag ggcttgggca ctgctatttt accttaagct ctctggagtc 35581
 attctaattt gtagccagtg ctaagggttg caatataagt gaatatattt cacgtatttg 35641
 tcaaacattg actgagtgcc cattatgtgc cagccattat aataggcact ggtgatccca 35701

cagtgaatca ggcacacaat gctgtcttca cggagcttgt tgtctagtgg gagagtcaaa 35761
 caaaagtgtat tatcaataat taagtgatta cagattgcaa taattacaat aagggtgata 35821
 aacaggttgc tataatatac aatagtattg cctttcacca gacatttctt aaagaggtaa 35881
 atcatctatc agacaccttt taaaaatctc atctaatttc aaaagtgtac ataaataatt 35941
 5 aagtgattac agtttgcaat aattacaatg aggggtgataa acaagttgct atgaaaagaga 36001
 aagatagcag agatctggct ttgagatggg ggtcaggga gactgctcca atagctgagt 36061
 tctaaagcta agaaggaact gagaaccttc acagaacgct ccaagtagaa gagaaagcac 36121
 actgaagact ctagggaaaag aggtctgctt gttgtaggaa ccgaaagaag gccaatgtgg 36181
 ctaggtgctg ggtagtgagg ggaaatggca caaggaaaaa atgaggttag agagatcagt 36241
 10 tgataccagt taatgttggg ccctaaacat taaccatggg aagtctttta gaattgattc 36301
 taattgcaat gaaaaccttt tgaaagattt taaaaagata tctgatagct gatttacctc 36361
 tctaagaaat gtctggtgaa agacaatacc atcacattgg agatggaaaa agatgaatgg 36421
 attctaaaaa cattctgaaa gtacattcaa aatgtttttc aggtagctta tgcaactaat 36481
 aaatagtggt ggcattctgg gtaagacaag ggaggagcag gcttgagctt tagggcaaga 36541
 15 aaggggtggg gagaagagct cagcactaaa atcatgtgtt ccatttgggg cacatcgagt 36601
 ctgagttgct atgagaccac caagtggaga tgccaagtaa atagtcagtt acatgaatct 36661
 ggagttcagt gaagaggtct agaagaaaga tgtatatattg ggcattattc ggatatagat 36721
 attatgtaaa gcaataaaaat tggatgagat cacctagggg gagaatgcac atagataaaa 36781
 actgacctag gaccacttca tgtctaaaac accaaaagca atgtcaacaa aagccaaaat 36841
 20 tgacaaatgg gatctaatta aactaaagag cttctgcaca gcaaaagaat cagagtgaac 36901
 aggcaaccga caaaatggaa gaaaattttc acaacctact catctgacaa agggctaata 36961
 tccagaatct acagtgaact caaacaaatt tacaagaaaa aaacaaacaa ccccatcaaa 37021
 aagtgggtga aggacatgaa cagacacttc tcaaaagaag acatttatgc agccaaaaaa 37081
 cacatgaaaa aatgctcacc atcactggcc atcagagaaa tgcaaatcaa aaccacaatg 37141
 25 agataccatc tcacaccagt tcaaatggca atcattaaaa agtcaggaaa caacaggtgc 37201
 tggagaggat ctggagaaat aggaacactt ttacactggt ggtgggacta taaactagtt 37261
 caaccattgg ggaagtcatt gtggcgattc ctgaggatc tagaactaga aataccattt 37321
 gaccagcca tccattact gggatatata ccaaaggact ataaatcatg ctgctataaa 37381
 gacacatgca catgtatggt tattgctggc ctattcacaa tagcaaagac ttggaaccaa 37441
 30 cctaaatgtc caacaatgat agactggatt aagaaaatgt ggcacatata cccatggaa 37501
 tactatgcag ccataaaaaa tgatgagttc atgtcctttg tagggacatg gatgaagctg 37561
 gaaaccatca tcctcagcaa actatcgcaa ggacaaaaaa ccaaaccacg catgttctca 37621
 cccatagggt ggaattgaac gatgagaaca catggacacg ggaaggggaa catcacacac 37681
 tggggactgt tgtgggtggg gaagaggggg gagggatagc attaggagat atacctaattg 37741
 35 ctaaattgacg agttattggg tgcagcacac cagcatggca catgtataca tatgtaccta 37801
 acctgcacat tgagcacacg taccctaaaa cttaaagtat aataataata aaataaaata 37861
 aaataaaaaa acaaaaattg atgtaggacc aattcctgaa gaacactgac agttaatttt 37921
 ttggtttagg agggaggaga gccagcaaac gacactgagt agcaatatcc aaagaaaaag 37981

aggaaaaagg aaaactggga gattatgagt gtcccagagg gaatgtttca agattaccat 38041
 cagcagtgag ctttgtgtaa aggtggcctc ctgtaataga ggtgcgggca ggagaaggca 38101
 gaatagggaa aaggggggtga aaaagcttcc ctcaagattt ataatacagt ggaagagaga 38161
 gacagagaga gagaaagaga aagagagaga gagaacttaa ggaggtagag gaagagagag 38221
 5 aaccaaaaaa gagggagctg agtatagaag caattagatt catagttttt agttgcggca 38281
 gtgatatttg agtgggggcc ttttatatat tccattctag gtgtttccca gttgatggga 38341
 gagggctcta cctagatctg catgtaaaag ggagtaggcc agctggcaga cttgacatgg 38401
 atcagtggta gaacatccta gcagttctgt gaatactctc tgagaatgac atgaaaggta 38461
 ttggttcagg ccttttgagg gtgataaaaa ccacaaaaat gtggacttat tgcaaattgt 38521
 10 atttgttacc atttgcaatg attataatta ttctcttata tataaggcttt cttacatata 38581
 gcttctctta cacatagcgt catgagttat gccttctctt acatatagtg tctgctatga 38641
 gttatgcaa gcagggcaaa caaaattgct gcctttcttt aaaaagagga cgctcctagt 38701
 atgggcctaa ttaattatga taattacagc tatggcatgg aacataagca cattcatata 38761
 cacaaagaca ataaaaataa gaacagttca aaacagaaac agttacatta tatatcagtt 38821
 15 tcagtatgaa taataccctg gactctgaaa tatgtctggg gcacattatt ctgtaattgg 38881
 tggtgaaaaa aaatctgcat cttatctcta cgccaatcct tatgaaggag ctgtttttca 38941
 ggagttcgag aaagagacac agggatgtcc agtcatcaaa gccgcagagc ctgaggaaga 39001
 ataaaggatt tgtggcagga aaaccagta atgaatgtat tctgctagtt tctcagcata 39061
 gaacttagaa aagaggccac agaaggaaaa gagaagtaat gttagacagc tgatgttggc 39121
 20 aatgggcaaa gaatttcatt ctcatatgca gggcagtggc taaaggggca gtgttgtgag 39181
 gaatcacatt ccaggtcatt catgtccagg gtgtggtagg aggactgtgt ttcatattatt 39241
 ttgtacatgg ccagctatg acctgtgca aggaaatgct taatcatttt ctatcacagt 39301
 tgcaaagaga ttcttatacct actcagaatg tacgtcttct cttctgtttc atttacagtc 39361
 acaaccaag tccttgcttt gacctcaaaa gcaccacttg atgtatccac ttcagaaaca 39421
 25 cacacagaca gtcacctct ttctatacct taccttctc cctcttcac tcaaaccacc 39481
 tacttccttt gcattcactc ttcttcagc ctgaaataat cttctccaa atatctacct 39541
 tctcactccc tcacttctct caagatacac ttaaagtta tattccctat gaggcctccc 39601
 ctggccatcc ctccccagcc ttctatccc cctctctgc tttattttat tctccttaat 39661
 atatatcaca ctctgataaa ccattgaatg tacttatcaa gttattgcct ctctctccct 39721
 30 ttccattgtc tccatcactg agagctctgt gaagaaaagg atttgtacct gtttcattta 39781
 gtgctgtacc ccagttccc acaacagcgc aacaggcact caataaatag ttgttgaata 39841
 agtgaataaa acagaagtag ctgcataatt tctggtaaca aatgatattc ttctgaaaat 39901
 gtcataattt cagacatatt tccgaaaata aattcaaatt agataaacat tgtattttta 39961
 gaccatttct tctttgcatt aatcactcct ctcaataata taacatttgt aaaacttagg 40021
 35 ttagatatgg gctcttcaac ttccattac agaagataaa gtgaaaaggc tagaccaat 40081
 ggtgttattc cttcatctac atctatcctt tggaaacaca tgaccaaatt gcttgccatc 40141
 acaatctcaa aatctaccct ttggtattaa ctacttcac ttgtccctct gtcccttttt 40201
 agatggtagc catcgggtctc tggagcagtg tagagtcaga acaacttcta tttggggaag 40261

5 aaatcattgc tgggtgacctt acttttcaatt actaactttc tagtgacatt tacataatth 40321
 tagagaaaat taacacctac acttgtaaag ttgtggcttt cccacaccta tttatcatct 40381
 ctcaatatct cttgaaaagg aaattatcaa tttatcattc tatattggca atgaaatgcc 40441
 cctaatatct gtcacctata agacaattga agatgatgtg ttgaaagctt tctgaaaatg 40501
 ctgatcatta ctttaaattg aattgaaatt ccagtttatt atttccaaaa atatgatctt 40561
 actgatcata ggataacatt tcataacatt tcagagattt cttccccttc gaggagccaa 40621
 acccatagga cctctggact cccacagatc ctggcagggg gttcccactc ataaaagcac 40681
 aggtgccctc agagtcattc agggatgaag aagcaaccct cattggccat gtcctacgtt 40741
 ccccatatag taaggactgg aggagaccag tgctcaattc tgcagtctca gacagctgtc 40801
 10 agaggagagt catgaatgtg cagtgtctag cacattacaa acgtttgtta attgactgat 40861
 cattcatggt gtgacagccc tataactcag ctatcctatt cagtcagaaa ttaactcagt 40921
 aatcaaagtc attaaaagga agaaaaaaaa aacctacagt accagacaga tgggtggggaa 40981
 atcagacaga tgaaaggaaa aatggctgta ggctcattgag taagacactg ggcagcaaaa 41041
 cctgggcctt gtgcctgggt atactccaca ttatagctcc agcaagggtt ggcaggattt 41101
 15 ccacagtcct ggcttattct aacctttctt gggagcagga gcagtgttg tcatagatagac 41161
 agacacataa ggaatctgtc caactggcac cgtgtgaatt tgggctcttg gtgtacatgg 41221
 ataactggga aaaagaggag agagacatgt aggactgac ctaccgtttg tgaagtcttg 41281
 ggcaagagta tgaatgaaaa cccacttctc tccccctgcc tggctccact gcacacagta 41341
 aagagcctca agcatagggt tgtggacatt gcaccatgta tccaagctct gaccatgcct 41401
 20 cttgaaacag ctattcctca gccaccctct gaccatggga ggaatgacc aggagaaatg 41461
 accacatagg tcttgaaaat gggctcaggg ctatttacga agtcaattcc ggggtcccag 41521
 gagtatggac taaaatgtga gtcaggcatg ccagatgggt atgttctatt gacttcaagg 41581
 attcctcatg ctgtgggaag gaacctctcc agaagagaga cagagcagaa ccctctaaat 41641
 gtggggcaca aagcaggagc cctctctgct ggattcaaa ggtcactact gaagagtgtg 41701
 25 ggttgagtct tattctcaca tcaactcatat cacttacaca cttcttttat agccttagca 41761
 gccatgccac aaagagaaac tcttcatgca tatcttttgg tccataagtc ataaatagtt 41821
 atccttgatc ccatgtcttt tttagagcca tggacagaga gaagcaaaaa tataccaagt 41881
 tcacactgag ttgtctccct tcatactct tagcagtcac tgaaagggtt tgagactcag 41941
 gctgggtttc tatcctctgt cctgaaacg acaacgttg cctcgtgatc aaccctagaa 42001
 30 tccagagcaa gatctcagac tgctctctct actaccagac agcacacttt gttttggggg 42061
 ctgtgtgctt gaaatattag ctatggcaaa aggctttgag tcttatgaca cccaagtaa 42121
 cttttacttt aggaatttga aatacagcct tgctgtaatg ctgtctcctt aacaaagcag 42181
 tacctttgaa atattttaaca acttgaaaag gaaaccgagc ttgaattttc ctttcaggtg 42241
 ctgaggaaat aatgtttcac ttctgtctga aattcaccat ctctcagac aaagaaggct 42301
 35 cttatggtaa aaggaatggc attttctcca caattttcga ataaaagata aagagaaaac 42361
 agcactgcag cctttttgtt aggatctaac aataaagaaa taatacgggt ttgccagggg 42421
 agagctctgg ttttaagctc agaatacaaa aataggctga caaaatttta caaaggaata 42481
 ttctcagcta ccactctgag gatggtagaa agtgaaatth caagaaaatt atattatttg 42541

attattttgat gatgattaag ctgattggcc agtcctatgt gaaattctaa agtagaagaa 42601
 atgtgatgtt gtctttttctg ctcacctctc ccttcattcc tacccecaaa tctctgcctc 42661
 tacccecaaac ccagcctgac ctttgggaaa ggaatggggg ctgtcacttg caccgtagct 42721
 cctcctgcct gcagtactct cccccacag tccagcctcc ctctgctcag ctaacttttc 42781
 5 ctatagctcc ttctggctac ttctaaggaa ccttcataa tgctgccac cctctcagta 42841
 acagcccctc tactatccct tgattacatg cactgtaatt ggtagtaat tgattttatg 42901
 tcctctgcca aattatactc catgagagca aaaatcatga gtattatctt taatgtttaa 42961
 atctccacaa ctatcccccata tataagtcta gagaataaat gaatgagtga attaataaat 43021
 agaaactcaa gccattatgt tgccactcct aggatatttg gattaacttt aactgaagag 43081
 10 taaaaagcat ttacctgtcc taaaggagac ataaaattag tgggagagta tttggagaaa 43141
 aaaaagacac tgtaatacat tcttttgtgt tgcctaccct gatgtagtca ggtgtccctg 43201
 atatggggtg ggctgaggat ttgaaataaa atacctaatt tgacattgta agtggaggaa 43261
 tcaggatttg aaaccaaatt ggtttgacct aaatcaagct attatgataa ggacttgcaa 43321
 gaaaaaagga atccataaaa accatatgaa tagctacca ttagtaagca tttatatgtg 43381
 15 tcaattacaa agccaaatgc aaatcatgct ttattttatat tagccacctt ttatagatga 43441
 agaaattgag acctgaatat taaaattgcc tactttttaca tagtaagtaa aggaatcagg 43501
 gtttgaacct aaattggttt cacctaagggt agaaaacct cccagcaagt ctctattaa 43561
 ctggaacctt attgtggtgg cctgagatat aacagtagct gtggaagcgc ttagagatcc 43621
 tggccatcct atgtgctcct gatctggctc ctctgccac ctgcttctgc tcctgtgcc 43681
 20 atccacccat ctggaagtct cccagtgtcc atcttcgggg gagacactca ccagagtctc 43741
 cagcttccag ccagtatgga gtgcccctgt cccacagcaa tctcaccgaa atcacagcta 43801
 catctgttaa aattaggcta ccaatgagt atagatgagg gggaaaaata ataatagtgt 43861
 actaaacaaa acaaatgttt atttttctca cacataaaaa tctagagggt gaagtccagg 43921
 gctggtccag aggtcccaag gatctgggat ttagactccc tctttcttgt tttccacag 43981
 25 catatggctt ccatttctgg ggccacattg gtccaaaatg tatgctgggg ctccagccat 44041
 tgcattcata tttcagccac aggaaggagg aagtggggaa gaaaggacag gccctaata 44101
 cctgtatagt tcaagaagac tatccgcgcc atacttccca accaccctta gttgaacaat 44161
 gctgtcttaa ttcaagacac tcacatgtct agccaaaaat ctgaattctg ttacaacaa 44221
 ggagaataga gatgtgcgcc acctcaatac ctcatccata gctacctttt cttttgtgca 44281
 30 gctgtggcca agtgaaagct gaaggagctg tggtaacct tctgaaggga ggctggggcc 44341
 tttcacaaga ggctgcatga ttgacattta tcctgcatgg cctgtgaagt acagagaaat 44401
 attttctctt gaagccacat catagcagtg gctgctttgt agcctgattc caccattatg 44461
 ctttaaaagt gcctagcaat tcagccttca catcatgcaa agaggaatat ctccagctct 44521
 ttgtaagatc agcttaattc taaccacctc cttacctccc actgcactcc tacacgcaca 44581
 35 cacaaatctt cttcactcag agcagaacca taaccaagc cctaccacct agagactgaa 44641
 gaatcaggct catgattaca aatatgcaat aattttttgt gtggataatg tcaatgggga 44701
 tgatggtaag agaattcctt ggtttacaca ttgacctct tcctgtccc ttacaatcag 44761
 gaaatatttg tcccaacacc ttgtttcttc tgttgcaggc tcctgtggct gtcttgtcat 44821

```

gatatattgttt gcctctgaag tgaaaatcca tcacctctca gaaaaaattg caaattataa 44881
agaagggact tatgtctaca aaacgcaaag tgaaaaatat accacctcat tctgggtcat 44941
tttcttttgc ttttttgttc attttctgaa tgggctccta atacgacttg ctggatttca 45001
gttccctttt gcaaaatcta aagacgcaga aacaactaat gtagctgcag atctaattgta 45061
5 ctgaaaggca aaccttttcta taattttaca agggagtaga cttgcttttg tcacttttag 45121
atgtgggttaa ttttgcataat ccttttagtc tgcataatatt aaagcatcag gacccttcgt 45181
gacaatgttt acaaattacg tactaaggat acaggctgga aagtaaggga agcagaagga 45241
aggctttgaa aagttgtttt atctgggtgg aaattgcttg acccaggtag tcaaaggcag 45301
ttgactagaa tcgacaaatt gttactccat atatatatat gtgtgtgtgt gtgtgtgtgt 45361
10 gtgtgtgttaa gatgtcttcc tatcaaaaag atatcaaagg cacatggaat atatttttaat 45421
aaaaacaaat aatatctcta atatatccac acatttgttg ccagatttca gaaaactgag 45481
ctgcaatcgc tttcctaaaa cagtagtgta ttaaatgaac atctataaaa tgtatcaaca 45541
cacattttaa aaaatttgtt taaagtatac tcttaggcca ggcgtggtga ctcacacctg 45601
taattccagc acttcaggag gccaaagggtg gaagatcatt tgagttcagg agttcgagtt 45661
15 acagcctggg caataaagtg agaccctgtc actaacaaaa ttaaaaaata aaataaatat 45721
aaaatatagg ctttaaaaaa gcatagtctt attaacctatg tctgttggtc aaaatctgca 45781
aactctaaaa gaagaaaaga agaaaaaacc aagcttaggg tatttttcct cccgtgcctg 45841
agtcccaatt acattcacga cagtactttc aatgaacata attgttagga ccactgagga 45901
atcatgaaaa atgatctctg cttagtacat ttgatgcaaa atgacttatt aggggctgtt 45961
20 tttctagcta tagtgtctcg agtactaata tgcaattatg aaaattatat taaatctggg 46021
attatgacgg tatcactgta tcactcttgg cttgttcttg ctgtcaccaa gcatgacca 46081
ggtcaacttt ttttttcccc tgaattacc atcaaattga tctgcagctg actaaaggcc 46141
acagctgagc ctggaactga cccttccttc atcctcaacc tgetgtcctc cagaaagcac 46201
caaggaaaaa gcagagaatg acagcaaca gatcactagg cctctgacca cagggtctga 46261
25 gtactcagca gccctcatat aatagggttg aaagtactcc ttaaaataaa aactgtttc 46321
cctttggaac tatttacaag gatgaaaca ccgtatacct gagaaataac ttgctctggt 46381
gtcaattcgc tattcgccag cagacatcag aacacaccga gtttccagat gctggttttt 46441
ccccttaaat caggaaatac acctggacaa tttctagaag actacaattc agtctagcca 46501
caaaggggat tttttttttt tggttaacagg ctagagcccg gttctgtaag tctttagctg 46561
30 aaatggtcca gtacaaaagc actggaaatg agtgggctag gaggacaagg accgtctcct 46621
gcgtgaggag ttggttgag gtccccaaagg ccaggtaccc cctgcactct tattggattc 46681
ctctctgtct tcttgaggtt ttgaaaaact ccttcgaaca ccaggctttt ttcttttagaa 46741
aacaagtctc caatcgttct ctgttccgta gaaagagaaa gaaaacctgg agcagctgct 46801
gaaaaatcta atgaggaact aagaggcaaa cccacca
35

```

A non-limiting example of a human wildtype CLRN1 genomic DNA sequence is SEQ ID NO: 9. The exons in SEQ ID NO: 9 are: nucleotide positions 1-544 (exon 1),

nucleotide positions 28764-29180 (exon 2), nucleotide positions 31239-31418 (exon 3), nucleotide positions 32481-32519 (exon 4), nucleotide positions 44799-46433 (exon 5), nucleotide positions 44799-44935 (exon 6), and nucleotide positions 46128-46837 (exon 7). The introns are located between each pair of these exons in SEQ ID NO: 9, i.e., at
 5 nucleotide positions 545-28763 (intron 1), nucleotide positions 29181-31238 (intron 2), nucleotide positions 31419-32480 (intron 3), nucleotide positions 32520-44798 (intron 4), and nucleotide positions 44936-46127 (intron 7).

Mouse CLRN1 Protein Isoform 1 (SEQ ID NO: 10)

10 MPSQQKKIIFCMAGVLSFLCALGVVTAVGTPLWVKATILCKTGALLVNASGKELDKFMGEMQYGLFHGEGV
 RQCGLGARPPFRSSRSMKERYSLYEDKGETAVFPDLVQAI PVSIHINIILFSMILVLTMTVGTAFFMYNAF
 GKPFTLHGPLGLYLVSFISGSCGCLVMILFASEVKVHRLSEKIANFKEGTYAYRTQENYTTSEFWVVFIC
 FVHF LNGLLIRLAGFQFPFTKSKETETTNVADLMY

Dog CLRN1 Protein (SEQ ID NO: 11)

15 MPNQKKVVFCTAGVLSFVCALGVVTALGTPLWIKATFLCKTGALLVNASGQELDKFMGEMQYGLFHGEGI
 RQCGLGARPPFRSLFPDLLKVIPVSIHVNILFSTILVLTMTVGTAFFMYNAFGKPFETLHGPLGLYLLSF
 ISGSCGCLVMILFASEVKIHHLSEKIANFYKGTQSEKYTTSEFWVVFICFLVHLLNGLLIRLAGFQF
 PFAKS KDTETTNVAADLMY

20

Vectors

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 CLRN1 protein, each of the
 25 encoded portions being at least 30 amino acids (e.g., between about 30 amino acids to about 202 amino acids, about 30 amino acids to about 200 amino acids, about 30 amino acids to about 180 amino acids, about 30 amino acids to about 170 amino acids, about 30 amino acids to about 160 amino acids, about 30 amino acids to about 150 amino acids, about 30 amino acids to about 140 amino acids, about 30 amino acids to about 130 amino
 30 acids, about 30 amino acids to about 120 amino acids, about 30 amino acids to about 110 amino acids, about 30 amino acids to about 100 amino acids, about 30 amino acids to about 90 amino acids, about 30 amino acids to about 80 amino acids, about 30 amino acids to about 70 amino acids, about 30 amino acids to about 60 amino acids, about 30 amino acids to about 50 amino acids, about 30 amino acids to about 40 amino acids,

about 60 amino acids to about 202 amino acids, about 60 amino acids to about 200 amino acids, about 60 amino acids to about 180 amino acids, about 60 amino acids to about 170 amino acids, about 60 amino acids to about 160 amino acids, about 60 amino acids to about 150 amino acids, about 60 amino acids to about 140 amino acids, about 60 amino acids to about 130 amino acids, about 60 amino acids to about 120 amino acids, about 60 amino acids to about 110 amino acids, about 60 amino acids to about 100 amino acids, about 60 amino acids to about 90 amino acids, about 60 amino acids to about 80 amino acids, about 60 amino acids to about 70 amino acids, about 90 amino acids to about 202 amino acids, about 90 amino acids to about 200 amino acids, about 90 amino acids to about 180 amino acids, about 90 amino acids to about 170 amino acids, about 90 amino acids to about 160 amino acids, about 90 amino acids to about 150 amino acids, about 90 amino acids to about 140 amino acids, about 90 amino acids to about 130 amino acids, about 90 amino acids to about 120 amino acids, about 90 amino acids to about 110 amino acids, about 90 amino acids to about 100 amino acids, about 100 amino acids to about 202 amino acids, about 100 amino acids to about 200 amino acids, about 100 amino acids to about 180 amino acids, about 100 amino acids to about 170 amino acids, about 100 amino acids to about 160 amino acids, about 100 amino acids to about 150 amino acids, about 100 amino acids to about 140 amino acids, about 100 amino acids to about 130 amino acids, about 100 amino acids to about 120 amino acids, about 90 amino acids to about 110 amino acids, about 120 amino acids to about 202 amino acids, about 120 amino acids to about 200 amino acids, about 120 amino acids to about 180 amino acids, about 120 amino acids to about 170 amino acids, about 120 amino acids to about 160 amino acids, about 120 amino acids to about 150 amino acids, about 120 amino acids to about 140 amino acids, about 120 amino acids to about 130 amino acids, about 150 amino acids to about 202 amino acids, about 150 amino acids to about 200 amino acids, about 150 amino acids to about 180 amino acids, about 150 amino acids to about 170 amino acids, about 150 amino acids to about 160 amino acids, about 170 amino acids to about 202 amino acids, about 170 amino acids to about 200 amino acids, about 170 amino acids to about 180 amino acids, about 190 amino acids to about 202 amino acids, or about 190 amino acids to about 200 amino acids) in length.

In some embodiments of these compositions, at least one of the coding sequences includes a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA (e.g., exons 1 and 2, or exons 5 and 6), and lacking the intronic sequence that naturally occurs between the two consecutive exons.

5 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
10 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 202 amino acids (e.g., or any of the subranges of this range described herein) in length.

15 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 CLRN1 genomic DNA (e.g., any of the exemplary introns in SEQ ID NO: 9 described herein), and wherein the two different segments overlap in sequence by at least 100 nucleotides (e.g., about 100 nucleotides to
20 about 10,000 nucleotides, 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
25 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 10,000 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,
30 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 10,000 nucleotides, about 400 nucleotides to
5 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
10 nucleotides to about 800 nucleotides, about 400 nucleotides to about 600 nucleotides,
about 600 nucleotides to about 10,000 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,
15 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 10,000 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
20 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 10,000 nucleotides, about 1,000 nucleotides to about
5,000 nucleotides, about 1,000 nucleotides to about 4,500 nucleotides, about 1,000
25 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 10,000
nucleotides, about 1,500 nucleotides to about 5,000 nucleotides, about 1,500 nucleotides
30 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 10,000 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 10,000 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 10,000 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 10,000 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 10,000 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 10,000 nucleotides about 4,500 nucleotides to about 5,000 nucleotides, or about 5,000 nucleotides to about 10,000 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 CLRN1 gene (e.g., any one or more of the exemplary exons in SEQ ID NO: 9 described herein).

In some embodiments, the number of different vectors in the composition is 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 CLRN1 protein. In some examples, the N-terminal portion of the CLRN1 gene is between about 30 amino acids to about 202 amino acids (or any of the subranges of this range described 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 examples, the second of the two different vectors includes a coding sequence that encodes a C-terminal portion of the CLRN1 protein. In some examples, the C-terminal portion of the CLRN1 protein is
 5 between 30 amino acids to about 202 amino acids (or any of the subranges of this range described above) in length. In some examples, the second vector further includes a polyadenylation signal sequence.

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
 10 comprising amino acid position 1 to any of the following: about amino acid position 202, about amino acid position 200, about amino acid 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
 15 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 CLRN1 protein (e.g., SEQ ID NO: 1, 3, 5, or 7).

In some examples where the number of different vectors in the composition is two, the N-terminal portion of the precursor CLRN1 protein can include a portion
 20 comprising amino acid position 1 to amino acid position 202, amino acid position 1 to about amino acid position 200, amino acid position 1 to about amino acid position 190, amino acid position 1 to about amino acid position 180, amino acid position 1 to about amino acid position 170, amino acid position 1 to about amino acid position 160, amino acid position 1 to about amino acid position 150, amino acid position 1 to about amino
 25 acid position 140, amino acid position 1 to about amino acid position 130, amino acid position 1 to about amino acid position 120, amino acid position 1 to about amino acid position 110, amino acid position 1 to about amino acid position 100, amino acid position 1 to about amino acid position 90, amino acid position 1 to about amino acid position 80, amino acid position 1 to about amino acid position 70, amino acid position 1 to about
 30 amino acid position 60, amino acid position 1 to about amino acid position 50, amino acid position 1 to about amino acid position 40, amino acid position 1 to about amino

acid position 30 of a wildtype CLRN1 protein (e.g., SEQ ID NO: 1, 3, 5, or 7).

As used herein, the term “vector” means a composition including a polynucleotide capable of carrying at least one exogenous nucleic acid fragment, e.g., 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)), a viral vector (e.g., any adenoviral vectors (e.g., pSV or pCMV vectors) or any retroviral vectors as described herein), and any Gateway® vectors. A vector can, e.g., include 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, a viral vector, etc.) that is capable of replicating when associated with the proper control elements. Thus, the term includes cloning and 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 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))).

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, 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.

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 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., a cochlear 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 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", 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 is 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.

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

In some embodiments, the AAV vector is selected from the group consisting 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%, 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 NO: 40, 41, 42, 43, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 or 60. 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 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-inactivating lentiviral vector as described 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.

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 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.

5 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.

10 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 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 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.

20 In some embodiments, the artificial chromosome(s) is a yeast artificial 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 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, 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.

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 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, 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 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 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 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 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 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 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 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 15,000 nucleotides, about 500 nucleotides to about 14,500 nucleotides, about 500 nucleotides to about 14,000 nucleotides, about 500 nucleotides to about 13,500 nucleotides, about 500 nucleotides to

92

93

1,000 nucleotides to about 9,000 nucleotides, about 1,000 nucleotides to about 8,500
nucleotides, about 1,000 nucleotides to about 8,000 nucleotides, about 1,000 nucleotides
to about 7,800 nucleotides, about 1,000 nucleotides to about 7,600 nucleotides, about
1,000 nucleotides to about 7,400 nucleotides, about 1,000 nucleotides to about 7,200
5 nucleotides, about 1,000 nucleotides to about 7,000 nucleotides, about 1,000 nucleotides
to about 6,800 nucleotides, about 1,000 nucleotides to about 6,600 nucleotides, about
1,000 nucleotides to about 6,400 nucleotides, about 1,000 nucleotides to about 6,200
nucleotides, about 1,000 nucleotides to about 6,000 nucleotides, about 1,000 nucleotides
to about 5,800 nucleotides, about 1,000 nucleotides to about 5,600 nucleotides, about
10 1,000 nucleotides to about 5,400 nucleotides, about 1,000 nucleotides to about 5,200
nucleotides, about 1,000 nucleotides to about 5,000 nucleotides, about 1,000 nucleotides
to about 4,800 nucleotides, about 1,000 nucleotides to about 4,600 nucleotides, about
1,000 nucleotides to about 4,400 nucleotides, about 1,000 nucleotides to about 4,200
nucleotides, about 1,000 nucleotides to about 4,000 nucleotides, about 1,000 nucleotides
15 to about 3,800 nucleotides, about 1,000 nucleotides to about 3,600 nucleotides, about
1,000 nucleotides to about 3,400 nucleotides, about 1,000 nucleotides to about 3,200
nucleotides, about 1,000 nucleotides to about 3,000 nucleotides, about 1,000 nucleotides
to about 2,600 nucleotides, about 1,000 nucleotides to about 2,400 nucleotides, about
1,000 nucleotides to about 2,200 nucleotides, about 1,000 nucleotides to about 2,000
20 nucleotides, about 1,000 nucleotides to about 1,800 nucleotides, about 1,000 nucleotides
to about 1,600 nucleotides, about 1,000 nucleotides to about 1,400 nucleotides, about
1,000 nucleotides to about 1,200 nucleotides, about 1,200 nucleotides to about 15,000
nucleotides, about 1,200 nucleotides to about 14,500 nucleotides, about 1,200 nucleotides
to about 14,000 nucleotides, about 1,200 nucleotides to about 13,500 nucleotides, about
25 1,200 nucleotides to about 13,000 nucleotides, about 1,200 nucleotides to about 12,500
nucleotides, about 1,200 nucleotides to about 12,000 nucleotides, about 1,200 nucleotides
to about 11,500 nucleotides, about 1,200 nucleotides to about 11,000 nucleotides, about
1,200 nucleotides to about 10,500 nucleotides, about 1,200 nucleotides to about 10,000
nucleotides, about 1,200 nucleotides to about 9,500 nucleotides, about 1,200 nucleotides
30 to about 9,000 nucleotides, about 1,200 nucleotides to about 8,500 nucleotides, about
1,200 nucleotides to about 8,000 nucleotides, about 1,200 nucleotides to about 7,800

95

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 nucleotides to about 5,600 nucleotides, about
5 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
10 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 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
15 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 15,000 nucleotides, about
1,600 nucleotides to about 14,500 nucleotides, about 1,600 nucleotides to about 14,000
nucleotides, about 1,600 nucleotides to about 13,500 nucleotides, about 1,600 nucleotides
to about 13,000 nucleotides, about 1,600 nucleotides to about 12,500 nucleotides, about
20 1,600 nucleotides to about 12,000 nucleotides, about 1,600 nucleotides to about 11,500
nucleotides, about 1,600 nucleotides to about 11,000 nucleotides, about 1,600 nucleotides
to about 10,500 nucleotides, about 1,600 nucleotides to about 10,000 nucleotides, about
1,600 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
25 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 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
30 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 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
5 nucleotides, about 1,600 nucleotides to about 4,000 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
10 1,600 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 15,000 nucleotides, about
1,800 nucleotides to about 14,500 nucleotides, about 1,800 nucleotides to about 14,000
nucleotides, about 1,800 nucleotides to about 13,500 nucleotides, about 1,800 nucleotides
15 to about 13,000 nucleotides, about 1,800 nucleotides to about 12,500 nucleotides, about
1,800 nucleotides to about 12,000 nucleotides, about 1,800 nucleotides to about 11,500
nucleotides, about 1,800 nucleotides to about 11,000 nucleotides, about 1,800 nucleotides
to about 10,500 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
20 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 nucleotides, about 1,800 nucleotides to about 6,800 nucleotides, about
25 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 nucleotides to about 5,400
nucleotides, about 1,800 nucleotides to about 5,200 nucleotides, about 1,800 nucleotides
30 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

98

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 nucleotides to about 2,200 nucleotides, about 2,200 nucleotides
to about 15,000 nucleotides, about 2,200 nucleotides to about 14,500 nucleotides, about
5 2,200 nucleotides to about 14,000 nucleotides, about 2,200 nucleotides to about 13,500
nucleotides, about 2,200 nucleotides to about 13,000 nucleotides, about 2,200 nucleotides
to about 12,500 nucleotides, about 2,200 nucleotides to about 12,000 nucleotides, about
2,200 nucleotides to about 11,500 nucleotides, about 2,200 nucleotides to about 11,000
nucleotides, about 2,200 nucleotides to about 10,500 nucleotides, about 2,200 nucleotides
10 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 nucleotides, about 6,200
nucleotides, about 6,000 nucleotides, about 5,800 nucleotides, about 5,600 nucleotides,
15 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
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
20 15,000 nucleotides, about 2,400 nucleotides to about 14,500 nucleotides, about 2,400
nucleotides to about 14,000 nucleotides, about 2,400 nucleotides to about 13,500
nucleotides, about 2,400 nucleotides to about 13,000 nucleotides, about 2,400 nucleotides
to about 12,500 nucleotides, about 2,400 nucleotides to about 12,000 nucleotides, about
2,400 nucleotides to about 11,500 nucleotides, about 2,400 nucleotides to about 11,000
25 nucleotides, about 2,400 nucleotides to about 10,500 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 nucleotides, about 2,400 nucleotides
to about 7,800 nucleotides, about 2,400 nucleotides to about 7,600 nucleotides, about
30 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 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
5 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
10 to about 3,800 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 15,000 nucleotides, about 2,600 nucleotides to about 14,500
15 nucleotides, about 2,600 nucleotides to about 14,000 nucleotides, about 2,600 nucleotides
to about 13,500 nucleotides, about 2,600 nucleotides to about 13,000 nucleotides, about
2,600 nucleotides to about 12,500 nucleotides, about 2,600 nucleotides to about 12,000
nucleotides, about 2,600 nucleotides to about 11,500 nucleotides, about 2,600 nucleotides
to about 11,000 nucleotides, about 2,600 nucleotides to about 10,500 nucleotides, about
20 2,600 nucleotides to about 10,000 nucleotides, about 2,600 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
25 to about 7,200 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 nucleotides to about 5,600
30 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 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
5 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 nucleotides to about 15,000
nucleotides, about 2,800 nucleotides to about 14,500 nucleotides, about 2,800 nucleotides
to about 14,000 nucleotides, about 2,800 nucleotides to about 13,500 nucleotides, about
10 2,800 nucleotides to about 13,000 nucleotides, about 2,800 nucleotides to about 12,500
nucleotides, about 2,800 nucleotides to about 12,000 nucleotides, about 2,800 nucleotides
to about 11,500 nucleotides, about 2,800 nucleotides to about 11,000 nucleotides, about
2,800 nucleotides to about 10,500 nucleotides, about 2,800 nucleotides to about 10,000
nucleotides, about 2,800 nucleotides to about 9,500 nucleotides, about 2,800 nucleotides
15 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
20 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 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
25 2,800 nucleotides to about 5,000 nucleotides, about 2,800 nucleotides to about 4,800
nucleotides, about 2,800 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
30 to about 3,400 nucleotides, about 2,800 nucleotides to about 3,200 nucleotides, about
2,800 nucleotides to about 3,000 nucleotides, about 3,000 nucleotides to about 15,000

102

to about 8,500 nucleotides, about 3,200 nucleotides to about 8,000 nucleotides, about
3,200 nucleotides to about 7,800 nucleotides, about 3,200 nucleotides to about 7,600
nucleotides, about 3,200 nucleotides to about 7,400 nucleotides, about 3,200 nucleotides
to about 7,200 nucleotides, about 3,200 nucleotides to about 7,000 nucleotides, about
5 3,200 nucleotides to about 6,800 nucleotides, about 3,200 nucleotides to about 6,600
nucleotides, about 3,200 nucleotides to about 6,400 nucleotides, about 3,200 nucleotides
to about 6,200 nucleotides, about 3,200 nucleotides to about 6,000 nucleotides, about
3,200 nucleotides to about 5,800 nucleotides, about 3,200 nucleotides to about 5,600
nucleotides, about 3,200 nucleotides to about 5,400 nucleotides, about 3,200 nucleotides
10 to about 5,200 nucleotides, about 3,200 nucleotides to about 5,000 nucleotides, about
3,200 nucleotides to about 4,800 nucleotides, about 3,200 nucleotides to about 4,600
nucleotides, about 3,200 nucleotides to about 4,400 nucleotides, about 3,200 nucleotides
to about 4,200 nucleotides, about 3,200 nucleotides to about 4,000 nucleotides, about
3,200 nucleotides to about 3,800 nucleotides, about 3,200 nucleotides to about 3,600
15 nucleotides, about 3,200 nucleotides to about 3,400 nucleotides, about 3,400 nucleotides
to about 15,000 nucleotides, about 3,400 nucleotides to about 14,500 nucleotides, about
3,400 nucleotides to about 14,000 nucleotides, about 3,400 nucleotides to about 13,500
nucleotides, about 3,400 nucleotides to about 13,000 nucleotides, about 3,400 nucleotides
to about 12,500 nucleotides, about 3,400 nucleotides to about 12,000 nucleotides, about
20 3,400 nucleotides to about 11,500 nucleotides, about 3,400 nucleotides to about 11,000
nucleotides, about 3,400 nucleotides to about 10,500 nucleotides, about 3,400 nucleotides
to about 10,000 nucleotides, about 3,400 nucleotides to about 9,500 nucleotides, about
3,400 nucleotides to about 9,000 nucleotides, about 3,400 nucleotides to about 8,500
nucleotides, about 3,400 nucleotides to about 8,000 nucleotides, about 3,400 nucleotides
25 to about 7,800 nucleotides, about 3,400 nucleotides to about 7,600 nucleotides, about
3,400 nucleotides to about 7,400 nucleotides, about 3,400 nucleotides to about 7,200
nucleotides, about 3,400 nucleotides to about 7,000 nucleotides, about 3,400 nucleotides
to about 6,800 nucleotides, about 3,400 nucleotides to about 6,600 nucleotides, about
3,400 nucleotides to about 6,400 nucleotides, about 3,400 nucleotides to about 6,200
30 nucleotides, about 3,400 nucleotides to about 6,000 nucleotides, about 3,400 nucleotides
to about 5,800 nucleotides, about 3,400 nucleotides to about 5,600 nucleotides, about

3,400 nucleotides to about 5,400 nucleotides, about 3,400 nucleotides to about 5,200
nucleotides, about 3,400 nucleotides to about 5,000 nucleotides, about 3,400 nucleotides
to about 4,800 nucleotides, about 3,400 nucleotides to about 4,600 nucleotides, about
3,400 nucleotides to about 4,400 nucleotides, about 3,400 nucleotides to about 4,200
5 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 15,000 nucleotides, about 3,600 nucleotides to about 14,500
nucleotides, about 3,600 nucleotides to about 14,000 nucleotides, about 3,600 nucleotides
to about 13,500 nucleotides, about 3,600 nucleotides to about 13,000 nucleotides, about
10 3,600 nucleotides to about 12,500 nucleotides, about 3,600 nucleotides to about 12,000
nucleotides, about 3,600 nucleotides to about 11,500 nucleotides, about 3,600 nucleotides
to about 11,000 nucleotides, about 3,600 nucleotides to about 10,500 nucleotides, about
3,600 nucleotides to about 10,000 nucleotides, about 3,600 nucleotides to about 9,500
nucleotides, about 3,600 nucleotides to about 9,000 nucleotides, about 3,600 nucleotides
15 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 nucleotides to about 7,000 nucleotides, about
3,600 nucleotides to about 6,800 nucleotides, about 3,600 nucleotides to about 6,600
20 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
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
25 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 nucleotides to about 4,000 nucleotides, about
3,600 nucleotides to about 3,800 nucleotides, about 3,800 nucleotides to about 15,000
nucleotides, about 3,800 nucleotides to about 14,500 nucleotides, about 3,800 nucleotides
30 to about 14,000 nucleotides, about 3,800 nucleotides to about 13,500 nucleotides, about
3,800 nucleotides to about 13,000 nucleotides, about 3,800 nucleotides to about 12,500

105

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 nucleotides, about 4,000 nucleotides
to about 5,200 nucleotides, about 4,000 nucleotides to about 5,000 nucleotides, about
5 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 15,000 nucleotides, about
4,200 nucleotides to about 14,500 nucleotides, about 4,200 nucleotides to about 14,000
nucleotides, about 4,200 nucleotides to about 13,500 nucleotides, about 4,200 nucleotides
10 to about 13,000 nucleotides, about 4,200 nucleotides to about 12,500 nucleotides, about
4,200 nucleotides to about 12,000 nucleotides, about 4,200 nucleotides to about 11,500
nucleotides, about 4,200 nucleotides to about 11,000 nucleotides, about 4,200 nucleotides
to about 10,500 nucleotides, about 4,200 nucleotides to about 10,000 nucleotides, about
4,200 nucleotides to about 9,500 nucleotides, about 4,200 nucleotides to about 9,000
15 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 nucleotides, about 4,200 nucleotides
to about 7,000 nucleotides, about 4,200 nucleotides to about 6,800 nucleotides, about
20 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 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
25 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 15,000 nucleotides, about 4,400 nucleotides
to about 14,500 nucleotides, about 4,400 nucleotides to about 14,000 nucleotides, about
4,400 nucleotides to about 13,500 nucleotides, about 4,400 nucleotides to about 13,000
30 nucleotides, about 4,400 nucleotides to about 12,500 nucleotides, about 4,400 nucleotides
to about 12,000 nucleotides, about 4,400 nucleotides to about 11,500 nucleotides, about

107

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 15,000 nucleotides, about 4,800 nucleotides to about 14,500 nucleotides, about 4,800 nucleotides to about 14,000 nucleotides, about 4,800 nucleotides to about 13,500 nucleotides, about 4,800 nucleotides to about 13,000 nucleotides, about 4,800 nucleotides to about 12,500 nucleotides, about 4,800 nucleotides to about 12,000 nucleotides, about 4,800 nucleotides to about 11,500 nucleotides, about 4,800 nucleotides to about 11,000 nucleotides, about 4,800 nucleotides to about 10,500 nucleotides, about 4,800 nucleotides to about 10,000 nucleotides, about 4,800 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 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 15,000 nucleotides, about 5,000 nucleotides to about 14,500 nucleotides, about 5,000 nucleotides to about 14,000 nucleotides, about 5,000 nucleotides to about 13,500 nucleotides, about 5,000 nucleotides to about 13,000 nucleotides, about 5,000 nucleotides to about 12,500 nucleotides, about 5,000 nucleotides to about 12,000 nucleotides, about 5,000 nucleotides to about 11,500 nucleotides, about 5,000 nucleotides to about 11,000 nucleotides, about 5,000 nucleotides to about 10,500 nucleotides, about 5,000 nucleotides to about 10,000 nucleotides, about 5,000 nucleotides to about 9,500 nucleotides, about 5,000 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 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 5,000 nucleotides to about 5,800 nucleotides, about 5,000 nucleotides to about 5,600
nucleotides, about 5,000 nucleotides to about 5,400 nucleotides, about 5,000 nucleotides
to about 5,200 nucleotides, about 5,200 nucleotides to about 15,000 nucleotides, about
5,200 nucleotides to about 14,500 nucleotides, about 5,200 nucleotides to about 14,000
nucleotides, about 5,200 nucleotides to about 13,500 nucleotides, about 5,200 nucleotides
10 to about 13,000 nucleotides, about 5,200 nucleotides to about 12,500 nucleotides, about
5,200 nucleotides to about 12,000 nucleotides, about 5,200 nucleotides to about 11,500
nucleotides, about 5,200 nucleotides to about 11,000 nucleotides, about 5,200 nucleotides
to about 10,500 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
15 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 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
20 5,200 nucleotides to about 6,600 nucleotides, about 5,200 nucleotides to about 6,400
nucleotides, about 5,200 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 15,000 nucleotides, about 5,400 nucleotides
25 to about 14,500 nucleotides, about 5,400 nucleotides to about 14,000 nucleotides, about
5,400 nucleotides to about 13,500 nucleotides, about 5,400 nucleotides to about 13,000
nucleotides, about 5,400 nucleotides to about 12,500 nucleotides, about 5,400 nucleotides
to about 12,000 nucleotides, about 5,400 nucleotides to about 11,500 nucleotides, about
5,400 nucleotides to about 11,000 nucleotides, about 5,400 nucleotides to about 10,500
30 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

110

111

to about 7,800 nucleotides, about 6,200 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
5 6,200 nucleotides to about 6,400 nucleotides, about 6,400 nucleotides to about 15,000
nucleotides, about 6,400 nucleotides to about 14,500 nucleotides, about 6,400 nucleotides
to about 14,000 nucleotides, about 6,400 nucleotides to about 13,500 nucleotides, about
6,400 nucleotides to about 13,000 nucleotides, about 6,400 nucleotides to about 12,500
nucleotides, about 6,400 nucleotides to about 12,000 nucleotides, about 6,400 nucleotides
10 to about 11,500 nucleotides, about 6,400 nucleotides to about 11,000 nucleotides, about
6,400 nucleotides to about 10,500 nucleotides, about 6,400 nucleotides to about 10,000
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
15 nucleotides, about 6,400 nucleotides to about 7,600 nucleotides, about 6,400 nucleotides
to about 7,400 nucleotides, about 6,400 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 15,000 nucleotides, about 6,600 nucleotides to about 14,500 nucleotides, about
20 6,600 nucleotides to about 14,000 nucleotides, about 6,600 nucleotides to about 13,500
nucleotides, about 6,600 nucleotides to about 13,000 nucleotides, about 6,600 nucleotides
to about 12,500 nucleotides, about 6,600 nucleotides to about 12,000 nucleotides, about
6,600 nucleotides to about 11,500 nucleotides, about 6,600 nucleotides to about 11,000
nucleotides, about 6,600 nucleotides to about 10,500 nucleotides, about 6,600 nucleotides
25 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
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
30 nucleotides, about 6,600 nucleotides to about 7,000 nucleotides, about 6,600 nucleotides
to about 6,800 nucleotides, about 6,800 nucleotides to about 15,000 nucleotides, about

113

114

to about 8,000 nucleotides, about 8,000 nucleotides to about 15,000 nucleotides, about
8,000 nucleotides to about 14,500 nucleotides, about 8,000 nucleotides to about 14,000
nucleotides, about 8,000 nucleotides to about 13,500 nucleotides, about 8,000 nucleotides
to about 13,000 nucleotides, about 8,000 nucleotides to about 12,500 nucleotides, about
5 8,000 nucleotides to about 12,000 nucleotides, about 8,000 nucleotides to about 11,500
nucleotides, about 8,000 nucleotides to about 11,000 nucleotides, about 8,000 nucleotides
to about 10,500 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
10 to about 15,000 nucleotides, about 8,500 nucleotides to about 14,500 nucleotides, about
8,500 nucleotides to about 14,000 nucleotides, about 8,500 nucleotides to about 13,500
nucleotides, about 8,500 nucleotides to about 13,000 nucleotides, about 8,500 nucleotides
to about 12,500 nucleotides, about 8,500 nucleotides to about 12,000 nucleotides, about
8,500 nucleotides to about 11,500 nucleotides, about 8,500 nucleotides to about 11,000
15 nucleotides, about 8,500 nucleotides to about 10,500 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 15,000
nucleotides, about 9,000 nucleotides to about 14,500 nucleotides, about 9,000 nucleotides
to about 14,000 nucleotides, about 9,000 nucleotides to about 13,500 nucleotides, about
20 9,000 nucleotides to about 13,000 nucleotides, about 9,000 nucleotides to about 12,500
nucleotides, about 9,000 nucleotides to about 12,000 nucleotides, about 9,000 nucleotides
to about 11,500 nucleotides, about 9,000 nucleotides to about 11,000 nucleotides, about
9,000 nucleotides to about 10,500 nucleotides, about 9,000 nucleotides to about 10,000
nucleotides, about 9,000 nucleotides to about 9,500 nucleotides, about 9,500 nucleotides
25 to about 10,000 nucleotides, about 10,000 nucleotides to about 15,000 nucleotides, about
10,000 nucleotides to about 14,500 nucleotides, about 10,000 nucleotides to about 14,000
nucleotides, about 10,000 nucleotides to about 13,500 nucleotides, about 10,000
nucleotides to about 13,000 nucleotides, about 10,000 nucleotides to about 12,500
nucleotides, about 10,000 nucleotides to about 12,000 nucleotides, about 10,000
30 nucleotides to about 11,500 nucleotides, about 10,000 nucleotides to about 11,000
nucleotides, about 10,000 nucleotides to about 10,500 nucleotides, about 10,500

nucleotides to about 15,000 nucleotides, about 10,500 nucleotides to about 14,500 nucleotides, about 10,500 nucleotides to about 14,000 nucleotides, about 10,500 nucleotides to about 13,500 nucleotides, about 10,500 nucleotides to about 13,000 nucleotides, about 10,500 nucleotides to about 12,500 nucleotides, about 10,500 nucleotides to about 12,000 nucleotides, about 10,500 nucleotides to about 11,500 nucleotides, about 10,500 nucleotides to about 11,000 nucleotides, about 11,000 nucleotides to about 15,000 nucleotides, about 11,000 nucleotides to about 14,500 nucleotides, about 11,000 nucleotides to about 14,000 nucleotides, about 11,000 nucleotides to about 13,500 nucleotides, about 11,000 nucleotides to about 13,000 nucleotides, about 11,000 nucleotides to about 12,500 nucleotides, about 11,000 nucleotides to about 12,000 nucleotides, about 11,000 nucleotides to about 11,500 nucleotides, about 11,500 nucleotides to about 15,000 nucleotides, about 11,500 nucleotides to about 14,500 nucleotides, about 11,500 nucleotides to about 14,000 nucleotides, about 11,500 nucleotides to about 13,500 nucleotides, about 11,500 nucleotides to about 13,000 nucleotides, about 11,500 nucleotides to about 12,500 nucleotides, about 12,000 nucleotides to about 15,000 nucleotides, about 12,000 nucleotides to about 14,500 nucleotides, about 12,000 nucleotides to about 14,000 nucleotides, about 12,000 nucleotides to about 13,500 nucleotides, about 12,000 nucleotides to about 13,000 nucleotides, about 12,000 nucleotides to about 12,500 nucleotides, about 12,500 nucleotides to about 15,000 nucleotides, about 12,500 nucleotides to about 14,500 nucleotides, about 12,500 nucleotides to about 14,000 nucleotides, about 12,500 nucleotides to about 13,500 nucleotides, about 12,500 nucleotides to about 13,000 nucleotides, about 13,000 nucleotides to about 15,000 nucleotides, about 13,000 nucleotides to about 14,500 nucleotides, about 13,000 nucleotides to about 14,000 nucleotides, about 13,000 nucleotides to about 13,500 nucleotides, about 13,500 nucleotides to about 15,000 nucleotides, about 13,500 nucleotides to about 14,500 nucleotides, about 13,500 nucleotides to about 14,000 nucleotides, about 14,000 nucleotides to about 15,000 nucleotides, about 14,000 nucleotides to about 14,500 nucleotides, or about 14,500 nucleotides to about 15,000 nucleotides (inclusive).

