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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 March 2023 (30.03.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/049931 A1

(51) International Patent Classification:
A61P 27/02 (2006.01) *C12N 15/113* (2010.01)
A61K 9/00 (2006.01) *C12N 15/86* (2006.01)

(21) International Application Number:
PCT/US2022/077100

(22) International Filing Date:
27 September 2022 (27.09.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/248,808 27 September 2021 (27.09.2021) US

(71) Applicants: **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; 412 Low Memorial Library, 535 West 116th Street, New York, New York 10027 (US). **THE BROAD INSTITUTE, INC.** [US/US]; 415 Main Street, Cambridge, Massachusetts 02142 (US). **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [US/US]; 17 Quincy Street, Cambridge, Massachusetts 02138 (US).

(72) Inventors: **QUINN, Peter M.J.**; 412 Low Memorial Library, 535 West 116th Street, New York, New York 10027 (US). **LOPES DA COSTA, Bruna**; 412 Low Memorial Library, 535 West 116th Street, New York, New York 10027 (US). **TSANG, Stephen H.**; 412 Low Memorial Library, 535 West 116th Street, New York, New York 10027 (US). **LIU, David R.**; c/o The Broad Institute, Inc., 415 Main Street, Cambridge, Massachusetts 02142 (US).

(74) Agent: **BARTON, Kelly A.**; Casimir Jones, S.C., 2275 Deming Way, Suite 310, Middleton, Wisconsin 53562 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU,

(54) Title: METHODS AND SYSTEMS FOR MODIFYING THE CRUMBS HOMOLOGUE-1 (CRB1) GENE

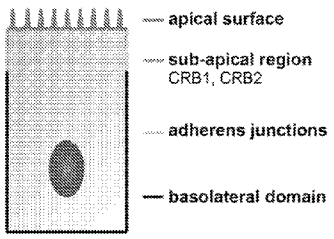

FIG. 1A

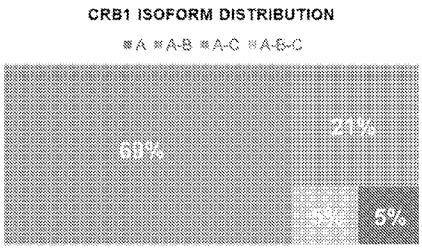

FIG. 1B

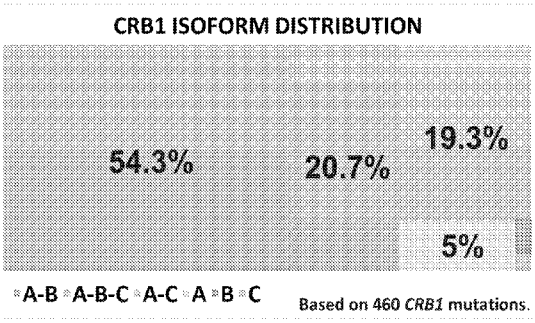

FIG. 1C

(57) Abstract: The present disclosure provides systems, methods, and compositions for modifying the crumbs homologue-1 gene. Particularly the present disclosure provides systems, methods, and compositions for prime editing insertion or correction of mutations in the crumbs homologue-1 gene. Modifying crumbs homologue-1 includes a Cas protein, reverse transcriptase, RNA polynucleotides, and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence.

WO 2023/049931 A1

[Continued on next page]

LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

Published:

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

METHODS AND SYSTEMS FOR MODIFYING THE *CRUMBS HOMOLOGUE-1 (CRB1)* GENE

FIELD

[0001] The present invention relates to systems, methods, and compositions for prime-editing modification of the crumbs homologue-1 gene.

CROSS-REFERENCE TO RELATED APPLICATIONS

[0002] This application claims the benefit of U.S. Provisional Application No. 63/248,808, filed September 27, 2021, the content of which is herein incorporated by reference in its entirety.

SEQUENCE LISTING STATEMENT

[0003] The contents of the electronic sequence listing titled COLUM-40034-601.xml (Size: 66,763 bytes; and Date of Creation: September 27, 2022) is herein incorporated by reference in its entirety.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0004] This invention was made with government support under grant numbers U01 EY030580, U54OD020351, R24EY028758, R24EY027285, 5P30EY019007, R01EY018213, R01EY024698, R01EY026682, and R21AG050437, awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND

[0005] The Crumbs complex (CRB) is crucial for cell polarity and epithelial tissue function, having an essential role during retinogenesis. Disruption of the CRB complex will interrupt the precise orchestration of spatiotemporal process during retinal development, such as cell fate choice, division, migration, and differentiation. This can cause retinal degeneration leading to impairment of retinal function and thus vision.