Provided herein are exemplary vectors that can be used in any of the compositions and methods described herein. See, e.g., FIGs. 1-24.

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 inner hair cell, a cochlear
5 outer hair cell, a retinal 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
10 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
15 integrated into an endogenous gene locus (e.g., a CLRN1 gene locus). In some embodiments, the vectors provided herein stably integrate into an endogenous defective CLRN1 gene locus, and thereby replace the defective CLRN1 gene with a nucleic acid encoding a functioning (e.g., wildtype) CLRN1 protein.

Various molecular biology techniques that can be used to introduce a mutation(s)
20 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.

25 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) signal, and a Kozak consensus sequence. Non-limiting examples of these control sequences are described herein. In some
30 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 CLRN1 gene). A promoter typically refers to, e.g., a nucleotide sequence to which an RNA
 5 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 CLRN1 protein (e.g., a human CLRN1 protein) can include a promoter and/or an enhancer. The
 10 vector encoding the N-terminal portion of the CLRN1 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
 15 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 an inner hair cell. In some examples, the
 20 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 -
 25 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
 30 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

5 (SAA), troponin I (TN I), duchenne muscular dystrophy, human immunodeficiency virus, and Gibbon Ape Leukemia 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 CAG promoter or a CAG/CBA

10 promoter. In some embodiments, the promoter is a CBA promoter, e.g., a CBA promoter comprising or consisting of SEQ ID NO: 18.

The term “constitutive” promoter refers to a nucleotide sequence that, when operably linked with a nucleic acid encoding a protein (e.g., a CLRN1 protein), causes RNA to be transcribed from the nucleic acid in a mammalian cell under most or all

15 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 inducible systems

25 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 dexamethasone

30 (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).

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).

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 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., *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, 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).

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 cochlear 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.

Enhancers

In some instances, a vector can include 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 CLRN1 protein). Enhancer sequences (50-1500 basepairs in length) generally 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 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: 17.

Poly(A) Signal

In some embodiments, any of the vectors provided herein can include a polyadenylation (poly(A)) signal 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 driven by the poly(A) signal sequence (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, 1994). In some embodiments, the poly(A) signal sequence is positioned 3' to the nucleic acid sequence encoding the C-terminus of the CLRN1 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 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 (or poly(A)) signal. 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 also can 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) signal sequence” or “polyadenylation signal sequence” is a sequence that triggers the endonuclease cleavage of an mRNA and the addition of a series of adenosines to the 3' end of the cleaved mRNA.

There are several poly(A) signal 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).

The poly(A) signal sequence can be AATAAA. The AATAAA sequence may be substituted with other hexanucleotide sequences with homology to AATAAA and that

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).

- 5 In some embodiments, the poly(A) signal sequence can be a synthetic polyadenylation site (see, e.g., the pCl-neo expression vector of Promega that is based on Levitt et al, *Genes Dev.* 3(7):1019-1025, 1989). In some embodiments, the poly(A) signal sequence is the polyadenylation signal of bovine growth hormone (CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCC
- 10 TTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGGAAAT TGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATG CGGTGGGCTCTATGG (SEQ ID NO: 20)). In some embodiments, the poly(A) signal sequence is the polyadenylation signal of soluble neuropilin-1 (sNRP)
- 15 (AAATAAAATACGAAATG (SEQ ID NO: 21)) (see, e.g., WO 05/073384). Additional examples of poly(A) signal sequences are known in the art.

Internal Ribosome Entry Site (IRES)

- In some embodiments, a vector encoding the C-terminal portion of the CLRN1
- 20 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 IRES is located (see, e.g., Pelletier and Sonenberg, *Mol. Cell. Biol.* 8(3):1103-1112, 1988).
- 25 There are several IRES sequences known to those skilled in the art, including those from, e.g., foot and mouth disease virus (FMDV), encephalomyocarditis virus (EMCV), human rhinovirus (HRV), cricket paralysis 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.,
- 30 *Genes Dev.* 15(13):1593-612, 2001.

In some embodiments, the IRES sequence that is incorporated into the vector that encodes the C-terminal portion of a CLRN1 protein is the foot and mouth disease virus (FMDV) 2A sequence. 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 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 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 tGFP (SEQ ID NO: 19). 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 hair cell, an ocular cell, such as a retinal cell) is detected by assays for beta-galactosidase activity. When 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 hair cell, an ocular cell, such as a retinal cell) may be measured by fluorescent techniques (e.g., fluorescent microscopy or FACS) or light production in a
 5 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.

10 **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, such as 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
 15 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
 20 expression of a CLRN1 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
 25 bases upstream of the start codon (AUG), and the start codon is followed by another "G". The 5' UTRs have also been known to form secondary structures that are involved in elongation factor binding.

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:
 30 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.

In some embodiments, a 5' UTR from a mRNA that is transcribed by a cell in the cochlea or retina 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 (in the RNA form) or thymidines (in the DNA form) 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 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-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 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: 12 or SEQ ID NO: 13. An exemplary human wildtype 5' UTR is or includes the sequence of SEQ ID NO: 14 or SEQ ID NO: 15.

In some embodiments of any of the compositions described herein, a 5' untranslated region (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 3'-UTR's 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, at least 70 contiguous, at least 75 contiguous, at least 80 contiguous, at least 85 contiguous, at least 90 contiguous, at least 100 contiguous, at least 105 contiguous, at least 110 contiguous, at least 115 contiguous, at least 120 contiguous, at least 125 contiguous, at least 130 contiguous, at least 135 contiguous, at least 140 contiguous, at least 145 contiguous, at least 150 contiguous, at least 155 contiguous, at least 160 contiguous, at least 165 contiguous, at least 170 contiguous, at least 175 contiguous, at least 180 contiguous, at least 185 contiguous, at least 190 contiguous, at least 195 contiguous, at least 200 contiguous, at least 205 contiguous, at least 210 contiguous, at least 215 contiguous, at least 220 contiguous, at least 225 contiguous, at least 230 contiguous, at least 235 contiguous, at least 240 contiguous, at least 245 contiguous, at least 250 contiguous, at least 255 contiguous, or at least 260 contiguous) nucleotides from anywhere within SEQ ID NO: 12 or SEQ ID NO: 13.

For example, a 5' UTR can include or consist of one or more of: nucleotide positions 1 to 291, nucleotide positions 1 to 290, nucleotide positions 1 to 280, nucleotide positions 1 to 270, nucleotide positions 1 to 260, nucleotide positions 1 to 250, nucleotide positions 1 to 240, nucleotide positions 1 to 230, nucleotide positions 1 to 220, nucleotide positions 1 to 210, nucleotide positions 1 to 200, nucleotide positions 1 to 190, nucleotide positions 1 to 180, nucleotide positions 1 to 170, nucleotide positions 1 to 160, nucleotide positions 1 to 150, nucleotide positions 1 to 140, nucleotide positions 1 to 130, nucleotide positions 1 to 120, nucleotide positions 1 to 110, 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 291, nucleotide positions 10 to 290, nucleotide positions 10 to 280, nucleotide positions 10 to 270, nucleotide positions 10 to 260, nucleotide positions 10 to 250, nucleotide positions 10 to 240, nucleotide positions 10 to 230, nucleotide positions 10 to 220, nucleotide positions 10 to 210, nucleotide positions 10 to 200, nucleotide positions 10 to 190, nucleotide positions 10 to 180, nucleotide positions 10 to 170,

nucleotide positions 10 to 160, nucleotide positions 10 to 150, nucleotide positions 10 to 140, nucleotide positions 10 to 130, nucleotide positions 10 to 120, nucleotide positions 10 to 110, 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 291, nucleotide positions 20 to 290, nucleotide positions 20 to 280, nucleotide positions 20 to 270, nucleotide positions 20 to 260, nucleotide positions 20 to 250, nucleotide positions 20 to 240, nucleotide positions 20 to 230, nucleotide positions 20 to 220, nucleotide positions 20 to 210, nucleotide positions 20 to 200, nucleotide positions 20 to 190, nucleotide positions 20 to 180, nucleotide positions 20 to 170, nucleotide positions 20 to 160, nucleotide positions 20 to 150, nucleotide positions 20 to 140, nucleotide positions 20 to 130, nucleotide positions 20 to 120, nucleotide positions 20 to 110, 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 291, nucleotide positions 30 to 290, nucleotide positions 30 to 280, nucleotide positions 30 to 270, nucleotide positions 30 to 260, nucleotide positions 30 to 250, nucleotide positions 30 to 240, nucleotide positions 30 to 230, nucleotide positions 30 to 220, nucleotide positions 30 to 210, nucleotide positions 30 to 200, nucleotide positions 30 to 190, nucleotide positions 30 to 180, nucleotide positions 30 to 170, nucleotide positions 30 to 160, nucleotide positions 30 to 150, nucleotide positions 30 to 140, nucleotide positions 30 to 130, nucleotide positions 30 to 120, nucleotide positions 30 to 110, 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 291, nucleotide positions 40 to 290, nucleotide positions 40 to 280, nucleotide positions 40 to 270, nucleotide positions 40 to 260, nucleotide positions 40 to 250, nucleotide positions 40 to 240, nucleotide positions 40 to 230, nucleotide positions 40 to 220, nucleotide positions 40 to 210, nucleotide positions 40 to 200, nucleotide positions 40 to 190, nucleotide positions 40 to 180, nucleotide positions 40 to 170, nucleotide positions 40 to 160, nucleotide positions 40 to 150, nucleotide positions 40 to

140, nucleotide positions 40 to 130, nucleotide positions 40 to 120, nucleotide positions 40 to 110, 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 291, nucleotide positions 50 to 290,

5 nucleotide positions 50 to 280, nucleotide positions 50 to 270, nucleotide positions 50 to 260, nucleotide positions 50 to 250, nucleotide positions 50 to 240, nucleotide positions 50 to 230, nucleotide positions 50 to 220, nucleotide positions 50 to 210, nucleotide positions 50 to 200, nucleotide positions 50 to 190, nucleotide positions 50 to 180, nucleotide positions 50 to 170, nucleotide positions 50 to 160, nucleotide positions 50 to

10 150, nucleotide positions 50 to 140, nucleotide positions 50 to 130, nucleotide positions 50 to 120, nucleotide positions 50 to 110, 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 291, nucleotide positions 60 to 290, nucleotide positions 60 to 280, nucleotide positions 60 to 270, nucleotide positions 60 to

15 260, nucleotide positions 60 to 250, nucleotide positions 60 to 240, nucleotide positions 60 to 230, nucleotide positions 60 to 220, nucleotide positions 60 to 210, nucleotide positions 60 to 200, nucleotide positions 60 to 190, nucleotide positions 60 to 180, nucleotide positions 60 to 170, nucleotide positions 60 to 160, nucleotide positions 60 to 150, nucleotide positions 60 to 140, nucleotide positions 60 to 130, nucleotide positions

20 60 to 120, nucleotide positions 60 to 110, nucleotide positions 60 to 100, nucleotide positions 60 to 90, nucleotide positions 60 to 80, nucleotide positions 60 to 70, nucleotide positions 70 to 291, nucleotide positions 70 to 290, nucleotide positions 70 to 280, nucleotide positions 70 to 270, nucleotide positions 70 to 260, nucleotide positions 70 to 250, nucleotide positions 70 to 240, nucleotide positions 70 to 230, nucleotide positions

25 70 to 220, nucleotide positions 70 to 210, nucleotide positions 70 to 200, nucleotide positions 70 to 190, nucleotide positions 70 to 180, nucleotide positions 70 to 170, nucleotide positions 70 to 160, nucleotide positions 70 to 150, nucleotide positions 70 to 140, nucleotide positions 70 to 130, nucleotide positions 70 to 120, nucleotide positions 70 to 110, nucleotide positions 70 to 100, nucleotide positions 70 to 90, nucleotide

30 positions 70 to 80, nucleotide positions 80 to 291, nucleotide positions 80 to 290, nucleotide positions 80 to 280, nucleotide positions 80 to 270, nucleotide positions 80 to

260, nucleotide positions 80 to 250, nucleotide positions 80 to 240, nucleotide positions
 80 to 230, nucleotide positions 80 to 220, nucleotide positions 80 to 210, nucleotide
 positions 80 to 200, nucleotide positions 80 to 190, nucleotide positions 80 to 180,
 nucleotide positions 80 to 170, nucleotide positions 80 to 160, nucleotide positions 80 to
 5 150, nucleotide positions 80 to 140, nucleotide positions 80 to 130, nucleotide positions
 80 to 120, nucleotide positions 80 to 110, nucleotide positions 80 to 100, nucleotide
 positions 80 to 90, nucleotide positions 90 to 291, nucleotide positions 90 to 290,
 nucleotide positions 90 to 280, nucleotide positions 90 to 270, nucleotide positions 90 to
 260, nucleotide positions 90 to 250, nucleotide positions 90 to 240, nucleotide positions
 10 90 to 230, nucleotide positions 90 to 220, nucleotide positions 90 to 210, nucleotide
 positions 90 to 200, nucleotide positions 90 to 190, nucleotide positions 90 to 180,
 nucleotide positions 90 to 170, nucleotide positions 90 to 160, nucleotide positions 90 to
 150, nucleotide positions 90 to 140, nucleotide positions 90 to 130, nucleotide positions
 90 to 120, nucleotide positions 90 to 110, nucleotide positions 90 to 100, nucleotide
 15 positions 100 to 291, nucleotide positions 100 to 290, nucleotide positions 100 to 280,
 nucleotide positions 100 to 270, nucleotide positions 100 to 260, nucleotide positions 100
 to 250, nucleotide positions 100 to 240, nucleotide positions 100 to 230, nucleotide
 positions 100 to 220, nucleotide positions 100 to 210, nucleotide positions 100 to 200,
 nucleotide positions 100 to 190, nucleotide positions 100 to 180, nucleotide positions 100
 20 to 170, nucleotide positions 100 to 160, nucleotide positions 100 to 150, nucleotide
 positions 100 to 140, nucleotide positions 100 to 130, nucleotide positions 100 to 120,
 nucleotide positions 100 to 110, nucleotide positions 110 to 291, nucleotide positions 110
 to 290, nucleotide positions 110 to 280, nucleotide positions 110 to 270, nucleotide
 positions 110 to 260, nucleotide positions 110 to 250, nucleotide positions 110 to 240,
 25 nucleotide positions 110 to 230, nucleotide positions 110 to 220, nucleotide positions 110
 to 210, nucleotide positions 110 to 200, nucleotide positions 110 to 190, nucleotide
 positions 110 to 180, nucleotide positions 110 to 170, nucleotide positions 110 to 160,
 nucleotide positions 110 to 150, nucleotide positions 110 to 140, nucleotide positions 110
 to 130, nucleotide positions 110 to 120, nucleotide positions 120 to 291, nucleotide
 30 positions 120 to 290, nucleotide positions 120 to 280, nucleotide positions 120 to 270,
 nucleotide positions 120 to 260, nucleotide positions 120 to 250, nucleotide positions 120

to 240, nucleotide positions 120 to 230, nucleotide positions 120 to 220, nucleotide positions 120 to 210, nucleotide positions 120 to 200, nucleotide positions 120 to 190, nucleotide positions 120 to 180, nucleotide positions 120 to 170, nucleotide positions 120 to 160, nucleotide positions 120 to 150, nucleotide positions 120 to 140, nucleotide positions 120 to 130, nucleotide positions 130 to 291, nucleotide positions 130 to 290, nucleotide positions 130 to 280, nucleotide positions 130 to 270, nucleotide positions 130 to 260, nucleotide positions 130 to 250, nucleotide positions 130 to 240, nucleotide positions 130 to 230, nucleotide positions 130 to 220, nucleotide positions 130 to 210, nucleotide positions 130 to 200, nucleotide positions 130 to 190, nucleotide positions 130 to 180, nucleotide positions 130 to 170, nucleotide positions 130 to 160, nucleotide positions 130 to 150, nucleotide positions 130 to 140, nucleotide positions 140 to 291, nucleotide positions 140 to 290, nucleotide positions 140 to 280, nucleotide positions 140 to 270, nucleotide positions 140 to 260, nucleotide positions 140 to 250, nucleotide positions 140 to 240, nucleotide positions 140 to 230, nucleotide positions 140 to 220, nucleotide positions 140 to 210, nucleotide positions 140 to 200, nucleotide positions 140 to 190, nucleotide positions 140 to 180, nucleotide positions 140 to 170, nucleotide positions 140 to 160, nucleotide positions 140 to 150, nucleotide positions 150 to 291, nucleotide positions 150 to 290, nucleotide positions 150 to 280, nucleotide positions 150 to 270, nucleotide positions 150 to 260, nucleotide positions 150 to 250, nucleotide positions 150 to 240, nucleotide positions 150 to 230, nucleotide positions 150 to 220, nucleotide positions 150 to 210, nucleotide positions 150 to 200, nucleotide positions 150 to 190, nucleotide positions 150 to 180, nucleotide positions 150 to 170, nucleotide positions 150 to 160, nucleotide positions 160 to 291, nucleotide positions 160 to 290, nucleotide positions 160 to 280, nucleotide positions 160 to 270, nucleotide positions 160 to 260, nucleotide positions 160 to 250, nucleotide positions 160 to 240, nucleotide positions 160 to 230, nucleotide positions 160 to 220, nucleotide positions 160 to 210, nucleotide positions 160 to 200, nucleotide positions 160 to 190, nucleotide positions 160 to 180, nucleotide positions 160 to 170, nucleotide positions 170 to 291, nucleotide positions 170 to 290, nucleotide positions 170 to 280, nucleotide positions 170 to 270, nucleotide positions 170 to 260, nucleotide positions 170 to 250, nucleotide positions 170 to 240, nucleotide positions 170 to 230, nucleotide positions 170 to 220, nucleotide

positions 170 to 210, nucleotide positions 170 to 200, nucleotide positions 170 to 190, nucleotide positions 170 to 180, nucleotide positions 180 to 291, nucleotide positions 180 to 290, nucleotide positions 180 to 280, nucleotide positions 180 to 270, nucleotide positions 180 to 260, nucleotide positions 180 to 250, nucleotide positions 180 to 240, nucleotide positions 180 to 230, nucleotide positions 180 to 220, nucleotide positions 180 to 210, nucleotide positions 180 to 200, nucleotide positions 180 to 190, nucleotide positions 190 to 291, nucleotide positions 190 to 290, nucleotide positions 190 to 280, nucleotide positions 190 to 270, nucleotide positions 190 to 260, nucleotide positions 190 to 250, nucleotide positions 190 to 240, nucleotide positions 190 to 230, nucleotide positions 190 to 220, nucleotide positions 190 to 210, nucleotide positions 190 to 200, nucleotide positions 200 to 291, nucleotide positions 200 to 290, nucleotide positions 200 to 280, nucleotide positions 200 to 270, nucleotide positions 200 to 260, nucleotide positions 200 to 250, nucleotide positions 200 to 240, nucleotide positions 200 to 230, nucleotide positions 200 to 220, nucleotide positions 200 to 210, nucleotide positions 210 to 291, nucleotide positions 210 to 290, nucleotide positions 210 to 280, nucleotide positions 210 to 270, nucleotide positions 210 to 260, nucleotide positions 210 to 250, nucleotide positions 210 to 240, nucleotide positions 210 to 230, nucleotide positions 210 to 220, nucleotide positions 220 to 291, nucleotide positions 220 to 290, nucleotide positions 220 to 280, nucleotide positions 220 to 270, nucleotide positions 220 to 260, nucleotide positions 220 to 250, nucleotide positions 220 to 240, nucleotide positions 220 to 230, nucleotide positions 230 to 291, nucleotide positions 230 to 290, nucleotide positions 230 to 280, nucleotide positions 230 to 270, nucleotide positions 230 to 260, nucleotide positions 230 to 250, nucleotide positions 230 to 240, nucleotide positions 240 to 291, nucleotide positions 240 to 290, nucleotide positions 240 to 280, nucleotide positions 240 to 270, nucleotide positions 240 to 260, nucleotide positions 240 to 250, nucleotide positions 250 to 291, nucleotide positions 250 to 290, nucleotide positions 250 to 280, nucleotide positions 250 to 270, nucleotide positions 250 to 260, nucleotide positions 260 to 291, nucleotide positions 260 to 290, nucleotide positions 260 to 280, nucleotide positions 260 to 270, nucleotide positions 270 to 291, nucleotide positions 270 to 290, nucleotide positions 270 to 280, nucleotide positions 280 to 291, or nucleotide positions 280 to 290, of SEQ ID NO: 12 or 13.

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 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%, or at least 99%) identical to SEQ ID NO: 12 or 13.

5 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 100 contiguous, at least 105 contiguous, at least 110 contiguous, at least 115 contiguous, at least 120 contiguous, at least 125 contiguous, at least 130 contiguous, at least 135 contiguous, at least 140 contiguous, at least 145 contiguous, at least 150 contiguous, at least 155 contiguous, at least 160 contiguous, at least 165 contiguous, at least 170 contiguous, at least 175 contiguous, at least 180 contiguous, at least 185 contiguous, at least 190 contiguous, at least 195 contiguous, at least 200 contiguous, at least 205 contiguous, at least 210 contiguous, at least 215 contiguous, at least 220 contiguous, at least 225 contiguous, at least 230 contiguous, at least 235 contiguous, at least 240 contiguous, at least 245 contiguous, at least 250 contiguous, at least 255 contiguous, at least 260 contiguous, at least 265 contiguous, at least 270 contiguous, at least 275 contiguous, at least 280 contiguous, at least 285 contiguous, at least 290 contiguous, at least 295 contiguous, at least 300 contiguous, at least 305 contiguous, at least 310 contiguous, at least 315 contiguous, at least 320 contiguous, at least 325 contiguous, at least 330 contiguous, at least 335 contiguous, at least 340 contiguous, at least 345 contiguous, at least 350 contiguous, at least 355 contiguous, at least 360 contiguous, at least 365 contiguous, at least 370 contiguous, at least 375 contiguous, at least 380 contiguous, at least 385 contiguous, at least 390 contiguous, at least 395 contiguous, at least 400 contiguous, at least 450 contiguous, at least 500 contiguous, at least 550 contiguous, at least 600 contiguous, at least 650 contiguous, at least 700 contiguous, at least 750 contiguous, at least 800 contiguous, at least 850 contiguous, at least 900 contiguous, at least 950 contiguous, at least 1000 contiguous, at least 1050 contiguous, at least 1100

contiguous, at least 1150 contiguous, at least 1200 contiguous, at least 1250 contiguous, at least 1300 contiguous, at least 1350 contiguous, at least 1400 contiguous, at least 1450 contiguous, at least 1500 contiguous, at least 1550 contiguous, at least 1600 contiguous, at least 1650 contiguous, at least 1700 contiguous, or at least 1750 contiguous)

5 nucleotides from anywhere within SEQ ID NO: 14 or 15.

For example, a 5' UTR can include or consist of one or more of: nucleotide positions 1 to 1773, nucleotide positions 1 to 1770, nucleotide positions 1 to 1750, nucleotide positions 1 to 1700, nucleotide positions 1 to 1650, nucleotide positions 1 to 1600, nucleotide positions 1 to 1550, nucleotide positions 1 to 1500, nucleotide positions 10 1 to 1450, nucleotide positions 1 to 1400, nucleotide positions 1 to 1350, nucleotide positions 1 to 1300, nucleotide positions 1 to 1250, nucleotide positions 1 to 1200, nucleotide positions 1 to 1150, nucleotide positions 1 to 1100, nucleotide positions 1 to 1050, nucleotide positions 1 to 1000, nucleotide positions 1 to 950, nucleotide positions 1 to 900, nucleotide positions 1 to 850, nucleotide positions 1 to 800, nucleotide positions 1 15 to 750, nucleotide positions 1 to 700, nucleotide positions 1 to 650, nucleotide positions 1 to 600, nucleotide positions 1 to 550, nucleotide positions 1 to 500, nucleotide positions 1 to 450, nucleotide positions 1 to 400, nucleotide positions 1 to 350, nucleotide positions 1 to 300, nucleotide positions 1 to 250, nucleotide positions 1 to 200, nucleotide positions 1 to 150, nucleotide positions 1 to 100, nucleotide positions 1 to 50, nucleotide position 1 20 to 25, nucleotide positions 25 to 1773, nucleotide positions 25 to 1770, nucleotide positions 25 to 1750, nucleotide positions 25 to 1700, nucleotide positions 25 to 1650, nucleotide positions 25 to 1600, nucleotide positions 25 to 1550, nucleotide positions 25 to 1500, nucleotide positions 25 to 1450, nucleotide positions 25 to 1400, nucleotide positions 25 to 1350, nucleotide positions 25 to 1300, nucleotide positions 25 to 1250, 25 nucleotide positions 25 to 1200, nucleotide positions 25 to 1150, nucleotide positions 25 to 1100, nucleotide positions 25 to 1050, nucleotide positions 25 to 1000, nucleotide positions 25 to 950, nucleotide positions 25 to 900, nucleotide positions 25 to 850, nucleotide positions 25 to 800, nucleotide positions 25 to 750, nucleotide positions 25 to 700, nucleotide positions 25 to 650, nucleotide positions 25 to 600, nucleotide positions 30 25 to 550, nucleotide positions 25 to 500, nucleotide positions 25 to 450, nucleotide positions 25 to 400, nucleotide positions 25 to 350, nucleotide positions 25 to 300,

nucleotide positions 25 to 250, nucleotide positions 25 to 200, nucleotide positions 25 to 150, nucleotide positions 25 to 100, nucleotide positions 25 to 50, nucleotide positions 50 to 1773, nucleotide positions 50 to 1770, nucleotide positions 50 to 1750, nucleotide positions 50 to 1700, nucleotide positions 50 to 1650, nucleotide positions 50 to 1600, nucleotide positions 50 to 1550, nucleotide positions 50 to 1500, nucleotide positions 50 to 1450, nucleotide positions 50 to 1400, nucleotide positions 50 to 1350, nucleotide positions 50 to 1300, nucleotide positions 50 to 1250, nucleotide positions 50 to 1200, nucleotide positions 50 to 1150, nucleotide positions 50 to 1100, nucleotide positions 50 to 1050, nucleotide positions 50 to 1000, nucleotide positions 50 to 950, nucleotide positions 50 to 900, nucleotide positions 50 to 850, nucleotide positions 50 to 800, nucleotide positions 50 to 750, nucleotide positions 50 to 700, nucleotide positions 50 to 650, nucleotide positions 50 to 600, nucleotide positions 50 to 550, nucleotide positions 50 to 500, nucleotide positions 50 to 450, nucleotide positions 50 to 400, nucleotide positions 50 to 350, nucleotide positions 50 to 300, nucleotide positions 50 to 250, nucleotide positions 50 to 200, nucleotide positions 50 to 150, nucleotide positions 50 to 100, nucleotide positions 100 to 1773, nucleotide positions 100 to 1770, nucleotide positions 100 to 1750, nucleotide positions 100 to 1700, nucleotide positions 100 to 1650, nucleotide positions 100 to 1600, nucleotide positions 100 to 1550, nucleotide positions 100 to 1500, nucleotide positions 100 to 1450, nucleotide positions 100 to 1400, nucleotide positions 100 to 1350, nucleotide positions 100 to 1300, nucleotide positions 100 to 1250, nucleotide positions 100 to 1200, nucleotide positions 100 to 1150, nucleotide positions 100 to 1100, nucleotide positions 100 to 1050, nucleotide positions 100 to 1000, nucleotide positions 100 to 950, nucleotide positions 100 to 900, nucleotide positions 100 to 850, nucleotide positions 100 to 800, nucleotide positions 100 to 750, nucleotide positions 100 to 700, nucleotide positions 100 to 650, nucleotide positions 100 to 600, nucleotide positions 100 to 550, nucleotide positions 100 to 500, nucleotide positions 100 to 450, nucleotide positions 100 to 400, nucleotide positions 100 to 350, nucleotide positions 100 to 300, nucleotide positions 100 to 250, nucleotide positions 100 to 200, nucleotide positions 100 to 150, nucleotide positions 150 to 1773, nucleotide positions 150 to 1770, nucleotide positions 150 to 1750, nucleotide positions 150 to 1700, nucleotide positions 150 to 1650, nucleotide positions 150 to 1600, nucleotide positions

150 to 1550, nucleotide positions 150 to 1500, nucleotide positions 150 to 1450,
nucleotide positions 150 to 1400, nucleotide positions 150 to 1350, nucleotide positions
150 to 1300, nucleotide positions 150 to 1250, nucleotide positions 150 to 1200,
nucleotide positions 150 to 1150, nucleotide positions 150 to 1100, nucleotide positions
5 150 to 1050, nucleotide positions 150 to 1000, nucleotide positions 150 to 950,
nucleotide positions 150 to 900, nucleotide positions 150 to 850, nucleotide positions 150
to 800, nucleotide positions 150 to 750, nucleotide positions 150 to 700, nucleotides
positions 150 to 650, nucleotide positions 150 to 600, nucleotide positions 150 to 550,
nucleotide positions 150 to 500, nucleotide positions 150 to 450, nucleotide positions 150
10 to 400, nucleotide positions 150 to 350, nucleotide positions 150 to 300, nucleotide
positions 150 to 250, nucleotide positions 150 to 200, nucleotide positions 200 to 1773,
nucleotide positions 200 to 1770, nucleotide positions 200 to 1750, nucleotide positions
200 to 1700, nucleotide positions 200 to 1650, nucleotide positions 200 to 1600,
nucleotide positions 200 to 1550, nucleotide positions 200 to 1500, nucleotide positions
15 200 to 1450, nucleotide positions 200 to 1400, nucleotide positions 200 to 1350,
nucleotide positions 200 to 1300, nucleotide positions 200 to 1250, nucleotide positions
200 to 1200, nucleotide positions 200 to 1150, nucleotide positions 200 to 1100,
nucleotide positions 200 to 1050, nucleotide positions 200 to 1000, nucleotide positions
200 to 950, nucleotide positions 200 to 900, nucleotide positions 200 to 850, nucleotide
20 positions 200 to 800, nucleotide positions 200 to 750, nucleotide positions 200 to 700,
nucleotides positions 200 to 650, nucleotide positions 200 to 600, nucleotide positions
200 to 550, nucleotide positions 200 to 500, nucleotide positions 200 to 450, nucleotide
positions 200 to 400, nucleotide positions 200 to 350, nucleotide positions 200 to 300,
nucleotide positions 200 to 250, nucleotide positions 250 to 1773, nucleotide positions
25 250 to 1770, nucleotide positions 250 to 1750, nucleotide positions 250 to 1700,
nucleotide positions 250 to 1650, nucleotide positions 250 to 1600, nucleotide positions
250 to 1550, nucleotide positions 250 to 1500, nucleotide positions 250 to 1450,
nucleotide positions 250 to 1400, nucleotide positions 250 to 1350, nucleotide positions
250 to 1300, nucleotide positions 250 to 1250, nucleotide positions 250 to 1200,
30 nucleotide positions 250 to 1150, nucleotide positions 250 to 1100, nucleotide positions
250 to 1050, nucleotide positions 250 to 1000, nucleotide positions 250 to 950,

nucleotide positions 250 to 900, nucleotide positions 250 to 850, nucleotide positions 250
 to 800, nucleotide positions 250 to 750, nucleotide positions 250 to 700, nucleotides
 positions 250 to 650, nucleotide positions 250 to 600, nucleotide positions 250 to 550,
 nucleotide positions 250 to 500, nucleotide positions 250 to 450, nucleotide positions 250
 5 to 400, nucleotide positions 250 to 350, nucleotide positions 250 to 300, nucleotide
 positions 300 to 1773, nucleotide positions 300 to 1770, nucleotide positions 300 to 1750,
 nucleotide positions 300 to 1700, nucleotide positions 300 to 1650, nucleotide positions
 300 to 1600, nucleotide positions 300 to 1550, nucleotide positions 300 to 1500,
 nucleotide positions 300 to 1450, nucleotide positions 300 to 1400, nucleotide positions
 10 300 to 1350, nucleotide positions 300 to 1300, nucleotide positions 300 to 1250,
 nucleotide positions 300 to 1200, nucleotide positions 300 to 1150, nucleotide positions
 300 to 1100, nucleotide positions 300 to 1050, nucleotide positions 300 to 1000,
 nucleotide positions 300 to 950, nucleotide positions 300 to 900, nucleotide positions 300
 to 850, nucleotide positions 300 to 800, nucleotide positions 300 to 750, nucleotide
 15 positions 300 to 700, nucleotides positions 300 to 650, nucleotide positions 300 to 600,
 nucleotide positions 300 to 550, nucleotide positions 300 to 500, nucleotide positions 300
 to 450, nucleotide positions 300 to 400, nucleotide positions 300 to 350, nucleotide
 positions 350 to 1773, nucleotide positions 350 to 1770, nucleotide positions 350 to 1750,
 nucleotide positions 350 to 1700, nucleotide positions 350 to 1650, nucleotide positions
 20 350 to 1600, nucleotide positions 350 to 1550, nucleotide positions 350 to 1500,
 nucleotide positions 350 to 1450, nucleotide positions 350 to 1400, nucleotide positions
 350 to 1350, nucleotide positions 350 to 1300, nucleotide positions 350 to 1250,
 nucleotide positions 350 to 1200, nucleotide positions 350 to 1150, nucleotide positions
 350 to 1100, nucleotide positions 350 to 1050, nucleotide positions 350 to 1000,
 25 nucleotide positions 350 to 950, nucleotide positions 350 to 900, nucleotide positions 350
 to 850, nucleotide positions 350 to 800, nucleotide positions 350 to 750, nucleotide
 positions 350 to 700, nucleotides positions 350 to 650, nucleotide positions 350 to 600,
 nucleotide positions 350 to 550, nucleotide positions 350 to 500, nucleotide positions 350
 to 450, nucleotide positions 350 to 400, nucleotide positions 400 to 1773, nucleotide
 30 positions 400 to 1770, nucleotide positions 400 to 1750, nucleotide positions 400 to 1700,
 nucleotide positions 400 to 1650, nucleotide positions 400 to 1600, nucleotide positions

400 to 1550, nucleotide positions 400 to 1500, nucleotide positions 400 to 1450,
nucleotide positions 400 to 1400, nucleotide positions 400 to 1350, nucleotide positions
400 to 1300, nucleotide positions 400 to 1250, nucleotide positions 400 to 1200,
nucleotide positions 400 to 1150, nucleotide positions 400 to 1100, nucleotide positions
5 400 to 1050, nucleotide positions 400 to 1000, nucleotide positions 400 to 950,
nucleotide positions 400 to 900, nucleotide positions 400 to 850, nucleotide positions 400
to 800, nucleotide positions 400 to 750, nucleotide positions 400 to 700, nucleotides
positions 400 to 650, nucleotide positions 400 to 600, nucleotide positions 400 to 550,
nucleotide positions 400 to 500, nucleotide positions 400 to 450, nucleotide positions 450
10 to 1773, nucleotide positions 450 to 1770, nucleotide positions 450 to 1750, nucleotide
positions 450 to 1700, nucleotide positions 450 to 1650, nucleotide positions 450 to 1600,
nucleotide positions 450 to 1550, nucleotide positions 450 to 1500, nucleotide positions
450 to 1450, nucleotide positions 450 to 1400, nucleotide positions 450 to 1350,
nucleotide positions 450 to 1300, nucleotide positions 450 to 1250, nucleotide positions
15 450 to 1200, nucleotide positions 450 to 1150, nucleotide positions 450 to 1100,
nucleotide positions 450 to 1050, nucleotide positions 450 to 1000, nucleotide positions
450 to 950, nucleotide positions 450 to 900, nucleotide positions 450 to 850, nucleotide
positions 450 to 800, nucleotide positions 450 to 750, nucleotide positions 450 to 700,
nucleotides positions 450 to 650, nucleotide positions 450 to 600, nucleotide positions
20 450 to 550, nucleotide positions 450 to 500, nucleotide positions 500 to 1773, nucleotide
positions 500 to 1770, nucleotide positions 500 to 1750, nucleotide positions 500 to 1700,
nucleotide positions 500 to 1650, nucleotide positions 500 to 1600, nucleotide positions
500 to 1550, nucleotide positions 500 to 1500, nucleotide positions 500 to 1450,
nucleotide positions 500 to 1400, nucleotide positions 500 to 1350, nucleotide positions
25 500 to 1300, nucleotide positions 500 to 1250, nucleotide positions 500 to 1200,
nucleotide positions 500 to 1150, nucleotide positions 500 to 1100, nucleotide positions
500 to 1050, nucleotide positions 500 to 1000, nucleotide positions 500 to 950,
nucleotide positions 500 to 900, nucleotide positions 500 to 850, nucleotide positions 500
to 800, nucleotide positions 500 to 750, nucleotide positions 500 to 700, nucleotides
30 positions 500 to 650, nucleotide positions 500 to 600, nucleotide positions 500 to 550,
nucleotide positions 550 to 1773, nucleotide positions 550 to 1770, nucleotide positions

550 to 1750, nucleotide positions 550 to 1700, nucleotide positions 550 to 1650,
nucleotide positions 550 to 1600, nucleotide positions 550 to 1550, nucleotide positions
550 to 1500, nucleotide positions 550 to 1450, nucleotide positions 550 to 1400,
nucleotide positions 550 to 1350, nucleotide positions 550 to 1300, nucleotide positions
5 550 to 1250, nucleotide positions 550 to 1200, nucleotide positions 550 to 1150,
nucleotide positions 550 to 1100, nucleotide positions 550 to 1050, nucleotide positions
550 to 1000, nucleotide positions 550 to 950, nucleotide positions 550 to 900, nucleotide
positions 550 to 850, nucleotide positions 550 to 800, nucleotide positions 550 to 750,
nucleotide positions 550 to 700, nucleotides positions 550 to 650, nucleotide positions
10 550 to 600, nucleotide positions 600 to 1773, nucleotide positions 600 to 1770,
nucleotide positions 600 to 1750, nucleotide positions 600 to 1700, nucleotide positions
600 to 1650, nucleotide positions 600 to 1600, nucleotide positions 600 to 1550,
nucleotide positions 600 to 1500, nucleotide positions 600 to 1450, nucleotide positions
600 to 1400, nucleotide positions 600 to 1350, nucleotide positions 600 to 1300,
15 nucleotide positions 600 to 1250, nucleotide positions 600 to 1200, nucleotide positions
600 to 1150, nucleotide positions 600 to 1100, nucleotide positions 600 to 1050,
nucleotide positions 600 to 1000, nucleotide positions 600 to 950, nucleotide positions
600 to 900, nucleotide positions 600 to 850, nucleotide positions 600 to 800, nucleotide
positions 600 to 750, nucleotide positions 600 to 700, nucleotides positions 600 to 650,
20 nucleotide positions 650 to 1773, nucleotide positions 650 to 1770, nucleotide positions
650 to 1750, nucleotide positions 650 to 1700, nucleotide positions 650 to 1650,
nucleotide positions 650 to 1600, nucleotide positions 650 to 1550, nucleotide positions
650 to 1500, nucleotide positions 650 to 1450, nucleotide positions 650 to 1400,
nucleotide positions 650 to 1350, nucleotide positions 650 to 1300, nucleotide positions
25 650 to 1250, nucleotide positions 650 to 1200, nucleotide positions 650 to 1150,
nucleotide positions 650 to 1100, nucleotide positions 650 to 1050, nucleotide positions
650 to 1000, nucleotide positions 650 to 950, nucleotide positions 650 to 900, nucleotide
positions 650 to 850, nucleotide positions 650 to 800, nucleotide positions 650 to 750,
nucleotide positions 650 to 700, nucleotide positions 700 to 1773, nucleotide positions
30 700 to 1770, nucleotide positions 700 to 1750, nucleotide positions 700 to 1700,
nucleotide positions 700 to 1650, nucleotide positions 700 to 1600, nucleotide positions

700 to 1550, nucleotide positions 700 to 1500, nucleotide positions 700 to 1450,
nucleotide positions 700 to 1400, nucleotide positions 700 to 1350, nucleotide positions
700 to 1300, nucleotide positions 700 to 1250, nucleotide positions 700 to 1200,
nucleotide positions 700 to 1150, nucleotide positions 700 to 1100, nucleotide positions
5 700 to 1050, nucleotide positions 700 to 1000, nucleotide positions 700 to 950,
nucleotide positions 700 to 900, nucleotide positions 700 to 850, nucleotide positions 700
to 800, nucleotide positions 700 to 750, nucleotide positions 750 to 1773, nucleotide
positions 750 to 1770, nucleotide positions 750 to 1750, nucleotide positions 750 to 1700,
nucleotide positions 750 to 1650, nucleotide positions 750 to 1600, nucleotide positions
10 750 to 1550, nucleotide positions 750 to 1500, nucleotide positions 750 to 1450,
nucleotide positions 750 to 1400, nucleotide positions 750 to 1350, nucleotide positions
750 to 1300, nucleotide positions 750 to 1250, nucleotide positions 750 to 1200,
nucleotide positions 750 to 1150, nucleotide positions 750 to 1100, nucleotide positions
750 to 1050, nucleotide positions 750 to 1000, nucleotide positions 750 to 950,
15 nucleotide positions 750 to 900, nucleotide positions 750 to 850, nucleotide positions 750
to 800, nucleotide positions 800 to 1773, nucleotide positions 800 to 1770, nucleotide
positions 800 to 1750, nucleotide positions 800 to 1700, nucleotide positions 800 to 1650,
nucleotide positions 800 to 1600, nucleotide positions 800 to 1550, nucleotide positions
800 to 1500, nucleotide positions 800 to 1450, nucleotide positions 800 to 1400,
20 nucleotide positions 800 to 1350, nucleotide positions 800 to 1300, nucleotide positions
800 to 1250, nucleotide positions 800 to 1200, nucleotide positions 800 to 1150,
nucleotide positions 800 to 1100, nucleotide positions 800 to 1050, nucleotide positions
800 to 1000, nucleotide positions 800 to 950, nucleotide positions 800 to 900, nucleotide
positions 800 to 850, nucleotide positions 850 to 1773, nucleotide positions 850 to 1770,
25 nucleotide positions 850 to 1750, nucleotide positions 850 to 1700, nucleotide positions
850 to 1650, nucleotide positions 850 to 1600, nucleotide positions 850 to 1550,
nucleotide positions 850 to 1500, nucleotide positions 850 to 1450, nucleotide positions
850 to 1400, nucleotide positions 850 to 1350, nucleotide positions 850 to 1300,
nucleotide positions 850 to 1250, nucleotide positions 850 to 1200, nucleotide positions
30 850 to 1150, nucleotide positions 850 to 1100, nucleotide positions 850 to 1050,
nucleotide positions 850 to 1000, nucleotide positions 850 to 950, nucleotide positions