[0006] Mutations in the Crumbs homologue-1 (CRB1) gene cause progressive and disabling autosomal recessive retinal dystrophies, including Leber congenital amaurosis 8 (LCA8) and retinitis pigmentosa 12 (RP12). Approximately 80,000 patients are affected worldwide, with a prevalence in the United States of 1 in 86,500. CRB1 mutations exhibit high phenotypic variability, with approximately 310 pathogenic variants reported. According to a meta-analysis, CRB1 gene mutations account for 2.7% and 10.1% of autosomal recessive retinitis pigmentosa

(RP) and Leber congenital amaurosis (LCA) cases, respectively, as well as an increasing number of juvenile macular dystrophy cases. LCA is a group of severe, infantile-onset retinal dystrophies that constitute more than 5% of all retinal dystrophies and is the most common cause of inherited blindness in childhood. Despite its prevalence and severity, the pathogenesis of CRB1 LCA remains unclear, and there is no treatment available to date.

[0007] The mouse and human retina contain three CRB1 isoforms, CRB1-A, CRB1-B, CRB1-C. CRB1-A and CRB1-B have predominately cell-type specific expression making the choice of gene augmentation strategy currently unclear. Due to the complexity of CRB1 cell-type specific retinal isoform diversity, there is a need for compositions and methods that can be used to treat a subject that has or is at risk of developing a disease characterized by Crumbs homologue 1 (CRB1) mutations.

SUMMARY

[0008] Provided herein are systems for modifying the crumbs homologue-1 (CRB1) gene. In some embodiments, the system comprises a Cas protein, or a nucleic acid encoding thereof; a reverse transcriptase, or a nucleic acid encoding thereof; one or more RNA polynucleotides comprising a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence; or one or more nucleic acids encoding thereof; and optionally, a nicking guide RNA (ngRNA), or a nucleic acid encoding thereof. In some embodiments, the RTT sequence encodes one or more base substitutions, deletions, or additions to modify the CRB1 gene sequence.

[0009] In some embodiments, the spacer sequence and the extension sequence are contained within a single RNA polynucleotide.

[0010] In some embodiments, the spacer sequence comprises any of SEQ ID NOs: 1, 17, 23-25, and 36-43. In some embodiments, the extension sequence comprises any of SEQ ID NOs: 44-56, the PBS comprises any of SEQ ID NOs: 4-6, 20-22, 30-31. In some embodiments, the RTT sequence comprises any of SEQ ID NOs: 2-3, 12-13, 14, 18-19, 28-29, and 54-57. In some embodiments, the ngRNA comprises any of SEQ ID NOs: 7-11, 15-16, 23-27, and 32-35.

[0011] In some embodiments, the spacer sequence comprises SEQ ID NO: 1, the PBS comprises any of SEQ ID NOs: 4-6, and the RTT sequence comprises any of SEQ ID NOs: 2-3 and 12-13. In some embodiments, the ngRNA comprises any of SEQ ID NOs: 7-11.

[0012] In some embodiments, the spacer sequence comprises SEQ ID NO: 1, the PBS comprises SEQ ID NO: 4, and the RTT sequence comprises any of SEQ ID NO: 14. In some embodiments, the ngRNA comprises any of SEQ ID NOs: 15-16.

[0013] In some embodiments, the spacer sequence comprises SEQ ID NO: 17, the PBS comprises any of SEQ ID NOs: 20-22, and the RTT sequence comprises any of SEQ ID NOs: 18-19 and 28-29. In some embodiments, the ngRNA comprises any of SEQ ID NOs: 23-27.

[0014] In some embodiments, the spacer sequence comprises any of SEQ ID NOs: 23-25, the PBS comprises any of SEQ ID NOs: 20, 22, and 30-31, and the RTT sequence comprises any of SEQ ID NOs: 54-57. In some embodiments, the ngRNA comprises any of SEQ ID NOs: 32-35.

[0015] In some embodiments, the spacer sequence comprises any of SEQ ID NOs: 36-43 and the extension sequence comprises any of SEQ ID NOs: 44-56.

[0016] In some embodiments, the Cas protein is Cas9 or a variant or fragment thereof. In some embodiments, the Cas protein is a Cas9 nickase. In some embodiments, the Cas protein comprises a Cas protein variant configured to target an expanded range of PAM sequences. In some embodiments, the Cas protein comprises a variant of the *Streptococcus pyogenes* Cas9 selected from xCas9, Cas9-VQR, SpG and SpRY.

[0017] In some embodiments, the Cas protein and the reverse transcriptase are contained within a single fusion protein.

[0018] In some embodiments, the CRB1 gene is a mutant CRB1 gene comprising one or more disease-causing mutations. In some embodiments, the RTT sequence corrects at least one disease-causing mutation in the CRB1 gene. In some embodiments, the CRB1 gene is a wild-type CRB1 gene. In some embodiments, the RTT sequence inserts at least one mutation in the CRB1 gene.

[0019] Provided herein are methods for modifying the crumbs homologue-1 (CRB1) gene comprising contacting a DNA encoding the CRB1 gene with a system as disclosed herein.

[0020] In some embodiments, the DNA encoding the CRB1 gene is in a cell. In some embodiments, the cell is a eukaryotic cell. In some embodiments, the cell is a human cell. In some embodiments, the cell is in vitro. In some embodiments, the cell is ex vivo. In some embodiments, the cell is in vivo.

[0021] In some embodiments, the contacting comprises introducing the system into the cell. In some embodiments, the introducing into the cell comprises administering to a subject.

[0022] Also provided are methods of treating or preventing a disease or disorder in a subject in need thereof. In some embodiments, the methods comprise administering a system as disclosed herein to the subject. In some embodiments, the disease or disorder is caused or mitigated by mutations in CRB1 gene. In some embodiments, the subject has one or more mutations in the CRB1 gene.

[0023] In some embodiments, the disease or disorder comprises a retinal dystrophy. In some embodiments, the disease or disorder comprises autosomal recessive retinitis pigmentosa (RP) or Leber congenital amaurosis (LCA).

[0024] In some embodiments, the system is configured for delivery to retinal cells. In some embodiments, the system is configured for delivery to photoreceptor cells and or Müller glial cells.

[0025] Other aspects and embodiments of the disclosure will be apparent in light of the following detailed description and related figures.

Brief Description of the Figures

[0026] FIG. 1A is a schematic of the retina, showing the canonical crumbs complex involvement in mediating apical polarity and promoting cell adhesion and interaction. FIGS. 1B and 1C show a visual depiction of the correlation between mutations in the CRB1 gene and its affected isoforms.

[0027] FIG. 2 is a schematic of the prime editing mechanism. The spacer anneals with its complementary strand of the DNA (1) directing the H840A SpCas9 nickase to nick the PAM-containing strand (black arrow) of the target DNA (2). The primer binding sequence (PBS) then hybridizes with the nicked DNA (3) initiating the elongation of the free 3′ end according to the reverse transcription template (RTT) sequence that carries the intended edit (4). The newly synthesized strand leads to either 3′ or 5′ flap excision. The excision of the 5′ flap is favored, and it leads to the heteroduplex formation (5). The replacement of the original sequence via endogenous DNA mismatch repair mechanism incorporates the desired mutation at the target site (6).

[0028] FIG. 3 is a graph of the editing efficiencies of the combination of pegRNA and nicking sgRNA. Prime editing delivered by nucleofection to p.(Gly1103Arg) patient iPSCs. Editing efficiencies calculated by ICE analysis.

[0029] FIGS. 4A and 4B show Sanger sequences of bulk iPSCs from two compound heterozygous p.(Cys948Tyr) patients (FIG. 4A) and two homozygous p.(Gly1103Arg) patients (FIG. 4B), before and after prime editing.

[0030] FIG. 5 shows next-generation sequencing (NGS) of bulk cells before and after prime editing for the correction of the p.(Gly1103Arg) mutation for Patient 2 (FIG. 5A) and Patient 1 (FIG. 5B). Unedited sequences are SEQ ID NOs: 58 and 59, edited sequence is SEQ ID NO: 60, and edited + indels are SEQ ID NOs: 61 and 62.

[0031] FIG. 6 is a table of the most prevalent CRB1 mutations.

[0032] FIG. 7 is a graph of the editing efficiencies of the indicated combination of pegRNA and nicking sgRNA for the correction of the c.2843G>A *CRB1* mutation. Prime editing components were delivered by nucleofection to heterozygous p.(Cys948Tyr) patient iPSCs. Editing efficiencies were calculated by NGS. Positive editing was considered that at or above the 50% threshold determined by NGS from control (unedited patient cells).

[0033] FIGS. 8A and 8B shows installation of the CRB1 mutation in control iPSCs and HEK293 cells. FIG. 8A shows prime editing for knock-in of the c.2843G>A, p.(Cys948Tyr) *CRB1* mutation in HEK293 cells. FIG. 8B shows prime editing for knock-in of the c.3307G>A, p.(Gly1103Arg) *CRB1* mutation into control iPSC lines. Wild-type sequence is SEQ ID NO: 63. c.2843G>A sequence is SEQ ID NO: 64.

[0034] FIGS. 9A-9D show dual AAV delivery of split prime editor. FIG. 9A is a schematic of an exemplary design of an AAV split intein system of prime editing using two different vectors for the expression of the prime editor. FIGS. 9B-9D are schematics of PE-C (C-terminal portion) and two versions of PE-N (N-terminal portion) for insertion of the mutation or correction of the mutation, respectively.