850 to 900, nucleotide positions 900 to 1773, nucleotide positions 900 to 1770,
nucleotide positions 900 to 1750, nucleotide positions 900 to 1700, nucleotide positions
900 to 1650, nucleotide positions 900 to 1600, nucleotide positions 900 to 1550,
nucleotide positions 900 to 1500, nucleotide positions 900 to 1450, nucleotide positions
5 900 to 1400, nucleotide positions 900 to 1350, nucleotide positions 900 to 1300,
nucleotide positions 900 to 1250, nucleotide positions 900 to 1200, nucleotide positions
900 to 1150, nucleotide positions 900 to 1100, nucleotide positions 900 to 1050,
nucleotide positions 900 to 1000, nucleotide positions 900 to 950, nucleotide positions
950 to 1773, nucleotide positions 950 to 1770, nucleotide positions 950 to 1750,
10 nucleotide positions 950 to 1700, nucleotide positions 950 to 1650, nucleotide positions
950 to 1600, nucleotide positions 950 to 1550, nucleotide positions 950 to 1500,
nucleotide positions 950 to 1450, nucleotide positions 950 to 1400, nucleotide positions
950 to 1350, nucleotide positions 950 to 1300, nucleotide positions 950 to 1250,
nucleotide positions 950 to 1200, nucleotide positions 950 to 1150, nucleotide positions
15 950 to 1100, nucleotide positions 950 to 1050, nucleotide positions 950 to 1000,
nucleotide positions 1000 to 1773, nucleotide positions 1000 to 1770, nucleotide
positions 1000 to 1750, nucleotide positions 1000 to 1700, nucleotide positions 1000 to
1650, nucleotide positions 1000 to 1600, nucleotide positions 1000 to 1550, nucleotide
positions 1000 to 1500, nucleotide positions 1000 to 1450, nucleotide positions 1000 to
20 1400, nucleotide positions 1000 to 1350, nucleotide positions 1000 to 1300, nucleotide
positions 1000 to 1250, nucleotide positions 1000 to 1200, nucleotide positions 1000 to
1150, nucleotide positions 1000 to 1100, nucleotide positions 1000 to 1050, nucleotide
positions 1050 to 1773, nucleotide positions 1050 to 1770, nucleotide positions 1050 to
1750, nucleotide positions 1050 to 1700, nucleotide positions 1050 to 1650, nucleotide
25 positions 1050 to 1600, nucleotide positions 1050 to 1550, nucleotide positions 1050 to
1500, nucleotide positions 1050 to 1450, nucleotide positions 1050 to 1400, nucleotide
positions 1050 to 1350, nucleotide positions 1050 to 1300, nucleotide positions 1050 to
1250, nucleotide positions 1050 to 1200, nucleotide positions 1050 to 1150, nucleotide
positions 1050 to 1100, nucleotide positions 1100 to 1773, nucleotide positions 1100 to
30 1770, nucleotide positions 1100 to 1750, nucleotide positions 1100 to 1700, nucleotide
positions 1100 to 1650, nucleotide positions 1100 to 1600, nucleotide positions 1100 to

1550, nucleotide positions 1100 to 1500, nucleotide positions 1100 to 1450, nucleotide
positions 1100 to 1400, nucleotide positions 1100 to 1350, nucleotide positions 1100 to
1300, nucleotide positions 1100 to 1250, nucleotide positions 1100 to 1200, nucleotide
positions 1100 to 1150, nucleotide positions 1150 to 1773, nucleotide positions 1150 to
5 1770, nucleotide positions 1150 to 1750, nucleotide positions 1150 to 1700, nucleotide
positions 1150 to 1650, nucleotide positions 1150 to 1600, nucleotide positions 1150 to
1550, nucleotide positions 1150 to 1500, nucleotide positions 1150 to 1450, nucleotide
positions 1150 to 1400, nucleotide positions 1150 to 1350, nucleotide positions 1150 to
1300, nucleotide positions 1150 to 1250, nucleotide positions 1150 to 1200, nucleotide
10 positions 1200 to 1773, nucleotide positions 1200 to 1770, nucleotide positions 1200 to
1750, nucleotide positions 1200 to 1700, nucleotide positions 1200 to 1650, nucleotide
positions 1200 to 1600, nucleotide positions 1200 to 1550, nucleotide positions 1200 to
1500, nucleotide positions 1200 to 1450, nucleotide positions 1200 to 1400, nucleotide
positions 1200 to 1350, nucleotide positions 1200 to 1300, nucleotide positions 1200 to
15 1250, nucleotide positions 1250 to 1773, nucleotide positions 1250 to 1770, nucleotide
positions 1250 to 1750, nucleotide positions 1250 to 1700, nucleotide positions 1250 to
1650, nucleotide positions 1250 to 1600, nucleotide positions 1250 to 1550, nucleotide
positions 1250 to 1500, nucleotide positions 1250 to 1450, nucleotide positions 1250 to
1400, nucleotide positions 1250 to 1350, nucleotide positions 1250 to 1300, nucleotide
20 positions 1300 to 1773, nucleotide positions 1300 to 1770, nucleotide positions 1300 to
1750, nucleotide positions 1300 to 1700, nucleotide positions 1300 to 1650, nucleotide
positions 1300 to 1600, nucleotide positions 1300 to 1550, nucleotide positions 1300 to
1500, nucleotide positions 1300 to 1450, nucleotide positions 1300 to 1400, nucleotide
positions 1300 to 1350, nucleotide positions 1350 to 1773, nucleotide positions 1350 to
25 1770, nucleotide positions 1350 to 1750, nucleotide positions 1350 to 1700, nucleotide
positions 1350 to 1650, nucleotide positions 1350 to 1600, nucleotide positions 1350 to
1550, nucleotide positions 1350 to 1500, nucleotide positions 1350 to 1450, nucleotide
positions 1350 to 1400, nucleotide positions 1400 to 1773, nucleotide positions 1400 to
1770, nucleotide positions 1400 to 1750, nucleotide positions 1400 to 1700, nucleotide
30 positions 1400 to 1650, nucleotide positions 1400 to 1600, nucleotide positions 1400 to
1550, nucleotide positions 1400 to 1500, nucleotide positions 1400 to 1450, nucleotide

positions 1450 to 1773, nucleotide positions 1450 to 1770, nucleotide positions 1450 to 1750, nucleotide positions 1450 to 1700, nucleotide positions 1450 to 1650, nucleotide positions 1450 to 1600, nucleotide positions 1450 to 1550, nucleotide positions 1450 to 1500, nucleotide positions 1500 to 1773, nucleotide positions 1500 to 1770, nucleotide positions 1500 to 1750, nucleotide positions 1500 to 1700, nucleotide positions 1500 to 1650, nucleotide positions 1500 to 1600, nucleotide positions 1500 to 1550, nucleotide positions 1550 to 1773, nucleotide positions 1550 to 1770, nucleotide positions 1550 to 1750, nucleotide positions 1550 to 1700, nucleotide positions 1550 to 1650, nucleotide positions 1550 to 1600, nucleotide positions 1600 to 1773, nucleotide positions 1600 to 1770, nucleotide positions 1600 to 1750, nucleotide positions 1600 to 1700, nucleotide positions 1600 to 1650, nucleotide positions 1650 to 1773, nucleotide positions 1650 to 1770, nucleotide positions 1650 to 1750, nucleotide positions 1650 to 1700, nucleotide positions 1700 to 1773, nucleotide positions 1700 to 1770, nucleotide positions 1700 to 1750, nucleotide positions 1750 to 1773, or nucleotide positions 1750 to 1770, of SEQ ID NO: 14 or 15.

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 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%, or at least 99%) identical to SEQ ID NO: 14 or 15.

Human 5' UTR of CLRN1 (SEQ ID NO: 12)

AGGAGATACTTGAAGGCAGTTTGAAAGACTTGTTTTACAGATTCTTAGTCCAAAGATTTCCAATTAGGGAG
AAGAAGCAGCAGAAAAGGAGAAAAGCCAAGTATGAGTGATGATGAGGCCCTTCATCTACTGACATTTAACCT
GGCGAGAACCGTCGATGGTGAAGTTGCCTTTTCAGCTGGGAGCTGTCCGTTTCAGCTTCCGTAATAAATGCA
GTCAAAGAGGCAGTCCCTTCCCATTGCTCACAAAGGTCTTGTTTTGAACCTCGCCCTCACAGAAGCCGTT
TCTCATC

Human 5'UTR of CLRN1 (SEQ ID NO: 13)

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGC
CTCAGTGAGCGAGCGAGCGCGCAGAGAGGAGTGGCCAACCTCCATCACTAGGGGTTTCCT

Human 3' UTR of CLRN1 (SEQ ID NO: 14)

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCA
AAGTTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

Human 3' UTR of CLRN1 (SEQ ID NO: 15)

AAGGCAAACCTTTCTATAATTTTACAAGGGAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTTAATTTTG
 CATATCCTTTTAGTCTGCATATATTAAAGCATCAGGACCCCTTCGTGACAATGTTTACAAATTACGTACTAA
 5 GGATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTTATCTGGTGGGAAATTGC
 TTGACCCAGGTAGTCAAAGGCAGTTGACTAGAATCGACAAATTGTTACTCCATATATATATATGTGTGTGT
 GTGTGTGTGTGTGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGATATCAAAGGCACATGGAATATATTTTAA
 TAAAAACAAATAATATCTCTAATATATCCACACATTTGTTGCCAGATTTTCAGAAAACAGAGCTGCAATCGC
 TTTCTTAAACAGTAGTGTATTAAATGAACATCTATAAAATGTATCAACACACATTTTAAAAAATTTGTTT
 10 AAAGTATACTCTTAGGCCAGGCGTGGTGACTCACACCTGTAATTCCAGCAGCTTCAGGAGGCCAAGGTGGGA
 AGATCATTTGAGTTCAGGAGTTCGAGTTACAGCCTGGGCAATAAAGTGAGACCCGTGCTACTAACAAAATTA
 AAAAATAAAATAAATAAATAAATATAGGCTTTAAAAAAGCATAGTCTTATTAACCATGTCTGTTGGTCAAAA
 TCTGCAAACTCTAAAAAGAAGAAAAGAAGAAAAACCAAGCTTAGGGTATTTTTCTCCCGTGCCTGAGTCC
 CAATTACATTACAGCAGTACTTTCAATGAACATAATTGTTAGGACCACTGAGGAATCATGAAAAATGATC
 15 TCTGCTTAGTACATTTGATGCAAAATGACTTATTAGGGGCTGTTTTCTAGCTATAGTGTCTCGAGTACTA
 ATATGCAATTATGAAAAATTATATTAAATCTGGGATTATGACGGTATCACTGTATCATCTTGGTCTTGTCT
 GGCTGTCACCAAGCATGACCCAGGTCAACTTTTTTTTTTCCCTGAATTACCCATCAAATTGATCTGCAGCT
 GACTAAAGGCCACAGCTGAGCCTGGAAGTACCCTTCCTTCATCCTCAACCTGCTGTCTCCAGAAAAGCAC
 CAAGGAAAAAGCAGAGAATGACAGCAAACAGATCACTAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAG
 20 CCCTCATATAATAGGTTTGAAAGTACTCCTTAAATAAAACACTGTTTCCCTTTGGAACATTTTACAAGGA
 TGAAACAACCGTATACCTGAGAAATAACTTGCTCTGGTGTCAATTTCGCTATTCGCCAGCAGACATCAGAAC
 ACACCGAGTTTCAGATGCTGGTTTTTTCCCTTAAATCAGGAAATACACCTGGACAATTTCTAGAAGACTA
 CAATTCACTCTAGCCACAAAGGGGATTTTTTTTTTTTGGTAACAGGCTAGAGCCCGGTTCTGTAAAGTCTTT
 AGCTGAAATGGTCCAGTACAAAAGCACTGGAAATGAGTGGGCTAGGAGGACAAGGACCGTCTCCTGCGTGA
 25 GGAGTTGGTTGGAGGTCCCCAAGGCCAGGTACCCCTGCACTCTTATTGGATTCTCTCTGCTCTTCTTGGGA
 GTTTTGAAAAACTCCTTCGAACACCAGGCTTTTTTTCTTTAGAAAACAAGTCTCCAATCGTTCTCTGTTCCG
 TAGAAAGAGAAAAGAAAACCTGGAGCAGCTGCTGAAAAATCTAATGAGGAACTAAGAGGCAAACCCACCA

In some embodiments, the introduction, removal, or modification of 3' UTR
 30 AREs can be used to modulate the stability of an mRNA encoding a CLRN1 protein. In
 other embodiments, AREs can be removed or mutated to increase the intracellular
 stability and thus increase translation and production of a CLRN1 protein.

In other embodiments, non-ARE sequences may be incorporated into the 5' or 3'
 UTRs. In some embodiments, introns or portions of intron sequences may be
 35 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.

In some embodiments of any of the vectors described herein, the vector includes a chimeric intron sequence (SEQ ID NO: 16).

5

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
10 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
15 embodiments, the cell is a specialized cell of the cochlea. In some embodiments, the cell is a cochlear hair cell, such as a cochlear inner hair cell or a cochlear out hair cell. In some embodiments, the cell is an ocular cell (e.g. a retinal cell, a retinal ganglion cell, an amacrine cell, a horizontal cell, a bipolar cell, a photoreceptor cell).

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

Methods

Also provided herein are methods that include: introducing into a cochlea of a
25 mammal (e.g., a human) a therapeutically effective amount of any of the compositions described herein.

Also provided herein are methods of increasing expression of a full-length CLRN1 protein in a mammalian cell, that include: introducing any of the compositions described herein into the mammalian cell.

30 Also provided herein are methods of increasing expression of a full-length CLRN1 protein in an inner hair cell, an outer hair cell, or both, 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 herein are methods of increasing expression of a full-length CLRN1 protein in an eye of a mammal (e.g., a human) that include: intraocularly
5 administering to the eye of the mammal a therapeutically effective amount of any of the compositions described herein.

Also provided herein are methods of treating hearing loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions described herein into the cochlea of the subject.

10 Also provided herein are methods of treating vision loss in a subject identified as having a defective CLRN1 gene that include: administering a therapeutically effective amount of any of the compositions described herein into the eye of the subject.

In some embodiments of any of these methods, the mammal has been previously identified as having a defective CLRN1 gene (e.g., a CLRN1 gene having a mutation that
15 results in a decrease in the expression and/or activity of a CLRN1 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 CLRN1 gene. Some embodiments of any of these methods can further include detecting a mutation in a CLRN1 gene in a subject. Some embodiments of any of the methods can
20 further include identifying or diagnosing a subject as having hearing loss and/or vision 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
25 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
30 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 these methods, two or more doses of any of the compositions described herein are introduced or administered into the eye 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 eye (e.g., intraocular space) 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 eye of the mammal or subject found not to have a vision within a normal range (e.g., as determined using any test for vision known in the art).

In some embodiments of any of the methods described herein, the composition can be formulated for intra-ocular administration. In some embodiments of any of the methods described herein, the compositions described herein can be administered via intra-ocular administration or local administration.

5 In some embodiments, intra-ocular administration can be performed using any of the methods described herein or known in the art.

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
10 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,
15 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
20 mammal has or is at risk of developing hearing loss and/or vision loss (e.g., Usher syndrome type III, retinitis pigmentosa). In some embodiments of any of the methods described herein, the subject or mammal has been previously identified as having a mutation in a CLRN1 gene. In some embodiments of any of the methods described herein, the subject or mammal has any of the mutations in a CLRN1 gene that are
25 described herein or are known in the art to be associated with hearing loss and/or vision 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 CLRN1 gene (e.g., via genetic testing). In some embodiments of any of the methods described herein, the
30 subject or human has been identified as having a mutation in a CLRN1 gene and has been diagnosed with hearing loss and/or vision loss (e.g., Usher syndrome type III, retinitis

pigmentosa). In some embodiments of any of the methods described herein, the subject or human has been identified as having hearing loss and/or vision loss (e.g., Usher syndrome type III, retinitis pigmentosa).

In some embodiments, successful treatment of hearing loss (e.g., Usher syndrome
5 type III) 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).

In some embodiments, successful treatment of vision loss can be determined in a
10 subject using any of the conventional functional vision tests known in the art. Non-limiting examples of functional retinal and vision tests are acuity testing, intraocular pressure (IOP) testing, and an electroretinogram (ERG).

Also provided herein are methods of increasing expression of an active CLRN1 protein (e.g., a full-length CLRN1 protein) in a mammalian cell that include introducing
15 any of the compositions described herein into the mammalian cell. In some embodiments of these methods, the mammalian cell is a cochlear hair cell (e.g., an inner hair cell, an outer hair cell) or an ocular cell (e.g., a retinal cell). In some embodiments of these methods, the mammalian cell is a human cell (e.g., a human cochlear hair cell). In some embodiments of these methods, the mammalian cell is *in vitro*. In some embodiments of
20 these methods, the mammalian cell is in a mammal. In some embodiments of these methods, the mammalian cell is originally obtained from a mammal and is cultured *ex vivo*. In some embodiments, the mammalian cell has previously been determined to have a defective CLRN1 gene.

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

An increase in expression of an active CLRN1 protein (e.g., a full-length CLRN1 protein) as described herein is, e.g., as compared to a control or to the level of expression of an active CLRN1 protein (e.g., a full-length CLRN1 protein) prior to the introduction
30 of the vector(s).

Methods of detecting expression and/or activity of CLRN1 are known in the art. In some embodiments, the level of expression of a CLRN1 protein can be detected directly (e.g., detecting CLRN1 protein or detecting CLRN1 mRNA). Non-limiting examples of techniques that can be used to detect expression and/or activity of CLRN1 directly include: real-time PCR, Western blotting, immunoprecipitation, immunohistochemistry, or immunofluorescence. In some embodiments, expression of a CLRN1 protein can be detected indirectly (e.g., through functional hearing tests, functional retinal and vision tests).

10 **Pharmaceutical Compositions and Kits**

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

5 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

10 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.

15 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

20 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 µg, at least 2 µg, at least 4 µg, at least 6 µg, at least 8 µg, at least 10 µg, at least

25 12 µg, at least 14 µg, at least 16 µg, at least 18 µg, at least 20 µg, at least 22 µg, at least 24 µg, at least 26 µg, at least 28 µg, at least 30 µg at least 32 µg, at least 34 µg, at least 36 µg, at least 38 µg, at least 40 µg, at least 42 µg, at least 44 µg, at least 46 µg, at least 48 µg, at least 50 µg, at least 52 µg, at least 54 µg, at least 56 µg, at least 58 µg, at least 60 µg, at least 62 µg, at least 64 µg, at least 66 µg, at least 68 µg, at least 70 µg, at least

30 72 µg, at least 74 µg, at least 76 µg, at least 78 µg, at least 80 µg, at least 82 µg, at least 84 µg, at least 86 µg, at least 88 µg, at least 90 µg, at least 92 µg, at least 94 µg, at least

96 µg, at least 98 µg, at least 100 µg, at least 102 µg, at least 104 µg, at least 106 µg, at least 108 µg, at least 110 µg, at least 112 µg, at least 114 µg, at least 116 µg, at least 118 µg, at least 120 µg, at least 122 µg, at least 124 µg, at least 126 µg, at least 128 µg, at least 130 µg, at least 132 µg, at least 134 µg, at least 136 µg, at least 138 µg, at least 140 µg, at least 142 µg, at least 144 µg, at least 146 µg, at least 148 µg, at least 150 µg, at least 152 µg, at least 154 µg, at least 156 µg, at least 158 µg, at least 160 µg, at least 162 µg, at least 164 µg, at least 166 µg, at least 168 µg, at least 170 µg, at least 172 µg, at least 174 µg, at least 176 µg, at least 178 µg, at least 180 µg, at least 182 µg, at least 184 µg, at least 186 µg, at least 188 µg, at least 190 µg, at least 192 µg, at least 194 µg, at least 196 µg, at least 198 µg, or at least 200 µg, 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.

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.

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.

Devices and Surgical Methods

5 Provided herein are therapeutic delivery systems for treating hearing loss and/or vision loss (e.g., Usher syndrome type III, retinitis pigmentosa). In one aspect, the therapeutic delivery systems include i) a 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
10 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., Usher syndrome type III). 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
15 administering intra-cochlearly a therapeutically effective 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
20 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 medical device capable of
25 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 includes a base and/or a reservoir capable of holding
30 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.

Also provided herein are surgical methods for treatment of vision loss (e.g., retinitis pigmentosa). In some embodiments, the methods include the steps of:

5 administering intra-ocularly a therapeutically effective amount of any of the compositions provided herein.

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
10 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 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
15 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

20 **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 a ratio of 1:1:2. The
25 AAV vectors used here express human CLRN1 or mouse CLRN1 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 CLRN1 constructs to test (i) a concatemerization-transplicing strategy, (ii) a hybrid intronic-homologous recombination-transplicing strategy, and (iii) an exonic homologous
30 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 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 μ l are recovered from the CsCl density gradient tube and purified through dialysis in 1 $\times$ PBS. The fractions are analyzed by dot blot to determine those containing rAAV genomes. The viral genome number (vg) of each preparation is 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

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

The AAV- CLRN1 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 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- CLRN1 to enter the cochlear perilymph of the scala tympani at a rate of approximately 1 μ L/min, but small enough to heal without surgical repair. The remaining portion of the microcatheter, proximal to the microneedle(s), is loaded with the AAV- CLRN1 /artificial perilymph formulation at a titer of approximately 1e13 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

5 AAV- CLRN1 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 made
10 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 deliver a total of approximately
15 300 nL of AAV- CLRN1 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
20 hydrochloride for 3 days.

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

Surgical Procedure

25 AAV- CLRN1 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 combination of pentobarbital sodium (Nembutal; 25 mg kg⁻¹, injected intraperitoneally),
30 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 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

5 perfluoroalkoxy-alkane-insulated silver wire electrode (203 μ m uncoated diameter) is inserted near the round window niche and glued to the bulla.

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

10 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.

AAV- CLRN1 at a maximum titer of 1×10^{14} 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

15 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 AAV- CLRN1 reservoir; (iv) the inlet to the integrated AAV- CLRN1 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- CLRN1 delivery is as follows: (i) an

20 internal AAV- CLRN1 -refresh loop is run, transferring AAV- CLRN1 from the AAV- CLRN1 reservoir into the main infuse-withdraw line; (ii) AAV- CLRN1 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- CLRN1 has been allowed to diffuse for some time, a volume of

25 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 ceilings are a thin (25.4 μ m), flexible, polyimide membrane. The pump chamber has a diameter of

30 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 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 .

5

Acute drug delivery in guinea pigs

The micropump is loaded with AAV- CLRN1 and artificial perilymph, and the cannula inserted into a cochleostomy made in the region of the cochlea between the 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- CLRN1 /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 delivered to the cochlea. After a 20 min wait time, approximately 10 μL of perilymph is withdrawn from the cochlea. AAV- CLRN1 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-treatment (n=4 per group) and their cochleae extracted. Extent of AAV transduction and CLRN1 expression along the organ of Corti is assessed via immunostaining with anti-CLRN1 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

AAV- CLRN1 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 otoacoustic 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 (OHC) function. Following baseline ABR and DPOAE measurements, 20 μL of AAV1- CLRN1

at titers of 1.0×10^{14} , 3.2×10^{13} , 1.0×10^{13} and 3.2×10^{12} 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 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 CLRN1 protein expression along the cochlear sensory epithelium. Antibodies against markers for hair cells (Myo7a), supporting cells (Sox2) and CLRN1 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 CLRN1 are counted within 200 um regions.

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 CLRN1 mRNA transcript.

Example 8: 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 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-CLRN1 prepared in artificial perilymph at a titer of 1×10^{13} vg/mL. The microcatheter is connected to a micromanipulator that infuses approximately 20 uL of the AAV-CLRN1 solution at a rate of approximately 1 uL / min. At the conclusion of the AAV-CLRN1 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.

Example 9: Non-Invasive Prenatal Testing of Maternal Blood to Detect CLRN1**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
 5 2,000 g for 20 minutes, followed by 3,220 g for 30 minutes, with 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 µL 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
 10 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 nested primers. Samples are
 15 prepared for sequencing by adding barcoded tags in a third 12-cycle round of PCR. The amplicons are then sequenced using an Illumina HiSeq sequencer. Genome sequence alignment is performed using commercially available software.

Example 10: Adenovirus (AAV) Trans-Splicing Strategy

20 At least two different nucleic acid vectors (e.g., AAV vectors) can be used to reconstitute an active CLRN1 gene (e.g., a full-length CLRN1 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.

25 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 CLRN1 protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the N-terminal portions of a CLRN1 protein described herein), and a splice
 30 donor sequence positioned at the 3' end of the first coding sequence. A second nucleic acid vector can include a splice acceptor sequence, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein (i.e., the entire portion of the CLRN1

protein that is not included in the N-terminal portion) positioned at the 3' end of the splice acceptor sequence (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the C-terminal portions of a CLRN1 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 CLRN1 protein (e.g., a full-length CLRN1 protein). When introduced into a mammalian cell (e.g., any of the mammalian cells described herein) splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes an active CLRN1 protein (e.g., a full-length CLRN1 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 CLRN1 gene that encodes a first portion of a CLRN1 protein (e.g., any of the CLRN1 coding sequences described herein) positioned 3' of the promoter, and a first splice donor sequence positioned at the 3' end of the first coding sequence. A second nucleic acid vector can include a first splice acceptor sequence, a second coding sequence of a CLRN1 gene that encodes a second portion of a CLRN1 protein positioned at the 3' end of the first splice acceptor sequence, and a second splice donor sequence positioned at the 3' end of the second coding sequence (e.g., any of the splicedonor sequences 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 splice donor sequence and the first splice acceptor sequence of the second nucleic acid vector). In some embodiments, the splice donor sequence of the first nucleic acid vector and the second splice donor sequence of the second nucleic acid vector are the same (e.g., any of the splice donor sequences described herein or known in the art). In some embodiments, the first splice donor sequence of the first nucleic acid vector and the second splice donor sequence of the second nucleic acid vector are different (e.g., any of the splice donor sequences described

herein or known in the art). A third nucleic acid vector will include a second splice acceptor sequence, a third coding sequence of a CLRN1 gene that encodes a third portion of a CLRN1 protein positioned at the 3' end of the second splice acceptor sequence, and a polyadenylation sequence positioned at the 3' end of the third coding sequence (e.g.,
5 any of the polyadenylation sequences described herein). In such methods where three nucleic acid vectors are used, the first splice donor sequence and the first splice acceptor sequence can assemble together (recombine) and the second splice donor sequence and the second splice acceptor sequence can assemble together (recombine), and the portion of CLRN1 protein encoded by the first, second, and third coding sequences do not
10 overlap, and when introduced into a mammalian cell (e.g., any of the mammalian cells described herein), splicing occurs between the first splice donor sequence and the first splice acceptor sequence, and between the second splice donor sequence and the second splice acceptor sequence, to form a recombined nucleic acid that encodes an active CLRN1 protein (e.g., a full-length CLRN1 protein). Based on the strategies provided
15 above, one skilled in the art would understand how to develop a strategy using four, five, or six different nucleic acid vectors.

In any of the examples of these methods, none of the amino acid sequences of the encoded portions overlap with any other encoded portion, and no single vector encodes an active CLRN1 protein (e.g., a full-length CLRN1 protein).

20 Each of the at least two different vectors includes a coding sequence that encodes a different portion of a CLRN1 protein, and each of the encoded portions can be at least 30 amino acids (e.g., between about 30 amino acids to about 1200 amino 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
25 and at least one intron of SEQ ID NO: 9 (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 three introns). In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a CLRN1 protein, and each of the
30 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-overlapping. In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of the amino acid sequence of SEQ ID NO: 3 (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: 3), provided that each of the encoded portions is non-overlapping with any other. In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of the amino acid sequence of SEQ ID NO: 5 (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: 5), provided that each of the encoded portions is non-overlapping with any other. In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of the amino acid sequence of SEQ ID NO: 7 (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: 7), provided that each of the encoded portions is non-overlapping with any other.

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, or an 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. Non-limiting examples of splice acceptor and/or donor 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 splice donor and acceptor sequences can be any endogenous intron splice donor/acceptor sequence of a gene (e.g., a CLRN1 gene). For example, the splice donor sequence can be: 5'-GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGA AACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCT-3' (SEQ ID NO: 22) and the splice acceptor sequence can be 5'-GATAGGCACCTATTGGTCTTACTG ACATCCACTTTGCCTTTCTCTCCACAG-3' (SEQ ID NO: 23) (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).

5 **Example 11: 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 CLRN1 gene (e.g., a full-length CLRN1 gene) within a cell following

10 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

15 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

20 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 CLRN1 gene (e.g., a full-length CLRN1 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

25 trans-splicing, an active CLRN1 gene (e.g., a full-length CLRN1 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

30 coding sequence that encodes an N-terminal portion of a CLRN1 protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a CLRN1 protein described herein

and/or any of the N-terminal portions of a CLRN1 protein described herein), a splice donor sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splice donor sequence. A second nucleic acid vector can include a second detectable marker gene, a splice acceptor sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the C-terminal portions of a CLRN1 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 sequences of the encoded portions do not overlap, and no single vector of the two different vectors encodes an active CLRN1 protein (e.g., a full-length CLRN1 protein). When introduced into a mammalian cell (e.g., any of the mammalian cells described herein) splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes an active CLRN1 protein (e.g., a full-length CLRN1 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 CLRN1 gene that encodes a first portion of a CLRN1 protein (e.g., any of the CLRN1 coding sequences described herein) positioned 3' of the promoter, a first splice donor 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 splice acceptor sequence positioned 3' of the second detectable marker gene, a second coding sequence of a CLRN1 gene that encodes a second portion of a CLRN1 protein positioned at the 3' end of the first splice acceptor sequence, a second splice donor sequence positioned at the 3' end of the second coding sequence (e.g., any of the splice donor sequences 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 splice donor sequence and the first splice acceptor sequence of the second nucleic acid

vector). In some embodiments, the splice donor sequence of the first nucleic acid vector and the second splice donor sequence of the second nucleic acid vector are the same (e.g., any of the splice donor sequences described herein or known in the art). In some embodiments, the first splice donor sequence of the first nucleic acid vector and the second splice donor sequence of the second nucleic acid vector are different (e.g., any of the splice donor sequences described herein or known in the art). A third nucleic acid vector can include a fourth detectable marker gene, a second splice acceptor sequence positioned 3' of the fourth detectable marker gene, a third coding sequence of a CLRN1 gene that encodes a third portion of a CLRN1 protein positioned at the 3' end of the second splice acceptor 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 splice donor sequence and the first splice acceptor sequence can assemble together (recombine) and the second splice donor sequence and the second splice acceptor sequence can assemble together (recombine), and the portions of CLRN1 protein encoded by the first, second, and third coding sequences do not overlap with each other, and when introduced into a mammalian cell (e.g., any of the mammalian cells described herein), splicing occurs between the first splice donor sequence and the first splice acceptor sequence, and between the second splice donor sequence and the second splice acceptor sequence, to form a recombined nucleic acid that encodes an active CLRN1 protein (e.g., a full-length CLRN1 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 splice acceptor sequence that is complementary to a splice donor 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 splice acceptor sequence that is complementary to the splice donor 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 splice donor sequence that is complementary to the splice acceptor 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
5 how to develop a strategy using four, five, or six vectors.

The CLRN1 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 CLRN1 protein, each of the encoded portions being, e.g., at least 30 amino acids (e.g.,
10 about 30 amino acids to about 1600 amino acids, or any of the other subranges of this range described herein).

In some embodiments, each of the at least two different vectors includes a coding sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding at least one exon and at least one intron of SEQ ID NO: 9 (e.g., at least
15 two exons and at least one intron, at least two exons and at least two 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 CLRN1 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 20%, up to 30%, up to 40%, up to 50%, up to 60%, up to 70% of SEQ ID NO: 1), provided that each of the encoded portions is non-overlapping with any other. In some embodiments, each of the at least two different vectors include a coding sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 3 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to
25 50%, up to 60%, up to 70% of SEQ ID NO: 3), provided that each of the encoded portions is non-overlapping with any other. In some embodiments, each of the at least two different vectors include a coding sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 5 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%, up to 60%, up to 70% of
30 SEQ ID NO: 5), provided that each of the encoded portions is non-overlapping with any other. In some embodiments, each of the at least two different vectors include a coding

sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 7 (e.g., up to 10%, up to 20%, up to 30%, up to 40%, up to 50%, up to 60%, up to 70% of SEQ ID NO: 7), provided that each of the encoded portions is non-overlapping with any other.

5 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 splice acceptor sequences and/or splice donor sequences are known in the art and have been described in Example 10.

10 **Example 12: Hybrid Vector Trans-Splicing Strategy using a F1 phage highly recombinogenic exogenous gene region (AK)**

At least two (e.g., two, three, four, five, or six) different nucleic acid vectors (e.g.,
15 AAV vectors) can also be used in any of the methods described herein to reconstitute an active CLRN1 gene (e.g., a full-length CLRN1 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,
20 incorporated in its entirety herein. As used herein, an F1 phage recombinogenic region (AK) will be used to allow coding sequence-independent recombination. The F1 phage recombinogenic region may be a 77 bp recombinogenic 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 recombinogenic region. Since the hybrid vector will be
25 constructed based on a trans-splicing vector as described in Example 10, a nucleic acid encoding an active CLRN1 protein (e.g., a full-length CLRN1 protein) may be generated using F1 phage recombinogenic region-induced recombination and trans-splicing. After trans-splicing, a nucleic acid encoding an active CLRN1 protein (e.g., a full-length CLRN1 protein) will be generated in a mammalian cell (e.g., any of the mammalian cells
30 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 CLRN1 protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the N-terminal portions of a CLRN1 protein described herein), a splice donor sequence positioned at the 3' end of the first coding sequence, and an F1 phage recombining region positioned 3' of the splice donor sequence. A second nucleic acid vector can include an F1 phage recombining region, a splice acceptor sequence positioned 3' of the F1 phage recombining region, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the C-terminal portions of a CLRN1 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 CLRN1 protein (e.g., a full-length CLRN1 protein). When introduced into a mammalian cell (e.g., any of the mammalian cells described herein) splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes an active CLRN1 protein (e.g., a full-length CLRN1 protein).

In another example, three different nucleic acid vectors will be used. A first nucleic acid vector can include a promoter sequence (e.g., any of the promoter sequences described herein), a first coding sequence that encodes a first portion of a CLRN1 protein (e.g., any of the CLRN1 coding sequences described herein) positioned 5' of the promoter, a first splice donor sequence positioned at the 3' end of the first coding sequence, and an F1 phage recombining region. A second nucleic acid vector can include an F1 phage recombining region, a first splice acceptor sequence positioned 3' of the F1 phage recombining region, a second coding sequence that encodes a second portion of a CLRN1 protein positioned at the 3' end of the first splice acceptor sequence, a second splice donor sequence positioned at the 3' end of the second coding sequence (e.g., any of the splice donor sequences 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 splice donor sequence and the first splice acceptor sequence of the second nucleic acid vector). In some embodiments, the splice donor sequence of the first nucleic acid vector and the second splice donor sequence of the second nucleic acid vector are the same (e.g., any of the splice donor sequences described herein or known in the art). In some embodiments, the first splice donor sequence of the first nucleic acid vector and the second splice donor sequence of the second nucleic acid vector are different (e.g., any of the splice donor sequences described herein or known in the art). A third nucleic acid vector can include an F1 phage recombinogenic region, a second splice acceptor sequence positioned 3' of the F1 phage recombinogenic region, a third coding sequence that encodes a third portion of a CLRN1 protein positioned at the 3' end of the second splice acceptor 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 splice donor sequence and the first splice acceptor sequence can assemble together (recombine) and the second splice donor sequence and the second splice acceptor sequence can assemble together (recombine), and the portion of CLRN1 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 splice donor sequence and the first splice acceptor sequence, and between the second splice donor sequence and the second splice acceptor sequence, to form a recombined nucleic acid that encodes an active CLRN1 protein (e.g., a full-length CLRN1 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 recombinogenic region and one of the different nucleic acid vectors may include a splice acceptor sequence that is complementary to a splice donor 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 include a splice acceptor sequence that is complementary to the splice donor 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 splice donor sequence that is complementary to the splice acceptor sequence in the second nucleic acid vector and the first nucleic acid vector will not
5 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.

The CLRN1 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
10 two different vectors include a coding sequence that encodes a different portion of a CLRN1 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 described herein).

In some embodiments, each of the at least two different vectors include a coding
15 sequence that encodes a different portion of a CLRN1 protein, each of the encoded portions encoding at least one exon and at least one intron of SEQ ID NO: 9 (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 three introns). In some embodiments, each of the at least two different
20 vectors includes a coding sequence that encodes a different portion of a CLRN1 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), provided 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
25 different portion of a CLRN1 protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 3 (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: 3), provided 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 CLRN1 protein, each of the
30 encoded portions encoding up to 80% of SEQ ID NO: 5 (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: 5), provided 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 CLRN1 protein, each of the encoded portions encoding up to 80% of SEQ ID NO: 7 (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: 7), provided 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 splice acceptor sequences and/or splice donor sequences are known in the art and have been described in Example 10.

Example 13. Hybrid Vector Trans-Splicing Strategy using Two CLRN1 Isoforms

At least two different nucleic acid vectors (e.g., AAV vectors) can be used to reconstitute an active CLRN1 gene (e.g., a full-length CLRN1 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 CLRN1 protein positioned 3' of the promoter (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the N-terminal portions of a CLRN1 protein described herein), and a splice donor sequence positioned at the 3' end of the first coding sequence.

A second nucleic acid vector can include a splice acceptor sequence, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein (i.e., the entire portion of the CLRN1 protein that is not included in the N-terminal portion) positioned at the 3' end of the splice acceptor sequence (e.g., any of the sizes of a portion of a CLRN1 protein described herein and/or any of the C-terminal portions of a CLRN1 protein described herein), and a polyadenylation signal 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 sequences of the two encoded portions do not overlap with each other; and no single vector of the two different vectors encodes an active CLRN1 protein (e.g., a full-length CLRN1 protein).

In some embodiments, each of the at least two different vectors includes a coding
5 sequence that encodes a different portion of a first isoform of the CLRN1 protein (e.g.,
SEQ ID NO: 3). In some embodiments, each of the at least two different vectors includes
a coding sequence that encodes a different portion of a CLRN1 protein, each of the
encoded portions encoding up to 80% of the amino acid sequence of SEQ ID NO: 3 (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
10 SEQ ID NO: 3), provided that each of the encoded portions is non-overlapping with any
other.

In some embodiments, one of the at least two different nucleic acid vectors further
includes a sequence that encodes a second isoform of the CLRN1 protein (e.g., SEQ ID
NO: 5). In some embodiments, each of the at least two different vectors includes a
15 coding sequence that encodes a different portion of a CLRN1 protein, each of the
encoded portions encoding up to 80% of the amino acid sequence of SEQ ID NO: 5 (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: 5), provided that each of the encoded portions is non-overlapping with any
other.

20 In some embodiments, each of the at least two different vectors includes a coding
sequence that encodes a different portion of a second isoform of the CLRN1 protein. In
some embodiments, one of the at least two different nucleic acid vectors further includes a
sequence that encodes a first isoform of the CLRN1 protein.

When introduced into a mammalian cell (e.g., any of the mammalian cells
25 described herein), splicing occurs between the splice donor sequence and the splice
acceptor sequence, thereby forming a recombined nucleic acid that encodes an active
CLRN1 protein (e.g., a full-length CLRN1 protein).

Non-limiting examples of such vectors are shown in FIGs. 1, 2, 4, 7-10, and 12-
24.

30

Example 14. CLRN1 Expression in HEK293FT Cells

HEK293FT cells were transfected with exemplary CLRN vectors. 48 hours post-transfection, HEK293FT cell lysates were prepared and CLRN1 protein expression was determined by Western blot. As shown in Figures 25 to 31, CLRN1 protein was detected in all tested samples. This result confirmed that CLRN1 protein can be expressed by exemplary CLRN1 vectors described herein.

Figure 32 showed that HEK293FT cells transfected with CLRN1-6eGFP vector expressed high levels of GFP as soon as 72 hours post-transfection. At 24 hours post-transfection, few HEK293FT cells expressed GFP following transfection with CLRN1-e6GFP vector at MOI 8.41E+04 and 2.53E+05. At 72 hours post-transfection, most HEK293FT cells transfected with CLRN1-e6GFP vector at MOI 2.53E+05 expressed GFP, while some HEK293FT cells transfected with CLRN1-e6GFP vector at MOI 8.41E+04 expressed GFP. As shown in Figure 33, CLRN1 was expressed at high levels in HEK293FT cells transfected with CLRN-0 vector, CLRN-3 vector, and CLRN-13 vector.

Example 15. CLRN1 Expression P2 Cochlear Mouse Explants

P2 cochlear explants from WT mice were infected 16 hours after plating and were harvested for RNA and immunofluorescence 72 hours after infection. As shown in Figure 34, CLRN1 was efficiently expressed in cochlear explants. As shown in Figure 35, outer hair cells (OHC) and inner hair cells (IHC) of P2 cochlear explants express Myo7a when transfected with CLRN-0 vector (1.3E10 VG/cochlea), CLRN-3 (9.9E09 VG/cochlea) and CLRN-13 (1.0E10 VG/cochlea). Transfection with either vector did not disrupt the structural integrity of OHCs or IHCs of the cochlea. Thus, Figure 35 shows lack of toxicity of CLRN1 constructs with viable and organized outer hair cells (OHC), inner hair cells (IHC) and stereociliary bundles. As shown in Figure 36, eGFP expression with CLRN1-3'UTR appeared to specify the transduction compared to CAG promoter alone. Cochlear explants infected with 1E09 AAV/Anc80.CAG.eGFP expressed Myo7a in the inner hair cells. Cochlear explants infected with 1E09 AAV/Anc80.CAG.eGFP.CLRN-3'UTR expressed Myo7a in both the OHC and IHC.

Higher transduction and co-expression of CLRN1 was seen in OHCs of cochlear explants infected with AAV/Anc80.CAG.eGFP.CLRN-3'UTR. GFP expression in cochlear explants infected with AAV/Anc80.CAG.eGFP.CLRN-3'UTR was restricted to OHCs.

OTHER EMBODIMENTS

5 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 the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

10 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. Section headings and any descriptions of materials, methods, and examples are illustrative only and not intended to be limiting.