[0035] FIG. 10 shows prime editing correction of the c.3307G>A, p.(Gly1103Arg) mutation using a split prime editor.

[0036] FIG. 11A is a graph of the distribution of CRB1 Pathogenic Alleles From the CRB1 Leiden Open Variation Database. Analysis of 1259 individual pathogenic or likely pathogenic CRB1 alleles in the Leiden Open Variation Database. The most prevalent mutation types include G>A (28.4%), C>T(12.6%), T>C (11.5%) and deletions (12.2%). FIG. 11B is a graph of the editable alleles by variation consequence from the CRB1 Leiden Open Variation Database. Missense, termination, and deletion represent the predominant editable CRB1 variants.

[0037] FIG. 12 shows prime editing design for the top ten most prevalent *CRB1* mutations. CRB1 sequence fragments to be edited are from top to bottoms SEQ ID NOs: 65-74, respectively. The spacer and 3' extension sequences can be found in Table 1: c.614 3' extension is SEQ ID NO: 53; c.2843 3' extension is SEQ ID NO: 44; c.2688 3' extension is SEQ ID NO: 48; c.2234 3' extension is SEQ ID NO: 45; c.1576 3' extension is SEQ ID NO: 51; c.3307 3' extension is SEQ ID NO: 52; c.613_619del 3' extension is SEQ ID NO: 49; c.2401 3' extension is SEQ ID NO: 46; c.2290 3' extension is SEQ ID NO: 47; and c.498_506del 3' extension is SEQ ID NO: 50.

Detailed Description

[0038] Crumbs (Crb) is a large transmembrane protein initially discovered at the apical membrane of *Drosophila* epithelial cells. The human CRB1 gene is mapped to chromosome 1q31.3, and contains 12 exons, has 12 identified transcript variants so far, three CRB family members, and over 210 kb genomic DNA. Canonical CRB1 (CRB1-A) is a large transmembrane protein consisting of multiple epidermal growth factor (EGF) and laminin-globular like domains in its extracellular N-terminus. The intracellular C-terminal domain contains a FERM and a conserved glutamic acid-arginine-leucine-isoleucine (ERLI) PDZ binding motives. An alternative transcript of CRB1, CRB1-B, was suggested to have significant extracellular domain overlap with canonical CRB1 while bearing unique 5′ and 3′ domains. In mammals, CRB1 is a member of the Crumbs family together with CRB2 and CRB3.

[0039] The disclosed systems, compositions, and methods advance methods to correct or install CRB1 mutations. As shown herein, the disclosed systems, compositions, and methods corrected CRB1 patient mutations and were able to achieve high editing efficiencies.

[0040] Section headings as used in this section and the entire disclosure herein are merely for organizational purposes and are not intended to be limiting.

Definitions

[0041] The terms "comprise(s)," "include(s)," "having," "has," "can," "contain(s)," and variants thereof, as used herein, are intended to be open-ended transitional phrases, terms, or words that do not preclude the possibility of additional acts or structures. As used herein, comprising a certain sequence or a certain SEQ ID NO usually implies that at least one copy of said sequence is present in recited peptide or polynucleotide. However, two or more copies are

also contemplated. The singular forms "a," "and" and "the" include plural references unless the context clearly dictates otherwise. The present disclosure also contemplates other embodiments "comprising," "consisting of," and "consisting essentially of," the embodiments or elements presented herein, whether explicitly set forth or not.

[0042] For the recitation of numeric ranges herein, each intervening number there between with the same degree of precision is explicitly contemplated. For example, for the range of 6-9, the numbers 7 and 8 are contemplated in addition to 6 and 9, and for the range 6.0-7.0, the number 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, and 7.0 are explicitly contemplated.

[0043] Unless otherwise defined herein, scientific, and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. The meaning and scope of the terms should be clear; in the event, however of any latent ambiguity, definitions provided herein take precedent over any dictionary or extrinsic definition. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular.

[0044] As used herein, the terms "administering," "providing," and "introducing," are used interchangeably herein and refer to the placement into a subject by a method or route which results in at least partial localization a desired site. Administration can be by any appropriate route which results in delivery to a desired location in the subject.

[0045] The term "contacting" as used herein refers to bring or put in contact, to be in or come into contact. The term "contact" as used herein refers to a state or condition of touching or of immediate or local proximity.

[0046] The term "gene" refers to a DNA sequence that comprises control and coding sequences necessary for the production of an RNA having a non-coding function (e.g., a ribosomal or transfer RNA), a polypeptide, or a precursor of any of the foregoing. The RNA or polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired activity or function is retained. Thus, a "gene" refers to a DNA or RNA, or portion thereof, that encodes a polypeptide or an RNA chain that has functional role to play in an organism. For the purpose of this disclosure, it may be considered that genes include regions that regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. Accordingly, a gene includes, but is not necessarily limited to, promoter sequences, terminators, translational regulatory sequences

such as ribosome binding sites and internal ribosome entry sites, enhancers, silencers, insulators, boundary elements, replication origins, matrix attachment sites, and locus control regions.

[0047] As used herein, a "nucleic acid" or a "nucleic acid sequence" refers to a polymer or oligomer of pyrimidine and/or purine bases, preferably cytosine, thymine, and uracil, and adenine and guanine, respectively (See Albert L. Lehninger, Principles of Biochemistry, at 793-800 (Worth Pub. 1982)). The present technology contemplates any deoxyribonucleotide, ribonucleotide, or peptide nucleic acid component, and any chemical variants thereof, such as methylated, hydroxymethylated, or glycosylated forms of these bases, and the like. The polymers or oligomers may be heterogenous or homogenous in composition and may be isolated from naturally occurring sources or may be artificially or synthetically produced. In addition, the nucleic acids may be DNA or RNA, or a mixture thereof, and may exist permanently or transitionally in single-stranded or double-stranded form, including homoduplex, heteroduplex, and hybrid states. In some embodiments, a nucleic acid or nucleic acid sequence comprises other kinds of nucleic acid structures such as, for instance, a DNA/RNA helix, peptide nucleic acid (PNA), morpholino nucleic acid (see, e.g., Braasch and Corey, Biochemistry, 41(14): 4503-4510 (2002)) and U.S. Pat. No. 5,034,506), locked nucleic acid (LNA; see Wahlestedt et al., Proc. Natl. Acad. Sci. U.S.A., 97: 5633-5638 (2000)), cyclohexenyl nucleic acids (see Wang, J. Am. Chem. Soc., 122: 8595-8602 (2000)), and/or a ribozyme. Hence, the term "nucleic acid" or "nucleic acid sequence" may also encompass a chain comprising non-natural nucleotides, modified nucleotides, and/or non- nucleotide building blocks that can exhibit the same function as natural nucleotides (e.g., "nucleotide analogs"); further, the term "nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin, which may be single or double-stranded, and represent the sense or antisense strand. The terms "nucleic acid," "polynucleotide," "nucleotide sequence," and "oligonucleotide" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof.

[0048] As used herein, the term "preventing" refers to partially or completely delaying onset of a disease, disorder and/or condition; partially or completely delaying onset of one or more symptoms, features, or manifestations of a particular disease, disorder, and/or condition; partially or completely delaying progression from a particular disease, disorder and/or condition; and/or

decreasing the risk of developing pathology associated with the disease, disorder, and/or condition.

[0049] As used herein, "treat," "treating," and the like means a slowing, stopping, or reversing of progression of a disease or disorder. The term also means a reversing of the progression of such a disease or disorder. As such, "treating" means an application or administration of the methods or devices described herein to a subject, where the subject has a disease or a symptom of a disease, where the purpose is to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve, or affect the disease or symptoms of the disease.

[0050] A "subject" or "patient" may be human or non-human and may include, for example, animal strains or species used as "model systems" for research purposes, such a mouse model as described herein. Likewise, patient may include either adults or juveniles (e.g., children). Moreover, patient may mean any living organism, preferably a mammal (e.g., human or non-human) that may benefit from the administration of devices and systems contemplated herein. Examples of mammals include, but are not limited to, any member of the Mammalian class: humans, non-human primates such as chimpanzees, and other apes and monkey species; farm animals such as cattle, horses, sheep, goats, swine; domestic animals such as rabbits, dogs, and cats; laboratory animals including rodents, such as rats, mice, and guinea pigs, and the like. Examples of non-mammals include, but are not limited to, birds, fish, and the like. In one embodiment of the methods herein, the mammal is a human.

[0051] A "vector" or "expression vector" is a replicon, such as plasmid, phage, virus, or cosmid, to which another DNA segment, e.g., an "insert," may be attached or incorporated so as to bring about the replication of the attached segment in a cell.

[0052] The term "wild-type" refers to a gene or a gene product that has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designated the "normal" or "wild-type" form of the gene. In contrast, the term "modified," "mutant," or "polymorphic" refers to a gene or gene product that displays modifications in sequence and or functional properties (e.g., altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

[0053] Preferred methods and materials are described below, although methods and materials similar or equivalent to those described herein can be used in practice or testing of the present disclosure. All publications, patent applications, patents and other references mentioned herein are incorporated by reference in their entirety. The materials, methods, and examples disclosed herein are illustrative only and not intended to be limiting.