SEQUENCE LISTING**SEQ ID NO: 1 - Human Full-length Wildtype CLRN1 Protein Isoform D**

5 MPSQQKKIIFCMAGVFSFACALGVVTALGTPLWIKATVLCKTGALLVNASGQELDKFMGEMQYGLFHGEGV
 RQCGLGARPFRESFFPDLLKAIPVSIHVNILFSAILIVLTMVGTAFFMYNAFGKPFETLHGPLGLYLLSF
 ISVALWLPATRHQAQSGCGLVMILFASEVKIHLSEKIANANYKEGTYVYKTQSEKYTTSFWVIFCFFVHF
 LNGLLIRLAGFQFPFAKSKDAETTNVAADLMY

10 SEQ ID NO: 2 - Human Wildtype CLRN1 Isoform D cDNA

atgcccaagc caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg
 tgccctcgga gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg
 caaaacggga gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga
 aatgcagtac gggccttttcc acggagaggg tgtgaggcag tgtgggttg gagcaaggcc
 15 ctttcggttc tcattttttc cagatttgct caaagcaatc ccagttagca tccacgtcaa
 tgtcattctc ttctctgcca tccttattgt gttaaccatg gtggggacag ccttcttcat
 gtacaatgct tttggaaaac cttttgaaac tctgcatggc cccctagggc tgtacctttt
 gagcttcatt tcagttgccc tttggctgcc agctaccagg caccaggctc aaggctcctg
 tggctgtctt gtcattgatat tgtttgcctc tgaagtgaat atccatcacc tctcagaaaa
 20 aattgcaaat tataaagaag ggacttatgt ctacaaaacg caaagtgaat aatataccac
 ctcatctctg gtcattttct tttgcttttt tgttcatttt ctgaatgggc tcctaatacg
 acttgctgga tttcagttcc cttttgcaaa atctaaagac gcagaaacaa ctaatgtagc
 tgcagatcta atgtactga

25 SEQ ID NO: 3 - Human Full-length Wildtype CLRN1 Protein Isoform A

MPSQQKKIIFCMAGVFSFACALGVVTALGTPLWIKATVLCKTGALLVNASGQELDKFMGEMQYGLFHGEGV
 RQCGLGARPFRESFFPDLLKAIPVSIHVNILFSAILIVLTMVGTAFFMYNAFGKPFETLHGPLGLYLLSF
 ISGSCGCLVMILFASEVKIHLSEKIANANYKEGTYVYKTQSEKYTTSFWVIFCFFVHFLNGLLIRLAGFQF
 PFAKS KDAETTNVAADLMY

30

SEQ ID NO: 4 - Human Wildtype CLRN1 Isoform A cDNA

atgcccaagc caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg
 tgccctcgga gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg
 caaaacggga gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga
 35 aatgcagtac gggccttttcc acggagaggg tgtgaggcag tgtgggttg gagcaaggcc
 ctttcggttc tcattttttc cagatttgct caaagcaatc ccagttagca tccacgtcaa
 tgtcattctc ttctctgcca tccttattgt gttaaccatg gtggggacag ccttcttcat
 gtacaatgct tttggaaaac cttttgaaac tctgcatggc cccctagggc tgtacctttt

gagcttcatt tcaggctcct gtggctgtct tgtcatgata ttgtttgcct ctgaagtga
aatccatcac ctctcagaaa aaattgcaaa ttataaagaa gggacttatg tctacaaaac
gcaaagtga aaatatacca cctcattctg ggtcattttc ttttgctttt ttgttcattt
tctgaatggg ctccataatac gacttgctgg atttcagttc ccttttgcaa aatctaaaga
5 cgcagaaaca actaatgtag ctgcagatct aatgtactga

SEQ ID NO: 5 - Human Full-length Wildtype CLRN1 Protein Isoform C

MQALQQQPVPFDLLKAI PVSIHVNVILFSAILIVLTMVGTAFFMYNAFGKPFETLHGPLGLYLLSFISGSC
GCLVMILFASEVKIHHLSEKIANYPEGTYVY KTQSEKYTTSEFWLTKGHS

10

SEQ ID NO: 6 - Human Wildtype CLRN1 Isoform C cDNA

atgcaggc cctgcagcag caaccagttt ttccagattt gctcaaagca atcccagtg
gcatccacgt caatgtcatt ctcttctctg ccacccattat tgtgttaacc atggtgggga
cagccttctt catgtacaat gcttttggaa aaccttttga aactctgcat ggtcccctag
15 ggctgtacct tttgagcttc atttcaggct cctgtggctg tcttgtcatg atattgtttg
cctctgaagt gaaaatccat cacctctcag aaaaaattgc aaattataaa gaagggactt
atgtctacaa aacgcaaagt gaaaaatata ccacctcatt ctggctgact aaaggccaca gctga

SEQ ID NO: 7 - Human Full-length Wildtype CLRN1 Protein Isoform E

MPSQQKKIIFCMAGVFSFACALGVVTALGTPLWIKATVLCCTGALLVNASGQELDKFMGEMQYGLFHGEGV
RQCGLGARPPRFSCYFLDPFMGLPTGVPHLLSLPCSTSCRREHTSERVQEPAGCFSAVRSKLHAGPAAATS
FSRFAQSNPSEHPRQCHS LLCHPYCVNHGGDSLLHVQCFWKT

SEQ ID NO: 8 - Human Wildtype CLRN1 Isoform E cDNA

atgccaagc caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg
tgccctcgga gttgtgacag ccttggggac accgttgtgg atcaaagcca ctgtcctctg
caaaacggga gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga
aatgcagtac gggcttttcc acggagaggg tgtgaggcag tgtgggttgg gagcaaggcc
ctttcggttc tcatgtatt ttcttgacct cttcatggga ctcccaacag gggtagccca
30 ttactcagc ctgccctgct caacctcttg caggaggag cacacgagt aacgagtga
ggaaccagct ggctgcttta gtgtgtgag gagtaaacct catgcaggcc ctgcagcagc
aaccagtttt tccagatttg ctcaaagcaa tccagtgag catccacgtc aatgtcattc
tcttctctgc catccttatt gtgttaacca tggtagggac agccttcttc atgtacaatg
cttttggaac acccttttga

35

SEQ ID NO: 9 - Human Wildtype CLRN1 Genomic sequence

aggagatact tgaaggcagt ttgaaagact tgttttacag attcttagtc caaagatttc 61
 caattagggga gaagaagcag cagaaaagga gaaaagccaa gtatgagtga tgatgaggcc 121
 ttcactact gacatttaac ctggcgagaa ccgtcgatgg tgaagttgcc ttttcagctg 181
 5 ggagctgtcc gttcagcttc cgtaataaat gcagtcaaag aggcagtccc ttcccattgc 241
 tcacaaaggt cttgtttttg aacctcgccc tcacagaagc cgtttctcat catgccaagc 301
 caacagaaga aaatcatttt ttgcatggcc ggagtgttca gttttgcatg tgccctcgga 361
 gttgtgacag ccttggggac accgttgttg atcaaagcca ctgtcctctg caaaacggga 421
 gctctgctcg tcaatgcctc agggcaggag ctggacaagt ttatgggtga aatgcagtac 481
 10 gggcttttcc acggagaggg tgtgaggcag tgtgggttgg gagcaaggcc ctttcggttc 541
 tcatgtaagt agcaattgca tttgagttat ttaatgcttt aggcagactc ttcccagtg 601
 tgcgaggaat tatatttgag aattttgccg tgtttactgc aggacttttt aaatcggtgt 661
 gaaccatag aaaaacctat gactctgagc aatttcttct tcctagtttt tattatttta 721
 tacttgcttt ttattataat atagagttaa ttcattgtta cataattaag gtttttgaa 781
 15 atattggcaa ttaagatgct taagtattaa tatttatgta aaaaattatg gagtcttttt 841
 aaaaaagtaa acttggggaa ataggaaagc tgtaaagaat gatctttatg ctttttggtc 901
 tttataaaaa gaaccaaggt catgggctcc gtatttaacc aggttgccac ctttctcatg 961
 attttgtttc ctgctcccca ctccctccca ttattcctgc taagaccttt cctgctgcta 1021
 aatattcagt tttcattttt aactaatttg gaatcatttg gctatagaaa tttaaaatga 1081
 20 tctgctgtgc taactgggaa agaaatggat gcctatttag tatagaacat tttaaactga 1141
 ttgacctgca aatcatgtag agaatatgag agagattttc ttgttgatgat ttttgatgaa 1201
 tggaagtgt atccacagta tttataacct gtttatctta agaagagaat ttttaaaaat 1261
 taccatgtga ataggcaact cattaaatga aaattaatag gaagtcattt gttatatctc 1321
 ttacaacaca cattcagaag ttattattat ttcagaaggg ctggtttgga acaaccttat 1381
 25 gaagacacag tcagtaaatt actgcataaa tctactcttca ggaaaggagg ttaccaactg 1441
 aagcatttaa aatgaattat tattttgccc aggttttttt tttctttcta gtataggtag 1501
 aaggctaaat taattgaatt tattattaac atatgcagtg cctaattaaa tttcagtgtc 1561
 ggtctattta tatttctgca acattcetta tatcttctta gcagtcattg gacaccaacc 1621
 ttcagctcac ataggttact aagtgatatg aattttcata gggctccaga aaattttcaa 1681
 30 gaattggttg ttagcttttt aattgatgaa gtggatacca gttcttttca ctgaatggct 1741
 tttattcatt aaggtaatgg ggctgttaga gttgcttagt tttcctgggg aagggggaag 1801
 aagaaaacaa agcagaatgt catgtgatat gcaactgtat taaaaaacg aaaaggaaaa 1861
 aagttgagag agatgattta accgtgagtc accggcagcc aaagcgtgag taaagcttct 1921
 cacagatgaa tttagacaaa agcggagaag gtactggtga attttctgga gcctttacat 1981
 35 tttctacagt gaaatggaga taaactttac tcatgccata ggacatgttt caaaacaata 2041
 ataagatgtt ttctgaacac ttactacata ctaagcactt tatatgcttt gtctcattta 2101
 atccttacac agccacattc ttctgggggt tagcgaatga tttttgtggt tgtgtcctat 2161
 gcttgtcctg tctaaggatg aagttgttct aattgggtgc ccctcctttt gctttctgtg 2221

aggacttgca gaactggtgg ggtttaaaca gtaccctcac ttatctcaca gaatttcatt 2281
 agctcccaga tacccttgac attctcccc tagcctagtg aagaaaatct tccatttact 2341
 tgttcattct gcagtgcacag ctccatcaat atacaataga ctatacatat taagtgtact 2401
 gtatatacta tacatgttaa aaatctcatt catttttggtg aggcccagct aagaatactt 2461
 5 acagtagagc tttttttttt ttccctaagca taaaagtatc tttttcaatg cagcatgaga 2521
 cagagttggg aaaacccaaa taaatagatc caatggactc cccaaagagg ataatttca 2581
 tttaaataaa caccctctc agtggttaaaa ctttctaate aacatgcctt tgggacacat 2641
 tgcaccctca aagtttacac tcccattgca acgcagcttt gtggttcacg ttttttccat 2701
 tcagaatgtc attaccctgt caatgatgtt tcatcaacgt ttgcttggat gagaatcctc 2761
 10 tgatattctt cctgatagaa atgtataagc cctgttcata taaatgaata aaagatctaa 2821
 ctttactttc tcagtagtgg cttccttgga gcaaaaagca gggacctcca gagagctcag 2881
 gtggatgact cttttctgtt tcttccagag ctcaacttac aattagtgc caattcattt 2941
 ccagaatgt cttctttctt attgtgcctt tagaaagtta ttaagcaaac atttgaattc 3001
 acagaatctt accagtgtaa gaggaatgga aaaggtaact tatcaaggta acaatcactt 3061
 15 cgtggccagt tttttcggct cactgcaact acccctcctg ggttcaagcg attctcctgc 3121
 ctacgcctcc caagtagctg ggactacagg cgtgcaccac catgcccagc taattttttt 3181
 tttgtatttt tagtaaagac agggtttcac catgtgggccc aggctggtct catggcaagt 3241
 tttctttgtg ttgtcatgtt attatcaatt aataggaatt tatatttcag ttctgttagg 3301
 tggataaaca ctattttgca tacctaaatg tttcatttat atcagcactg gccataaaaa 3361
 20 atatactata agcaggccgg gtgcagtggc tcacgcctgt aatcccaact tttgggaggg 3421
 caacactttg ggaggacaca gggtcaggag atcgagacca tcctggctaa cctggtgaaa 3481
 tcccgctctc actaaaaata caaaaaatta gccgggctg gcgggggccc ctgtagtccc 3541
 agtacttgg gaggtgagg caggagaatg gcacgaacct gggaggtgga gcttgcatgtg 3601
 agccgagatc tcgccactgc actccagcct tggtgacaga gcaagactct gtctcaaaaa 3661
 25 aaaaaaaaaa aaaaaaaaaa aaaaaatata tatatatata tatatattac aagccacaag 3721
 ccacatatgt acttttaaat gtccctagtag ccatattaga aaacaaaagt aaaaaggaat 3781
 agatgaaatt aatttttaag atttttttta acccaagata tccaaaatat tatcatttta 3841
 acatataatt aaaaatttat tgagatgttt tacattgttt ttttttttct tgacgctgtc 3901
 ttggaaatct agagtgtatt ttacatttac attacatcta attcagtcta gccacatttc 3961
 30 acatgctcat ttttttgtgt gtggttttat ggagcagaga gtttaatagg caagaaagaa 4021
 aagagaaggc agaagaaaat ggctccctg tacagagacg cgggggtggg gcgctccaaa 4081
 gccaaaagag gaggtcccta agtatggtag acaccagcca ggaatatatg cagtgtctgg 4141
 aggaggggat gtctgatttg catagggtca catgctcatt tttatggcta ctgtattgg 4201
 cagtacagat ttagatgggg atttgctctg aacaagttgg tgattgatgg tgttcatatt 4261
 35 ttaattgaat ttttctggg ttctgtctga cacttgatg tgtgttagtt tcatgtgcaa 4321
 tgcttggtgc acttttaaaa ccagatatata tggcataaca tgagaatgaa aaatggacca 4381
 gaaaaatagt ttggcaatgt agtcatgttt gttcctatta aatgttcctt attgaccact 4441
 ctatctcttt taattataac aagaatctgc cctgccagca tgcccagtta cgctgggaaa 4501

acttctgcct catttactct ggctgattct catccactta tgggtcagtg gttcattttc 4561
 tagaggtcac cagcattcat acctagcata caatttcatt cattataatg aaggaatgtt 4621
 ttcccttcaa agagacacaa ctagtgggct taatttttct tgatatgtca cctgtaaaat 4681
 tttaatgatg atgttttaac tctaaatgta gccatcaaga caaaaactgc aaattttgag 4741
 5 cctcagtggtg tgtgggtggg tttctgtttc ggtaatttga aacattgcag aatatcatca 4801
 aaatatgata cccaagaatc atatggtatc aatcattcct agatatactg atctattcat 4861
 tgccaagata gttcaatgag ctggcaaaaa catatggaat tattttctta aaatgtgaaa 4921
 aataaaattht aaccaatcat gtatcacagc ttgcaacttt agtcatactt tgaaaagcat 4981
 tttaatthtg acctcatgat tgaaaattta taaaaagctg aacagaaatt agttcacttc 5041
 10 atattthtga aaagcagagt ttctthtacta atgaggcat ttgacccaaa ttggagagaa 5101
 aatgthtgaat ctacttctgt gagcaagcag gtggcttctc aaacacatgt tgggatgaaa 5161
 tggthtgggc tcagggtctc agtgctgtc actgagagtt ggcaactctc atctccatgg 5221
 tctctccaa gtgtgactct tgtctcttgc tgacctgacc tgcccaagt gactcactgg 5281
 tcatgacct gcacacctg cgtctctctc atcacctgc cgatggcaga gctacaaaagg 5341
 15 tctthtgatg agctctgtct gatatctgt gthtccccct atggtctgtg tggaagcagg 5401
 tatgggggtg tgtaaagggg aagcctatga agthcatctg caaagactac ctggttaggg 5461
 gaggagagga agaagctata tgcaccttht caccagcaag catgggtctc tctgccttht 5521
 agcttagggg tctgtthtgc tagtctcact cacctattaa aacagtccag caaatagagg 5581
 tthtghtht ctcccattht aaaaaagca attaatggaa tagaagataa taatgtatga 5641
 20 gaagcactat tgtgaaagaa aaaccttcaa ctctctcag ccaaataatt gcttctcct 5701
 ctgtccttht cagaccttht tgtthtctc attatttht aacaactata ttataaatat 5761
 ttgagaatgt gthtthtctc gcaaggagat cttaatcccc aagaaggcag gagctgtgta 5821
 ttattcatcc cagtgtctc tcaaaccagg gcctcagaca gtgcatggc caaagaggta 5881
 cttaataaac gthtggtaaa ctgacactat tgaaattaa caacctggat ttgaggthtgg 5941
 25 tctctgccac tcacaagaga ttacgttht agaaaattc atcacttht tgattthtctc 6001
 ttctthtcatc tgtatatggg agacgatact aggtgatttht tgacatctc caccgthttaa 6061
 atgctctgtg atctatacaa cgagggtctc gctgtthtctg acaagthtctc tccagcttht 6121
 cagthtgcata acctthtctaa tcttagtcac atgatgactc cactgacaga tthtthtggcca 6181
 ccatcattag acatgctgag ttacgtgtgc thtthtctc atctcaaaa ctcatgatttht 6241
 30 ttaaagthttht ctgaaatatht taccatttht caggatccag atggatttht tgaccaaagt 6301
 ggatgthttht thtctctccc attacaatht thtactthttht gtgtgggaaa ttgcatgtta 6361
 aagaaaggga aattgaagaa tgggatgctt tggaattctg gcaagatgga ttagthtgggt 6421
 ccagaaagta ggggcagcca caaataccga aataaatgag atcgtattat tgagaaagca 6481
 caaatggaag aaggtcaaaa gcaagaagaa gctgacatcc tgcttctcctc aattthtthtct 6541
 35 thtctctgtht thtcaagaaa ctctcttht aagcctthtct gaaaaactcc actthtcttaa 6601
 atctaacttht ttaaactgat aatggcaaga agthtagaat gaccaatatht thtaacactc 6661
 caagagthttht tgtthtthttht tgtthtthtga ctgggtctca ctctgtctc cagcctggag 6721
 tgcagthtgggt tgatcacagc tcaactgcagc ctthtacctc cctgccaag gaatctctcc 6781

acctcagcct cctgagtaga agggaccaca ggcattgtgcc actacacctg gctaaactttt 6841
 aaaaaatttt tttgtagaga tgggggttgcc caggcagggtc tcaaactgct gggctcaagc 6901
 aatcctcctg cattggcctc ccaaagtgtc ggaattatgg gcataagcca ctacgcctga 6961
 cctctccaag ggcattcttt acccagaaga ggaacttgcc agaacttatc ctccaattgg 7021
 5 tgaggaatat ggagaaaatg actttaagca aaggaacttc tggttctgcc tacctaattcc 7081
 agaaaaagaa gttttatttc tcccttcccc tagtaactat cttcccatat tcacataaaa 7141
 aagtacagaa tcaacattgt tcaagaatta taattttact tgtaagcaca tgtgcacacg 7201
 cacaccata taccttcctt ccctttaaat catcccacac cctaatagta gtaaaatcat 7261
 tgacccgagc atacctggga gaggaagagg agtctgacag gggcagggtc taagtggcac 7321
 10 tcctggaact taaccctggt gtatatgaac ttacctatt gaaggatgac tcctcaactg 7381
 ttctcacaat ttgctgctct gctttctttt ctaatttctg aagggtgactc atcttcccca 7441
 aggactttca gacttctcag aagaaaaaaa tattgggtgg gtctctgcca ctggcaaaag 7501
 attagacttt gagaatcata aaagtatata agtatatact cattaatatt gaattactat 7561
 aattaatatt atgatattga tataatgata gaatgatatt gataaaagca atattcaata 7621
 15 atgaatatta tttcagctgc ccacttattg ggtgcctcat aggtgccagg cattttgtat 7681
 gtattatcta caacccttac atgggacata ttatgatgct gtttctcttg aagaaatatg 7741
 gaaactggaa acagagaggt caccacaatt ttccaaagtc acatagctaa taggtagcag 7801
 acttgggatt caaattcata tgcatatggt aaatcatgct cttcctctgc tacattttgc 7861
 ccccttagaa tatgaaaaag ggatacaaag agatgaagaa aatatgtaag attatccttc 7921
 20 aatttcaact tcttttttaa agtttttttt attatacttt aaattctgaa atacatgtgc 7981
 agaacatgca ggtttgttac ataggtatac acgtgccatg atggtttggt gcacccatca 8041
 acctgtcatc tacattaggt atttttccta atgctatccc tcccctagtc cccacccac 8101
 cgacaagccc cggtgtgtga tattcccctc cctgtgtcca tgtgttctca ttgttcaact 8161
 cccacttatg agtgagaaca tgcgggtgtt gggtttctgt tcctgcgtca gtttgcgtgag 8221
 25 aatgatggtt tccagcttca ttcattgttc tgcaaaaggc atgaactcat tttatggctg 8281
 cataccattc tatggtatac atgtgccaca ttttcttaat ctagtctatc attgatgggc 8341
 atttggctta gttccaagtc cttgctattg tgaacagtgc tgcaataaac atatgtgggc 8401
 atatgtcttt atagtagaat gatattataat cctttgggta tagaccagc aatgggattg 8461
 ctgggtcaaa tggcatttct ggttctaggt ccttcaggaa ttgccacact gtcttcacac 8521
 30 atagctgaac tagtttacac tcccaccaac agtgtaaaag cgttcctatt tctccacatc 8581
 ctctccaaca tctgttgctt cctgactttt taatgattgc cattctaatt ggagtgagat 8641
 ggtatctcag tgtggttttg attagcattt ctttaatgac aagtgatgat gagctttttt 8701
 tcatgtttgt tggccgtata aatgtcttct tttgagaagt gtccgttcat atcctttgcc 8761
 cactttttaa cgggggtttt tcttgtaaat ttgtttaact tccttgtaga ttctggatat 8821
 35 tagtcctttg tcagatgggt agattgcaaa aattttcttc cattctgcag gttgccggtt 8881
 cactctgata atagtttctt ttgctgcgca gaagtttttt tagtttaatt agatccatt 8941
 tgtcaatttt ggcttttggt gccattgctt ttgggtgtta gtcataagat ctttgccac 9001
 gcctatgtcc tgaatggtaa tgccataggtt ttcttctagg atttttatgg ttttaggtct 9061

tatgtttaaa tctttaatcc atcttgagtt aatTTTTgta taaggtataa ggaaggggtc 9121
 cagtttagtt ttctgcatat ggctagccag ttttcccaac accatttatt aaatagggaa 9181
 tcctttcccc gttgcttggt tttgtcaggt ttgtcaaaga gcagatgggt gtagatgtgt 9241
 ggcattatTT ctgaggcctc tgttctgttt cattgggtctc tgtatctgtt ttgatacaag 9301
 5 taccatgctg ttttggttac tgtagacttg tagtataatt tgaagtcagg tagcgttaata 9361
 cctccagttt tgttctTTTT gcttaggatt gtcttgacta ttcaggctct tttttggttc 9421
 catatgaaat ttaaagtagt tttttctaata tctgtgaaga aagtcaatgg tagcttgatg 9481
 ggaaaagcat tgaatctata agttactttg ggcagtatgg ccattttcat gatattaatt 9541
 ctccctatcc gtgagcatgg aatgtttttc catttggttg tgtcctttct tatttccttg 9601
 10 agcagtagtt tgtagttctc cttgaagagg tccatcacat cccttgtaag ttgtattcct 9661
 aggtattttg ttctctttgt agcaattgtg aatggaagtt cactcataag tttgctctct 9721
 gtttgtctgt tattggtgta taggaatgct tgtgattttt gcacattgat tttgtatcct 9781
 gagactttgc cgaagttgct tatcagctta aggagatttt gggctgagac gacagggttt 9841
 tctaaatata caatcatctc atctgcaaac agagacaatt tgacttcctc tcttcctatt 9901
 15 tgaatacgtt ttatttcttt ctcttgcttg tttgccctgg ccagaacttc caatactgtg 9961
 ttgaatagga gtgttcacca cctattttta gaatagtatt gaagcctcac aaaagctggt 10021
 tctcatgtaa ccatctgaga atatttggtc ttatgacttg aattcattca ttgccttttt 10081
 atttcacatt ttagtgatcc tgatgtctaa atcttaatct ttgatccttg caaggtaaaa 10141
 tagccaagtc aagcctgttt aataatattg gttgaggaag tcacatgctt atgatcaatc 10201
 20 tttgggttat gtaattatat taccttaatg ttggcagttt aggtgtaagg cagagatatc 10261
 tgatcacatg tgtggttagc taatttaaga tcaactgcaa ctaaaatctt catggtatga 10321
 tcttcaaagt tagctacttt gaccacagca atgatttcac cacagcaatt aacaaaatgg 10381
 cagactcttt cctgaggtgg catgaacagt tttaaaacaa agtcaaggac caaaagaaaa 10441
 gcaggcacat ggcatttgat tcatttcaaa aactagtatt gtattaagag ccaaggggat 10501
 25 agaattgtag catcaattaa aatcttggtt gaaaaaaaat aaaaacaaac gtccattttt 10561
 atctctcaaa tatattaggg ttttcataaa gttataggtg tattttttaa aaaacaaaac 10621
 tcatatacat tagactgaaa aatttgcttg tcatctcatc atgcagctaa atgcaattgt 10681
 ctatggcgag atacactctt attagaggta tgatagcact aactaatagc aaactttgta 10741
 cctggttagtc taatttatgc agggttcata tttcgctccc tctcagcatg ctgtaactgt 10801
 30 ggcaaagcca ttctcaggag ttattacccc aacataatca taccctgtg gattaggagc 10861
 agttaaagtg gtctgttat cagagacaca tatgtgccac agccgctgtc atccctaagc 10921
 caccgtgggt gattaagact cactgatggg actacctctg aataggcttg agtgaggtgg 10981
 atacacttca gctgagagaa attcaggtaa ggggctgaga aaatcagatt ttgaatgggt 11041
 ttatcatacc atcaggtctc cttttaagtg ctggggctcat ggatttcatt caacctgacc 11101
 35 acatagcctt tggaagcttg gctcaatacc tgagtgtgag attatgggtg atattaaagg 11161
 agatgaatgc attgagctga tgtcagagaa tgtcttttac tggattttca taatgactgg 11221
 ctgcagatgg gctgagggga aagtcagatc aagagcattt ctgtaagaag aaagaaatct 11281
 tccctcttat tctctttcaa gaaaatgaat agctgagcac caagaggcca aatacttttt 11341

taaaaaacac atccttttat gtagagaagg acaggttgag acgaaaaaca ggacctctga 11401
 aactgtcttt atagctttaa tctaggaaga aattgcggca ccattgctga catcattatc 11461
 agaggctgcc ttagttctga gagcttcaca gatggccttt tctgcatttt acatctggcc 11521
 agatgaaggc aaaaagggtg atgaaaccaa aactattaga tcagtgggtcc ataactctgg 11581
 5 ctgcacttta gaatcatctg gggagtcctt aaaactactg atcctgaggc tccatcccag 11641
 accaattgaa tcaggatctc tgggggtggg acctgagcat tggaatgctt taaaagtcc 11701
 tcagggactc taatgtgcag ccaaggctga gaatgactga ggtggatggt ggccaagaca 11761
 ggtgaggcca agagttagaa gcccttactg ttaggggaagg cagaagccac aaggaggagg 11821
 ggagggggaa ggagcagtat ttagcatttc ttccacacag ttgggggggt ttctgatcaa 11881
 10 aattaggctg ggatctttcg ctccatttt catgaaggct ttatgctttg tccacagcca 11941
 cctggcctca gggcaatgag caatctgatt gatgaatttt cagtaagaaa ctgagcacac 12001
 ttggctctca gccccagtgg ctccctctgt ttccaaatct gccaccagt tacaggagcc 12061
 tgctcaccaa ctggttgggt taaagaagtc ggctctgtcc ttggcaggga ggcttcagc 12121
 tgtctggccc tgtctgtgac tgcgtgggtg aagctgccta atttggggaa cttgatggaa 12181
 15 gatctaagct atgttctcta acagttttat cagaaataaa gttaactttt gacctccatc 12241
 tgctgtctc ctgtggaagg gctctgtctc ttccaaagag actctcaggg gtctctctta 12301
 gaggtgtgtt atcagtccea agatcatttt agaccagcca ttacagagga tgtccaagaa 12361
 attgcacaag ggaaatagaa gtaagaatga gagcaatata tcagatagta aggaaagtat 12421
 gtgactggtt tcaataaagt aattaaccat aaatccaaag ttctctgatt agttatagca 12481
 20 atagtcatcc acctgggtcag gaaaaaaaaa gtaaaagact gagctgcaag atagaaagtt 12541
 tgctgaatt ccacgtactt caaggactac tatggaggat tccttgggtc caatgcagag 12601
 acatgtactc tgaccaccca tgggccaaat cctcttacc caccctcagt gattcctcct 12661
 gggacctcac tacattgagt tcttacattc cccactcttt tcgggggaaa tataaccctt 12721
 ctgctcttct catgatgtta aaaatattta gacaattact taaaatttac aacaatcttc 12781
 25 cacataaatg atctcactta tgtttcttgt gagctctgtg atttatttct attatcatct 12841
 atgctgatag ttaaggaaac tgagatatcg agaccaacat gcatagtaaa tggcaaaacc 12901
 agatgaattt taaactttcc taactccaaa agccacatcc tgcccaaccc gccatgctgc 12961
 cgctcagtta atgcctggct gtttgtctcc ccattggccc ctctacccat tgtgctttga 13021
 tggcacactg tattccaact gcctgagact cctgggttaat gccattacgt acagacttag 13081
 30 gttgaattta ctaggatttt taagatttgt aggataagag atatgactgt tagactggaa 13141
 tcagcaatag ataaaagggt aacaaagttt cagaataaaa tataatagaa aaccccagca 13201
 gatagaagta aatgaatgg taagacaact gaaatgaaag ccatacattt agaaatatca 13261
 aataaacac agattaatgg cagacaataa aggaacatac ttagcagtta gtaaaaacac 13321
 tttttacctt attcttatta cctccagta accttttttt tttttttttt tgagacaaaag 13381
 35 tctcgctgtg tcgtccaggc tggagggtag tgatgtgatc tcggctcact gcaacttctg 13441
 cctcttgggc tgaagagatt ctctgcccc agcctcctga gtagctggga ctataggggc 13501
 ccatcactgc acctggctaa tttttgtatt ttagtagag acggggtttc accatgttgg 13561
 ccaggctggt cttgaactcc tggcctcagg tgatccacc gcctcagcct cccaaagtgc 13621

tgagattaca agcgtgagcc accaagcccg gctgtaacct attgaaaata acattacttg 13681
 acatgtgaga caaattatatt gtaagttaaa gagtttatgt gccttcaatg ctcccaccct 13741
 tccctcccct aaaaagatta ttgagtgcc atgatgtgcc taagccttct ggtagactct 13801
 gagaatgtga agagagttag aggatacttg ttcaacaacc cagaatctag cagaagtgtt 13861
 5 ttgaaatgac cagccacttg ggagagctat gctagcatat cattcaggat gggctgggtc 13921
 atgttttaaat aacaagtaag tttgaacttt ataggtttga aacaaaaaag atgtgtttct 13981
 tgctcacagt tcatatTTTT tggagattgg ctggagcttc attccaggat actgccacca 14041
 tctggaatgt tgacagttac cgtaacagag aaagagttgt ggaggctgtc aactggcaa 14101
 ttaaattcgc tgacctggaa ctgacacatg ccacctccac ttatatTTTA ttggccaagg 14161
 10 cttgttacac agccacatct aacttcagag aggtcaagat gagcaaatcc taccacgagc 14221
 tatagatgga tgagtagcac tcattattat cacaattatt ttatttgTTA taaaaactcc 14281
 aagaaggaga ggtactctat gtgaggaata ggcataaggaa aatcagaggg aagttcttaa 14341
 agatgagcta gatttcacta gatggcattg aagaaacatt ccaagtaaT gaacagcata 14401
 agcaaTgca tgaagaactt attgtggTcc tagtacagac gatgcgtgag agtgtggagg 14461
 15 agggaaaagg taagactgga gaggaaggca ggaaccagaa caggacagac tgtgttTact 14521
 gtcaggcagt taagtctatc ttgtaggcaa catgaagcct tTTaaggaa gcaagcagct 14581
 atgtgacagg acagaagatg gattggaaga aaatcaaaag agaagtggg accagtcata 14641
 aggtcctcc aatatttggg gatccaaacc aaagcactgg cagagaaata gaaaggaaagg 14701
 gaaatatttc caaaatattc aaaacagaaa ccaagagaac ttgatgacag aagatgaacc 14761
 20 caggtttcta ctgaatggac agttgttaca ttctctgaga taaggaaTac agaaggaaaa 14821
 agttgagagg gaatatgaaa ttattttTaa tcatgtTaa tttgatcact tgtggaacat 14881
 caaccagaga tgtccatcag ttacaccaat ttgtaggTtt tgagagaggc ctgaggtaga 14941
 ggagagaag tgggaatcag cagattagcg gcagtagttg gaaccataac tgaagatgag 15001
 atttcccagg gaggggagtt gaacaggaag gtagacaaag gctgagtgca gtggctcata 15061
 25 cctgtaatcc cagcactttg ggaggctgca gtgggtagat cacaacaag atcaggagtt 15121
 caagaccagc ctggccaaca tggTgaaacc ccgtctctac taaaaaaaca caaaaattag 15181
 ccgggcgtgg tggcacacac ttgtaatccc agcctcagga ggctgaggca ggagaattgc 15241
 ttgaaccgg gaggcagagg ttgcagtgag ccaagatcgt gccattgcac tccagcctgg 15301
 gcgacagagc gagactccgt cccaccgac cccTaaaaa agaaagtaga caaagatgtg 15361
 30 tcctagaaaa cactaatatg aatgggtaga gaggggagac ttgtcaagga gatcgagggg 15421
 agagtcagt aggtgaaacc ttagaggata gaccttcacc aaggaaTaa cagtcaacag 15481
 gcctaaatgc cacagagatg tcaaTgaga gacactggaa atggTtcttt gaatattaca 15541
 gccagaacat cacaggagac catttccaaa gcagtcttga tgaagtggTg gggaagggag 15601
 gtccctgaag gagctaagga gggactggga gatcatgacc cagagataag tgttgCaggt 15661
 35 ataaaaggga aagactgaga tcaggaaata gccacaggat ccctaaggcc aggtctcggg 15721
 tggggTggTg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tataaaatgg gaagaccttg 15781
 agtataattc tgtactaaga agatagagca gtagagagga aaaggctgaa gacagatggg 15841
 agggtaaata gtgaattaga aaggtctgtg caaagatgag aaaggatgag tctgaaagca 15901

cagggagaag ggccagcctt ggacaggaga aaacatttct ttcactaaga ataaagggaa 15961
 ggttgggtat acaccacaaa gaaggtgtag atgggtggca atggagtctt gtcaagttga 16021
 gggcattcca tgatagcctc acctttctct gtgaaatgag agagttagt ttacaagata 16081
 tttctgagca cttcataagc caggttcata gactgaaaga agtggaggtg gagtgtgata 16141
 5 ggtccttaag aataggggaa gtttggaata tctgacaagg gacagagagg aaaaaactag 16201
 aaaaggcttt gcagaatgtg ggccacaga tcagaggcta aggggtcccc atttgtgcag 16261
 gagagtaagg gcagggaggc aaggctcaca gcccaaaata tagagcccct caccaaatga 16321
 ctgcaggagg gcagctttcc tatgagagca tccctatcac tgttttctact ccgagtcatt 16381
 aacttacgac ttactcagct ctgtttcgta atagcagact cgagtaatga gggtagtaca 16441
 10 gcctctctct gcatgccaa gtagtcagcg tggatttctt ttttcgcttt ctctctcctg 16501
 tggcttaggt gccttctgtt ctgctaccag gatagagaac ccagtgacta gtttcttcta 16561
 gctctctttt tctgactagg tatcttgtca gaaatttctg cttaccagac ttcattggaga 16621
 gggaatcaag ctttgaatca ggggtgaaaa agtagagctt aatatatata ttacaaaatg 16681
 ccactcacgt tcttgagggt acctgtatc tataccacaa ctagcattct tttagaaaagc 16741
 15 accattaccg aagtaatccc tttctggga attcacccaa aaaggttatt ccacttatc 16801
 ccccatctcc aaaataaaaa agaaaaatgt gtgtgcttag agatgttctt ggaagcatga 16861
 gctgtaatac tgaaccaata gaagacgaac taacagattg cagggcatcc gtttgcaaaa 16921
 aaacttatgc agttatctaa atgatagtta tgaagacaat atgtataaca cattatatc 16981
 tgagtgaata gaacagaagg tgatttcaaa actgcattgg gataatagta atataggaaa 17041
 20 tagtgagatg aacaaagatt tcaaaggaac aaaaataaag aaaaattttg ctttttatat 17101
 tggtaggcgt atgggtgaaa ttccaatttt aattttaatt tcatagttat aaaattgttc 17161
 ataaaaaaag gtccccata gacagttggg ctttgggaca aactaacaga aacagagtag 17221
 gaagaaaatt catcttcctt caatccccct tctctgttta aaacaaaaca aaaaagagct 17281
 ttgtcatgtt caggtgtgca acgaattctt tttccaaatc tggaacttta catctgctat 17341
 25 taaacaggag tcagtttcca tgtaacatgt tgacaatccc ccaagtgtgt tggaataatt 17401
 ttttttaatg aggagatttg aaattccatt tcaattgcc aactgcctct ttcaacttct 17461
 aaaaacaaag taaaacaaaa caaaaacaca ctgggtccta tcacccctc ttgctactac 17521
 tattttatct ccattgccct gaattctttc caaacttctt tccaccagc tttgatttgt 17581
 tttcagtcgg gtttattgag gcataattta cgtacagtaa aattcatcct tcttagattt 17641
 30 agaggtctat gcattttgac taatgcata tggcttataa tgactaccac aacaaagata 17701
 tagaacacac ccatcactcc cgggttctcc ctgtcccttt tttgggtccat ctctctcct 17761
 acccccaacc ccttggaac cactgatctg ttttctgtcc ttatcatttt gctttttcca 17821
 ggatgttgta tataaggaat catgcagcat gcagcctctc gagtctgact tctccagtt 17881
 agcacagtta tttaagatct atccgagtta ttgtgagtag cagcatcttt tttattgctg 17941
 35 actagtattt catcacatgg atgggccaca acttgtttat ctgttcacct gtcaatggat 18001
 actaagttgt ttccagtttt tggcaaatat gaataaagta aacatttgca tacagatttt 18061
 tgtgtggaca catgttttca attctcttag gtaataactg agcaatggga ttgtgggtt 18121
 atatgttaag tctatgttca gttttctaag ttctgaaaca attggatatc tatatgcaaa 18181

aatacaaaat taaccttgac tcaaacttgt accatacaca aaaaataacc tgaaagagat 18241
cacagggtgta aatgttaaacc tagaactaaa aacttcaagg agagaaacat agcagaaaaat 18301
atgtgtgacc ttgcattagg caaagacttc ttagatttga catttgaaac atgatacata 18361
aaaaatcttg ataaattgga gttcataaaa ataagaacta ctcttcaaaa gacactgata 18421
5 agagaatgaa aatacaagcc acagacagag aaaatatttg tgaatttctt atctgataaa 18481
gggtttgtat ccagcatgca taaagacttt tcaaaactca ataataaaca atccaataag 18541
aaatgcacaa aagatacaaa cacatcatcg aaaatgacct atgaaaggca aataagccca 18601
caaaaaaatt ctcaacatca ttagacactt gggaaatgca aattaaaaca aactgagat 18661
aaactacata tctattaaat ggctaccact ttaaaaacct gtcaagtgcc agccagaatg 18721
10 tggaacaagt aggactctct tacattgcta gtgggaaggt gaatggtaca gccacttttg 18781
aaaacactcc gcagcttctt atagttacac ataggacctg ggggtagagg atggattgac 18841
tgtaaagggg caaaacttcc tggtaggggt ggggaaagtt tttagaaggg tgaaattggt 18901
ctatatcttg attattgtgg tggttacaca actgtctgca tttgtcaaaa ctcacagaac 18961
tgtactactaa aaaggatat ttttatgctc tgctaatttt actttaatct taaaaatagg 19021
15 aaggaaaaaa taaaatcaat gccactgtgc gactttgggc aagttacttc acttctctgt 19081
gccttggctt tttgaaatct atacattaag gataaaataa taccttcctc atactgttag 19141
aattaaatgt gctaatttat gttttatata tataaagtac ttggccgggt gtggtgactc 19201
acacctgtaa tcccagcact gtgggaggcc gaggtgggca gatcacttga ggacaggagt 19261
tcaagactag cctggtcaac atggtgaaac cctgtctcta ctaaaaatac aaaaattagc 19321
20 tgggcttgggt ggtgcgtgct tgtaatccca gctacttgag tggctgaggt gggaggatca 19381
cttgaactcc agagggtggag gctgcagtgg gctccacca ctgcctgggt gacagagcta 19441
gactccatct cgaagaaaaa ataagtactt gaaacatagc aaatgtttta taattattgg 19501
cttttttttc ttattgttat tacttgcatt attgctgttt gaagaagttt ggtgataatg 19561
gagagaaaag ggcaattagg ggtctgggat ggtttaagta tgaggagacc gagacacatt 19621
25 gactcagagt gaagaaatca aagataagag aatgaaagaa agggagggtta attctgaacg 19681
cacagacaaa gttatgttac taacgtggca tcggctgtgt gttgtaataa ataactccct 19741
ttcacttgct aatagctaat caaatacttt ctggagacca gaagtgactt gctggatcaa 19801
tgacaactcc tccacagatc aaaatgttca aatccttttt ctgtgttgta atcctaaatc 19861
taaaagaaca gagagaccaa gcaaactctac ctcccaacat cattaaagtg acaactctca 19921
30 gtattttattt gaatggtctg ctctcagctt caaccaagga aaagtcaaat tagtgttggt 19981
cagaaaacag aagggtgtta cgagagttct ggctggctcat tacagacttg gggatttttg 20041
attaaaagaa gaagaagaag aagaaacctg ataaagtgt aatatagcaa gcagggatta 20101
gtgtcctgct gggcatgtt ctcacaacag tgagaatttc agagatttca taagaattaa 20161
actgctccac atgacaattt attttacctt ctggcttttc caggaggcaa atcagtgcaa 20221
35 cttctttctg cttttgtttc aatttggtta caaccctcaa ttttaggaca ggctaaacct 20281
agccacccta tcagagatga tgaagtagcc atctttttta cagggtggga gatgaatgga 20341
atcagggttt gtttgtttgt ttgtttaata actgctagta aaaaccaagt caatagctga 20401
ctgagtgtaa gggaggctcc agaaggcagg ttattgtagt atagatgtga ctcgacttat 20461

gatgatgtta cttcccgata aaccgcgcat aagttgaaat atcgttaaat tgaaaatgct 20521
 ttttaatacac cgaatctacc gaacatcata gcttagctca gcctacgtta aatgtgctca 20581
 ggacgcacat tgccctacagc tgagccaaat cacctggcaa cacaggacac tgtagagtat 20641
 cgggtgctgg cccttggtgat gctgtgactg actgggagct gcgcttagtg cctctaccca 20701
 5 gcattgagag tttcttatcg cttattacta gcctgggaaa agaccaaaat tcaaaaactca 20761
 aagtgcggtt tctaccgaat gcttataact ttcagaccat catgatgttg aaaaatcgaa 20821
 ccatcgtagg ttgggatcca tcctataaga cgagctacac tgccggaagt gtaagactgc 20881
 tatgctgccg gaagatgggg catagtggac aactgcaagt cctgacaaca ggaggtcagc 20941
 atctgcgacc tttaacatcc acattgacac taccacagtc ttccaaacag agctgatgat 21001
 10 atagtttggg tgctgctccct gcccaaatct catgtcgaat cgtaatcccc agtggttgag 21061
 gtggggcctg gtgggaggtg attgggtcat gggggcagag ttcttatgaa tggtttagca 21121
 cgggtcccccc ttggtactgt atagtgagtg agttctcatg cgatctggtt gtttaaaagt 21181
 gtgtggcacc tccccctctct ctctttctct tactctggcc atgtgaagtg ttggtcccg 21241
 ctttgccctc caccatgatt gttaaattcc cagaggtctc cctagaagct aatgctgcc 21301
 15 cgtacagcct ggagaactgt gagccaatta aacctcatct ctttttaaat taccagtct 21361
 caggccgggc gcggtgactg acacctgtaa tctcagcact ttgggaggct caggcaggaa 21421
 gatgatttga ggtcaggagt tcgagaccag cctggccaac atggtgaaac ccatctctta 21481
 ctgaaaatat aaaaattagc caggcatggt ggccgggtgcc tgtaatccca gctacttggg 21541
 aggctgaggc aggagaatcg cttgaacctg ggagtcagag gttgcagaga gccaaaatgg 21601
 20 agccactgta ctccagcctg ggcaatggag tgagaccctg tctcaaaaaa tatatatata 21661
 ttaccagac tcaggatatt ctttacctga gactatgaga gaatggacta atatagctga 21721
 agaattttat tttattttta aaaaactttt acgtttgggg gtacctgtaa aagtctgtta 21781
 cataggtaaa ctctgtcat gaggatttgt tgtacagatt ctttcctgct cctccccctc 21841
 ctcccaccct ccacctcaa gaagatccca gtgtctgttg tttccttctt tgtgttcgta 21901
 25 agttctcatc atgtagctcc cacgtataag tgagaacatg cagtatttgg tttctgttcc 21961
 ctgtgttagt ttgctaagga tgatagctc caactccatc tatcttcctg caaaagacat 22021
 gatctcattc atttttattg ctgcatagta ttccatggtg tatatgtacc acattttctt 22081
 tatccagtct gtcattgatg ggcatttagg ttgattctgt gtcttcagaa ttgtgaatag 22141
 tgctgcaacg aacattcggt tgcttggtgc tttatagtag aatgatttct attcttctgg 22201
 30 tagtaatggg attgctgggt caaatggctg ttctgctttt agctctttgc agaataacca 22261
 tactgctttc cacagtgggt gaactaattt acactccac taacagtgt taagtgttcc 22321
 cttttctctg caaccttgcc agcctctgtt atcttttgac tttttaataa aaaccattct 22381
 aattagtgtg atggatattt attgttggtt tgatttgcac ttctctaata atcagtgatg 22441
 ttgagctttt tttcatgttc gttggctgca ggtacatctt cttttgaaaa gtgtctgctc 22501
 35 atgtcctttg cccacttttt aatgggggtt tttttctctt gtaaatttaa gttcctcata 22561
 gatgctgggt attagacctt tgtcagatgt atagcttgca aatattttct cccattctgt 22621
 aggttgtctg cttactcttt tgattgtttc ttttaccatg cagaagctcc taagtttaat 22681
 tagatcccat ttgtcaattt ttgcttttgt tgcaattgct tttgggtgtc ttgtcatgaa 22741

atctttgcca ggtcctatgt ccagaatgat attgcctagg ttgtcttcta gggtttttat 22801
 agttttgggt tttacattta aatctttaat ccatcttgag ttgatttttg tgtttggtgt 22861
 aaggaagggg tccagtttca atattctgca tatggctagc cagttatccc agcattattt 22921
 attgagtaag gagtatctcc tccgttgctt gtttttccca ggtttggtga agatcagatg 22981
 5 gttgtaggtg tgtggcctta ttttggggct ctctatcctg ttcatttggt ctatgtgcct 23041
 gtttttgtag cagtaccatt ctgttttggt tactgtagcc ctatagcata tttcaaagtt 23101
 gggtaacatg atgcctccag ctttattctt tttgcttaga attaccttgg ccatttgggc 23161
 tctttttggt accatatgaa gtttaaaata gtttttttct tagttatgtg aagaatgtcg 23221
 ttggtaattt gataggaata acatgtaatg atattgattc ttcctatcca tgagcatggg 23281
 10 atgtttttcc atttgtttgt gtcttctctg atttcttcaa gcagtgtttt gtaactcata 23341
 ttgtagagat tattcacctc cttgcttagc tgtattccta ggtattgtat tctttctgta 23401
 gtaattgtga atgggattgc ttttctgatt tggccctcag cttggtattg ttggtgtata 23461
 ggaatgctag tgattttttg taccctgaga ctttctgaa gttatttatc agctgaagga 23521
 gcttttgggc tgagactagg gggtttttta gatatagaat catgttctct gcaaacagat 23581
 15 ttagtttgac ttcctctctt cctacttgga tgccctttat ttcttctctt tgctgatatt 23641
 ccctggccag gacttccagt accatgttga ataggcgtgg tgagagaggg cattcttgtc 23701
 ttgtgccagt tttcaggag aatgcttcca cttttgccc attcagtacg atgtttgtgg 23761
 tggtttgta tatatggcta ttattatttt gaggtgtgtt cttttaatac ctagtattatt 23821
 gacagttttt aacatgaagc agtgtttaat tttattaaaa gtcttttctg cctctgttga 23881
 20 gatagtcatg tggctttttgt ctttagttct gtttatgtga tgaatcacat ttgttgattt 23941
 ctttatgttg gaccaacctt gcacccagg gatgaagcct acttgattgt ggtggttttag 24001
 ctttttgata tactactgga ttcagtttgc aagtattttg ttgaggattt ttgcattgat 24061
 gttcatcacg gatatcggcc tgaagtttct ttttttggtg tgtctctgtc aggttttggt 24121
 atcagaatga tgctggcctc ctagaatgag ttggggagga gttcctcctc ctcaattttt 24181
 25 ttggaatagg ttctgtagga atggtaccag ctcttcttta tacatctggg agagtttggc 24241
 tgtgaagcca tcaggtcctg ggatttttta gttggtaggg tatttattac tgattcccta 24301
 aatagaccga taatgatttt agaagtggag tcggtttttt cctgggtccag tcttggaag 24361
 gtgtatgtat ccaggaattt atttagctct tctaggtttt ctagtgttg tgcatatggg 24421
 tgttcatagt agtttctgat ggttggtttt atttccgtgg gatcagtggg aacattctct 24481
 30 tcatcatttc tttttttttt tttttttttt ttttttgaga cggagtctcg ctctgtcgcc 24541
 caggctggag tgcaagtggc cgaatctggc tcaactgcaag ctccgcctcc cgggttcacg 24601
 ccattctcct gcctcagcct ccgagtagc tgggactaca ggcgcccgt accacgtccg 24661
 gctaattttt tgtattttta gtagagacgg ggtttcacg tgtagccag gatggtctcg 24721
 atctcctgac ctctgtgatc gcccgcctcg gcctcccaaa gtgctgggat tacaggcgtg 24781
 35 agccaccgcg ccggtcccat ttctaattgt gtttatttga atcctctctc ttctcttctt 24841
 tattaggcta gctagtggcc tatctatctt attaatTTTT tcaaaaaacc agctcctgga 24901
 tttcttgatc ttttgaatgg ttttcatgt atcaatcctt cagttcagct ctgatttttg 24961
 ttatttcttg tcttggtgta gctttggggg tgacttggtc ttgcttctct aattctttca 25021

gttctgatgt tagtttggtta gtttgagatc taactttttg atgtggacat ttagtgctat 25081
 aaatttaact cttaacactg ccttagctgt gtcccagaga gtctggtatg ttgtatcttt 25141
 gttctcatta gtttgaaaaa acttcttgat ttctgtctta atttaattat ttatccaaag 25201
 tcattcagga acatgttggt taatttccat gtaattgcat ggttttgagc gattttctta 25261
 5 gtcttgactt ctatttttat tgtaccgtgg tctgaggggtg tttgatatga ctttggttct 25321
 tttgcatttg ctgaggattg ttttatgtcc aattatgtgg ttgatttttag agtatgtgcc 25381
 atgtggtgat gagaagaatg tatattctgt tggttttggg tacagagttc tgtagaggtc 25441
 tattagatcc atttgggtcca atgttgagtt cagatcctga atatctttgc taatttcctg 25501
 cctccatgat ctaatactgt cagtaaagca ctgaagtctc ctactactat tgtgtgggag 25561
 10 tctatgtctc tttataggtc tctaagaact tgctttatga atctgggtgc ttctgtgttg 25621
 gatgcatata tathtaggat agttagatct tcttggtgaa ttgaaccctt taccattatg 25681
 taacgccctt ctttgtcttt ttttttcttt gttgggttga agtcttcttt gtctgaaatt 25741
 aggattgcaa cccctgcttt tttctgtttt ctgtttgctt ggtagatttt cctccatccc 25801
 tttattttgg acctatgggt gtccttacat attttatctt tatctatcca tccagccatc 25861
 15 cagccatcca ttcacccgta tcatttttaa ccaataagga cttttaaaag cgcaaccaca 25921
 acaccattaa cataaccaat aaaatctata acaatgataa aatatcatct aatactcagt 25981
 ccatgtccaa ttttccctcg ctatctcaaa atcgtcttct tagaaatggt ctgttcaaat 26041
 gagatcacat ggacagagga aggcgaacct cacactgtgg ggactgttgt ggggtggggg 26101
 gaggggggag gगतगगगगगग cctaatgcta gatgacaagt tagtgggtgc 26161
 20 agcgaccag catggcacat gtatatgtat gtaactaacc tgcacaatgt gcacatgtac 26221
 cctaaaactt aaagtataat aataaaaaat aaaaaataaa aaataaaaag tgaaaaaaga 26281
 aaaaaaaaaa aaaaaagaaat ggtctgttca aatcacaaac cagattcaga aacaatagcc 26341
 atacattaca ttttattaat atgtctctta aatttctttt aatctattac agtctttgga 26401
 atttttatgt ctctgtttat ccttccaatt attaaaaaaa aagtattttt gtattcattg 26461
 25 aatagacaat gcttgcagaa aagtaaaaaa aaaaaattt agtacaaaaa ggtacatagt 26521
 gagctgttcc ttagtctccc tccccagaag caatgttacc acttttgtac aaatagtctc 26581
 tgcctagaca cacatgccag tccctaaggt ggctgtaaca aggtgggtta gagtgagaac 26641
 atgaattcaa attcctatta tgccactcac taagtataaa tcttgggtcat ggtacatgcc 26701
 tctgtgctc agtttttaaat aatggtacct acctcatagg gctgttgaga gaattaaatc 26761
 30 agataagtgc ttaaataact attaatattt attattattc acattccctt ttggcttttt 26821
 tcccaaatag cagagtgggtg cacatatgtc ttcatthatt tggcttggtt tttcacctca 26881
 catcacattt tgatgaataa ttccatacat gttgttatag atttgcttca ttctttgtaa 26941
 tcattgacta atattccatt gtatgaatat gctactgcta aacatgtacg ttatttccaa 27001
 cctcttatta tcaagaaatg ctgcaatgaa tctccttgta atactttagt ggattcatgt 27061
 35 gcaaaaaatat tcataggata aaatcctgaa agttaaatgt ctgagctaaa gggtatgtgc 27121
 attttaatgc tttatagatt gccagctgc ctcaaaggag gttataacaa ttactactcc 27181
 caagaaaaat gcacaagggt ccccatthcc ccatacccta gctaacacag gatattgcta 27241
 aatgctttca tctttgtaa catgatgtat tgaaaatggt atctcaaagt tttaatgtgc 27301

atttttctga ttgtgagaag ataaaggaaa tgtagtataa ctaaacaatca gcagtcaaat 27361
 gacctggcca tgactcctga gtgaggacac tggtaaacac catcaggatc caaacacctc 27421
 tgtatttacg aagaggatgc tccctattgg atagcactaa gcttatttca tgtagtata 27481
 tatgtagtta gttaattaca tccagcgggtg gcaaagggtt tgttctgacc caatgaaact 27541
 5 ttctctcctg gcccccttcc agcatgtggt caggagtaga gtgttgtggc catgaggcat 27601
 gcatttgtac agatgactac ttactcctcc ttgaaacatt tttttccatt tgcttccctg 27661
 ctgtctcact catgggtctg ctctaattc acaaatcact cttttcccag tcttcttgggt 27721
 tgggttttct cctcttctgt gcttgtagac atgggggagc cccagggtt ctctcttgaa 27781
 ctacagcttc tccctgggtt catctccttg ggatgtctgt tccaatgggt ttaaatacta 27841
 10 gggctaggac tagggagagg tcattgaggc acttagcaca taaagtttaa ggaaacattc 27901
 tttatcaggc ccatgcaagt gcaggactgg ccctggaggg tgactaccac cttacctttt 27961
 ccaccctagg caccttgttt gccttaccct agccccagtc ctctttaaca cccagtcctt 28021
 ctccctggac ctccagaaac ataaatccta tttgacattt ctacttgag gttaataagg 28081
 aactcaaacy taaaatatct aagacagaac tcttgcatca ctcccatcc tggggcccaa 28141
 15 gcctgtctct tctactagtc tatctcagtt aacagcatca ccatttattc agttgctcag 28201
 gacaaaaaat ttgaagtaat ccttgactct tctttttttt ttttgagac ggattctcac 28261
 tctgttgccc aggctggagt gcagtgggtg gaccttggct cactgcaacc tctgcctcct 28321
 gggttcaagc aattctcctg cctcagtctt ctgacctcgt gatccactcg cctcggcctc 28381
 ccaaagtgtc gggattacag gcgtgagcca ccgcaccagg cctggattct tttttttttt 28441
 20 tttaaacaag tcctatcttc catctccaaa atgtatccca aatctgacaa cctctccac 28501
 cgtaggccag ccccatctc tccctctga aaatagcctc ccttagatct ctggacattt 28561
 gttcttcccc accccttgt gatcactatt cagcattcag aatgatcttt taatattatg 28621
 aaggagactg tgttctctc ctacttaaaa ttctctagtg gcttcctaataa aatttagaa 28681
 taaaaagcca actcctcgcc atggtcacca ggccaggatc cagtgggtac aacaatctgt 28741
 25 tcccagcaca ataatccac ttctgcctcc ccaccattca ccaggctcca ctgcaactggc 28801
 tcattcctgc ctgagttgtg cttcttttcc ctggaaagtt ctttctgtag atctttaaaag 28861
 ggtgatttcc ttctcaacat tcaggtgtca gctcgtttca ctttctgac catgcccac 28921
 cactcagaga tcaactcaaaa tcccattacc ctattttatt tctccatcat atgtatcact 28981
 atctgaaact atcttgttgt tgatgcaggc tattttcttg accccttcat gggactccca 29041
 30 acaggggtac cccatttact cagcctgcc tgctcaacct cttgcaggag ggagcacacg 29101
 agtgaacgag tgcaggaacc agctggctgc tttagtgtg tgaggagtaa actccatgca 29161
 ggccctgcag cagcaaccag gtaggggtgc ctgcaacccc agggcccccag aggtgtgtt 29221
 acaatgctct cgtagctctg ccatctgtgg acagcagtg gttgtcagct cagtggcccc 29281
 tttgcttcat catgtagggt ggtgcctc tgctgtgag ggcaaagggt cagggtgaca 29341
 35 gtcttttttg gtaccacaaa tttgtgcac ctgaattctt gtttggtgcc caagaagaat 29401
 ggggtcacac agatgaactg aaggatggtg aatgcagaga attagcaatg aaagtggctc 29461
 tcagcagaga gagaagctga aaagggaatg ggaagggcag gtcactctcc cctgaagtca 29521
 agtcacatct ctccaatgtc cagccaccat ctctgaagtc aagtttctc tctctgatgt 29581

ccagccactt ctctctctcta ctggctgagt ctgggggtatt tataggcaga ggataggtgg 29641
 tggggcaggc catacataat tttggaaaag gcaacattct attggtaaaa agacattatt 29701
 cataaagaac caattgggaa agagcgggca cacagggatg gaagttctca ctttgggctg 29761
 caggtttcag gcttttcagc tcaaaagtga ggtttttcca gggacctgcc ccgtctgcct 29821
 5 aaaatttcta cacttctgtc attgataccg ctggataaca gctctcctat aaagttcagg 29881
 ggcttaaaac aaaaatttac taattcaccg ggacagttct tgatagggtc tctcctaact 29941
 gttgcagta catagcaact gggttggagt catcttttct gggcttgaca tccaggacag 30001
 cttcttcctt tgtgtgtctg gtgcctcagt gctcctccag gcagcctttc tctccagaag 30061
 agtagcctgg acttcttggc aactcaagct tccaaaagaa aaaaaagcag agactgctgg 30121
 10 ttctcttata aaacaggcct ggaactggca caatgtactt ctgctgccat attcaggagg 30181
 tcaaagcaat cgaaggccaa tccaagttca aaggcattgg agaaaaatga gaagtgtcac 30241
 ttaaatgaga agtgacacat gcgtaaaagg gggaaaagca ttgattgtgg ccatttttgg 30301
 agataagcta tcacgtttat ttgtttgttt gcttttcta ggtctgtctt ctcccattag 30361
 gttataagct ctgtgagaca acaggaatct tgtccatctt gttttatggc tctacttcca 30421
 15 acacctagaa taatgcctgg cacatagtag gtgctcagt aataacttaa gcacttgata 30481
 catgtttggg gaactaaaat gaacagaatt aaacttcccc agattgggtcc tgccagattt 30541
 gctgatgcca agcatgctga tgccaccca gatgaaagaa gccctaaaat gtagggtttc 30601
 gctttctctg caaacaagaa aaacttgccc tgaacacaaa atctagaaat agatttggcg 30661
 tgttttctac attgaaatat ttcccgtagt accagaaatt attttccac agctttgtgc 30721
 20 tacattaaaa tattgaagtt gactgaaaat atctccattc tttaatcttg gtgtagacta 30781
 gaaacaattt ttttgtaaca aagtaaatat gaaaacttcc taatatttga actccccaga 30841
 tatccccaga tatctccaaa cttaaaatat cattgcaagt taagataaat ttttttaaat 30901
 gactaccgag aaaggtcatt aaaggcttgg ttattaaaaat gtacagattt gggttataaa 30961
 gccagaactt atttgtttta atcattacat atgaccaagc acagaaaata aattacctca 31021
 25 aatctcctct ttgctaattt ttactggtaa actctataaa atgacccat ctttaaacct 31081
 ttttgtaaac cccttataag ttagtaagtg agaatgtatt catcagaagg attttagtga 31141
 tgtttgaaat taaaaagag agatttgatt tttaaaatta tacttgaga ctactgctaa 31201
 tgaaacttct tctaacccta gttttgtctt atcttcagtt tttccagatt tgctcaaagc 31261
 aatcccagtg agcatccacg tcaatgtcat tctcttctct gccatcctta ttgtgttaac 31321
 30 catgggtggg acagccttct tcatgtacaa tgcttttggg aaaccttttg aaactctgca 31381
 tgggtccccta gggctgtacc ttttgagctt catttcaggt aagtacaaaa ttctacctct 31441
 gaagacaaat gtgcttttca atatgtcaaa aagaccgtct acctaaatat aaagttataa 31501
 tcttaacata tatacatgga tgcacactgt agtattatac ataataacaa aaagtgtgga 31561
 aataccacac ttgctcaaca gtagggaatt caataaatac tttatggaca tctatatgaa 31621
 35 taactgtgat gctgacatta aattatattt ttgaagatgt aatcaagagg ataaacgctt 31681
 gtctcaaaaa gttacatggg aaaagcagta tgtaaaactta tataaacatt gtgaacctaa 31741
 tttgattata tatataaaat atagggaata tacatataaa atacatatat gtatatatgg 31801
 gggctatata tatatatata tgaaagagat agatagatag atagatagat agatagatag 31861

acagacagac agacaatata gactccaatc tgttggctgt ggttgtctct gacccatgac 31921
 actatggggg actttttattt ttgctcatac ttttcaatat ttcttagtgt tcaataatgt 31981
 gctattattt atacataata ataaaaataa ataaatggca tcaaaaaaga gtaaagggcc 32041
 agtgttccgc ccacatatga gcagccatat tcaagcctgt agacactttg tgtagcctaa 32101
 5 tgctaggtgt atctgggcaa ggataaactc taaagccaga aattagttca tcaataaaca 32161
 tgtgtacttc aatagctagg gctgatggaa aagaatataa aaccagctct gtgccaaatg 32221
 gtgcttacta tctgcaagtg ggagaaggag aaagacgaga aaatgaaaaa tgtgtgtata 32281
 atttatatgt agctgttctg taggagatct ctgacttcac ccatttctaa ctttgcaaaa 32341
 agatccaaca ctttgtcaga ttcttgggag gcaagtaatt ttattgatgg tttcatggag 32401
 10 ggatacagaa cgataacaac tcacacaaaag caaacaatgt aatgaaaatc tctattcgac 32461
 tgtttctttt tctctgaag ttgccctttg gctgccagct accaggcacc aggtcaagg 32521
 tactttcttg ctcttgacac tactcccttc tctcatacaa ttcaaccca accacaaacg 32581
 tgtatagatc tctctctcta taaaacaaag gcctgtagtt aacaggaggt cacttgcagt 32641
 gtagcctctg ttcatgttta cttgtgcaca ctgcttaggg tctcaccca tccacattct 32701
 15 gctaatacaca ttattcacc atccaatgta gatctctcca gtggagattc tgctaataatt 32761
 ttcttttagat ttgtcacaag tatataatac agttttaaat tgtacagatg attatcctat 32821
 aacagaagag tctaagcctt ttacatcttt gtatctctaa cgaaagcatt cagcatgaag 32881
 ctctgtacac agcagacaat tcaatatgaa tttgtctgact tgaaacagca agcctagaaa 32941
 ggagatgta acttggctac ttagacagaa caggtttcag caatcagaat tcagatgaca 33001
 20 tggaactggt agaacaggcg ctttgaagca ataggacatg agccagttag gagaggatg 33061
 gaatatcata aacaaaggcc aagggtcttg caatcagagc tgaagagcca agagcacagg 33121
 ctgagggtgt gggcagactg aatgagaaag tgattcaatc acatgtgaaa gtccagatga 33181
 gaagagagag ttgggattac ttctgtctac caaacatcca aaaccaaaca ggtggatccg 33241
 ggtggtgtgc tgttttactg atgaccatta cacagaatth taacagaagg aatgtaaagc 33301
 25 agtggttctc aaactggagt tcccagatga gcagcgtctg catcatctgg aaacctgata 33361
 aagcagcaaa ttctcaggcc ctaccccaga cccactgaat cagaaacttg ggggtctggg 33421
 ggaagatggc catctgtatt ttaacaatct cccttcagga gattctgagg ctggctcaag 33481
 tttgaactac aggtagttag ttcaacagggt gttggtggac tgacaaacaa aaagagactc 33541
 cgaggtaact ccaagatggg aatgtcagaa agcagctacc acccctaggg cttgggggaa 33601
 30 ccaacaaaag agtttgcatg tgccagaacc tagaatcttg aggagaggcc ccagagcatt 33661
 gtgtctcaga ccttgaggac ttggcactgg gacaccatga ggggtttctg ggtgtgggaa 33721
 agggctggaa actccccagt tgctgccacc agggagaact acaagtgaag tggaaggtgt 33781
 gggcctttct cccttttctt ttctctctct ctctctcccc ctggtgctca tgtttgacag 33841
 aaaacagctg aaaaggcaga actagtttgg ggagtcttga cctggcatca taaagcagag 33901
 35 aaaagcaaag ctggagtga ggtgagacac aacagctcat tagcagcaac agcgtctag 33961
 cgctccagct tctgaatgaa attctgaaga acagcgact tggaagacaa attatttgac 34021
 agttctgaca gacgacaaa ctaacagcat ttgaaaagca agatgactca gagaatacag 34081
 aatttaatcc aaatcccaga tcctatctct gcctttggcc aggcctaag caaggagaac 34141

ctgagcacac aaatatatgc aggaatgatac aggacctgtg cctgcattct attctgtctc 34201
 acccaccttc aaatttgttg taaaaacatg ggctcaataa aggtttgtga atcagggaag 34261
 gaagagaagg ggagaaagga agggaaggag ccagctccag atctgtgtct tgcagaggat 34321
 aaaggccagt gtttttagat caccagtggt ttttctaagc cccaataact tattttgaaa 34381
 5 tatcaaatg ttcaataact aaaaaaaaaa ccgttacaac aataaaacat gtttgagggt 34441
 cagatggact ggcagtttgt gacctctggg gataaacagg tcactttgga atcacagact 34501
 tcctcattcc ccttaaactc catatggtac ccagaagccc ttggaacttt ggaagggtgt 34561
 tattcacagt tgtaatgtcc atgcagaccc tggctctaag acccaattgt gtaagggtag 34621
 gttttagacc cttatcccaa acattctaag tgtgagccaa tgcgtcacac actcagaggc 34681
 10 cagagactgt attgggggtcc tttatttcac gtacgagtca cattccatta agagacccca 34741
 gaagtcagct ctcttcact gactggttct cttcccttgt ttctcttgcc aatgtgtgct 34801
 gccaggtgg cagtgtcac tgtcagcaga gaagaaaaat gctttcctcc ttggacctct 34861
 tttctctttt tctcctccct actcacatc aggttcctta agcttccccg ctcttgtgc 34921
 tgaagtcatt ctatggtcat ttcttcaact gtctacttcc ctgctggatg ggcacccaag 34981
 15 acttggcatc ctggggcatg tagaaagggg aaagggaagg gaaggaaaa gaagtcctcc 35041
 caattgtcta tctggacctt tccacactgc ccagagtagt gctatgggca tctccttatg 35101
 tctcccgatg tgggtcatgc cagacctgc aggtagaaaa ggaaagaaag caacccattg 35161
 gaccaggcca gcaaaggctt cagtcacaca gctgggtcat acttatggct tcatattctg 35221
 ttgcctcttg aaccagacat ttcttccact ctcataacct ccagtttagc tcgtattcct 35281
 20 cagcattctc catgtaatat tgttgcatga aacctatgca agtcagccaa tttgctcttt 35341
 ctcatcttgt catttataaa ttgatgctga gaagtctttt tccaagggtt tttagtaata 35401
 cttcatcat ccccatagt tcattttggg cagtgattcg cttctttgac tgtacattag 35461
 aatcatctga agaactttct aaaactactg atctcaggtc tcacacaaca ctaattaaat 35521
 tatagtctct ggtggggtag ggcttgggca ctgctatttt accttaagct ctctggagtc 35581
 25 attctaattt gttagccagt ctaagggttg caatataagt gaatatattt cacgtatttg 35641
 tcaaacattg actgagtgc cattatgtgc cagccattat aataggcact ggtgatccca 35701
 cagtgaatca ggcacacaat gctgtcttca cggagcttgt tgtctagtgg gagagtcaaa 35761
 caaaagtgt tatcaataat taagtgatta cagattgcaa taattacaat aagggtgata 35821
 aacaggttgc tataatatac aatagtattg cttttacca gacatttctt aaagaggtaa 35881
 30 atcatctatc agacaccttt taaaaatctc atctaatttc aaaagtgtac ataaataatt 35941
 aagtgattac agtttgcaat aattacaatg aggggtgataa acaagttgct atgaaagaga 36001
 aagatagcag agatctggct ttgagatggg ggtcaggga gactgtcca atagctgagt 36061
 tctaaagcta agaagggaact gagaaccttc acagaacgtc ccaagtagaa gagaagcac 36121
 actgaagact ctagggaaaag aggtctgctt gttgtaggaa ccgaaagaag gccaatgtgg 36181
 35 ctaggtgctg ggtagtgagg ggaaatggca caaggaaaa atgaggttag agagatcagt 36241
 tgataccagt taatgttgga ccctaaacat taacctgggt aagtctttta gaattgattc 36301
 taattgcaat gaaaaccttt tgaaagattt taaaaagata tctgatagct gatttacctc 36361
 tctaagaaat gtctggtgaa agacaatacc atcacattgg agatggaaaa agatgaatgg 36421

attctaaaaa cattctgaaa gtacattcaa aatgtttttc aggtagctta tgcaactaat 36481
 aaatagtggg ggcattctgg gtaagacaag ggaggagcag gcttgcagtt tagggcaaga 36541
 aaggggtggg gagaagagct cagcactaaa atcatgtgtt ccatttgggg cacatcgagt 36601
 ctgagttgct atgagaccac caagtggaga tgccaagtaa atagtgcagtt acatgaatct 36661
 5 ggagttcagt gaagaggtct agaagaaaga tgtatatattg ggcattattc ggatatagat 36721
 attatgtaaa gcaataaaat tggatgagat cacctagggg gagaatgcac atagataaaa 36781
 actgacctag gaccacttca tgtctaaaac accaaaagca atgtcaacaa aagccaaaat 36841
 tgacaaatgg gatctaatta aactaaagag cttctgcaca gcaaaagaat cagagtgaac 36901
 aggcaaccca caaaatggaa gaaaattttc acaacctact catctgacaa agggctaata 36961
 10 tccagaatct acagtgaact caaacaaatt tacaagaaaa aaacaaacaa ccccatcaaa 37021
 aagtgggtga aggacatgaa cagacacttc tcaaaagaag acatttatgc agccaaaaaa 37081
 cacatgaaaa aatgctcacc atcactggcc atcagagaaa tgcaaatcaa aaccacaatg 37141
 agataccatc tcacaccagt tcaaatggca atcattaaaa agtcaggaaa caacaggtgc 37201
 tggagaggat ctggagaaat aggaacactt ttacactggt ggtgggacta taaactagtt 37261
 15 caaccattgg ggaagtcatt gtggcgattc ctcagggatc tagaactaga aataccattt 37321
 gaccagcca tccattact gggatatata ccaaaggact ataaatcatg ctgctataaa 37381
 gacacatgca catgtatgtt tattgcgga ctattcacia tagcaaagac ttggaaccaa 37441
 cctaaatgtc caacaatgat agactggatt aagaaaatgt ggcacatata cccatggaa 37501
 tactatgcag ccataaaaaa tgatgagttc atgtcctttg tagggacatg gatgaagctg 37561
 20 gaaaccatca tcctcagcaa actatcgcaa ggacaaaaaa ccaaaccacg catgttctca 37621
 cccataggtg ggaattgaac gatgagaaca catggacacg ggaaggggaa catcacacac 37681
 tggggactgt tgtgggtgg gaagaggggg gagggatagc attaggagat atacctaagt 37741
 ctaaagacg agttattggg tgcagcacac cagcatggca catgtataca tatgtaccta 37801
 acctgcacat tgagcacacg taccctaaaa cttaaagtat aataataata aaataaaata 37861
 25 aaataaaaaa acaaaaattg atgtaggacc aattcctgaa gaacactgac agttaatttt 37921
 ttggtttagg aggaggagaa gccagcaaac gacactgagt agcaatatcc aaagaaaaag 37981
 aggaaaaag aaaactggga gattatgagt gtcccagagg gaatgtttca agattaccat 38041
 cagcagttag ctttgtgtaa aggtggcctc ctgtaataga ggtgcgggca ggagaaggca 38101
 gaatagggaa aaggggtgga aaaagcttcc ctcaagattt ataatacagt ggaagagaga 38161
 30 gacagagaga gagaaagaga aagagagaga gagaacttaa ggaggtagag gaagagagag 38221
 aacaaaaaaa gagggagctg agtatagaag caattagatt catagttttt agttgcggca 38281
 gtgatatttg agtgggggcc ttttatatat tccattctag gtgtttccca gttgatggga 38341
 gaggtctta cctagatctg catgtaaaag ggagtaggac agctggcaga cttgacatgg 38401
 atcagtggta gaacatccta gcagttctgt gaatactctc tgagaatgac atgaaaggta 38461
 35 ttggttcagg ctttttgag gtgataaaaa ccacaaaaat gtggacttat tgcaaatgt 38521
 atttgttacc atttgcaatg attataatta ttctcttata tataggcttt cttacatata 38581
 gcttctctta cacatagcgt catgagttat gccttctctt acatatagtg tctgctatga 38641
 gttatgcaa gcagggcaaa caaaattgct gcctttcttt aaaaagagga cgctcctagt 38701

atgggcctaa ttaattatga taattacagc tatggcatgg aacataagca cattcatata 38761
 cacaaagaca ataaaataaa gaacagttca aaaacagaac agttacatta tatatcagtt 38821
 tcagtatgaa taataccctg gactctgaaa tatgtctggg gcacattatt ctgtaattgg 38881
 tggtgaaaaa aaatctgcat cttatctcta cgccaatcct tatgaaggag ctgtttttca 38941
 5 ggagttcgag aaagagacac agggatgtcc agtcatcaaa gccgcagagc ctgaggaaga 39001
 ataaaggatt tgtggcagga aaaccagta atgaatgtat tctgctagtt tctcagcata 39061
 gaacttagaa aagaggccac agaaggaaaa gagaagtaat gttagacagc tgatgttggc 39121
 aatgggcaaa gaatttcatt ctcatatgca gggcagtgcc taaaggggca gtgtgtgag 39181
 gaatcacatt ccaggtcatt catgtccagg gtgtggtagg aggactgtgt ttcatatttatt 39241
 10 ttgtacatgg ccagctatg acctgtgca aggaaatgct taatcatttt ctatcacagt 39301
 tgcaaagaga ttcttacct actcagaatg tacgtcttct cttctgtttc atttacagtc 39361
 acaaccaag tccttgcttt gacctcaaa gcaccacttg atgtatccac ttcagaaaca 39421
 cacacagaca gtcacacctt ttctatccct taccttctc cctctttcac tcaaaccacc 39481
 tacttctttt gcattcactc ttcttccagc ctgaaataat ctctctccaa atatctacct 39541
 15 tctcactccc tcaattctct caagatacac ttaaagtgtta tattccctat gaggcctccc 39601
 ctggccatcc ctccccagcc ttctatccc cctctctgc tttattttat tctccttaat 39661
 atatatcaca ctctgataaa ccattgaatg tacttatcaa gttattgcct ctctctccct 39721
 ttccattgtc tccatcactg agagctctgt gaagaaaagg atttgtagct gtttcattta 39781
 gtgctgtacc ccagttccc acaacagcgc aacaggcact caataaatag ttgttgaata 39841
 20 agtgaataaa acagaagtag ctgcataatt tctggtaaca aatgatattc ttctgaaaat 39901
 gtcataattt cagacatatt tccgaaaata aattcaaatt agataaacat tgtattttta 39961
 gaccatttct tctttgcatt aatcactcct ctcaataata taacatttgt aaaacttagg 40021
 ttagatatgg gctcttcaac ttccattac agaagataaa gtgaaaaggc tagaccaat 40081
 ggtgttattc ctcatctac atctatcctt tggaaacaca tgaccaaatt gcttgccatc 40141
 25 acaatctcaa aatctaccct ttggtattaa ctacttcac ttgtccctct gtcccttttt 40201
 agatggtagc catcgggtctc tggagcagtg tagagtcaga acaacttcta tttggggaag 40261
 aaatcattgc tggtgacctt actttcaatt actaactttc tagtgacatt tacataattt 40321
 tagagaaaat taacacctac acttgtaaag ttgtggcttt cccacaccta ttatcatct 40381
 ctcaatattc cttgaaaagg aaattatcaa ttatcattc tatattggca atgaaatgcc 40441
 30 cctaataatct gtcacctata agacaattga agatgatgtg ttgaaagctt tctgaaaatg 40501
 ctgatcatta ctttaaatgg aattgaaatt ccagtttatt atttccaaa atagatctt 40561
 actgatcata ggataacatt tcataacatt tcagagattt cttccccttc gaggagccaa 40621
 acccatagga cctctggact cccacagatc ctggcagggg gttcccactc ataaaagcac 40681
 aggtgccctc agagtcattc agggatgaag aagcaacct cattggccat gtcctacgtt 40741
 35 ccccatatag taaggactgg aggagaccag tgctcaattc tgcagtctca gacagctgtc 40801
 agaggagagt catgaatgtg cagtgtctag cacattacaa acgtttgtta attgactgat 40861
 cattcatggt gtgacagccc tataactcag ctatcctatt cagtcagaaa ttaactcagt 40921
 aatcaaagtc attaaaagga agaaaaaaaa aacctacagt accagacaga tgggtggggaa 40981

atcagacaga tgaaaggaaa aatggctgta ggtcattgag taagacactg ggcagcaaaa 41041
 cctgggcctt gtgcctggtt atactccaca ttatagctcc agcaaggttt ggcaggattt 41101
 ccacagtcct ggcttattct aacctttctt gggagcagga gcagtgttg tcagatagac 41161
 agacacataa ggaatctgtc caactggcac cgtgtgaatt tgggctcttg gtgtacatgg 41221
 5 ataactggga aaaagaggag agagacatgt aggactgac ctaccgtttg tgaagtcttg 41281
 ggcaagagta tgaatgaaaa cccacttctc ttccctgccc tggctccact gcacacagta 41341
 aagagcctca agcatagggtg tgtggacatt gcaccatgta tccaagctct gaccatgcct 41401
 cttgaaacag ctattcctca gccaccctct gaccatggga ggaatgacct aggagaaatg 41461
 accacatagg tcttgaaaaat gggctcaggg ctatttacga agtcaattcc ggggtcccag 41521
 10 gagtatggac taaaatgtga gtcaggcatg ccagatgggt atgttctatt gacttcaagg 41581
 attcctcatg ctgtgggaag gaacctctcc agaagagaga cagagcagaa ccctctaaat 41641
 gtggggcaca aagcaggagc cctctttgct ggattcaaag ggtcactactg gaagagtgtg 41701
 ggttgagtct tattctcaca tcaactcatat cacttacaca cttcttttat agccttagca 41761
 gccatgccac aaagagaaac tcttcatgca tatcttttgg tccataagtc ataaatagtt 41821
 15 atccttgatc ccatgtcttt tttagagcca tggacagaga gaagcaaaaa tataccaagt 41881
 tcacactgag ttgtctccct tcatatctct tagcagtcac tgaaagggtta tgagactcag 41941
 gctgggtttc tatcctctgt cctgaaacg acaacgttga cctcgtgatc aaccctagaa 42001
 tccagagcaa gatctcagac tgtcctctct actaccagac agcacacttt gttttggggg 42061
 ctgtgtgctt gaaatattag ctatggcaaa aggcctttgag tcttatgaca cccaagtaa 42121
 20 cttttacttt aggaatttga aatacagcct tgctgtaatg ctgtctcctt aacaaagcag 42181
 tacctttgaa atattttaaca acttgaaaag gaaaccgagc ttgaattttc ctttcagggtg 42241
 ctgaggaaat aatgtttcac ttctgtctga aattcaccat ctctcagac aaagaaggct 42301
 cttatggtaa aaggaatggc attttctcca caattttcga ataaaagata aagagaaaac 42361
 agcactgcag cctttttgtt aggatctaac aataaagaaa taatacgggt ttgccagggg 42421
 25 agagctctgg ttttaagctc agaatacaaa aataggctga caaaatttta caaaggaata 42481
 ttctcagcta ccactctgag gatggtagaa agtgaaattt caagaaaatt atattatttg 42541
 attatttgat gatgattaag ctgattggcc agtcctatgt gaaattctaa agtagaagaa 42601
 atgtgatgtt gtcttttctg ctcacctctc ccctcattcc tacccecaaa tctctgcctc 42661
 tacccecaac ccagcctgac ctttgggaaa ggaatggggg ctgtcacttg caccgtagct 42721
 30 cctcctgcct gcagtactct ccccccacag tccagcctcc ctctgctcag ctaacttttc 42781
 ctatagctcc ttctggctac ttctaaggaa ctttcataa tgctgccac cctctcagta 42841
 acagcccctc tactatccct tgattacatg cactgtaatt ggttagtaat tgattttatg 42901
 tcctctgcca aattatactc catgagagca aaaatcatga gtattatctt taatgtttaa 42961
 atctccacaa ctatcccccata tataagtcta gagaataaat gaatgagtga attaatgaat 43021
 35 agaaactcaa gccattatgt tgccactcct aggatatttg gattaacttt aactgaagag 43081
 taaaaagcat ttacctgtcc taaaggagac ataaaattag tgggagagta tttggagaaa 43141
 aaaaagacac tgtaatacat tcttttgtgt tgcctaccct gatgtagtca ggtgtccctg 43201
 atatgggggtg ggctgaggat ttgaaataaa atacctaatt tgacattgta agtggaggaa 43261

tcaggatttg aaaccaaatt ggtttgacct aaatcaagct attatgataa ggacttgcaa 43321
 gaaaaaagga atccataaaa accatatgaa tagctaccaaa ttagtaagca tttatatgtg 43381
 tcaattacaa agccaaatgc aaatcatgct ttattttatat tagccacctt ttatagatga 43441
 agaaattgag acctgaatat taaaattgcc tactttttaca tagtaagtaa aggaatcagg 43501
 5 gtttgaaccc aaattgggtt cacctaaggt agaaaacccat cccagcaagt ctctatttaa 43561
 ctggaaccct attgtggtgg cctgagatat aacagtagct gtggaagcgc tgtagagtcc 43621
 tggccatcct atgtgctcct gatctgggtcc ctcttgccac ctgcttctgc tccctgtgcc 43681
 atccacccat ctggaagtct cccagtgtcc atcttcgggg gagacactca ccagagtttc 43741
 cagcttccag ccagtatgga gtgcccctgt cccacagcaa tctcaccgaa atcacagcta 43801
 10 catctgttaa aattaggcta ccaatgagtg atagatgagg gggaaaaata ataatagtgt 43861
 actaaacaaa acaaatgttt atttttctca cacataaaaa tctagagggt gaagtcagg 43921
 gctggtccag aggctccaag gatctgggat ttagactccc tctttcttgt ttttccacag 43981
 catatggctt ccatctctgg ggccacattg gtccaaaatg tatgtctggg ctccagccat 44041
 tgcattccata tttcagccac aggaaggagg aagtggggaa gaaaggacag gcccctaata 44101
 15 cctgtatagt tcaagaagac tatcccgcc atacttccca accaccetta gttgaacaat 44161
 gctgtcttaa ttcaagacac tcacatgtct agccaaaaat ctgaattctg ttacaacaa 44221
 ggagaataga gatgtgcgcc acctcaatac ctcatccata gctacctttt cctttgtgca 44281
 gctgtggcca agtgaaaagt gaaggagctg tggtaaccct tctgaaggga ggctggggcc 44341
 tttcacaaga ggctgcatga ttgacattta tctgcatgg cctgtgaagt acagagaaat 44401
 20 attttctctt gaagccacat catagcagtg gctgctttgt agcctgattc caccattatg 44461
 cctttaaagt gcctagcaat tcagccttca catcatgcaa agaggaatat ctccagttc 44521
 ttgtaagatc agcttaattc taaccacctc cttacctccc actgcactcc tacacgcaca 44581
 cacaaatctt cttcactcag agcagaacca taaccgaagc cctaccacct agagactgaa 44641
 gaatcaggct catgattaca aatatgcaat aattttttgt gtggataatg tcaatgggga 44701
 25 tgatggtaag agaattcctt ggtttacaca ttgacctct tccctgtccc ttacaatcag 44761
 gaaatatttg tcccaacacc ttgtttcttc tgttgcaggc tctgtgggt gtcttgtcat 44821
 gatattgttt gcctctgaag tgaaaatcca tcacctctca gaaaaaattg caaattataa 44881
 agaagggact tatgtctaca aaacgcaaag tgaaaaatat accacctcat tctgggtcat 44941
 tttcttttgc ttttttgttc attttctgaa tgggctccta atacgacttg ctggatttca 45001
 30 gttccctttt gcaaaatcta aagacgcaga aacaactaat gtagctgcag atctaattgta 45061
 ctgaaaggca aacctttcta taattttaca agggagtaga cttgctttgg tcaacttttag 45121
 atgtggttaa ttttgcataat ccttttagtc tgcataatatt aaagcatcag gacccttcgt 45181
 gacaatgttt acaaaattacg tactaaggat acaggctgga aagtaaggga agcagaaggga 45241
 aggctttgaa aagttgtttt atctgggtgg aaattgcttg acccaggtag tcaaaggcag 45301
 35 ttgactagaa tcgacaaatt gttactccat atatatatat gtgtgtgtgt gtgtgtgtgt 45361
 gtgtgtgtaa gatgtcttcc tatcaaaaag atatcaaagg cacatggaat atattttaat 45421
 aaaaacaaat aatatctcta atatatccac acatttggtt ccagatttca gaaaactgag 45481
 ctgcaatcgc tttcctaata cagtagtgta ttaaatgaac atctataaaa tgtatcaaca 45541

cacatttttaa aaaatttggtt taaagtatac tcttaggccca ggcggtggtga ctcacacctg 45601
 taattccagc acttcaggag gccaaaggtgg gaagatcatt tgagttcagg agttcgagtt 45661
 acagcctggg caataaagtg agaccctgtc actaacaaaa ttaaaaaata aaataaatat 45721
 aaaatatagg ctttaaaaaa gcatagtctt attaacatg tctgttggtc aaaatctgca 45781
 5 aactctaaaa gaagaaaaga agaaaaaacc aagcttaggg tatttttcct cccgtgcctg 45841
 agtcccaatt acattcacga cagtactttc aatgaacata attgttagga ccactgagga 45901
 atcatgaaaa atgatctctg cttagtacat ttgatgcaaa atgacttatt aggggctggt 45961
 tttctagcta tagtgtctcg agtactaata tgcaattatg aaaattatat taaatctggg 46021
 attatgacgg tatcactgta tcactcttgg cttgttctgg ctgtcaccaa gcatgaccca 46081
 10 ggtcaacttt ttttttcccc tgaattacc atcaaattga tctgcagctg actaaaggcc 46141
 acagctgagc ctggaactga ccttctcttc atcctcaacc tgctgtcctc cagaaagcac 46201
 caaggaaaaa gcagagaatg acagcaaaca gatcactagg cctctgacca cagggtgctga 46261
 gtactcagca gccctcatat aatagggttg aaagtactcc ttaaaataaa aactgtttc 46321
 cctttggaac tattttacaag gatgaaacaa ccgtatacct gagaaataac ttgctctggt 46381
 15 gtcaattcgc tattcgccag cagacatcag aacacaccga gtttccagat gctggttttt 46441
 ccccttaaat caggaaatac acctggacaa tttctagaag actacaattc agtctagcca 46501
 caaaggggat tttttttttt tggtaacagg ctagagcccg gttctgtaag tcttttagctg 46561
 aaatggtcca gtacaaaagc actggaaatg agtgggctag gaggacaagg accgtctcct 46621
 gcgtgaggag ttggttggag gtccccaagg ccaggtagcc cctgcactct tattggattc 46681
 20 ctctctgtct tcttggagtt ttgaaaaact ccttcgaaca ccaggctttt ttcttttagaa 46741
 aacaagtctc caatcgttct ctgttccgta gaaagagaaa gaaaacctgg agcagctgct 46801
 gaaaaatcta atgaggaact aagaggcaaa cccacca

SEQ ID NO: 10 - Mouse CLRN1 Protein Isoform 1

25 MPSQQKKIIFCMAGVLSFLCALGVVTA VGTPLWKATILCKTGALLVNASGKELDKFMGEMQYGLFHGEGV
 RQCGLGARPFREFSSRS MKERYSLYEDKGETAVFPDLVQAI PVSIHINIILFSMILVVLTMVGTAF FMYNAF
 GKPFTLHG PLGLYLVSFISGSCGLVMILFASEVKVHRLSEKIANFKEGTYAYRTQENYTT SFWVFIC
 FFVHF LNGLLIRLAGFQFPFTKSKETETTNVADLMY

30 SEQ ID NO: 11 - Dog CLRN1 Protein

MPNQKKVVFCTAGVLSFVCALGVVTALGTPLWKATFLCKTGALLVNASGQELDKFMGEMQYGLFHGEGI
 RQCGLGARPFREFSLFPDLLKVIPVSIHVNILFSTILVVLTMVGTAF FMYNAFGKPFTLHG PLGLYL LLSF
 ISGSCGLVMILFASEVKIHLSEKIAN YKEGTYAYKTQSEKYTT SFWVFICFLVHLLNGLLIRLAGFQF
 PFAKS KDTETTNVAADLMY

35

SEQ ID NO: 12 - 5' UTR

40 AGGAGATACTTGAAGGCAGTTTGAAAGACTTGTTTTACAGATTCTTAGTCCAA
 AGATTTCCAATTAGGGAGAAGAAGCAGCAGAAAAGGAGAAAAGCCAAGTAT
 GAGTGATGATGAGGCCTTCATCTACTGACATTTAACCTGGCGAGAACCGTCG

199

GATGCTGGTTTTTCCCCTTAAATCAGGAAATACACCTGGACAATTTCTAGAAG
 ACTACAATTCAGTCTAGCCACAAAGGGGATTTTTTTTTTTTGGTAACAGGCTA
 GAGCCCGGTTCTGTAAAGTCTTTAGCTGAAATGGTCCAGTACAAAAGCACTGG
 AAATGAGTGGGCTAGGAGGACAAGGACCGTCTCCTGCGTGAGGAGTTGGTTG
 5 GAGGTCCCAAGGCCAGGTACCCCTGCACTCTTATTGGATTCTCTCTGTCT
 TCTTGGAGTTTTGAAAACTCCTTCGAACACCAGGCTTTTTTCTTTAGAAAAC
 AAGTCTCCAATCGTTCTCTGTTCCGTAGAAAGAGAAAGAAAACCTGGAGCAG
 CTGCTGAAAAATCTAATGAGGAACTAAGAGGCAAACCCACCA

10 **SEQ ID NO: 16 – Chimeric intron**

GGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCCTCGCGCCG
 CCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGAC
 GGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGACGGCTCGTTTC
 15 TTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCG
 GGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGCGTGCGGAGCGCC
 GCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGG
 GCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCC
 CGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTG
 20 CGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTCTGGGCTGTAACCCCC
 CCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCCGGCTTCGGGTGCGGG
 GCTCCGTGCGGGGCGTGCGCGGGGGCTCGCCGTGCCGGGCGGGGGGTGGCG
 GCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCCTCGGGCCGGGGAGGGCTC
 GGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGC
 25 GAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTT
 CCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGAGGCGCCGCGCACCC
 CCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGG
 GCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCGTCCCCTTCTCCCTCTCCAG
 CCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGGGGGGGACGGGGCAGG
 30 GCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACC
 ATGTTTCATGCCTTCTTCTTTTCCTACAG

SEQ ID NO: 17 – CMV Enhancer

GACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTT
 CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
 TGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTT
 CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTT
 ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACG
 40 CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGT
 ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTACCATG

45

SEQ ID NO: 18 – CBA promoter

TCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCA
 CCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCG
 5 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGG
 GCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCC
 GAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCG
 AAGCGCGCGGCGGGCG

10 SEQ ID NO: 19 –tGFP

MESDEGLPAMEIECRITGTLNGVEFELVGGGEGTPEQGRMTNKMKSTKGALTF
 SPYLLSHVMGYGFYHFGTYPSGYENPFLHAINNGGYTNTRIEKYEDGGVLHVSFS
 YRYEAGRVIGDFKVMGTGFPEDSVIFTDKIIRSNATVEHLHPMGDNDLDGSFTRT
 15 FSLRDGGYYSSVVDSHMHFKSAIHPSILQNGGPMFAFRRVEEDHSNTELGIVEYQ
 HAFKTPDADAGEE

SEQ ID NO: 20 – bovine growth hormone poly-A

CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCT
 TGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCTTAATAAAATGAGGAAATT
 GCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
 AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATG
 20 CGGTGGGCTCTATGG

25

SEQ ID NO: 21 – polyadenylation signal of soluble neuropilin-1

AAATAAAATACGAAATG

30

SEQ ID NO: 22 – Splice donor sequence

GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAACTGGGCT
 TGTCGAGACAGAGAAGACTCTTGCGTTTCT

35 SEQ ID NO: 23 - Splice acceptor sequence

GATAGGCACCTATTGGTCTTACTG ACATCCACTTTGCCTTTCTCTCCACAG

SEQ ID NO: 24 – 3'ITR

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCC
 40 CGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGC
 AGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCT

45

5

10

15

30

45

CAAAGGGGATTTTTTTTTTTTGGTAACAGGCTAGAGCCCGGTTCTGTAAGTCT
 TTAGCTGAAATGGTCCAGTACAAAAGCACTGGAAATGAGTGGGCTAGGAGG
 ACAAGGACCGTCTCCTGCGTGAGGAGTTGGTTGGAGGTCCCCAAGGCCAGGT
 ACCCCCTGCACTCTTATTGGATTCTCTCTGTCTTCTTGGAGTTTTGAAAACT
 5 CCTTCGAACACCAGGCTTTTTCTTTAGAAAACAAGTCTCCAATCGTTCTCTG
 TTCCGTAGAAAGAGAAAGAAAACCTGGAGCAGCTGCTGAAAAATCTAATGA
 GGA ACTAAGAGGCCAAACCCACCA

SEQ ID NO: 30 – FP

10 KRKRRG

SEQ ID NO: 31 – T2A

EGRGSLLTCGDVEENPGP

15 SEQ ID NO: 32 – eGFP

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLP
 VPWPTLVTTLYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYK
 TRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQKNGI
 KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRD
 20 HMLLEFVTAAGITLGMDELYK

SEQ ID NO: 33 – Myc

EQKLISEEDLEQKLISEEDLEQKLISEEDL

25 SEQ ID NO: 34 – HA

YPYDVPDYA

SEQ ID NO: 35 – 3x FLAG tag

DYKDHDGDYKDHDIDYKDDDDK

30

SEQ ID NO: 36 – 3'UTR 1357

AAGGCAAACCTTTCTATAATTTTACAAGGGAGTAGACTTGCTTTGGTCACTTT
 TAGATGTGGTTAATTTTGCATATCCTTTTAGTCTGCATATATTAAAGCATCAG
 GACCCTTCGTGACAATGTTTACAAATTACGTACTAAGGATACAGGCTGGAAA
 35 GTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTTATCTGGTGGGAAAT
 TGCTTGACCCAGGTAGTCAAAGGCAGTTGACTAGAAATCGACAAATTGTTACT
 CCATATATATATATGTGTGTGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGA
 TATCAAAGGCACATGGAATATATTTTAATAAAAACAAATAATATCTCTAATA
 TATCCACACATTTGTTGCCAGATTTTCAGAAAACCTGAGCTGCAATCGCTTTCCT
 40 AAAACAGTAGTGTATTAAATGAACATCTATAAAATGTATCAACACACATTTT
 AAAAAATTTGTTTAAAGTATACTCTTAGGCCAGGCGTGGTGACTCACACCTGT
 AATTCCAGCACTTCAGGAGGCCAAGGTGGGAAGATCATTTGAGTTCAGGAGT
 TCGAGTTACAGTCTGGGCAATAAAGTGAGACCCTGTCACTAACAAAATTAAA
 AAATAAAATAAATATAAAATATAGGCTTTAAAAAAGCATAGTCTTATTAACC
 45 ATGTCTGTTGGTCAAAATCTGCAAACTCTAAAAGAAGAAAAGAAGAAAAA
 CCAACGTTAGGGTATTTTTCCTCCCGTGCCTGAGTCCCAATTACATTCACGAC