Prime Editing

[0054] Prime editing is a double-strand break (DSB)-independent clustered-regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system that can ameliorate both transition and transversion mutations in addition to small deletions and insertions. Generally, a prime editing guide RNA (pegRNA) is used in conjunction with a prime editor, e.g., a H840A Streptococcus pyogenes Cas9 (spCas9) nickase linked to an optimized Moloney murine leukemia virus (MMLV) reverse transcriptase (RT).

[0055] pegRNAs are similar to standard single-guide RNAs (sgRNAs) but differ due to a 3′ extension sequence comprising a primer binding site (PBS) and an adjacent reverse transcription template (RTT) sequence. The primer binding site hybridizes with the bases upstream of the prime editor generated nick, while the RTT encodes the information of the intended edits and directs reverse transcription. Together, the prime editor and the pegRNA form the prime editing 2 strategy (PE2). The Cas9 nickase is guided to the DNA target site by the pegRNA. After nicking by Cas9, the reverse transcriptase uses the pegRNA to template reverse transcription of the desired edit, directly polymerizing DNA onto the nicked target DNA strand. The edited DNA strand replaces the original DNA strand, creating a heteroduplex containing one edited strand and one unedited strand. Lastly, the prime editor guides resolution of the heteroduplex to favor copying the edit onto the unedited strand, completing the process. Once the prime editor incorporates the edit into one strand, there is a mismatch between the original sequence on one strand and the edited sequence on the other strand. In some embodiments, an additional nicking guide RNA (ngRNA) is used to nick the non-edited strand, directing DNA repair enzymes to use the edited strand as a template to remake the mismatched strand. The prime editor, the pegRNA, and ngRNA form prime editing 3 (PE3) strategies.

[0056] Disclosed herein are methods and systems for modifying the sequence of the human crumbs homologue-1 gene. The phrase "modifying" as used herein when referencing a nucleic acid sequence, refers to modifying at least one feature of a nucleic acid sequence. Nucleic acid

modifications include, for example, deletion, addition, or insertion of one or more nucleotides to the nucleic acid sequence.

[0057] In some embodiments, the methods and systems comprise a sequence-specific nuclease, or a nucleic acid encoding thereof; an RNA-dependent DNA polymerase, or a nucleic acid encoding thereof; one or more RNA polynucleotides comprising a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence, or one or more nucleic acids encoding thereof; and optionally, a nicking guide RNA (ngRNA), or a nucleic acid encoding thereof. In some embodiments, the RTT sequence encodes one or more nucleotides to modify the crumbs homologue-1 gene. In some embodiments, the RTT sequence is configured to correct any of those mutations found in the table in FIG. 6. In some embodiments, the RTT encodes a wild-type sequence of the crumbs homologue-1 gene.

[0058] In some embodiments, systems and methods described herein may be used to insert one or more defects in the crumbs homologue-1 gene. In some embodiments, the RTT sequence encodes a sequence of the crumbs homologue-1 gene having any of those mutations found in the table in FIG. 6 In such cases, the target sequence encodes a wild-type or normal version of the gene, and the disclosed compositions and systems comprise a nucleic acid molecule which encodes a disease-causing version of the gene. For example, the disclosed composition and system may be used to install mutations for disease modeling in cells and organisms.

[0059] In some embodiments, the methods comprise contacting the DNA encoding the crumbs homologue-1 gene with the disclosed system. In some embodiments, the methods comprise contacting the DNA encoding the crumbs homologue-1 gene with a Cas9 protein; a reverse transcriptase; one or more RNA polynucleotides comprising a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence; and optionally, a nicking guide RNA (ngRNA), wherein the RTT sequence encodes one or more nucleotides to correct or insert a mutation in the crumbs homologue-1 gene. In some embodiments, the RTT encodes a wild-type sequence of the crumbs homologue-1 gene. In some embodiments, the RTT sequence encodes a sequence of the crumbs homologue-1 gene having any of those mutations found in the table in FIG. 6

[0060] In some embodiments, the DNA encoding the crumbs homologue-1 gene is in a cell. In some embodiments, the cell is a eukaryotic cell. In some embodiments, the cell is a human

cell. In some embodiments, the cell is in vitro. In some embodiments, the cell is ex vivo. In some embodiments the cell is in vivo. In some embodiments, the DNA encoding the crumbs homologue-1 gene is genomic DNA. In some embodiments, the crumbs homologue-1 gene is the human crumbs homologue-1 gene.

[0061] In some embodiments, contacting the DNA encoding the crumbs homologue-1 gene comprises introducing into the cell: a nucleic acid encoding a Cas9 protein, a nucleic acid encoding a reverse transcriptase; and a nucleic acid encoding one or more RNA polynucleotides comprising a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence; and, optionally, a nucleic acid encoding a nicking guide RNA (ngRNA). In some embodiments, a single nucleic acid encodes the Cas9 protein and the reverse transcriptase. In some embodiments, a single nucleic acid encodes the spacer sequence, the extension sequence and, optionally, the ngRNA.