AGTACTTTCAATGAACATAATTGTTAGGACCACTGAGGAATCATGAAAAATG
 ATCTCTGCTTAGTACATTTGATGCAAAATGACTTATTAGGGGCTGTTTTCTA
 GCTATAGTGTCTCGAGTACTAATATGCAATTATGAAAATTATATTAAATCTGG
 GATTATGACGGTATCACTGTATCATCTTGGTCTTGTTCTGGCTGTCACCAAGC
 5 ATGACCCAGGTCAACTTTTTTTTTTCCCTGAATTACCCATCAAATTGATCTGC
 AGCTGACTAAAGGCCACAGCTGAGCCTGGAAGTACCCTTCCTTCATCCTCA
 ACCTGCTGTCCTCCAGAAAGCACCAAGGAAAAAGCAGAGAATGACAGCAAA
 CAGATCACTAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAGCCCTCATAT
 AATAGGTTTGAAAGTACTCCTTAAAATAAAACACTGTTTCCCTTTGGAAGTAT
 10 TTACAAGGATGAAACAACCGTATACCTGAGAAATAACTTGCTCTGGTGTCAA
 TTCGCTATTCGCCAGCAGACATCAGAACACACCGAGTTTCCAGATGCT

SEQ ID NO: 37 - 3'UTR 1406

AAGGCCAAACCTTTCTATAATTTTACAAGGGAGTAGACTTGCTTTGGTCACTTT
 15 TAGATGTGGTTAATTTTGCATATCCTTTTAGTCTGCATATATTAAAGCATCAG
 GACCCTTCGTGACAATGTTTACAAATTACGTACTAAGGATACAGGCTGGAAA
 GTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTTATCTGGTGGGAAAT
 TGCTTGACCCAGGTAGTCAAAGGCAGTTGACTAGAATCGACAAATTGTTACT
 CCATATATATATATATGTGTGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGA
 20 TATCAAAGGCACATGGAATATATTTTAATAAAAAACAAATAATATCTCTAATA
 TATCCACACATTTGTTGCCAGATTTTCAAGAACTGAGCTGCAATCGCTTTCCT
 AAAACAGTAGTGTATTAAATGAACATCTATAAAATGTATCAACACACATTTT
 AAAAAATTTGTTTAAAGTATACTCTTAGGCCAGGCGTGGTGACTCACACCTGT
 AATTCCAGCACTTCAGGAGGCCAAGGTGGGAAGATCATTTGAGTTCAGGAGT
 25 TCGAGTTACAGCCTGGGCAATAAAGTGAGACCCTGTCACTAACAAAATTAAA
 AAATAAAATAAATATAAAATATAGGCTTTAAAAAAGCATAGTCTTATTAACC
 ATGTCTGTTGGTCAAAATCTGCAAACTCTAAAAGAAGAAAAGAAGAAAAAA
 CCAACGTTAGGGTATTTTTCCTCCCGTGCCTGAGTCCCAATTACATTCACGAC
 AGTACTTTCAATGAACATAATTGTTAGGACCACTGAGGAATCATGAAAAATG
 30 ATCTCTGCTTAGTACATTTGATGCAAAATGACTTATTAGGGGCTGTTTTCTA
 GCTATAGTGTCTCGAGTACTAATATGCAATTATGAAAATTATATTAAATCTGG
 GATTATGACGGTATCACTGTATCATCTTGGTCTTGTTCTGGCTGTCACCAAGC
 ATGACCCAGGTCAACTTTTTTTTTTCCCTGAATTACCCATCAAATTGATCTGC
 AGCTGACTAAAGGCCACAGCTGAGCCTGGAAGTACCCTTCCTTCATCCTCA
 35 ACCTGCTGTCCTCCAGAAAGCACCAAGGAAAAAGCAGAGAATGACAGCAAA
 CAGATCACTAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAGCCCTCATAT
 AATAGGTTTGAAAGTACTCCTTAAAATAAAACACTGTTTCCCTTTGGAAGTAT
 TTACAAGGATGAAACAACCGTATACCTGAGAAATAACTTGCTCTGGTGTCAA
 TTCGCTATTCGCCAGCAGACATCAGAACACACCGAGTTTCCAGATGCTGGTTT
 40 TTCCCTTAAATCAGGAAATACACCTGGACAATTTCTAGAAGAC

SEQ ID NO: 38 – CLRN1 antibody epitope
 EKIANYKEGTYVYKTQSEKY

45 **SEQ ID NO: 39 – CLRN-0**

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCTGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 5 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 10 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 15 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 20 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 25 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCGCCCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 CTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 30 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 35 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTACGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 40 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 GGTGGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 45 CTGCCTGGTCATGATTCTGTTGCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC

206

SEQ ID NO: 40 – CLRN-1

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCTGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 5 GACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAGTTCATAGCCCATATA
 TGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 10 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGGCGG
 15 GCGGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 20 TTAATGACGGCTCGTTTCTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 25 TGCCTGCGGGGTGTGTGCGTGGGGGGGGTGAGCAGGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 30 CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 35 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATTGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTACGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 40 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTTCAGATTCAGCTTCTTCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 GGTGCGAACCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 45 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC

GAGAAGATCGCCAACTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGCAGCG
 5 GAGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCG
 GACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCGATCTGCTGAAAGC
 CATTCCTGTCAGCATCCATGTCAACGTTCATCCTCTTCTCTGCCATCCTCATTGT
 CCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGT
 TTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGC
 10 TCTTGC GGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAAATTCACCA
 CCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACG
 CAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTTAAG
 AGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTT
 GCCCCCTCCCCCGTGCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTT
 15 TCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTAT
 TCTGGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAA
 TAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAATTCAGCTG
 ACGTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCT
 CTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAGGTCGCCCCGACGCC
 20 CGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCA
 GG

SEQ ID NO: 41 – CLRN-2eGFP

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
 25 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTTCTGCGGGCCGACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 30 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 35 ACCCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGGCGG
 GGCGGGGGCGAGGCGGAGAGGTGCGGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCCTTCGCCCCGTGCCCGCT
 40 CCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 45 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC

TCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGGC
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG
 5 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 10 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTTCAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 15 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 20 GGTCCGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACTACAAAGAGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 25 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGCAGCG
 GAGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCG
 GACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCCGATCTGCTGAAAGC
 CATTCCTGTGAGCATCCATGTCAACGTATCCTCTTCTCTGCCATCCTCATTGT
 30 CCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGT
 TTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGC
 TCTTGCGGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAATTCACCA
 CCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACG
 CAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTGGAT
 35 CAGGCGAAGGCAGGGGATCTCTGCTGACCTGCGGAGATGTGCAAGAGAATCC
 TGGACCAatggtgagcaagggcgaggagctgtcacggggtgtgcccacctggtcgagctggacggcgacgtaaa
 cggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgacctgaagttcatctgcaccacc
 ggcaagctgcccgtgcccgtgcccaccctcgtagccacctgacctacggcgtgcagtgttcagccgctaccccgaccacat
 gaagcagcagcacttctcaagtcgcccacatgccgaaggctacgtccaggagcgcaccatcttctcaaggacgacggcaact
 40 acaagaccgcgcccaggtgaagttcgagggcgacacctggtgaaccgcatcgagctgaaggcagcacttcaaggagg
 acggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagcagaagaacgg
 catcaaggtgaactcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaacaccccc
 atcggcgacggccccgtgctgctgcccgaacacactacctgagcaccagtcgcccctgagcaagaccccaacgagaag
 cgcgatcacatggtcctgctggagttcgtgaccgcccggcgatcactctcggcacggagctgtacaagTAAGAGC
 45 TCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC
 CTCCCCCGTGCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCT

AATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG
 GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGC
 AGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAATTCAGCTGACGT
 GCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGC
 5 GCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGG
 CTTTGGCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 42 – CLRN-3

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCGGG
 10 CGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 15 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 20 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCT
 25 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 30 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGCGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
 35 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCCGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GCGCGCCGCGCACCCCCCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 40 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTAGGAGATACTTGAAGGCAGTTTGAAAGAC
 45 TTGTTTTACAGATTCTTAGTCCAAAGATTTCGAATTAGGGAGAAGAAGCAGC
 AGAAAAGGAGAAAAGCCAAGTATGAGTGATGATGAGGCCTTCATCTACTGAC

ATTTAACCTGGCGAGAACCGTCGATGGTGAAGTTGCCTTTTCAGCTGGGAGCT
 GTCCGTTTCAGCTTCCGTAATAAAATGCAGTCAAAGAGGCAGTCCCTTCCCATTG
 CTCACAAAGGTCTTGTTTTTGAACCTCGCCCTCACAGAAGCCGTTTCTCATCG
 CCACCATGCCTAGCCAGCAGAAGAAAATCATCTTCTGCATGGCCGGCGTGT
 5 CAGCTTCGCCTGTGCTCTGGGAGTTGTGACAGCCCTGGGAACCCCTCTGTGGA
 TCAAAGCCACAGTGCTGTGCAAGACAGGCGCCCTGCTGGTTAATGCCTCTGG
 CCAAGAGCTGGACAAGTTCATGGGCGAGATGCAGTACGGCCTGTTCCATGGC
 GAAGGCGTCAGACAGTGTGGCCTGGGAGCCAGACCTTTTCAGATTCAGCTTCT
 TCCCAGACCTGCTGAAGGCTATCCCCGTGTCCATCCACGTGAACGTGATCCTG
 10 TTCAGCGCCATCCTGATCGTGCTGACAATGGTCGGAACCGCCTTCTTCATGTA
 CAACGCCTTCGGCAAGCCCTTCGAGACACTGCATGGACCTCTGGGCCTGTAC
 CTGCTGAGCTTTATCAGCGGCAGCTGTGGCTGCCTGGTCATGATTCTGTTTCG
 CAGCGAAGTGAAGATCCACCACCTGAGCGAGAAGATCGCCAACTACAAAGA
 GGGCACCTACGTCTACAAGACCCAGTCCGAGAAGTACACCACCAGCTTTTGG
 15 GTTATCTTCTTCTGTTTCTTCGTGCACTTCCTGAACGGCCTGCTGATCAGACTG
 GCCGGCTTCCAGTTTCCATTCGCCAAGAGCAAGGACGCCGAAACCACAAACG
 TGGCCGCCGATCTGATGTACTAAGAGCTCAAGGCAAACCTTTCTATAATTTTA
 CAAGGGAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTTAATTTTGCATATC
 CTTTTAGTCTGCATATATTAAAGCATCAGGACCCTTCGTGACAATGTTTACAA
 20 ATTACGTACTAAGGATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCT
 TTGAAAAGTTGTTTTATCTGGTGGGAAATTGCTTGACCCAGGTAGTCAAAGGC
 AGTTGACTAGAATCGACAAATTGTTACTCCATATATATATATGTGTGTGTGTG
 TGTGTAAGATGTCTTCCTATCAAAAAGATATCAAAGGCACATGGAATATATTT
 TAATAAAAACAAATAATATCTCTAATATATCCACACATTTGTTGCCAGATTTT
 25 AGAAAAGTCTGAGCTGCAATCGCTTTCCTAAAACAGTAGTGTATTAAATGAACA
 TCTATAAAATGTATCAACACACATTTTAAAAAATTTGTTTAAAGTATACTCTT
 AGGCCAGGCGTGGTGACTCACACCTGTAATTCCAGCACTTCAGGAGGCCAAG
 GTGGGAAGATCATTTGAGTTCAGGAGTTCGAGTTACAGTCTGGGCAATAAAG
 TGAGACCCTGTCATAACAAAATTAATAAATAAATAAATAAATAAATAAATAGG
 30 CTTTAAAAAAGCATAGTCTTATTAACCATGTCTGTTGGTCAAAATCTGCAAAC
 TCTAAAAGAAGAAAAGAAGAAAAAACCAACGTTAGGGTATTTTTCCTCCCGT
 GCCTGAGTCCCAATTACATTCACGACAGTACTTCAATGAACATAATTGTTAG
 GACCACTGAGGAATCATGAAAAATGATCTCTGCTTAGTACATTTGATGCAAA
 ATGACTTATTAGGGGCTGTTTTTCTAGCTATAGTGTCTCGAGTACTAATATGC
 35 AATTATGAAAATTATATTAAATCTGGGATTATGACGGTATCACTGTATCATCT
 TGGTCTTGTTCTGGCTGTCACCAAGCATGACCCAGGTCAACTTTTTTTTTTCCCC
 TGAATTACCCATCAAATTGATCTGCAGCTGACTAAAGGCCACAGCTGAGCCT
 GGAAGTACCCCTTCCTTCATCCTCAACCTGCTGTCCTCCAGAAAGCACCAAGG
 AAAAAGCAGAGAATGACAGCAAACAGATCACTAGGCCTCTGACCACAGGTG
 40 CTGAGTACTCAGCAGCCCTCATATAATAGGTTTGAAAGTACTCCTTAAAATAA
 AACACTGTTTCCCTTTGGAAGTATTTACAAGGATGAAACAACCGTATACCTGA
 GAAATAACTTGCTCTGGTGTCAATTCGCTATTCGCCAGCAGACATCAGAACA
 CACCGAGTTTCCAGATGCTCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGC
 CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTC
 45 CTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTC
 TGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATA

GCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAATTCAGCTGAC
 GTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTG
 CGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGG
 GCTTTGCCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

5

SEQ ID NO: 43 – CLRN-4

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 10 GACTAGTTATTAATAGTAATCAATTACGGGGTTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 15 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGGCGG
 20 GGCGGGGGCGAGGCGGAGAGGTGCGGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 25 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 30 TGCCTGCGGGGTGTGTGCGTGGGGGGGGTGAGCAGGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 35 CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 40 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTTCACTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 45 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG

CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 GGTCGGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 5 CTGCCTGGTCATGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGCAGCG
 10 GAGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCG
 GACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCCGATCTGCTGAAAGC
 CATTCTGTGAGCATCCATGTCAACGTATCCTCTTCTCTGCCATCCTCATTGT
 CCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGT
 TTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGC
 15 TCTTGCGGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAATTCACCA
 CCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACG
 CAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTTAAG
 AGCTCAAGGCAAACCTTTCTATAATTTTACAAGGGAGTAGACTTGCTTTGGTC
 ACTTTTAGATGTGGTTAATTTTGCATATCCTTTTAGTCTGCATATATTAAAGCA
 20 TCAGGACCCTTCGTGACAATGTTTACAAATTACGTACTAAGGATACAGGCTG
 GAAAGTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTTATCTGGTGGG
 AAATTGCTTGACCCAGGTAGTCAAAGGCAGTTGACTAGAATCGACAAATTGT
 TACTCCATATATATATATATATGTGTGTGTGTGTGTAAGATGTCTTCCTATCAA
 AAGATATCAAAGGCACATGGAATATATTTTAAATAAAAACAAATAATATCTCT
 25 AATATATCCACACATTTGTTGCCAGATTTTCAGAAAAGTCTGCAATCGCTT
 TCCTAAAACAGTAGTGATTAAATGAACATCTATAAAATGTATCAACACACA
 TTTTAAAAAATTTGTTTAAAGTATACTCTTAGGCCAGGCGTGGTGACTCACAC
 CTGTAATTCCAGCACTTCAGGAGGCCAAGGTGGGAAGATCATTGAGTTCAG
 GAGTTCGAGTTACAGCCTGGGCAATAAAGTGAGACCCTGTCTACTAACAAAAT
 30 TAAAAAATAAAATAAATATAAAATATAGGCTTTAAAAAAGCATAGTCTTATT
 AACCATGTCTGTTGGTCAAAATCTGCAAACTCTAAAAGAAGAAAAGAAGAAA
 AAACCAACGTTAGGGTATTTTTTCTCCCGTGCCTGAGTCCCAATTACATTAC
 GACAGTACTTTCAATGAACATAATTGTTAGGACCACTGAGGAATCATGAAAA
 ATGATCTCTGCTTAGTACATTTGATGCAAAATGACTTATTAGGGGCTGTTTTT
 35 CTAGCTATAGTGTCTCGAGTACTAATATGCAATTATGAAAATTATATTAAATC
 TGGGATTATGACGGTATCACTGTATCATCTTGGTCTTGTTCTGGCTGTCACCA
 AGCATGACCCAGGTCAACTTTTTTTTTTCCCCTGAATTACCCATCAAATTGATC
 TGCAGCTGACTAAAGGCCACAGCTGAGCCTGGAAGTACCCCTTCCTTCATCCT
 CAACCTGCTGTCTCCAGAAAGCACCAAGGAAAAAGCAGAGAATGACAGCA
 40 AACAGATCACTAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAGCCCTCAT
 ATAATAGGTTTGAAAGTACTCCTTAAATAAAACACTGTTTCCCTTTGGAAGT
 ATTTACAAGGATGAAACAACCGTATACCTGAGAAATAACTTGCTCTGGTGTG
 AATTCGCTATTTCGCCAGCAGACATCAGAACACACCGAGTTTCCAGATGCTGG
 TTTTTCCCCTTAAATCAGGAAATACACCTGGACAATTTCTAGAAGACAGGCCT
 45 CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCT
 TGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCTTAATAAAATGAGGAAATT

GCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
 AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATG
 CGGTGGGCTCTATGGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGG
 AACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACT
 5 GAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCCGGGCGGCCT
 CAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 44 – CLRN-6eGFP

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCCGGG
 10 CGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 15 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 20 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 25 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCCGGCGGCTGTGAGCGC
 30 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGCGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
 35 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCCGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCGCACCCCCCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 40 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCatggtgagcaaggcgaggagctgttcaccggggtg
 45 gtgccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgaggcgaggcgatgccacc
 tacggcaagctgacctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccacctgtgaccacctgacctac

ggcgtgcagtgtctcagccgctaccccgaccacatgaagcagcagcacttctcaagtccgccatgcccgaaggctacgtcca
 ggagcgcaccatcttctcaaggacgacggcaactacaagacccgcgccgaggtgaagtcgagggcgacaccctgggtgaac
 cgcacgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccac
 aacgtctatatcatgcccgcacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcg
 5 tgcagctcgcgcaccactaccagcagaacacccccatcggcgacggcccgctgtgctgcccgcacaaccactacctgagcac
 ccagtcgcacctgagcaaagaccccaacgagaagcgcgatcacatggctctgtgagttcgtgaccgcccgggatcact
 ctcggcatggacgagctgtacaagTAATAAGAGCTCAAGGCAAACCTTTCTATAATTTTAC
 AAGGGAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTTAATTTTGCATATCC
 TTTTAGTCTGCATATATTAAGCATCAGGACCCTTCGTGACAATGTTTACAAA
 10 TTACGTACTAAGGATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTT
 TGAAAAGTTGTTTTATCTGGTGGGAAATTGCTTGACCCAGGTAGTCAAAGGC
 AGTTGACTAGAATCGACAAATTGTTACTCCATATATATATATGTGTGTGTGTG
 TGTGTGTGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGATATCAAAGGCACA
 TGAATATATTTTAATAAAAAACAAATAATATCTCTAATATATCCACACATTTG
 15 TTGCCAGATTTTCAAAAAGTGAAGCTGCAATCGCTTTCCTAAAACAGTAGTGTA
 TTAAATGAACATCTATAAAATGTATCAACACACATTTTAAAAAATTTGTTTAA
 AGTATACTCTTAGGCCAGGCGTGGTGACTCACACCTGTAATTCCAGCACTTCA
 GGAGGCCAAGGTGGGAAGATCATTTGAGTTCAGGAGTTCGAGTTACAGCCTG
 GGCAATAAAGTGAGACCCTGTCACTAACAAAATTAATAAATAAATAAATAT
 20 AAAATATAGGCTTTAAAAAAGCATAGTCTTATTAACCATGTCTGTTGGTCAAA
 ATCTGCAAACCTCTAAAAGAAAGAAAAGAAAGAAAAACCAACGTTAGGGTATT
 TTTCTCCCGTGCCTGAGTCCCAATTACATTCACGACAGTACTTTCAATGAAC
 ATAATTGTTAGGACCACTGAGGAATCATGAAAAATGATCTCTGCTTAGTACA
 TTTGATGCAAAATGACTTATTAGGGGCTGTTTTTCTAGCTATAGTGTCTCGAG
 25 TACTAATATGCAATTATGAAAATTATATTAATCTGGGATTATGACGGTATCA
 CTGTATCATCTTGGTCTTGTCTGGCTGTCACCAAGCATGACCCAGGTCAACT
 TTTTTTTTCCCCTGAATTACCCATCAAATTGATCTGCAGCTGACTAAAGGCCA
 CAGCTGAGCCTGGAAGTGAACCTTCCTTCATCCTCAACCTGCTGTCTCCTCAGA
 AAGCACCAAGGAAAAAGCAGAGAATGACAGCAAACAGATCACTAGGCCTCT
 30 GACCACAGGTGCTGAGTACTCAGCAGCCCTCATATAATAGGTTTGAAAGTAC
 TCCTTAAAAATAAAACACTGTTTCCCTTTGGAAGTATTTACAAGGATGAAACAA
 CCGTATACCTGAGAAATAACTTGCTCTGGTGTCAATTCGCTATTCGCCAGCAG
 ACATCAGAACACACCGAGTTTCCAGATGCTGGTTTTTCCCCTTAAATCAGGAA
 ATACACCTGGACAATTTCTAGAAGACTACAATTCAGTCTAGCCACAAAGGGG
 35 ATTTTTTTTTTTTTGGTAACAGGCTAGAGCCCGGTTCTGTAAAGTCTTTAGCTGAA
 ATGGTCCAGTACAAAAGCACTGGAAATGAGTGGGCTAGGAGGACAAGGACC
 GTCTCCTGCGTGAGGAGTTGGTTGGAGGTCCCCAAGGCCAGGTACCCCCTGC
 ACTCTTATTGGATTCTCTCTGTCTTCTTGGAGTTTTGAAAACTCCTTCGAAC
 ACCAGGCTTTTTTCTTTAGAAAACAAGTCTCCAATCGTTCTCTGTTCCGTAGA
 40 AAGAGAAAGAAAACCTGGAGCAGCTGCTGAAAAATCTAATGAGGAACATAAG
 AGGCAAACCCACCACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCT
 CCCCCGTGCCTTCCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTCTCTAA
 TAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGG
 GGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAG
 45 GCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAATTCAGCTGACGTGC
 CTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGC

GCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCCGGGCT
TTGCCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 45 –CLRN-7eGFP

5 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGGCGTCTGGG
CGACCTTTGGTGCCTCGCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
10 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
15 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGG
GGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
GGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
20 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
TTAATGACGGCTCGTTTCTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
25 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTCGCGAGGGGAGCGC
GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
TGCCTGCGGGGTGTGTGCGTGCGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCAGTTGCTGAGCACGGCC
30 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCGCCCTCG
GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
35 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
CCCTTCTCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
GGGGGACGGGGCAGGGCGGGGTTCCGGCTTCTGGCGTGTGACCGGCGGCTCTA
GAGCCTCTGCTAACCATGTTTCATGCCTTCTTTCTTTCTACAGCTCCTGGGCA
40 ACGTGCTGGTTATTGTGACCGGTGCCACCA_{tggtgagcaaggcgaggagctgttcaccggggtg}
_{gtgccatcctggtcagctggacggcgacgtaaacggccacaagtcagcgtgtccggcgaggcgaggcgatgccacc}
_{tacggcaagctgacctgaagttcatctgaccaccggcaagctgccctgcccaccctcgtgaccacctgacctac}
_{ggcgtgcagtgtctcagccgtaccccgaccacatgaagcagcagcactcttcaagtcgccaatgccgaaggctacgtcca}
_{ggagcgcaccatcttctcaaggacgacggcaactacaagaccgcgccgaggtgaagtcgaggcgacacctgtgtaac}
45 _{cgcacgagctgaaggcgatcgaactcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccac}
_{aacgtctatatcatgcccagacaagcagaagaacggcatcaaggtgaactcaagatccgcacaacatcgaggacggcagcg}

tgcagctcgccgaccactaccagcagaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcac
 ccagtcgcccctgagcaaagaccccaacgagaagcgcgatcacatggctctgctggagttcgtgaccgcccggggtact
 ctggcatggacgagctgtacaagTAAGAGCTCAAGGCCAAACCTTTCTATAATTTTACAAG
 GGAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTTAATTTTGCATATCCTTTT
 5 AGTCTGCATATATTAAGCATCAGGACCCTTCGTGACAATGTTTACAAATTAC
 GTACTAAGGATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTTTGAA
 AAGTTGTTTTATCTGGTGGGAAATTGCTTGACCCAGGTAGTCAAAGGCAGTTG
 ACTAGAATCGACAAATTGTTACTCCATATATATATATGTGTGTGTGTGTGTGT
 GTGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGATATCAAAGGCACATGGA
 10 ATATATTTTAATAAAAACAAATAATATCTCTAATATATCCACACATTTGTTGC
 CAGATTTTCAGAAACTGAGCTGCAATCGCTTTCCTAAAACAGTAGTGTATTA
 AATGAACATCTATAAAATGTATCAACACACATTTTAAAAAATTTGTTTAAAGT
 ATACTCTTAGGCCAGGCGTGGTGACTCACACCTGTAATTCCAGCACTTCAGGA
 GGCCAAGGTGGGAAGATCATTTGAGTTCAGGAGTTCGAGTTACAGCTGTGCC
 15 TTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCT
 GGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGGAAATTGCATCG
 CATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACA
 GCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGG
 GCTCTATGGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGGAACCCC
 20 TAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCC
 GGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGA
 GCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 46 – CLRN-8

25 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCGGG
 CGACCTTTGGTTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACCTCCATCACTAGGGGTTCTGCGGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCC
 30 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 35 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGGCGAGGCGGAGAGGTGCGGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGGCGGCCCTATAAAAAGC
 40 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 45 TGTGCGTGGGGAGCGCCGCTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TCGGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC

GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCGTGCGGGGTGTGTGCGTGGGGGGGGTGAGCAGGGGGGTGTGGGCGCGGGCG
 GTCGGGCTGTAAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 5 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGGCGGGCCCCCGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGGCGCCG
 10 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTG
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGACGGGGCAGGGCGGGGTTCCGGCTTCTGGCGTGTGACCGGCGGGCTCTA
 GAGCCTCTGCTAACCATTGTCATGCCTTCTTCTTTTCTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 15 CATCTTCTGCATGGCCGGCGTGTTTCACTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 20 TCCATCCACGTGAACGTGATCCTGTTTCAAGCGCCATCCTGATCGTGCTGACAAT
 GGTGCGAACC GCCTTCTTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACCTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC
 25 GAGAAGTACACCACCAGCTTTTGGGTTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGCAGCG
 GAGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCG
 GACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCCGATCTGCTGAAAGC
 30 CATTCCTGTGAGCATCCATGTCAACGTATCCTCTTCTCTGCCATCCTCATTGT
 CCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGT
 TTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGC
 TCTTGCGGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAATTCACCA
 CCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACG
 35 CAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTTAAG
 AGCTCGAAGATCATTTGAGTTCAGGAGTTCGAGTTACAGCCTGGGCAATAAA
 GTGAGACCCTGTCATAACAAAATTAATAAATAAATAAATAAATAAATAAATAG
 GCTTTAAAAAAGCATAGTCTTATTAACCATGTCTGTTGGTCAAATCTGCAAA
 CTCTAAAAGAAGAAAAGAAGAAAAAACCAACGTTAGGGTATTTTTCTCCCG
 40 TGCCTGAGTCCCAATTACATTCACGACAGTACTTCAATGAACATAATTGTTA
 GGACCACTGAGGAATCATGAAAAATGATCTCTGCTTAGTACATTTGATGCAA
 AATGACTTATTAGGGGCTGTTTTTCTAGCTATAGTGTCTCGAGTACTAATATG
 CAATTATGAAAATTATATTAATCTGGGATTATGACGGTATCACTGTATCATC
 TTGGTCTTGTCTGGCTGTACCAAGCATGACCCAGGTCAACTTTTTTTTTCCC
 45 CTGAATTACCCATCAAATTGATCTGCAGCTGACTAAAGGCCACAGCTGAGCC
 TGGAAC TGACCTTCCTTCATCCTCAACCTGCTGTCCTCCAGAAAGCACCAAG

GAAAAAGCAGAGAATGACAGCAAACACTGTGCCTTCTAGTTGCCAGCCATCT
 GTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCAC
 TGTCTTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTC
 ATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGG
 5 AAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAAT
 TCAGCTGACGTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCAC
 TCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCGC
 CGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCT
 GCCTGCAGG

10

SEQ ID NO: 47 – CLRN-9

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCGGG
 CGACCTTTGGTGCCTCGCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTTCTGCGGCCGCACGCGTGACATTGATTATT
 15 GACTAGTTATTAATAGTAATCAATTACGGGGTTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 20 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGGCGG
 25 GGCAGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCT
 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 30 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCAGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 35 TGCCTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGGAGCGCCGGCGG
 40 CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 45 GGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTTTCTTCTACAGCTCCTGGGCA

ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTTACGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 5 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTACGCGCCATCCTGATCGTGCTGACAAT
 GGTCGGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 10 CTGCCTGGTCATGATTCTGTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACTACAAAGAGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGCAGCG
 15 GAGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCG
 GACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCGATCTGCTGAAAGC
 CATTCCTGTGAGCATCCATGTCAACGTATCCTCTTCTCTGCCATCCTCATTGT
 CCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGT
 TTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGC
 20 TCTTGGCGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAATTACCA
 CCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACG
 CAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTTAAG
 AGCTCAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAGCCCTCATATAATA
 GGTGTTGAAAGTACTCCTTAAAATAAAACACTGTTTCCCTTTGGAAGTATTAC
 25 AAGGATGAAACAACCGTATACCTGAGAAATAACTTGCTCTGGTGTCAATTCG
 CTATTTCGCCAGCAGACATCAGAACACACCGAGTTTCCAGATGCTGGTTTTTCC
 CCTTAAATCAGGAAATACACCTGGACAATTTCTAGAAGACTACAATTCAGTC
 TAGCCACAAAGGGGATTTTTTTTTTTTGGTAACAGGCTAGAGCCCGGTTCTGT
 AAGTCTTTAGCTGAAATGGTCCAGTACAAAAGCACTGGAAATGAGTGGGCTA
 30 GGAGGACAAGGACCGTCTCCTGCGTGAGGAGTTGGTTGGAGGTCCCCAAGGC
 CAGGTACCCCTGCACTCTTATTGGATTCTCTCTGTCTTCTTGGAGTTTTGAA
 AACTCCTTCGAACACCAGGCTTTTTTCTTTAGAAAACAAGTCTCCAATCGTT
 CTCTGTTCCGTAGAAAGAGAAAGAAAACCTGGAGCAGCTGCTGAAAAATCTA
 ATGAGGAACTAAGAGGCAAACCCACCACTGTGCCTTCTAGTTGCCAGCCATC
 35 TGTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCA
 CTGTCCTTTCTTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGT
 CATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGG
 GAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAA
 TTCAGCTGACGTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCA
 40 CTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGC
 CCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGC
 TGCCTGCAGG

SEQ ID NO: 48 – CLRN-10

45 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCCGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT

GGCCAACTCCATCACTAGGGGTTCTGCGGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTTCATTAGTTCATAGCCCATATA
 TGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 5 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 10 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCT
 15 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 20 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 25 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 CTGTGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGACCCCCCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 30 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 35 CATCTTCTGCATGGCCGGCGTGTTTCAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 40 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 GGTGCGAACCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACCTACAAAGAGGGCACCTACGTCTACAAGACCCAGTCC
 45 GAGAAGTACACCACAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCAGTTTCCATTGCCAAGA

GCAAGGACGCCGAAACCAACAAACGTGGCCGCCGATCTGATGTACGGATCCTA
 TCCCTATGATGTGCCAGACTATGCTAAGCGGAAGAGAAGAGGCCGAAGGCAG
 AGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCGGACCTATGCAG
 GCTCTGCAGCAGCAGCCAGTGTTCCCGATCTGCTGAAAGCCATTCCTGTGAG
 5 CATCCATGTCAACGTCATCCTCTTCTCTGCCATCCTCATTGTCTCACTATGGT
 TGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGTTTGAAACCCTGC
 ACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGCTCTTGCGGCTGC
 CTCGTGATGATCCTCTTTGCCTCTGAAGTCAAAATTCACCACCTGTCTGAGAA
 AATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACGCAGAGCGAAAA
 10 GTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTGGATCCGACTACAAA
 GACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
 ACAAGTAAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCAT
 CTGTTGTTTGCCCCCTCCCCCGTGCTTCCTTGACCCTGGAAGGTGCCACTCCC
 ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTG
 15 TCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGG
 GAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAA
 TTCAGCTGACGTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCA
 CTCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCGC
 CCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGC
 20 TGCCTGCAGG

SEQ ID NO: 49 – CLRN-10myc

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
 CGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 25 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 30 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 35 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 40 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 45 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG

GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG
 5 GGGCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 10 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTTCAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 15 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 GGTCCGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 20 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACTACAAAGAGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 25 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGATCCG
 AGCAGAAACTCATCTCTGAAGAAGATCTGGAACAAAAGTTGATTTTCAGAAGA
 AGATCTGGAACAGAAGCTCATCTCTGAGGAAGATCTGAAGCGGAAGAGAAG
 AGGCGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCC
 CGGACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCCGATCTGCTGAAA
 30 GCCATTCTGTGATCATCCATGTCAACGTATCCTCTTCTCTGCCATCCTCATT
 GTCCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCC
 GTTTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCG
 GCTCTTGCGGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAATTCAC
 CACCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAAA
 35 CGCAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTGG
 ATCCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTAC
 AAGGATGACGATGACAAGTAAGAGCTCGCTGATCAGCCTCGACTGTGCCTTC
 TAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCCTTGACCCTGGA
 AGGTGCCACTCCCACTGTCCTTTTCTAATAAAATGAGGAAATTGCATCGCATT
 40 GTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA
 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTC
 TATGGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGGAACCCCTAGT
 GATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGC
 GACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCCGGGCGGCCTCAGTGAGCGA
 45 GCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 50- CLRN-10NF

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCTGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 5 GACTAGTTATTAATAGTAATCAATTACGGGGTTCATTAGTTCATAGCCCATATA
 TGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCC
 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 10 TCAATGACGGTAAATGGCCCCGCTGGCATTATGCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGGCGG
 15 GCGGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 20 TTAATGACGGCTCGTTTCTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 25 TGCCTGCGGGGTGTGTGCGTGGGGGGGGTGAGCAGGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 30 CTGTCGAGGCGCGGCGAGCCGACGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 35 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATTGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTACGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 40 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTTCAGATTCAGCTTCTTCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTACGCGCCATCCTGATCGTGCTGACAAT
 GGTCCGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 45 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC

GAGAAGATCGCCAACTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGATCCTA
 5 TCCCTATGATGTGCCAGACTATGCTAAGGGCGAAGGCAGAGGCAGCCTGCTT
 ACATGTGGCGACGTGGAAGAGAACCCCGGACCTATGCAGGCTCTGCAGCAGC
 AGCCAGTGTTCCCGATCTGCTGAAAGCCATTCCCTGTCAGCATCCATGTCAAC
 GTCATCCTCTTCTCTGCCATCCTCATTGTCTCTCACTATGGTTGGAACGGCCTTT
 TTTATGTATAATGCCTTTGGGAAGCCGTTTGAAACCCTGCACGGACCCCTGGG
 10 ACTCTATCTCCTGAGCTTCATCTCCGGCTCTTGCGGCTGCCTCGTGATGATCCT
 CTTTGCCTCTGAAGTCAAAATTCACCACCTGTCTGAGAAAATTGCTAATTACA
 AAGAAGGGACATACGTTTACAAAACGCAGAGCGAAAAGTATACGACCAGCT
 TCTGGCTGACCAAGGGCCACTCTGGATCCGACTACAAAGACCATGACGGTGA
 TTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGAGCTC
 15 GCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCT
 CCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTTCTCTAA
 TAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGG
 GGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAG
 GCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAATTCAGCTGACGTGC
 20 CTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGC
 GCTCGCTCGCTCACTGAGGCCGGGCGACCAAGGTGCGCCGACGCCCGGGCT
 TTGCCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 51 – CLRN-11

25 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCCGGG
 CGACCTTTGGTTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 30 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 35 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGGCGAGGCGGAGAGGTGCGGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGGCGGCCCTATAAAAAGC
 40 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 45 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TCGGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC

GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCGTGCGGGGTGTGTGCGTGCGGGGGGTGAGCAGGGGGTGTGGGCGCGGGCG
 GTCGGGCTGTAAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 5 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGGCGGCCCCCGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGGCGCCG
 10 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
 GGGGACGGGGCAGGGCGGGGTTCCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATTGTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 15 CATCTTCTGCATGGCCGGCGTGTTTCACTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 20 TCCATCCACGTGAACGTGATCCTGTTTCAAGCGCCATCCTGATCGTGCTGACAAT
 GGTCCGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACCTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC
 25 GAGAAGTACACCACCAGCTTTTGGGTTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGATCCTA
 TCCCTATGATGTGCCAGACTATGCTAAGCGGAAGAGAAGAGGCGAAGGCAG
 AGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCCCGGACCTATGCAG
 30 GCTCTGCAGCAGCAGCCAGTGTTCCCCGATCTGCTGAAAGCCATTCTGTGAG
 CATCCATGTCAACGTCATCCTCTTCTCTGCCATCCTCATTGTCTCACTATGGT
 TGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCCGTTTGAAACCCTGC
 ACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCGGCTCTTGCGGCTGC
 CTCGTGATGATCCTCTTTGCCTCTGAAGTCAAATTCACCACCTGTCTGAGAA
 35 AATTGCTAATTACAAAGAAGGGACATACGTTTACAAAACGCAGAGCGAAAA
 GTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTGGATCCGACTACAAA
 GACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
 ACAAGTAAGAGCTCAAGGCAAACCTTTCTATAATTTTACAAGGGAGTAGACT
 TGCTTTGGTCACTTTTAGATGTGGTTAATTTTGCATATCCTTTTAGTCTGCATA
 40 TATTAAAGCATCAGGACCCTTCGTGACAATGTTTACAAATTACGTACTAAGG
 ATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTT
 ATCTGGTGGGAAATTGCTTGACCCAGGTAGTCAAAGGCAGTTGACTAGAATC
 GACAAATTGTTACTCCATATATATATATATGTGTGTGTGTGTGTAAGATGTCT
 TCCTATCAAAAAGATATCAAAGGCACATGGAATATATTTTAATAAAAAACAAA
 45 TAATATCTCTAATATATCCACACATTTGTTGCCAGATTCAGAAAACCTGAGCT
 GCAATCGCTTTCCTAAAACAGTAGTGTATTAAATGAACATCTATAAAATGTAT

SEQ ID NO: 52 – CLRN-11myc

227

TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGC GG GCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 5 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGC GTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 10 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGGCGGCCCCCGAGCGCCGGCGG
 CTGTGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGC GCCCGCCGCACCCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGCTC
 15 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTTCAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 20 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 25 GGTGGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCA ACTACAAAGAGGGCACCTACGTCTACAAGACCCAGTCC
 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 30 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGATCCG
 AGCAGAAACTCATCTCTGAAGAAGATCTGGAACAAAAGTTGATTTTCAAGA
 AGATCTGGAACAGAAGCTCATCTCTGAGGAAGATCTGAAGCGGAAGAGAAG
 AGGCGAAGGCAGAGGCAGCCTGCTTACATGTGGCGACGTGGAAGAGAACCC
 35 CGGACCTATGCAGGCTCTGCAGCAGCAGCCAGTGTTCCCCGATCTGCTGAAA
 GCCATTCTGTGAGCATCCATGTCAACGTCATCCTCTTCTCTGCCATCCTCATT
 GTCCTCACTATGGTTGGAACGGCCTTTTTTATGTATAATGCCTTTGGGAAGCC
 GTTTGAAACCCTGCACGGACCCCTGGGACTCTATCTCCTGAGCTTCATCTCCG
 GCTCTTGCGGCTGCCTCGTGATGATCCTCTTTGCCTCTGAAGTCAAAATTCAC
 40 CACCTGTCTGAGAAAATTGCTAATTACAAAGAAGGGACATACGTTTACAAA
 CGCAGAGCGAAAAGTATACGACCAGCTTCTGGCTGACCAAGGGCCACTCTGG
 ATCCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTAC
 AAGGATGACGATGACAAGTAAGAGCTCAAGGCAAACCTTTCTATAATTTTAC
 AAGGGAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTTAATTTTGCATATCC
 45 TTTTAGTCTGCATATATTAAAGCATCAGGACCCTTCGTGACAATGTTTACAAA
 TTACGTACTAAGGATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTT

TGAAAAGTTGTTTTATCTGGTGGGAAATTGCTTGACCCAGGTAGTCAAAGGC
 AGTTGACTAGAAATCGACAAATTGTTACTCCATATATATATATATGTGTGTGTG
 TGTGTAAGATGTCTTCCTATCAAAAAGATATCAAAGGCACATGGAATATATTT
 TAATAAAAACAAATAATATCTCTAATATATCCACACATTTGTTGCCAGATTTT
 5 AGAAAAGTGAAGCTGCAATCGCTTTCCTAAAACAGTAGTGTATTAAATGAACA
 TCTATAAAATGTATCAACACACATTTTAAAAAATTTGTTTAAAGTATACTCTT
 AGGCCAGGCGTGGTGACTCACACCTGTAATTCCAGCACTTCAGGAGGCCAAG
 GTGGGAAGATCATTGAGTTCAGGAGTTCGAGTTACAGCCTGGGCAATAAAG
 TGAGACCCTGTCACATAACAAAATTAATAAAATAAAATAAATATAAAATATAGG
 10 CTTTAAAAAAGCATAGTCTTATTAACCATGTCTGTTGGTCAAAATCTGCAAAC
 TCTAAAAGAAGAAAAGAAGAAAAAACCAACGTTAGGGTATTTTTCCTCCCGT
 GCCTGAGTCCCAATTACATTCACGACAGTACTTCAATGAACATAATTGTTAG
 GACCACTGAGGAATCATGAAAAATGATCTCTGCTTAGTACATTTGATGCAAA
 ATGACTTATTAGGGGCTGTTTTTCTAGCTATAGTGTCTCGAGTACTAATATGC
 15 AATTATGAAAATTATATTAAATCTGGGATTATGACGGTATCACTGTATCATCT
 TGGTCTTGTCTGGCTGTCACCAAGCATGACCCAGGTCAACTTTTTTTTTTCCCC
 TGAATTACCCATCAAATTGATCTGCAGCTGACTAAAGGCCACAGCTGAGCCT
 GGAAGTGAACCTTCCCTCATCCTCAACCTGCTGTCCTCCAGAAAGCACCAAGG
 AAAAAGCAGAGAATGACAGCAAACAGATCACTAGGCCTCTGACCACAGGTG
 20 CTGAGTACTCAGCAGCCCTCATATAATAGGTTTGAAAGTACTCCTTAAAATAA
 AACACTGTTTCCCTTTGGAAGTATTTACAAGGATGAAACAACCGTATACCTGA
 GAAATAACTTGCTCTGGTGTCAATTCGCTATTCGCCAGCAGACATCAGAACA
 CACCGAGTTTCCAGATGCTGGTTTTTCCCCTTAAATCAGGAAATACACCTGGA
 CAATTTCTAGAAGACAGGCCTCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTT
 25 TGCCCCCTCCCCCGTGCCTTCCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCT
 TTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCAATTCTA
 TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGACA
 ATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAGCTTGAATTCAGCT
 GACGTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTC
 30 TCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGC
 CCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC
 AGG

SEQ ID NO: 53 – CLRN-11NF

35 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
 CGACCTTTGGTGCCTCGCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCCCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 40 CAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 45 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC

GGGGGGGGGGGGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 5 CCGCGCCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 10 TGC GG GCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGC GTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 15 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GCGCCCGCCGCACCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 20 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGT
 CCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
 GGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 25 CATCTTCTGCATGGCCGGCGTGTTCAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTG
 30 TCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGACAAT
 GGTCCGAACCGCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACA
 CTGCATGGACCTCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGG
 CTGCCTGGTCATGATTCTGTTCCGCCAGCGAAGTGAAGATCCACCACCTGAGC
 GAGAAGATCGCCAACTACAAAGAGGGGCACCTACGTCTACAAGACCCAGTCC
 35 GAGAAGTACACCACCAGCTTTTGGGTATCTTCTTCTGTTTCTTCGTGCACTTC
 CTGAACGGCCTGCTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGA
 GCAAGGACGCCGAAACCACAAACGTGGCCGCCGATCTGATGTACGGATCCTA
 TCCCTATGATGTGCCAGACTATGCTAAGGGCGAAGGCAGAGGCAGCCTGCTT
 ACATGTGGCGACGTGGAAGAGAACCCCGGACCTATGCAGGCTCTGCAGCAGC
 40 AGCCAGTGTTCCCGATCTGCTGAAAGCCATTCTGTCAGCATCCATGTCAAC
 GTCATCCTCTTCTGCCATCCTCATTGTCCTCACTATGGTTGGAACGGCCTTT
 TTTATGTATAATGCCTTTGGGAAGCCGTTTGAAACCCTGCACGGACCCCTGGG
 ACTCTATCTCCTGAGCTTCATCTCCGGCTCTTGCGGCTGCCTCGTGATGATCCT
 CTTTGCCTCTGAAGTCAAAATTCACCACCTGTCTGAGAAAATTGCTAATTACA
 45 AAGAAGGGACATACGTTTACAAAACGCAGAGCGAAAAGTATACGACCAGCT
 TCTGGCTGACCAAGGGCCACTCTGGATCCGACTACAAAGACCATGACGGTGA

TTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAAGAGCTC
 AAGGCAAACCTTTCTATAATTTTACAAGGGAGTAGACTTGCTTTGGTCACTTT
 TAGATGTGGTTAATTTTGCATATCCTTTTAGTCTGCATATATTAAAGCATCAG
 GACCCTTCGTGACAATGTTTACAAATTACGTACTAAGGATACAGGCTGGAAA
 5 GTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTTATCTGGTGGGAAAT
 TGCTTGACCCAGGTAGTCAAAGGCAGTTGACTAGAATCGACAAATTGTTACT
 CCATATATATATATATGTGTGTGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGA
 TATCAAAGGCACATGGAATATATTTTAATAAAAAACAAATAATATCTCTAATA
 TATCCACACATTTGTTGCCAGATTTTCAGAAAAGTCTGAGCTGCAATCGCTTTCT
 10 AAAACAGTAGTGTATTAAATGAACATCTATAAAATGTATCAACACACATTTT
 AAAAAATTTGTTTAAAGTATACTCTTAGGCCAGGCGTGGTGACTCACACCTGT
 AATTCAGCACTTCAGGAGGCCAAGGTGGGAAGATCATTGAGTTCAGGAGT
 TCGAGTTACAGCCTGGGCAATAAAGTGAGACCCTGTCACTAACAAAATTAAA
 AAATAAAATAAATATAAAATATAGGCTTTAAAAAAGCATAGTCTTATTAACC
 15 ATGTCTGTTGGTCAAAATCTGCAAACCTCTAAAAGAAGAAAAGAAGAAAAA
 CCAACGTTAGGGTATTTTTCTCCCGTGCCTGAGTCCCAATTACATTCACGAC
 AGTACTTTCAATGAACATAATTGTTAGGACCACTGAGGAATCATGAAAAATG
 ATCTCTGCTTAGTACATTTGATGCAAAATGACTTATTAGGGGCTGTTTTCTA
 GCTATAGTGTCTCGAGTACTAATATGCAATTATGAAAATTATATTAAATCTGG
 20 GATTATGACGGTATCACTGTATCATCTTGGTCTTGTTCTGGCTGTCACCAAGC
 ATGACCCAGGTCAACTTTTTTTTTCCCTGAATTACCCATCAAATTGATCTGC
 AGCTGACTAAAGGCCACAGCTGAGCCTGGAAGTACCCTTCCTTCATCCTCA
 ACCTGCTGTCCTCCAGAAAGCACCAAGGAAAAAGCAGAGAATGACAGCAAA
 CAGATCACTAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAGCCCTCATAT
 25 AATAGGTTTGAAAGTACTCCTTAAATAAAACACTGTTTCCCTTTGGAAGTAT
 TTACAAGGATGAAACAACCGTATACCTGAGAAATAACTTGCTCTGGTGTCAA
 TTCGCTATTTCGCCAGCAGACATCAGAACACACCCGAGTTTCCAGATGCTGGTTT
 TTCCCTTAAATCAGGAAATACACCTGGACAATTTCTAGAAGACAGGCCTCT
 GTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTG
 30 ACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGGAAATTGC
 ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAG
 GACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCG
 GTGGGCTCTATGGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGGAA
 CCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGA
 35 GGCCGGGCGACCAAAGGTGCGCCGACGCCGGGCTTTGCCCCGGGCGGCCTCA
 GTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 54 – CLRN-12

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
 40 CGACCTTTGGTGCCTCGCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 45 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG

TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTTTAATTATTTTGTGCAGCGATGGGGGGC
 5 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 10 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 15 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 20 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 25 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTAGGAGATACTTGAAGGCAGTTTGAAAGAC
 TTGTTTTACAGATTCTTAGTCCAAAGATTTCCAATTAGGGAGAAGAAGCAGC
 30 AGAAAAGGAGAAAAGCCAAGTATGAGTGATGATGAGGCCTTCATCTACTGAC
 ATTTAACCTGGCGAGAACCGTCGATGGTGAAGTTGCCTTTTCAGCTGGGAGCT
 GTCCGTTACAGTTCCGTAATAAATGCAGTCAAAGAGGCAGTCCCTTCCCATTG
 CTCACAAAGGTCTTGTTTTTGAACCTCGCCCTCACAGAAGCCGTTTCTCATCG
 CCACCATGCCTAGCCAGCAGAAGAAAATCATCTTCTGCATGGCCGGCGTGTT
 35 CAGCTTCGCCTGTGCTCTGGGAGTTGTGACAGCCCTGGGAACCCCTCTGTGGA
 TCAAAGCCACAGTGCTGTGCAAGACAGGCGCCCTGCTGGTTAATGCCTCTGG
 CCAAGAGCTGGACAAGTTTCATGGGCGAGATGCAGTACGGCCTGTTCCATGGC
 GAAGGCGTCAGACAGTGTGGCCTGGGAGCCAGACCTTTCAGATTCAGCTTCT
 TCCCAGACCTGCTGAAGGCTATCCCCGTGTCCATCCACGTGAACGTGATCCTG
 40 TTCAGCGCCATCCTGATCGTGCTGACAATGGTTCGGAACCGCCTTCTTCATGTA
 CAACGCCTTCGGCAAGCCCTTCGAGACACTGCATGGACCTCTGGGCCTGTAC
 CTGCTGAGCTTTATCAGCGGCAGCTGTGGCTGCCTGGTTCATGATTCTGTTTCG
 CAGCGAAGTGAAGATCCACCACCTGAGCGAGAAGATCGCCAACCTACAAAGA
 GGGCACCTACGTCTACAAGACCCAGTCCGAGAAGTACACCACCAGCTTTTGG
 45 GTTATCTTCTTCTGTTTCTTCGTGCACTTCTGAACGGCCTGCTGATCAGACTG
 GCCGGCTTCAGTTTCCATTCGCCAAGAGCAAGGACGCCGAAACCACAAACG

TGGCCGCCGATCTGATGTACGGATCCTATCCCTATGATGTGCCAGACTATGCT
 TAAGAGCTCAAGGCAAACCTTTCTATAATTTTACAAGGGAGTAGACTTGCTTT
 GGTCACTTTTAGATGTGGTTAATTTTGCATATCCTTTTAGTCTGCATATATTAA
 AGCATCAGGACCCTTCGTGACAATGTTTACAAATTACGTACTAAGGATACAG
 5 GCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTTTGAAAAGTTGTTTTATCTGG
 TGGGAAATTGCTTGACCCAGGTAGTCAAAGGCAGTTGACTAGAATCGACAAA
 TTGTTACTCCATATATATATATGTGTGTGTGTGTGTGTGTAAAGATGTCTTCCTATC
 AAAAAGATATCAAAGGCACATGGAATATATTTTAATAAAAAACAAATAATATC
 TCTAATATATCCACACATTTGTTGCCAGATTTTCAAGAAAAGTCTGAGCTGCAATCG
 10 CTTTCCTAAAACAGTAGTGTATTAAATGAACATCTATAAAATGTATCAACACA
 CATTTTAAAAAATTTGTTTAAAGTATACTCTTAGGCCAGGCGTGGTGACTCAC
 ACCTGTAATTCAGCACTTCAGGAGGCCAAGGTGGGAAGATCATTGAGTTC
 AGGAGTTCGAGTTACAGTCTGGGCAATAAAGTGAGACCCTGTCACTAACAAA
 ATTAATAAAATAAAATAAATATAAAATATAGGCTTTAAAAAAGCATAGTCTTA
 15 TTAACCATGTCTGTTGGTCAAAATCTGCAAACTCTAAAAGAAGAAAAGAAGA
 AAAAACCAACGTTAGGGTATTTTTCCTCCCGTGCCTGAGTCCCAATTACATTC
 ACGACAGTACTTTCAATGAACATAATTGTTAGGACCACTGAGGAATCATGAA
 AAATGATCTCTGCTTAGTACATTTGATGCAAAATGACTTATTAGGGGCTGTTT
 TTCTAGCTATAGTGTCTCGAGTACTAATATGCAATTATGAAAATTATATTTAA
 20 TCTGGGATTATGACGGTATCACTGTATCATCTTGGTCTTGTTCTGGCTGTCACC
 AAGCATGACCCAGGTCAACTTTTTTTTTTCCCTGAATTACCCATCAAATTGAT
 CTGCAGCTGACTAAAGGCCACAGCTGAGCCTGGAAGTACCCCTTCCTTCATC
 CTCAACCTGCTGTCCTCCAGAAAGCACCAAGGAAAAAGCAGAGAATGACAG
 CAAACAGATCACTAGGCCTCTGACCACAGGTGCTGAGTACTCAGCAGCCCTC
 25 ATATAATAGGTTTGAAAGTACTCCTTAAAATAAAACACTGTTTCCCTTTGGAA
 CTATTTACAAGGATGAAACAACCGTATACCTGAGAAATAAAGTCTGCTGCTGGTG
 TCAATTCGCTATTCGCCAGCAGACATCAGAACACACCGAGTTTCCAGATGCT
 CTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCT
 TGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCTAATAAAATGAGGAAATT
 30 GCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGC
 AGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATG
 CGGTGGGCTCTATGGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGG
 AACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACT
 GAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCCGGGCGGCCT
 35 CAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

SEQ ID NO: 55 – CLRN-13

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 40 GGCCAACTCCATCACTAGGGGTTCCCTGCGGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCAATTAGTTCATAGCCCATATA
 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 45 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTATG

GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 GGGGGGGGGGGGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 5 GCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 10 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCCGGCGGCTGTGAGCGC
 TCGGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 15 TCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGGCGGCCCCCGAGCGCCGGCGG
 20 CTGTCGAGGCGCGGCGAGCCGACGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GCGCCCGCCGACCCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 25 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTTCAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 30 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCGGTTTCAGCTGCTACTTTCTGGACCCCTTCATGGGCCTGCCT
 ACCGGCGTTCCACATCTGCTGTCTCTGCCTTGACGACACAGCTGCAGAAGAG
 AGCACACCAGCGAGAGAGTGCAAGAGCCTGCCGGCTGTTTTTCTGCCGTGCG
 35 GTCTAAACTGCACGCCGACCTGCTGCCGCCACCAGCTTTTCTAGATTCGCCC
 AGAGCAACCCAGCGAGCACCTAGACAGTGTACAGCCTGCTGTGTACCCC
 CTACTGTGTGAATCACGGCGGCGATTCTCTGCTGCATGTGCAGTGCTTTTGGGA
 AAACCTTCGGCAGCGACTACAAGGACCACGACGGCGATTATAAGGATCACG
 ATATCGATTACAAGGACGATGACGACAAGGGCGAAGGCAGAGGCTCCCTGCT
 40 GACATGCGGAGATGTGGAAGAGAACCCCGGACCTATGCCTTCTCAGCAGAAA
 AAGATTATTTTCTGTATGGCTGGGGTGTTCCTTCGCTTGCGCCCTGGGTGTT
 GTTACCGCTCTCGGAACACCACTGTGGATTAAGGCTACCGTCCTGTGTAAAC
 CGGCGCTCTGCTCGTGAATGCCAGCGGACAAGAACTGGATAAGTTTATGGGA
 GAAATGCAATATGGGCTCTTTCACGGCGAGGGTGTAGACAGTGCGGACTCG
 45 GCGCTAGACCCTTCAGATTCAGCTTCTTCCAGACCTGCTGAAGGCTATCCCC
 GTGTCCATCCACGTGAACGTGATCCTGTTTCAGCGCCATCCTGATCGTGCTGAC

235

CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTTG
 5 TGTGCGTGCGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGCGGGGGGTGAGCAGGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 10 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGCGCGGGGCTCGCCGTGCC
 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGAGCGCCGGCGG
 CTGTCGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 15 GGCGCCGCCGCACCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGGACGGGGCAGGGCGGGGTTTCGGCTTCTGGCGTGTCACGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 20 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 CATCTTCTGCATGGCCGGCGTGTTACGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 25 CCAGACCTTTCCGGTTCAGCTGCTACTTTCTGGACCCCTTCATGGGCCTGCCT
 ACCGGCGTTCCACATCTGCTGTCTCTGCCTTGCAGCACCAAGCTGCAGAAGAG
 AGCACACCAGCGAGAGAGTGCAAGAGCCTGCCGGCTGTTTTTCTGCCGTGCG
 GTCTAAACTGCACGCCGACCTGCTGCCGCCACCAGCTTTTCTAGATTCGCCC
 AGAGCAACCCAGCGAGCACCTAGACAGTGTACAGCCTGCTGTGTACCCC
 30 CTACTGTGTGAATCACGGCGGCGATTCTCTGCTGCATGTGCAGTGCTTTTGGA
 AAACCTTCGGCAGCGGCGAAGGCAGAGGCTCCCTGCTGACATGCGGAGATGT
 CGAAGAGAACCCCGGACCTATGCCTTCTCAGCAGAAAAAGATTATTTTCTGT
 ATGGCTGGGGTGTTCTCCTTCGCTTGCGCCCTGGGTGTTGTTACCGCTCTCGG
 AACACCACTGTGGATTAAGGCTACCGTCCTGTGTAAAACCGGCGCTCTGCTC
 35 GTGAATGCCAGCGGACAAGAACTGGATAAGTTTATGGGAGAAATGCAATATG
 GGCTCTTTCACGGCGAGGGTGTTAGACAGTGCGGACTCGGCGCTAGACCCTT
 CAGATTCAGCTTCTTCCCAGACCTGCTGAAGGCTATCCCCGTGTCCATCCACG
 TGAACGTGATCCTGTTACGCGCCATCCTGATCGTGCTGACAATGGTCGGAACC
 GCCTTCTTCATGTACAACGCCTTCGGCAAGCCCTTCGAGACACTGCATGGACC
 40 TCTGGGCCTGTACCTGCTGAGCTTTATCAGCGGCAGCTGTGGCTGCCTGGTCA
 TGATTCTGTTTCGCCAGCGAAGTGAAGATCCACCACCTGAGCGAGAAGATCGC
 CAACTACAAAGAGGGCACCTACGTCTACAAGACCCAGTCCGAGAAGTACACC
 ACAAGCTTTTGGGTTATCTTCTTCTGTTTCTTCGTGCACTTCCTGAACGGCCTG
 CTGATCAGACTGGCCGGCTTCCAGTTTCCATTTCGCCAAGTCCAAGGACGCCG
 45 AAACCACAAACGTGGCCGCCGACCTGATGTACTAAGAGCTCAAGGCAAACCT
 TTCTATAATTTTACAAGGGAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTT

20

45

238

GGCCAACTCCATCACTAGGGGTTCTCGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 TGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 5 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 10 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 15 CCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGT
 TTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC
 GGGAGGGCCCTTTGTGCGGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGC
 20 TGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGC
 GGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGGAACAAAGGC
 TGCCTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCG
 GTCGGGCTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC
 CGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCC
 25 GGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCGCTCG
 GGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGG
 CTGTGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG
 AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGA
 GCGCCCGCCGACCCCCCTTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCG
 30 GCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC
 CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGG
 GGGGACGGGGCAGGGCGGGGTTCCGCTTCTGGCGTGTGACCGGCGGCTCTA
 GAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCA
 ACGTGCTGGTTATTGTGACCGGTGCCACCATGCCTAGCCAGCAGAAGAAAAT
 35 CATCTTCTGCATGGCCGGCGTGTTAGCTTCGCCTGTGCTCTGGGAGTTGTGA
 CAGCCCTGGGAACCCCTCTGTGGATCAAAGCCACAGTGCTGTGCAAGACAGG
 CGCCCTGCTGGTTAATGCCTCTGGCCAAGAGCTGGACAAGTTCATGGGCGAG
 ATGCAGTACGGCCTGTTCCATGGCGAAGGCGTCAGACAGTGTGGCCTGGGAG
 CCAGACCTTTCGGGTTAGCTGCTACTTCTGGACCCCTTCATGGGCCTGCCT
 40 ACCGGCGTTCCACATCTGCTGTCTCTGCCTTGACGACCAAGCTGCAGAAGAG
 AGCACACCAGCGAGAGAGTGCAAGAGCCTGCCGGCTGTTTTTCTGCCGTGCG
 GTCTAAACTGCACGCCGACCTGCTGCCGCCACCAGCTTTTCTAGATTCGCCC
 AGAGCAACCCAGCGAGCACCTAGACAGTGTACAGCCTGCTGTGTACCCC
 CTACTGTGTGAATCACGGCGGCGATTCTCTGCTGCATGTGCAGTGCTTTTGGGA
 45 AAACCTTCGGCAGCTAAGAGCTCAAGGCAAACCTTTCTATAATTTTACAAGG
 GAGTAGACTTGCTTTGGTCACTTTTAGATGTGGTTAATTTTGCATATCCTTTTA

GTCTGCATATATTAAAGCATCAGGACCCTTCGTGACAATGTTTACAAATTACG
 TACTAAGGATACAGGCTGGAAAGTAAGGGAAGCAGAAGGAAGGCTTTGAAA
 AGTTGTTTTATCTGGTGGGAAATTGCTTGACCCAGGTAGTCAAAGGCAGTTGA
 CTAGAATCGACAAATTGTTACTCCATATATATATATGTGTGTGTGTGTGTG
 5 TGTGTGTGTGTAAGATGTCTTCCTATCAAAAAGATATCAAAGGCACATGGAA
 TATATTTTAATAAAAAACAAATAATATCTCTAATATATCCACACATTTGTTGCC
 AGATTTTCAGAAAACCTGAGCTGCAATCGCTTTCCTAAAACAGTAGTGTATTAA
 ATGAACATCTATAAAATGTATCAACACACATTTTAAAAAATTTGTTTAAAGTA
 TACTCTTAGGCCAGGCGTGGTGACTCACACCTGTAATTCCAGCACTTCAGGAG
 10 GCCAAGGTGGGAAGATCATTTGAGTTCAGGAGTTCGAGTTACAGgctgatcagcctc
 gaCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTC
 CTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGGAAA
 TTGCATCGCATTGTCTGAGTAGGTGTCTATTCTATTCTGGGGGGTGGGGTGGGG
 CAGGACAGCAAGGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT
 15 GCGGTGGGCTCTATGGAAGCTTGAATTCAGCTGACGTGCCTCGGACCGCTAG
 GAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAC
 TGAGGCCGGGCGACCAAGGTGCGCCGACGCCCGGGCTTTGCCCGGGCGGCC
 TCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
 20 **SEQ ID NO: 59 – CLRN-17**
 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTTCT
 25 **SEQ ID NO: 60 – CLRN-18**
 CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCGTCGGG
 CGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGT
 GGCCAACTCCATCACTAGGGGTTTCTGCGGCCGCACGCGTGACATTGATTATT
 GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATA
 30 TGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCC
 CAACGACCCCCGCCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGC
 CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACG
 TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATG
 35 GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGG
 GTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCC
 ACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGC
 GGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGG
 GCGGGGCGGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTC
 40 CGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGC
 GAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCT
 CCGCGCCGCTCGCGCCGCCCCGCCCCGGCTCTGACTGACCGCGTTACTCCAC
 AGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTtagcgttggttaatgac
 ggcttggttctttctgtggtgcgtgaaagccttgaggggctccgggagctagagcctctgtaacctgttccttctttt
 45 cctacagtcctgggaacgtgctgttattgtctgtctcatcttttgcaagAGGAGATACTTGAAGGCAG
 TTTGAAAGACTTGTTTTACAGATTCTTAGTCCAAAGATTTCCAATTAGGGAGA

241

TTCTGTAAGTCTTTAGCTGAAATGGTCCAGTACAAAAGCACTGGAAATGAGT
GGGCTAGGAGGACAAGGACCGTCTCCTGCGTGAGGAGTTGGTTGGAGGTCCC
CAAGGCCAGGTACCCCTGCACTCTTATTGGATTCTCTGTCTTCTTGGAG
TTTTGAAAAACTCCTTCGAACACCAGGCTTTTTTCTTTAGAAAACAAGTCTCC
5 AATCGTTCTCTGTTCCGTAGAAAGAGAAAGAAAACCTGGAGCAGCTGCTGAA
AAATCTAATGAGGAACTAAGAGGCAAACCCACCACTGTGCCTTCTAGTTGCC
AGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAG
TAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAG
10 GATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAA
GCTTGAATTCAGCTGACGTGCCTCGGACCGCTAGGAACCCCTAGTGATGGAG
TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCCGGGCTTTGCCCCGGGCGGCCTCAGTGAGCGAGCGAGCG
CGCAGCTGCCTGCAGG
15

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 CLRN1 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 CLRN1 protein;

at least one of the coding sequences comprises a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive 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 CLRN1 protein.

2. The composition of claim 1, wherein each of the at least two different vectors includes a coding sequence that encodes a different portion of a first isoform of the CLRN1 protein.

3. The composition of claim 2, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

4. The composition of claim 3, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

5. The composition of claim 4, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

6. The composition of claim 2, wherein one of the at least two different nucleic acid vectors further comprises a sequence that encodes a second isoform of the CLRN1 protein.

7. The composition of claim 6, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

8. The composition of claim 7, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

9. The composition of claim 8, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

10. The composition of claim 1, wherein each of the at least two different vectors includes a coding sequence that encodes a different portion of a second isoform of the CLRN1 protein.

11. The composition of claim 10, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

12. The composition of claim 11, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

13. The composition of claim 12, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

14. The composition of claim 13, wherein one of the at least two different nucleic acid vectors further comprises a sequence that encodes a first isoform of the CLRN1 protein.

15. The composition of claim 14, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

16. The composition of claim 15, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

17. The composition of claim 16, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

18. The composition of any one of claims 1-17, wherein at least one of the at least two different vectors comprises a 5' untranslated region (UTR), a 3' UTR, or both.

19. The composition of claim 18, wherein the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12.

20. The composition of claim 19, wherein the 5' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12.

21. The composition of claim 20, wherein the 5' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

22. The composition of claim 21, wherein the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12.

23. The composition of claim 18, wherein the 5'UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 12.

24. The composition of claim 18, wherein the 3' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15.

25. The composition of claim 24, wherein the 3' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15.

26. The composition of claim 25, wherein the 3' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15.

27. The composition of claim 26, wherein the 3' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15.

28. The composition of claim 18, wherein the 3' UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 15.

29. 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.

30. 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).

31. 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.

32. The composition of claim 31, wherein each of the at least two different vectors is an AAV vector.

33. The composition of any one of claims 1-32, 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.

34. The composition of any one of claims 1-32, 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.

35. The composition of claim 34, wherein the overlapping amino acid sequence is between about 30 amino acid residues to about 202 amino acid residues in length.

36. The composition of any one of claims 1-32, 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 CLRN1 genomic DNA, and wherein the two different intron segments overlap in sequence by at least 100 nucleotides.

37. The composition of claim 36, wherein the two different intron segments overlap in sequence by 100 nucleotides to about 800 nucleotides.

38. The composition of any one of claims 1-37, wherein the entire nucleotide sequence of each of the at least two different vectors is between about 500 nucleotides to about 10,000 nucleotides in length.

39. The composition of claim 38, wherein the entire nucleotide sequence of each of the at least two different vectors is between 500 nucleotides to 5,000 nucleotides in length.

40. The composition of any one of claims 1-39, wherein the number of different vectors in the composition is two.

41. The composition of claim 40, wherein a first of the two different vectors comprises a coding sequence that encodes an N-terminal portion of the CLRN1 protein.

42. The composition of claim 41, wherein the N-terminal portion of the CLRN1 protein is between 30 amino acids to 202 amino acids in length.

43. The composition of claim 41, wherein the N-terminal portion of the CLRN1 protein is between 60 amino acids to 170 amino acids in length.

44. The composition of claim 41, wherein the first vector further comprises a 5'UTR sequence.

45. The composition of claim 44, wherein the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12.

46. The composition of claim 45, wherein the 5' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12.

47. The composition of claim 46, wherein the 5' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

48. The composition of claim 47, wherein the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12.

49. The composition of claim 44, wherein the 5'UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 12.

50. The composition of any one of claims 41-49, wherein the first vector further comprises one or both of a promoter and a Kozak sequence.

51. The composition of claim 50, wherein the first vector comprises a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter.

52. The composition of any one of claims 40-51, wherein the second of the two different vectors comprises a coding sequence that encodes a C-terminal portion of the CLRN1 protein.

53. The composition of claim 52, wherein the C-terminal portion of the CLRN1 protein is between 30 amino acids to 202 amino acids in length.

54. The composition of claim 53, wherein the C-terminal portion of the CLRN1 protein is between 60 amino acids to 170 amino acids in length.

55. The composition of any one of claims 52-54, wherein the second vector further comprises a polyadenylation signal sequence.

56. The composition of any one of claims 52-55, wherein the second vector further comprises a 3'UTR sequence.

57. The composition of claim 56, wherein the 3' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15.

58. The composition of claim 57, wherein the 3' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15.

59. The composition of claim 58, wherein the 3' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15.

60. The composition of claim 59, wherein the 3' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15.

61. The composition of claim 56, wherein the 3' UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 15.

62. A composition comprising a single nucleic acid vector, wherein the vector comprises one or both of (i) a first coding sequence encoding a first isoform of CLRN1 protein, and (ii) a second coding sequence encoding a second isoform of CLRN1 protein, wherein one or both of the first and second coding sequences comprises a nucleotide sequence spanning two consecutive exons of a CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive introns.

63. The composition of claim 62, wherein the single nucleic acid vector contains the first coding sequence and not the second coding sequence.

64. The composition of claim 62, wherein the single nucleic acid vector contains the second coding sequence and not the first coding sequence.

65. The composition of claim 62, wherein the single nucleic acid vector contains both the first coding sequence and the second coding sequence.

66. The composition of any one of claims 62, 63, and 65, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

67. The composition of claim 66, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

68. The composition of claim 67, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

69. The composition of any one of claims 62, 64, and 65, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

70. The composition of claim 69, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

71. The composition of claim 70, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

72. The composition of any one of claims 62-71, wherein the single nucleic acid vector further comprises a 5' untranslated region (UTR), a 3' UTR, or both.

73. The composition of claim 72, wherein the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12.

74. The composition of claim 73, wherein the 5' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12.

75. The composition of claim 74, wherein the 5' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

76. The composition of claim 75, wherein the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12.

77. The composition of claim 72, wherein the 5'UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 12.

78. The composition of claim 72, wherein the 3' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15.

79. The composition of claim 78, wherein the 3' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15.

80. The composition of claim 79, wherein the 3' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15.

81. The composition of claim 80, wherein the 3' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15.

82. The composition of claim 72, wherein the 3' UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 15.

83. The composition of claim 62, wherein the single nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector.

84. The composition of claim 62, wherein the single nucleic acid vector is a human artificial chromosome (HAC), yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC).

85. The composition of claim 62, wherein the single nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector.

86. The composition of claim 85, wherein the single nucleic acid vector is an AAV vector.

87. The composition of claim 62, wherein the single nucleic acid vector further comprises one or both of a promoter and a Kozak sequence.

88. The composition of claim 87, wherein the first vector comprises a promoter that is an inducible promoter, a constitutive promoter, or a tissue-specific promoter.

89. The composition of claim 87, wherein the single nucleic acid vector further comprises a polyadenylation signal sequence.

90. The composition of any one of claims 1-89, further comprising a pharmaceutically acceptable excipient.

91. A kit comprising a composition of any one of claims 1-90.

92. A kit of claim 91, further comprising a pre-loaded syringe comprising the composition.

93. A method comprising:
introducing into a cochlea of a mammal a therapeutically effective amount of the composition of any one of claims 1-90.

94. The method of claim 93, wherein the mammal is a human.

95. The method of claim 93 or 94, wherein the mammal has been previously identified as having a defective CLRN1 gene.

96. A method of increasing expression of a full-length CLRN1 protein in a mammalian cell, the method comprising introducing the composition of any one of claims 1-90 into the mammalian cell.

97. The method of claim 96, wherein the mammalian cell is a cochlear inner hair cell or a cochlear outer hair cell.

98. The method of claim 96, wherein the mammalian cell is a retinal cell.

99. The method of any one of claims 96-98, wherein the mammalian cell is a human cell.

100. The method of any one of claims 96-99, wherein the mammalian cell has previously been determined to have a defective CLRN1 gene.

101. A method of increasing expression of a full-length CLRN1 protein in an inner hair cell, an outer hair cell, or both, 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 1-90.

102. A method of increasing expression of a full-length CLRN1 protein in an eye of a mammal, the method comprising:

intraocularly administering to the eye of the mammal a therapeutically effective amount of the composition of any one of claims 1-90.

103. The method of claim 101 or 102, wherein the mammal has been previously identified as having a defective CLRN1 gene.

104. The method of any one of claims 101-104, wherein the mammal is a human.

105. A method of treating hearing loss in a subject identified as having a defective CLRN1 gene, the method comprising:

administering a therapeutically effective amount of a composition of any one of claims 1-90 into the cochlea of the subject.

106. A method of treating vision loss in a subject identified as having a defective CLRN1 gene, the method comprising:

administering a therapeutically effective amount of a composition of any one of claims 1-90 into the eye of the subject.

107. The method of claim 105 or 106, wherein the subject has Usher syndrome type III.

108. The method of any one claims 105-107, wherein the subject is a human.

109. The method of any one of claims 105-108, further comprising, prior to the administering step, determining that the subject has a defective CLRN1 gene.

110. 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 CLRN1 protein positioned 3' of the promoter, and a splice donor 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 splice acceptor sequence, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal 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 sequences of the two encoded portions do not overlap with each other; wherein no single vector of the two different vectors encodes a full-length CLRN1 protein; and

when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

111. The composition of claim 110, wherein the first coding sequence encodes an N-terminal portion of a first isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the first isoform of CLRN1 protein.

112. The composition of claim 111, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

113. The composition of claim 112, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

114. The composition of claim 113, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

115. The composition of claim 111, wherein the first nucleic acid vector of the second nucleic acid vector further comprises a sequence that encodes a second isoform of the CLRN1 protein.

116. The composition of claim 115, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

117. The composition of claim 116, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

118. The composition of claim 117, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

119. The composition of claim 110, wherein the first coding sequence encodes an N-terminal portion of a second isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the second isoform of CLRN1 protein.

120. The composition of claim 119, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

121. The composition of claim 120, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

122. The composition of claim 121, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

123. The composition of claim 119, wherein the first nucleic acid vector or the second nucleic acid vector further comprises a sequence that encodes a first isoform of the CLRN1 protein.

124. The composition of claim 123, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

125. The composition of claim 124, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

126. The composition of claim 125, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

127. The composition of any one of claims 110-126, wherein one or both of the first nucleic acid vector and the second nucleic acid vector comprises a 5' untranslated region (UTR), a 3' UTR, or both.

128. The composition of claim 127, wherein the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12.

129. The composition of claim 128, wherein the 5' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12.

130. The composition of claim 129, wherein the 5' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

131. The composition of claim 130, wherein the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12.

132. The composition of claim 127, wherein the 5'UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 12.

133. The composition of claim 127, wherein the 3' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15.

134. The composition of claim 133, wherein the 3' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15.

135. The composition of claim 134, wherein the 3' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15.

136. The composition of claim 135, wherein the 3' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15.

137. The composition of claim 127, wherein the 3' UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 15.

138. The composition of any one of claims 110-137, wherein each of the first nucleic acid vector and the second nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector.

139. The composition of any one of claims 110-137, wherein each of the first nucleic acid vector and the second nucleic acid vector is a human artificial chromosome (HAC), yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC).

140. The composition of any one of claims 110-137, wherein each of the first nucleic acid vector and the second nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector.

141. The composition of claim 140, wherein each of the first nucleic acid vector and the second nucleic acid vector is an AAV vector.

142. The composition of any one of claims 110-141, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons.

143. 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 CLRN1 protein positioned 3' of the promoter, a splice donor sequence positioned at the 3' end of the first coding sequence, and a first detectable marker gene positioned 3' of the splice donor sequence; and

a second nucleic acid vector of the two different nucleic acid vectors comprises a second detectable marker gene, a splice acceptor sequence positioned 3' of the second detectable marker gene, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal 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 sequences of the encoded portions do not overlap with each other; wherein no single vector of the two different vectors encodes a full-length CLRN1 protein; and

when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

144. The composition of claim 143, wherein the first coding sequence encodes an N-terminal portion of a first isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the first isoform of CLRN1 protein.

145. The composition of claim 144, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

146. The composition of claim 145, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

147. The composition of claim 146, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

148. The composition of claim 144, wherein the first nucleic acid vector of the second nucleic acid vector further comprises a sequence that encodes a second isoform of the CLRN1 protein.

149. The composition of claim 148, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

150. The composition of claim 149, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

151. The composition of claim 150, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

152. The composition of claim 143, wherein the first coding sequence encodes an N-terminal portion of a second isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the second isoform of CLRN1 protein.

153. The composition of claim 152, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

154. The composition of claim 153, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

155. The composition of claim 154, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

156. The composition of claim 152, wherein the first nucleic acid vector or the second nucleic acid vector further comprises a sequence that encodes a first isoform of the CLRN1 protein.

157. The composition of claim 156, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

158. The composition of claim 157, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

159. The composition of claim 158, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

160. The composition of any one of claims 143-159, wherein one or both of the first nucleic acid vector and the second nucleic acid vector comprises a 5' untranslated region (UTR), a 3' UTR, or both.

161. The composition of claim 160, wherein the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12.

162. The composition of claim 161, wherein the 5' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12.

163. The composition of claim 162, wherein the 5' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

164. The composition of claim 163, wherein the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12.

165. The composition of claim 160, wherein the 5'UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 12.

166. The composition of claim 160, wherein the 3' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15.

167. The composition of claim 166, wherein the 3' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15.

168. The composition of claim 167, wherein the 3' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15.

169. The composition of claim 168, wherein the 3' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15.

170. The composition of claim 160, wherein the 3' UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 15.

171. The composition of any one of claims 143-170, wherein each of the first nucleic acid vector and the second nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector.

172. The composition of any one of claims 143-170, wherein each of the first nucleic acid vector and the second nucleic acid vector is a human artificial chromosome (HAC), yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC).

173. The composition of any one of claims 143-170, wherein each of the first nucleic acid vector and the second nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector.

174. The composition of claim 173, wherein each of the first nucleic acid vector and the second nucleic acid vector is an AAV vector.

175. The composition of any one of claims 143-174, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons.

176. The composition of any one of claims 143-175, wherein the first or second detectable marker gene is alkaline phosphatase.

177. 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 CLRN1 protein positioned 3' to the promoter, a splice donor sequence positioned at the 3' end of the first coding sequence, and a F1 phage recombinogenic region positioned 3' to the splice donor sequence; and

a second nucleic acid vector of the two different nucleic acid vectors comprises a F1 phage recombinogenic region, a splice acceptor sequence positioned 3' of the F1 phage recombinogenic region, a second coding sequence that encodes a C-terminal portion of a CLRN1 protein positioned at the 3' end of the splice acceptor sequence, and a polyadenylation signal sequence positioned at the 3' end of the second coding sequence;

wherein each of the two encoded portions is at least 30 amino acid residues in length, wherein the amino acid sequences of the two encoded portions do not overlap with each other; wherein no single vector of the two different vectors encodes a full-length CLRN1 protein; and

when introduced into a mammalian cell, splicing occurs between the splice donor sequence and the splice acceptor sequence, thereby forming a recombined nucleic acid that encodes a full-length CLRN1 protein.

178. The composition of claim 177, wherein the first coding sequence encodes an N-terminal portion of a first isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the first isoform of CLRN1 protein.

179. The composition of claim 178, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

180. The composition of claim 179, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

181. The composition of claim 180, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

182. The composition of claim 178, wherein the first nucleic acid vector of the second nucleic acid vector further comprises a sequence that encodes a second isoform of the CLRN1 protein.

183. The composition of claim 182, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

184. The composition of claim 183, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

185. The composition of claim 184, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

186. The composition of claim 177, wherein the first coding sequence encodes an N-terminal portion of a second isoform of CLRN1 protein, and the second coding sequence encodes a C-terminal portion of the second isoform of CLRN1 protein.

187. The composition of claim 186, wherein the second isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 5.

188. The composition of claim 187, wherein the second isoform of the CLRN1 protein comprises SEQ ID NO: 5.

189. The composition of claim 188, wherein the second isoform of the CLRN1 protein consists of SEQ ID NO: 5.

190. The composition of claim 186, wherein the first nucleic acid vector or the second nucleic acid vector further comprises a sequence that encodes a first isoform of the CLRN1 protein.

191. The composition of claim 190, wherein the first isoform of the CLRN1 protein comprises a sequence that is at least 95% identical to SEQ ID NO: 3.

192. The composition of claim 191, wherein the first isoform of the CLRN1 protein comprises SEQ ID NO: 3.

193. The composition of claim 192, wherein the first isoform of the CLRN1 protein consists of SEQ ID NO: 3.

194. The composition of any one of claims 177-193, wherein one or both of the first nucleic acid vector and the second nucleic acid vector comprises a 5' untranslated region (UTR), a 3' UTR, or both.

195. The composition of claim 194, wherein the 5' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 12.

196. The composition of claim 195, wherein the 5' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 12.

197. The composition of claim 196, wherein the 5' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 12.

198. The composition of claim 197, wherein the 5' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 12.

199. The composition of claim 194, wherein the 5'UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 12.

200. The composition of claim 194, wherein the 3' UTR comprises at least 10 contiguous nucleotides from anywhere within SEQ ID NO: 15.

201. The composition of claim 200, wherein the 3' UTR comprises at least 20 contiguous nucleotides from anywhere within SEQ ID NO: 15.

202. The composition of claim 201, wherein the 3' UTR comprises at least 50 contiguous nucleotides from anywhere within SEQ ID NO: 15.

203. The composition of claim 202, wherein the 3' UTR comprises at least 80 contiguous nucleotides from anywhere within SEQ ID NO: 15.

204. The composition of claim 194, wherein the 3' UTR comprises a sequence that is at least 80% identical to SEQ ID NO: 15.

205. The composition of any one of claims 177-204, wherein each of the first nucleic acid vector and the second nucleic acid vector is a plasmid, a transposon, a cosmid, an artificial chromosome, or a viral vector.

206. The composition of any one of claims 177-204, wherein each of the first nucleic acid vector and the second nucleic acid vector is a human artificial chromosome (HAC), yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or a P1-derived artificial chromosome (PAC).

207. The composition of any one of claims 177-204, wherein each of the first nucleic acid vector and the second nucleic acid vector is a viral vector selected from an adeno-associated virus (AAV) vector, an adenovirus vector, a lentivirus vector, or a retrovirus vector.

208. The composition of claim 207, wherein each of the first nucleic acid vector and the second nucleic acid vector is an AAV vector.

209. The composition of any one of claims 177-204, wherein at least one of the coding sequences comprises a nucleotide sequence spanning two consecutive exons of CLRN1 genomic DNA, and lacking an intronic sequence between the two consecutive exons.

210. A kit comprising a composition of any one of claims 110-209.

211. A kit of claim 210, further comprising a pre-loaded syringe comprising the composition.

212. A method comprising:
introducing into a cochlea of a mammal a therapeutically effective amount of the composition of any one of claims 110-209.

213. The method of claim 212, wherein the mammal is a human.

214. The method of claim 212 or 213, wherein the mammal has been previously identified as having a defective CLRN1 gene.

215. A method of increasing expression of a full-length CLRN1 protein in a mammalian cell, the method comprising introducing the composition of any one of claims 110-209 into the mammalian cell.

216. The method of claim 215, wherein the mammalian cell is a cochlear inner hair cell or a cochlear outer hair cell.

217. The method of claim 215, wherein the mammalian cell is a retinal cell.

218. The method of any one of claims 215-217, wherein the mammalian cell is a human cell.

219. The method of any one of claims 215-218, wherein the mammalian cell has previously been determined to have a defective CLRN1 gene.

220. A method of increasing expression of a full-length CLRN1 protein in an inner hair cell, an outer hair cell, or both, 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 110-209.

221. A method of increasing expression of a full-length CLRN1 protein in an eye of a mammal, the method comprising:

intraocularly administering to the eye of the mammal a therapeutically effective amount of the composition of any one of claims 110-209.

222. The method of claim 220 or 221, wherein the mammal has been previously identified as having a defective CLRN1 gene.

223. The method of any one of claims 220-222, wherein the mammal is a human.

224. A method of treating hearing loss in a subject identified as having a defective CLRN1 gene, the method comprising:

administering a therapeutically effective amount of a composition of any one of claims 110-209 into the cochlea of the subject.

225. A method of treating vision loss in a subject identified as having a defective CLRN1 gene, the method comprising:

administering a therapeutically effective amount of a composition of any one of claims 110-209 into the eye of the subject.

226. The method of claim 224 or 225, wherein the subject has Usher syndrome type III.

227. The method of any one claims 224-226, wherein the subject is a human.

228. The method of any one of claims 224-227, further comprising, prior to the administering step, determining that the subject has a defective CLRN1 gene.