[0062] In some embodiments, introducing into the cell comprises administering to the subject.

1. Sequence-specific nuclease

[0063] Exemplary sequence-specific nucleases for use in the present invention include, but are not limited to, Cas proteins, Argonaute (Ago) proteins, zinc finger nucleases (ZFNs), and transcription activator-like effector nucleases (TALEN). In some embodiments, the sequence-specific nuclease is a Cas protein.

[0064] Cas proteins are described in further detail in, e.g., Haft et al., *PLoS Comput. Biol.*, 1(6): e60 (2005), incorporated herein by reference. The Cas protein may be any Cas endonuclease, or fragment or naturally-occurring or engineered variants thereof. In some embodiments, the Cas endonuclease is a Class 2 Cas endonuclease. In some embodiments, the Cas endonuclease is a Type V Cas endonuclease. In some embodiments, the Cas protein is Cas9, Cas12a, otherwise referred to as Cpf1, or Cas14. In one embodiment, the Cas9 protein is a wild-type Cas9 protein. In some embodiments, the Cas9 protein is a Cas9 variant.

[0065] The Cas9 protein can be obtained or derived from any suitable microorganism, and a number of bacteria express Cas9 protein orthologs or variants. In some embodiments, the Cas9 is from *Streptococcus pyogenes* or *Staphylococcus aureus*. Cas9 proteins of other species are known in the art (see, e.g., U.S. Patent Application Publication 2017/0051312, incorporated herein by reference) and may be used in connection with the present disclosure. The amino acid

sequences of Cas proteins from a variety of species are publicly available through the GenBank and UniProt databases.

[0066] In certain embodiments, a Cas nuclease can only cleave a target sequence if an appropriate PAM is present. See, for example Doudna et al., Science, 2014, 346(6213): 1258096, incorporated herein by reference. A PAM site is a nucleotide sequence in proximity to a target sequence. For example, PAM site may be a DNA sequence immediately following the DNA sequence targeted by the Cas protein. A PAM can be 5' or 3' of a target sequence. A PAM can be upstream or downstream of a target sequence. A PAM can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides in length. In certain embodiments, a PAM is between 2-6 nucleotides in length. Non-limiting examples of the PAM sequences include: CC, CA, AG, GT, TA, AC, CA, GC, CG, GG, CT, TG, GA, AGG, TGG, T-rich PAMs (such as TTT, TTG, TTC, etc.), NGG, NGA, NAG, and NGGNG, where "N" is any nucleotide.

[0067] In some embodiments, the Cas protein comprises a Cas variant configured to target an expanded or altered range of PAM sequences which may facilitate essentially PAMless cleavage. In some embodiments, the Cas protein comprises a variant of the *Streptococcus pyogenes* Cas9 enzyme selected from xCas9, Cas9-VQR, SpG and SpRY. See, for example, Walton et al., Science. 2020 Apr 17;368(6488):290-296, Hu, et al., Nature 2018; 556 (57-63), Kleinstiver et al., Nature 2015; 523(7561):481-5, Hu et al., Mol Plant 2016; 9, 43-945, incorporated herein by reference in their entirety.

[0068] In some embodiments, the Cas protein is a Cas9 nickase (Cas9n). Wild-type Cas9 has two catalytic nuclease domains facilitating double-stranded DNA breaks. A Cas9 nickase protein is typically engineered through inactivating point mutation(s) in one of the catalytic nuclease domains causing Cas9 to nick or enzymatically break only one of the two DNA strands using the remaining active nuclease domain. Cas9 nickases are known in the art (see, e.g., U.S. Patent Application Publication 2017/0051312, incorporated herein by reference) and include, for example, *Streptococcus pyogenes* with point mutations at D10 or H840.

2. RNA-dependent DNA polymerase

[0069] RNA-dependent DNA polymerases (e.g., reverse transcriptases) synthesize complementary DNA using RNA as a template. Any RNA-dependent DNA polymerase, or variant or truncation thereof or enzyme having RNA-dependent DNA polymerase activity can be utilized in the systems and methods herein. Exemplary RNA-dependent DNA polymerases

include retroviral reverse transcriptases, retrotransposon reverse transcriptases, bacterial reverse transcriptases, Tth DNA polymerase, Taq DNA polymerase, and Tma DNA polymerase. In some embodiments, the RNA-dependent DNA polymerases is a moloney murine leukemia virus (MMLV) reverse transcriptase

[0070] In some embodiments, the reverse transcriptase and the Cas protein comprise a fusion protein. The reverse transcriptase can be fused to the Cas protein in any orientation and may be separated from the Cas protein with an amino acid linker.