1/11

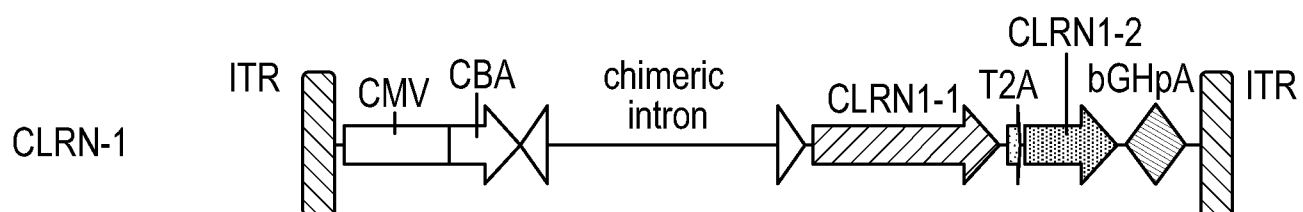


FIG. 1

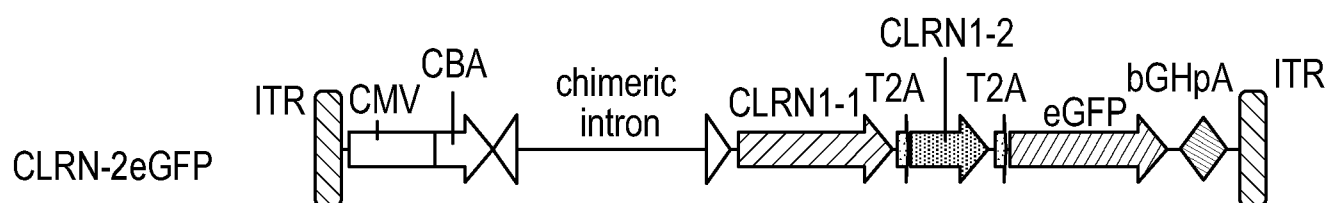


FIG. 2

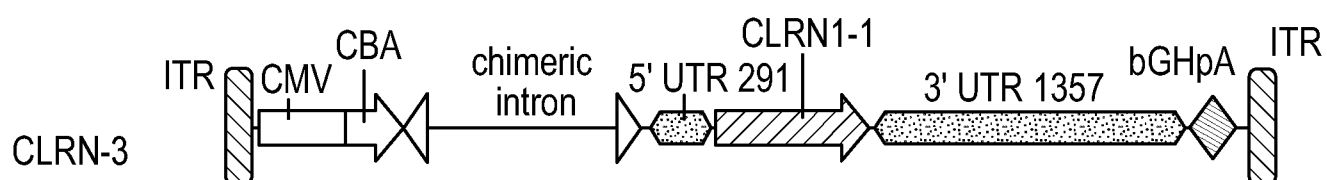


FIG. 3

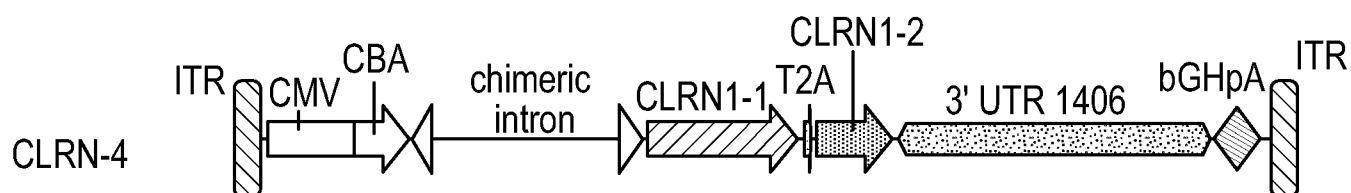


FIG. 4

2/11

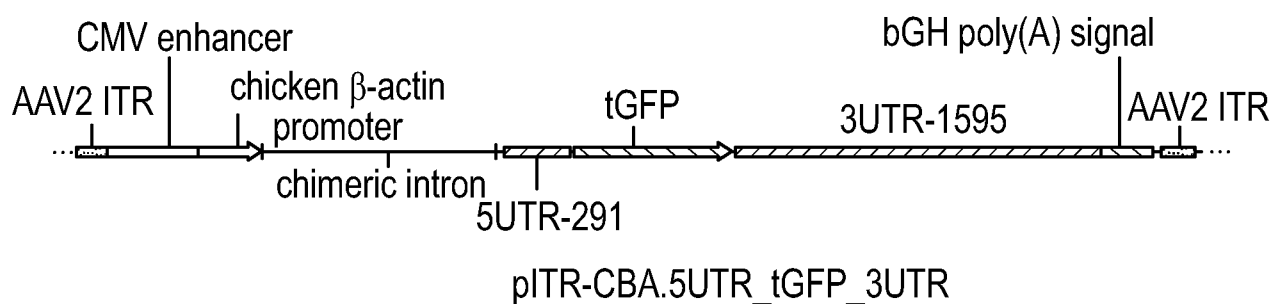


FIG. 5

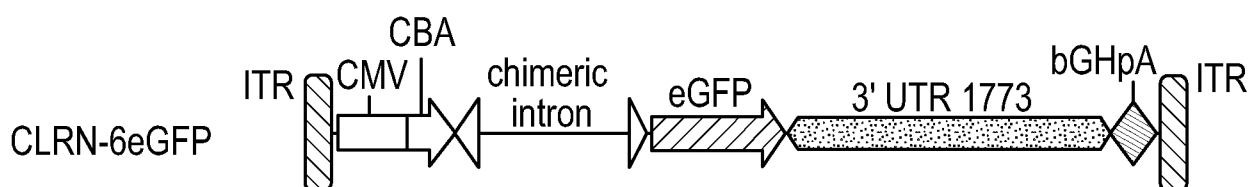


FIG. 6

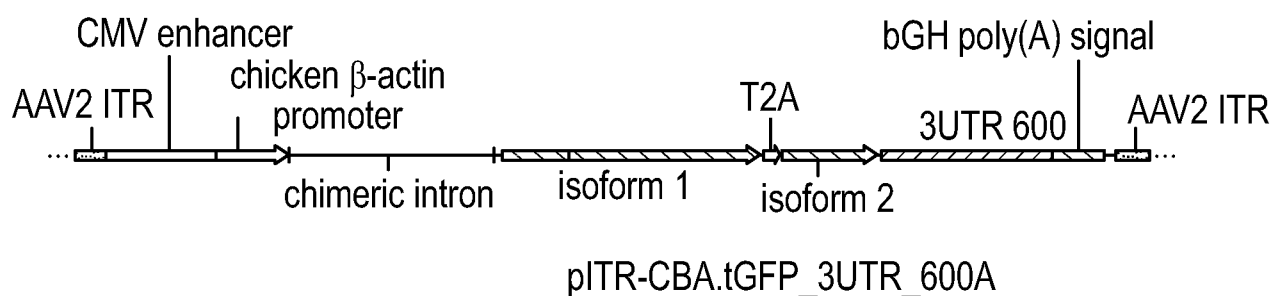


FIG. 7

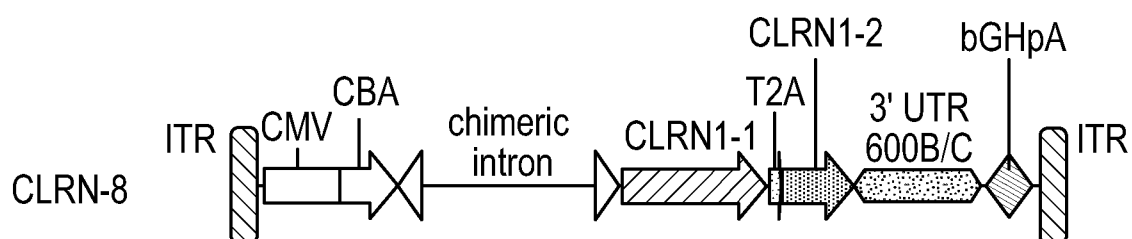


FIG. 8

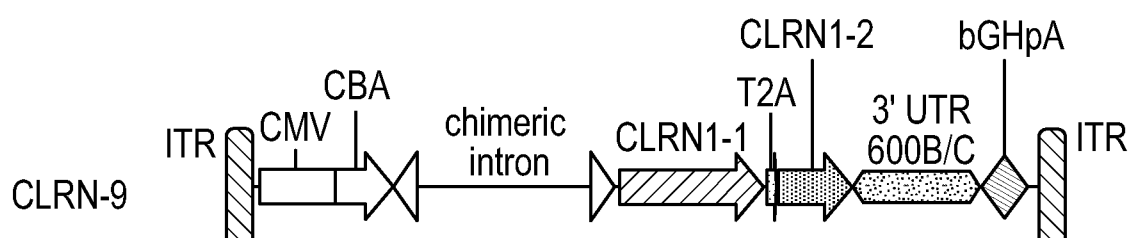


FIG. 9

3/11

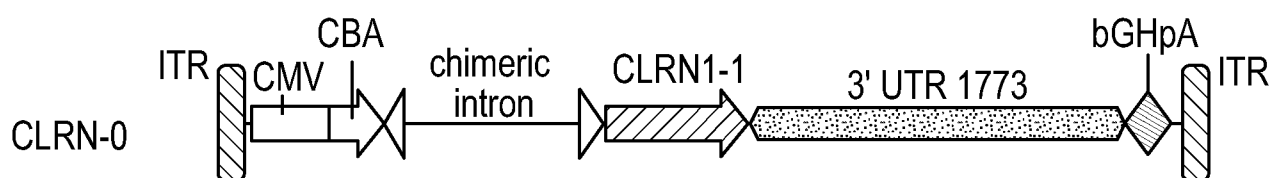


FIG. 10

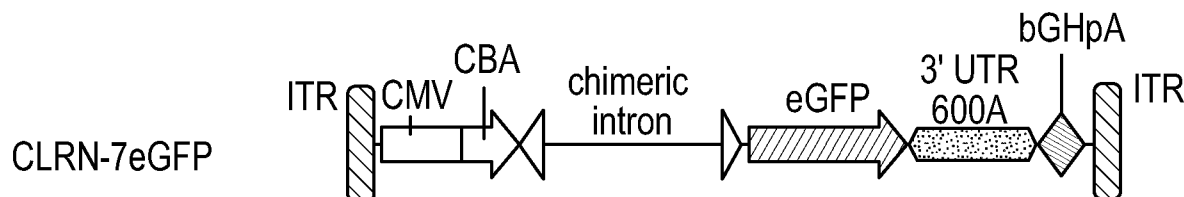


FIG. 11

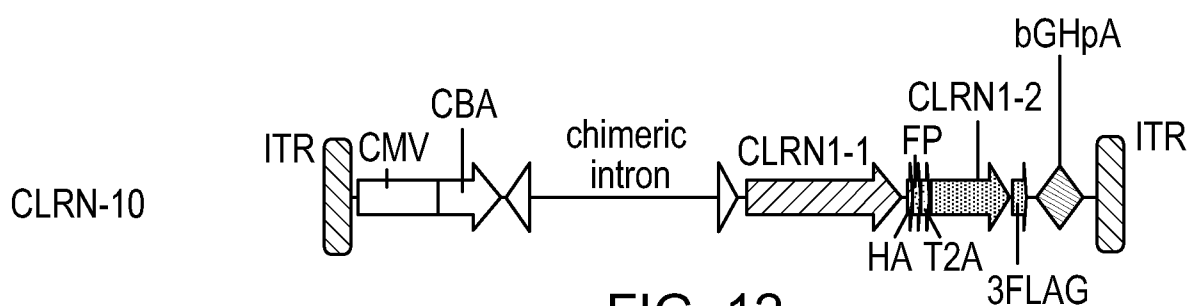


FIG. 12

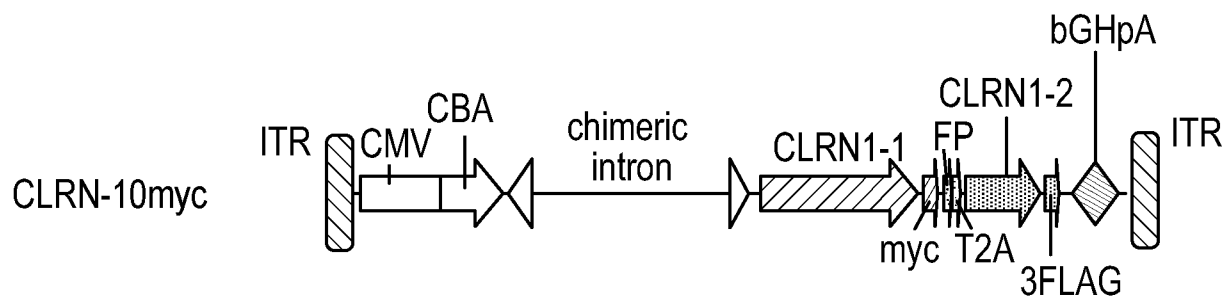


FIG. 13

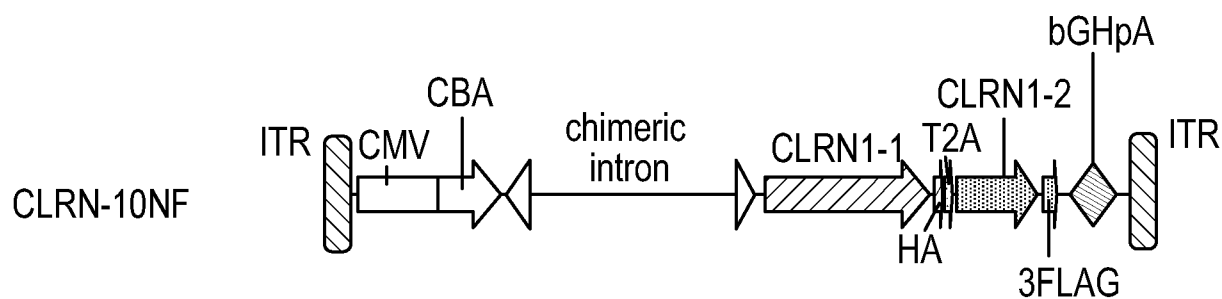


FIG. 14

4/11

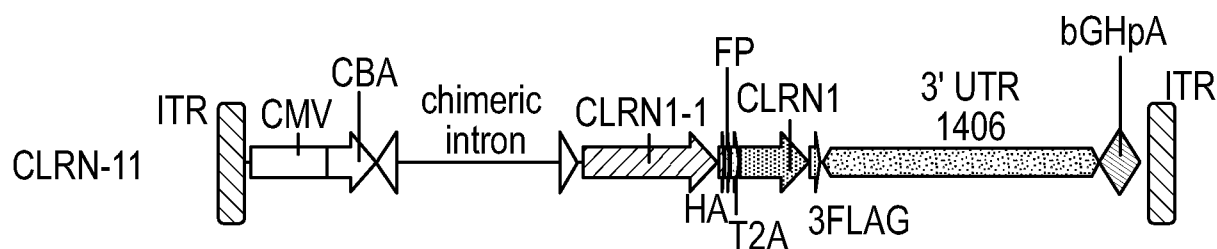


FIG. 15

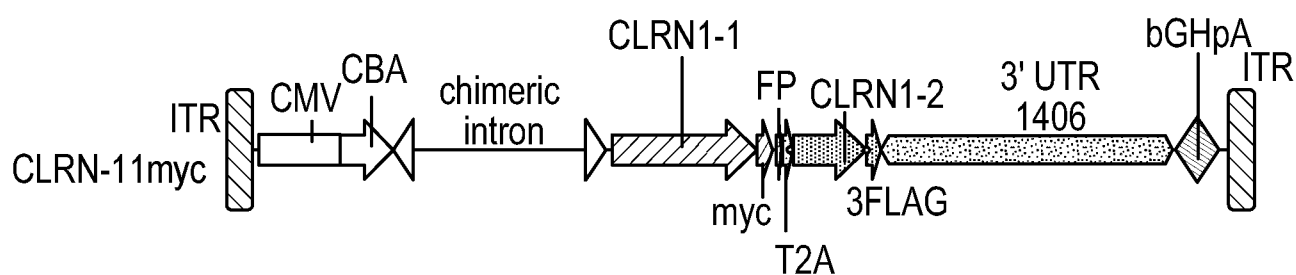


FIG. 16

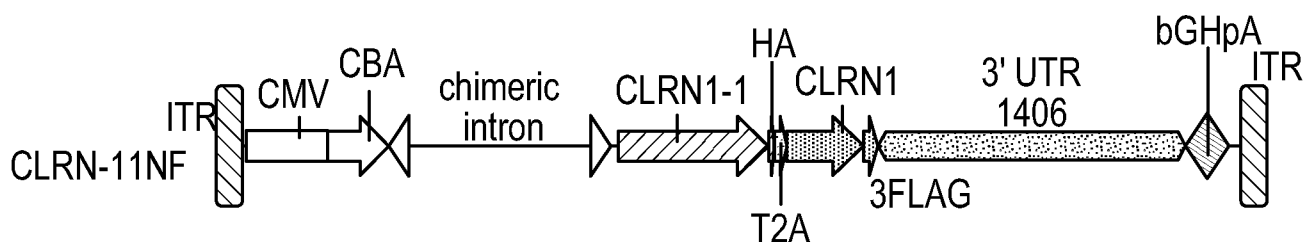


FIG. 17

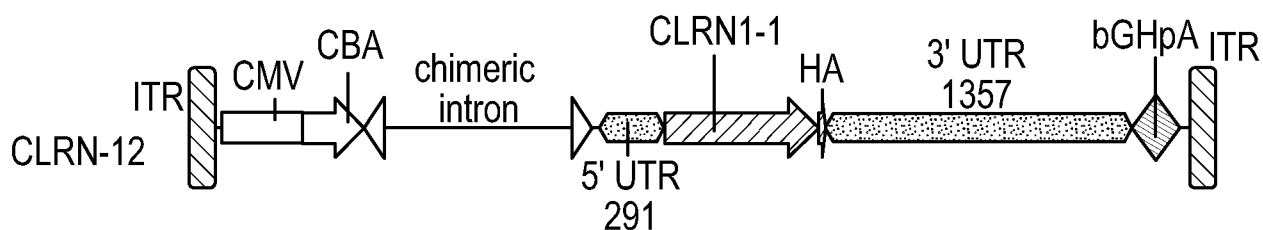


FIG- 18

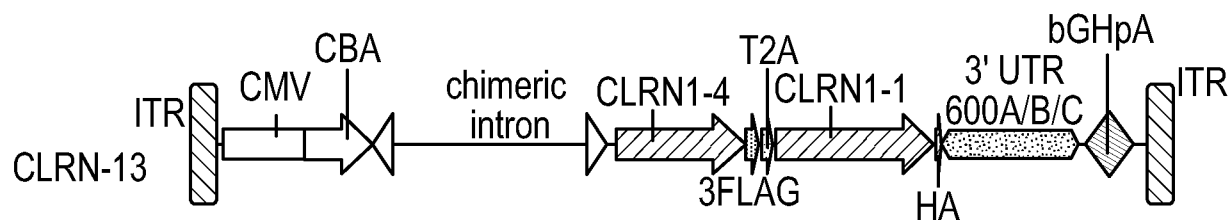


FIG. 19

5/11

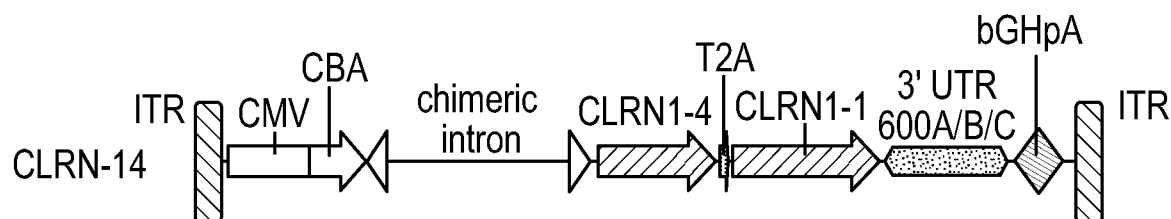


FIG. 20

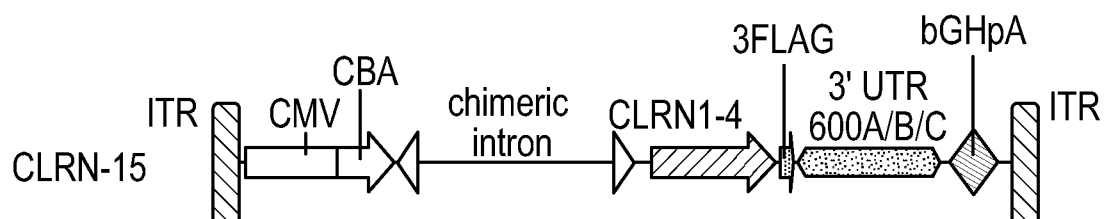


FIG. 21

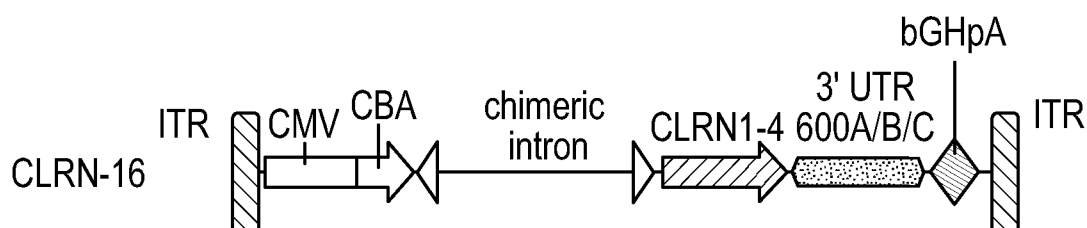


FIG. 22

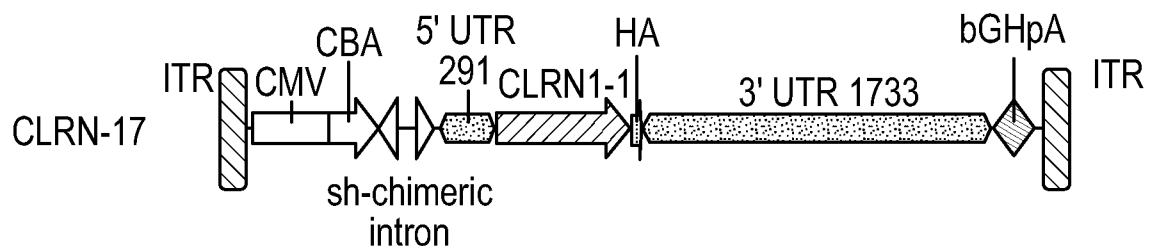


FIG. 23

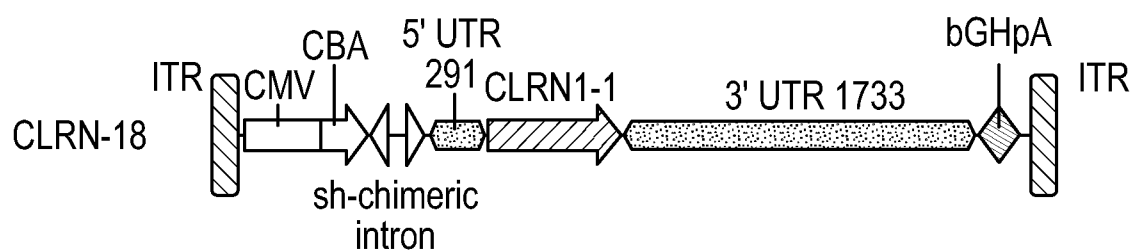


FIG. 24

6/11

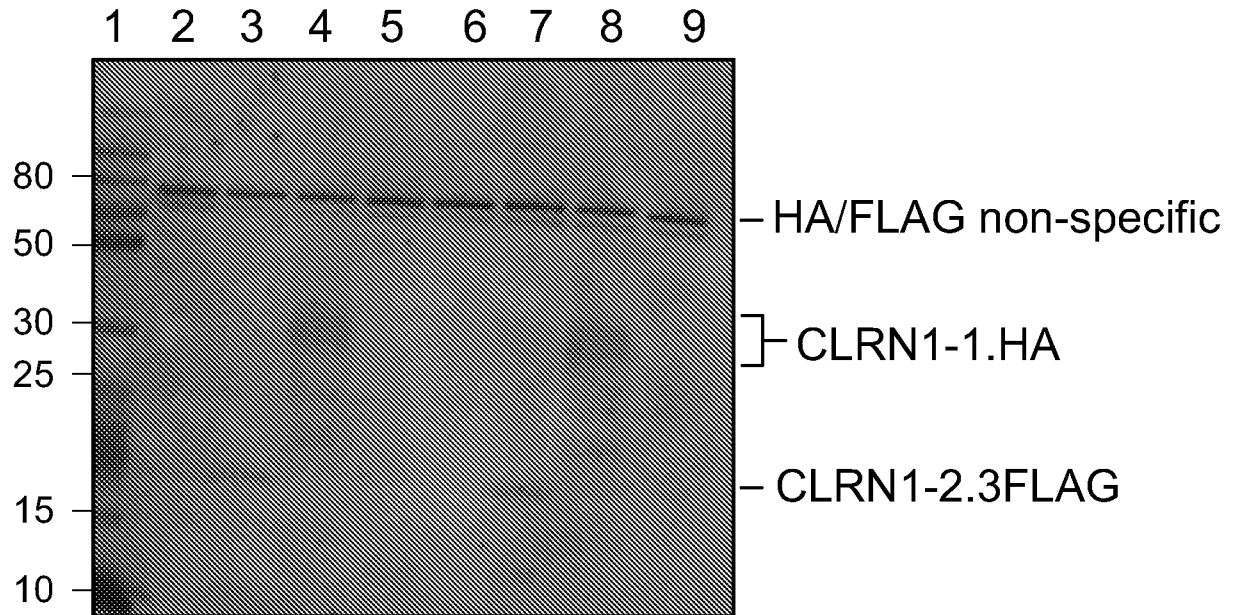


FIG. 25

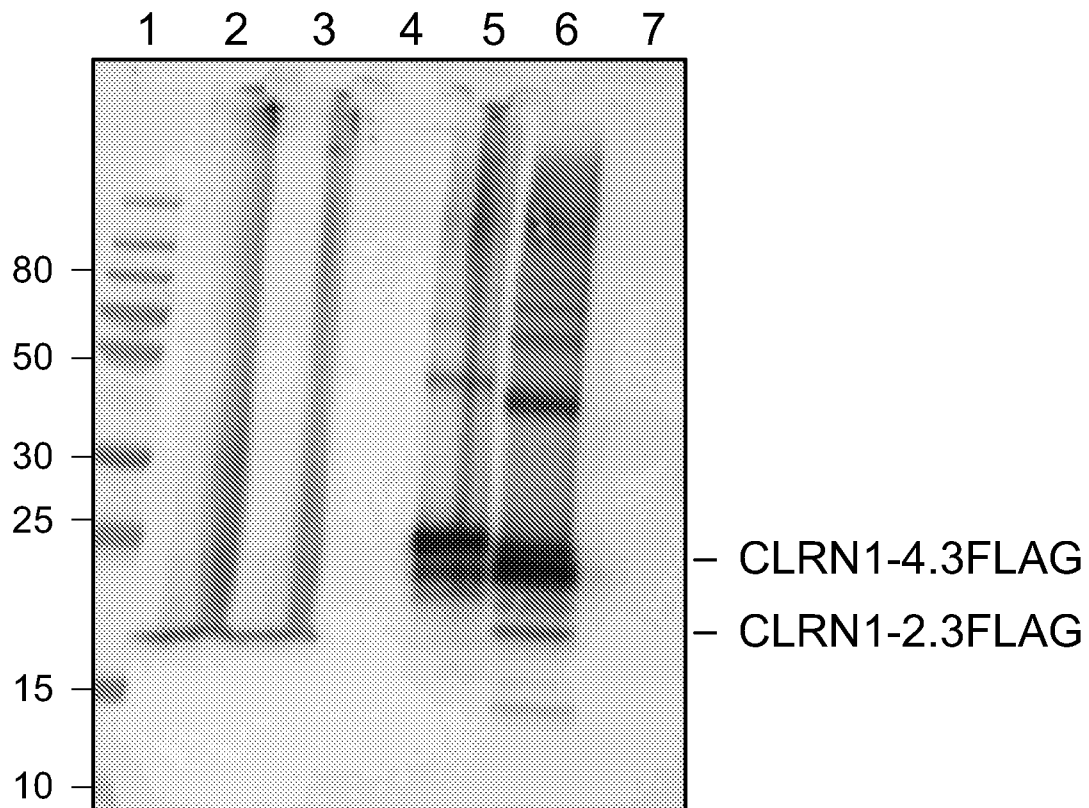


FIG. 26

7/11

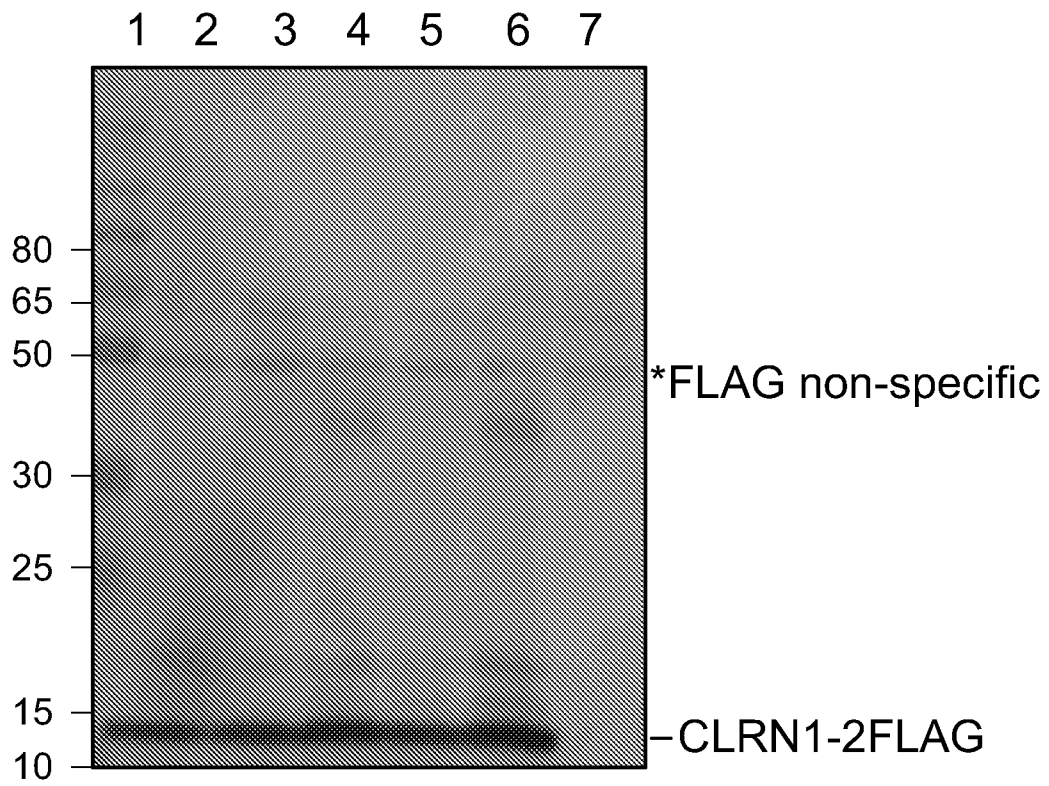


FIG. 27

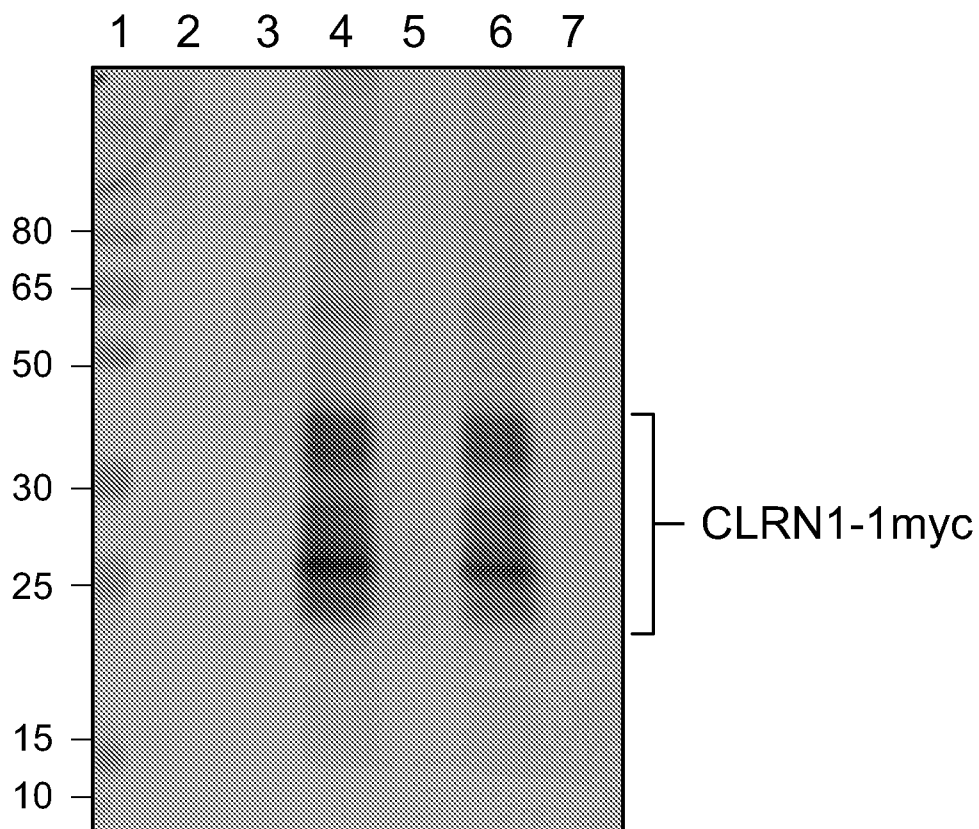


FIG. 28

8/11

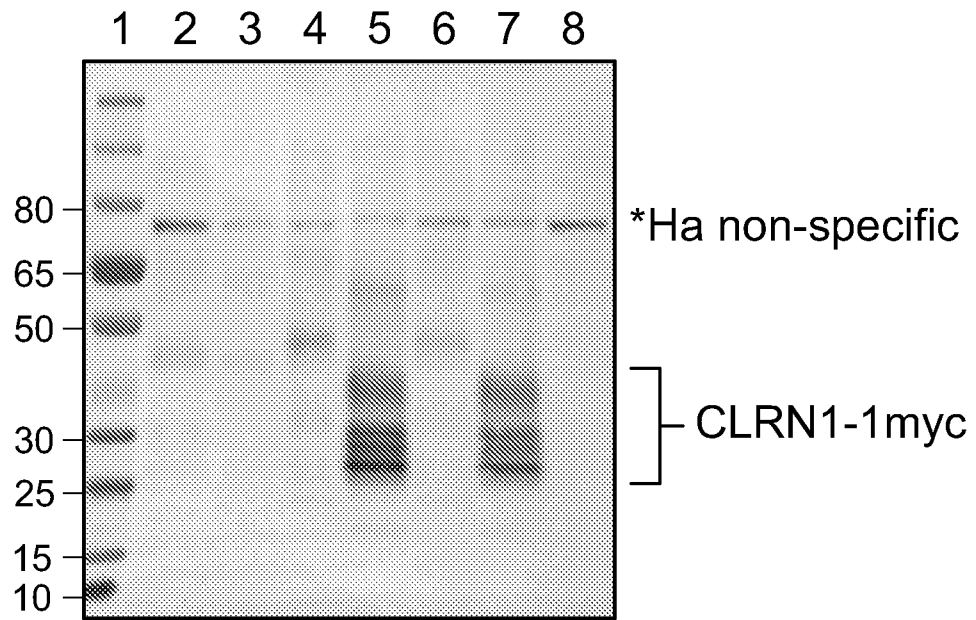


FIG. 29

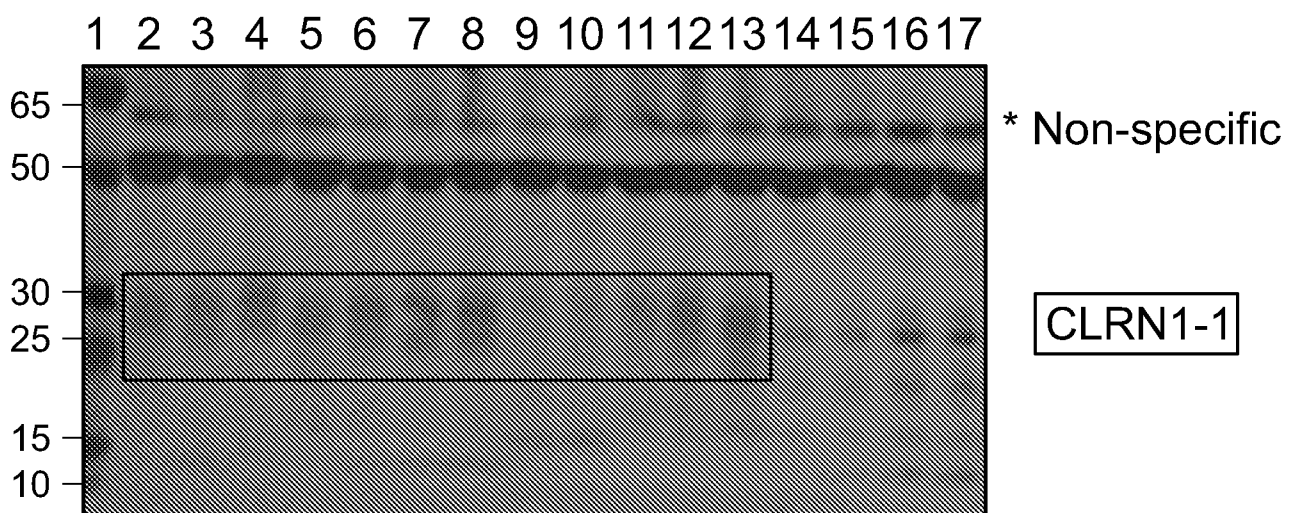


FIG. 30

9/11

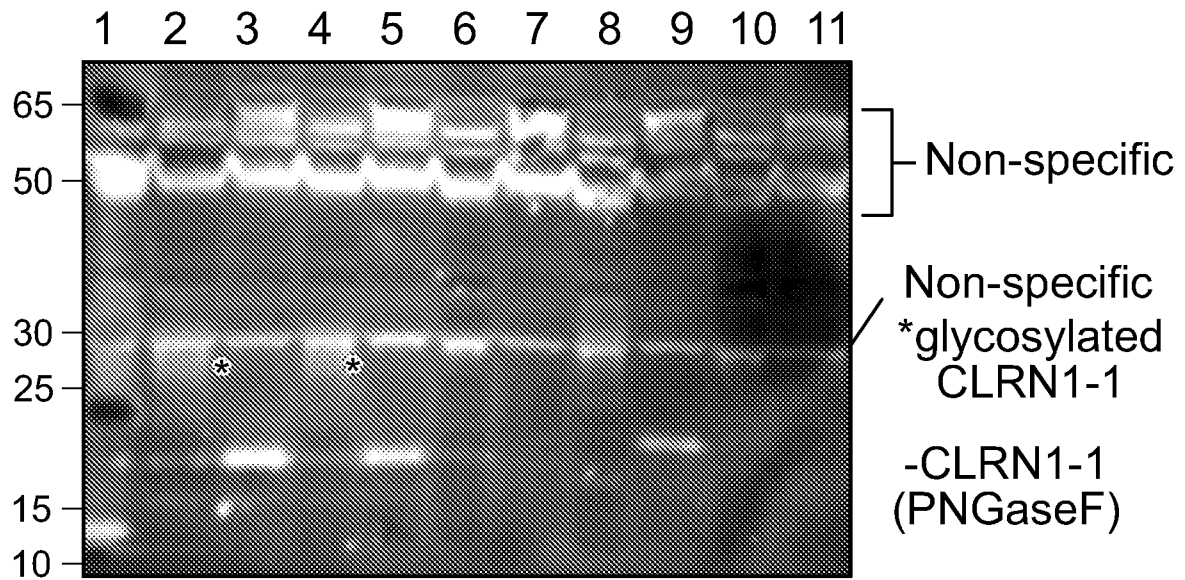


FIG. 31

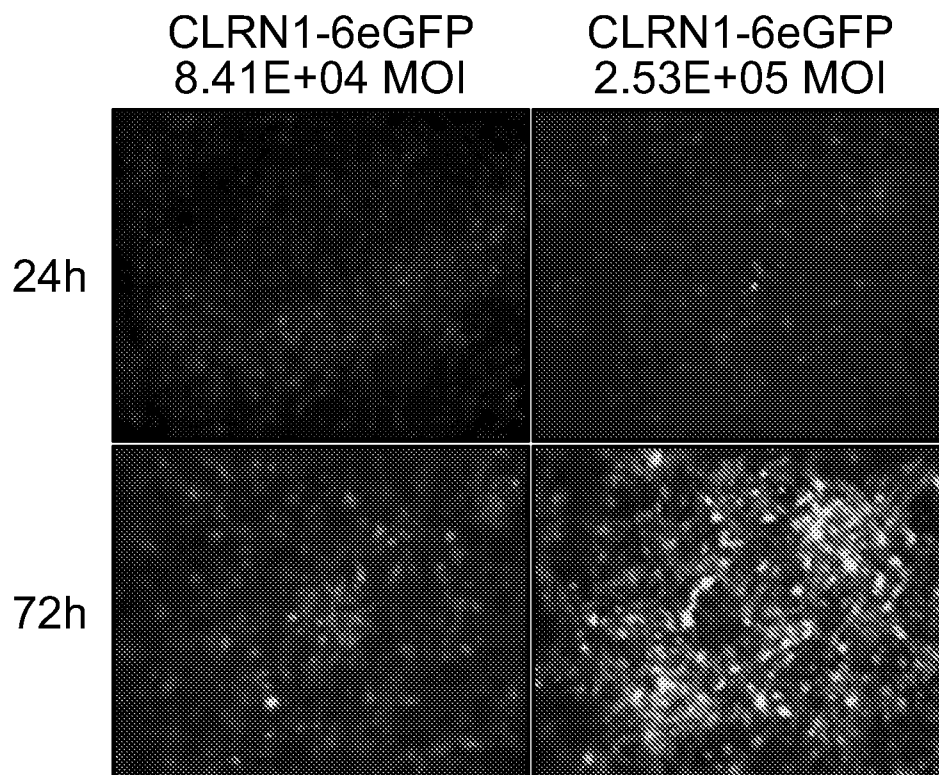


FIG. 32

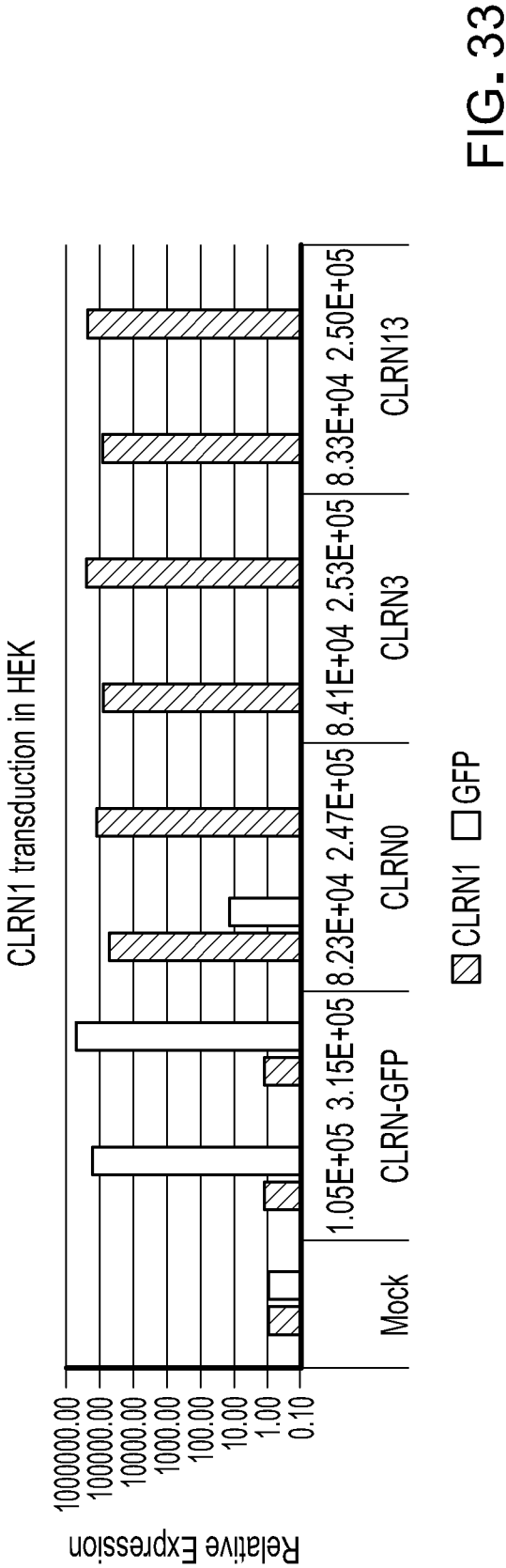


FIG. 33

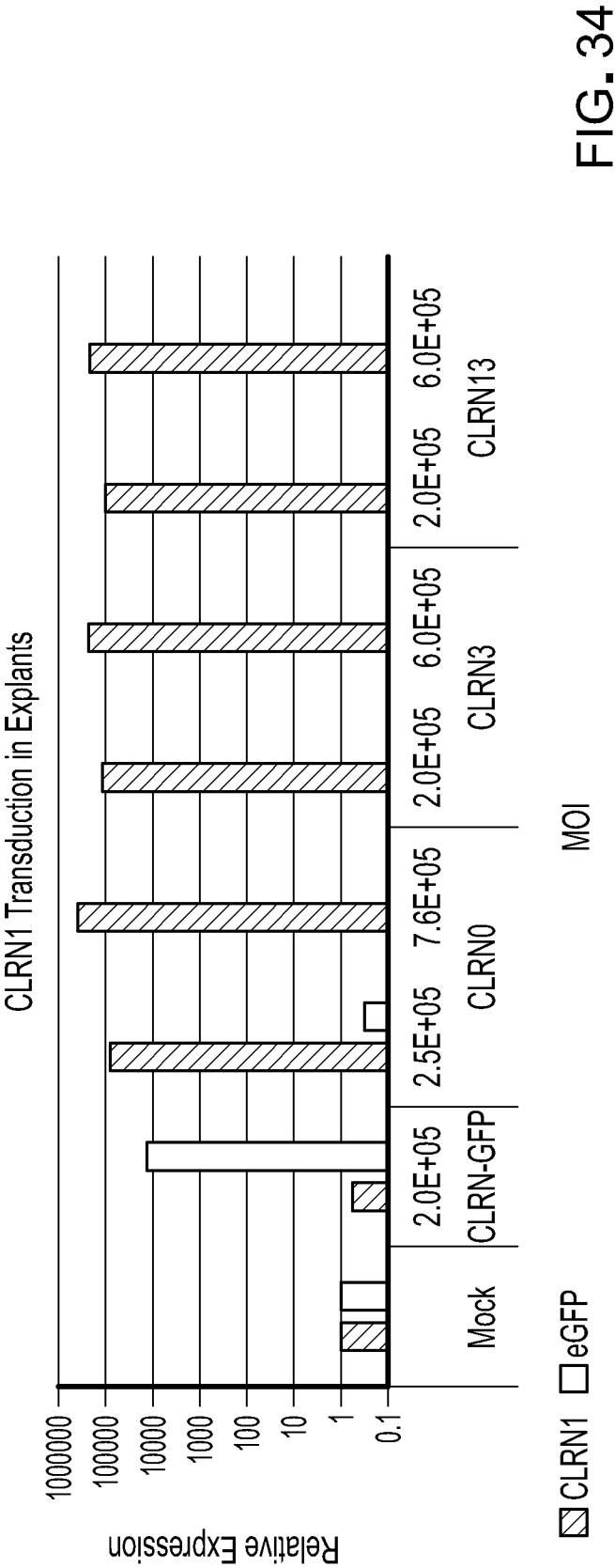


FIG. 34

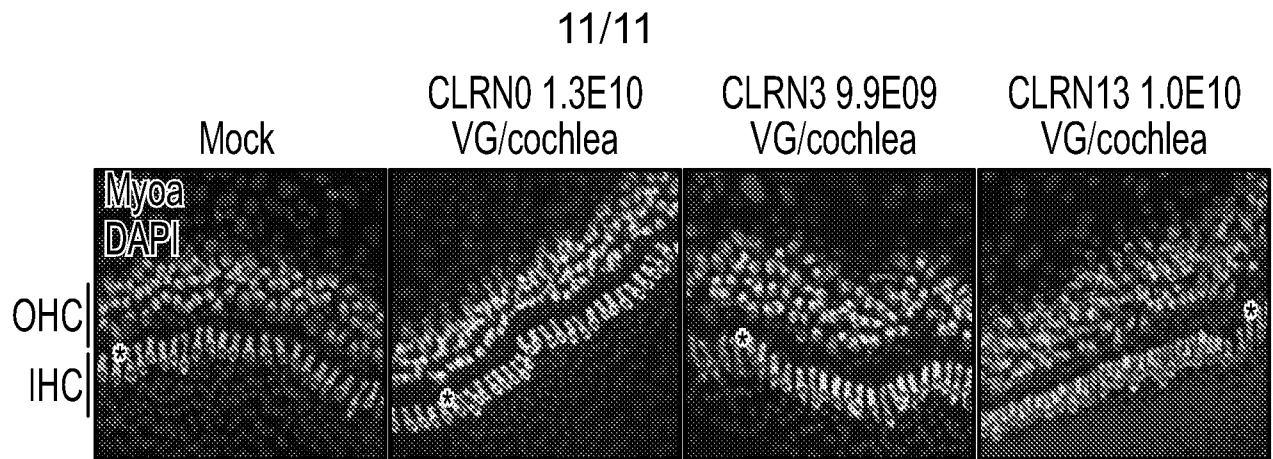


FIG. 35

Cochlear explants infected with 1E09 VG/cochlea AAVAnc80.CAG.eGFP.3'UTR

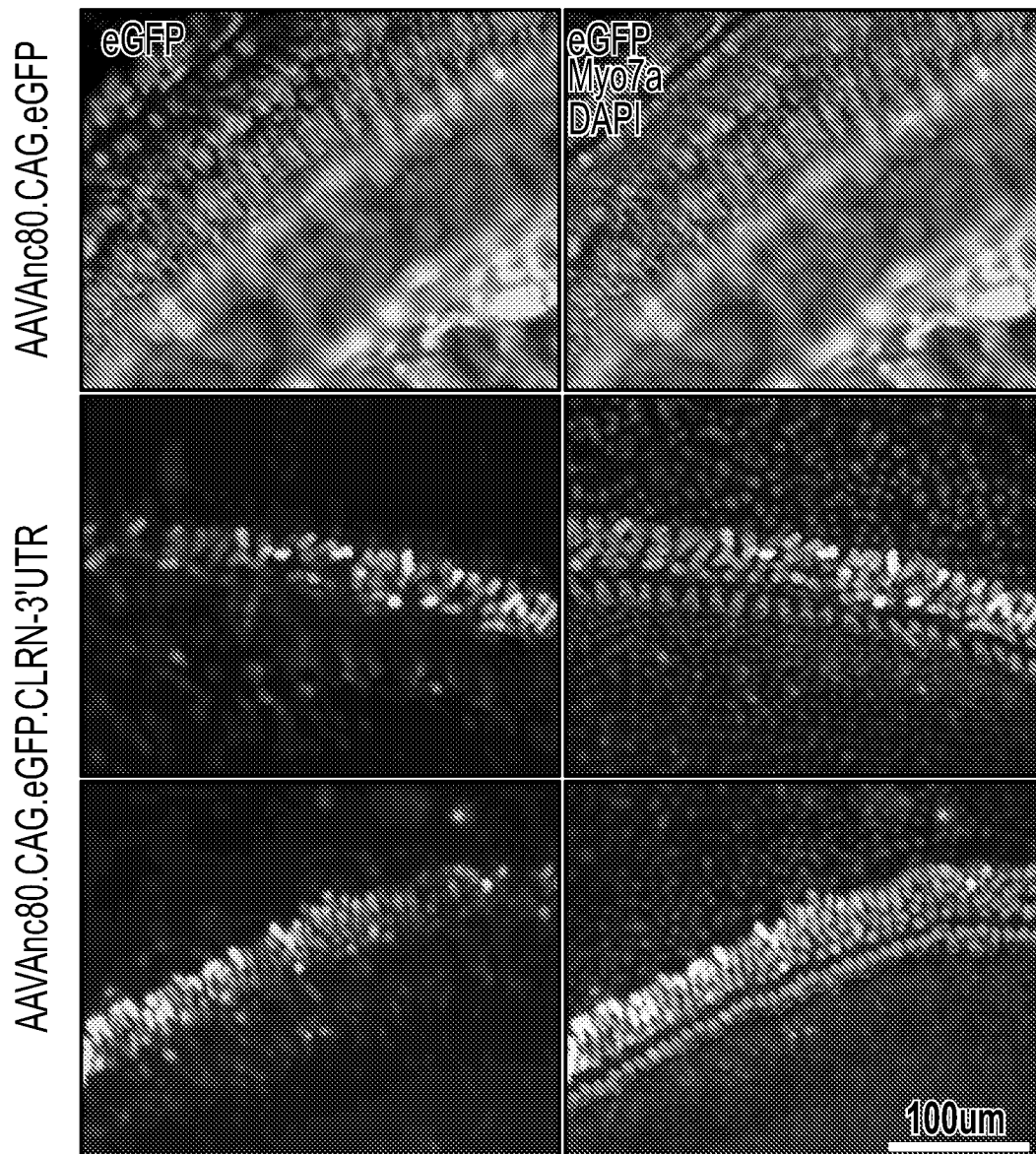


FIG. 36

SUBSTITUTE SHEET (RULE 26)

CLRN1 Transduction in Explants

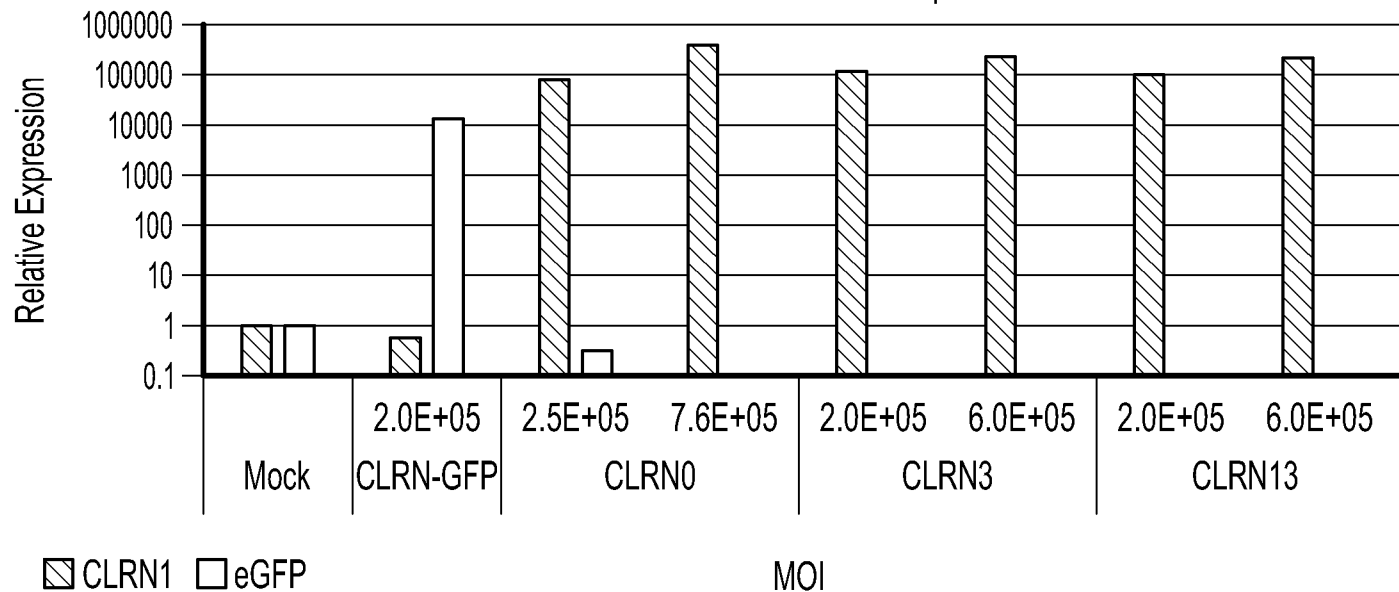


FIG. 34