[0071] In some embodiments, the prime editor, fusion protein comprising the reverse transcriptase and the Cas protein, is provided as a split-prime editor (e.g., a prime editor can in some cases be delivered as a split-prime editor, or a nucleic acid(s) encoding a split-prime editor) such that two separate proteins together form a functional prime editor. In some such cases the sequences that encode the two parts of the split-prime editor protein are present on the same vector. In some cases, they are present on separate vectors, e.g., as part of a vector system that encodes the prime editor, the gRNA(s), and systems thereof.

[0072] Exemplary methods and linkers include split-intein protein trans-splicing known in the art, e.g., for reconstituting the Cas9 protein. A split intein is any intein in which the N-terminal domain of the intein and the C-terminal domain of the intein are not directly linked via a peptide bond. Natural split inteins have been identified in cyanobacteria and archaea, but split inteins can also be created artificially by separating an intein's sequence into two pieces. For example, in some embodiments, a first nucleic acid may encode a first protein having the N-terminal portion of the prime editor fused to one half of a split intein and a second nucleic acid may encode a second protein having the C-terminal portion of the prime editor and the other half of the split intein.

3. pegRNA

[0073] The systems and methods disclosed herein include a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence, or one or more nucleic acids encoding thereof. In some embodiments, each of the spacer sequence, PBS, and RTT sequence are provided as a single prime editing guide RNA (pegRNA), or a nucleic acid encoding thereof. The spacer sequence directs the nuclease to bind to a DNA molecule having complementarity with the pegRNA, the PBS hybridizes with the

bases upstream of the nuclease generated nick, and the RTT encodes the information of the intended edits and directs reverse transcription.

[0074] "Complementarity" refers to the ability of a nucleic acid to form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick or other non-traditional types. A percent complementarity indicates the percentage of residues in a nucleic acid molecule, which can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence. Full complementarity is not necessarily required, provided there is sufficient complementarity to cause hybridization.

[0075] The spacer sequence and the extension sequence may be selected from any of the sequences disclosed herein or fragments thereof which lack one or more nucleotides from the 5' and/or 3' end. In some embodiments, the spacer sequence or the extension sequence comprises one, two, three, four, five, six, seven, eight, nine, or ten substitutions as compared to the sequences disclosed herein. In some embodiments, the spacer sequence or the extension sequence comprises one, two, three, four, five, six, seven, eight, nine, or ten additional nucleotides on the 5' and/or 3' end as compared to the sequences disclosed herein. The spacer sequence and the extension sequence may be optimized by the described additions or substitutions for use in a variety of cell types and methods.

[0076] The pegRNAs may comprise additional structural elements or sequences including a gRNA scaffold responsible for Cas9 binding, a transcription termination sequence that the 3' end of the molecule, and mutations or structural motifs that increase editing efficiency or enhance RNA stability or prevent RNA degradation. For example, the pegRNA may further comprise: a triple helix forming sequence (e.g., triple helix terminators from a long non-coding RNAs (lncRNAs), e.g., metastasis-associated lung adenocarcinoma transcript 1 (MALAT1)); a tRNA-like sequence; a pseudoknot (e.g., a modified prequeosine$_1$-1 riboswitch aptamer, (evopreQ$_1$) or the frameshifting pseudoknot from Moloney murine leukemia virus (MMLV)); and silent mutations near the intended edit (e.g., less than 10 bp away). See, for example, Nelson, et al. Nat Biotechnol. 2022 Mar;40(3):402-410, Chen, et al., Cell. 2021 Oct 28;184(22):5635-5652.e29, International Patent Publication No. WO2022067130, each of which is incorporated herein by reference in its entirety.

[0077] The additional structural elements or sequences may be present at any location in the pegRNA which does not interfere with the function of the spacer sequence, primer binding

sequence (PBS), and a reverse transcriptase template (RTT) sequence. In some embodiments, the additional structural elements or sequences are at the 3' end of the pegRNA.

4. Nicking guide RNA (ngRNA)

[0078] In some embodiments, the systems and methods comprise a nicking guide RNA (ngRNA) that complexes with the sequence-specific nuclease and introduces a nick in the non-edited DNA stand. In certain embodiments, the nick induced by using the ngRNA is on the opposite strand as the initial nick. In certain embodiments, the nick induced by using the ngRNA is on the same strand as the initial nick. Thus, the ngRNA sequence may target the same or different strand as the spacer sequence.

[0079] The ngRNA may be selected from any of the sequences disclosed herein or fragments thereof which lack one or more nucleotides from the 5' and/or 3' end. In some embodiments, the ngRNA comprises one, two, three, four, five, six, seven, eight, nine, or ten substitutions as compared to the sequences disclosed herein. In some embodiments, the ngRNA comprises one, two, three, four, five, six, seven, eight, nine, or ten additional nucleotides on the 5' and/or 3' end as compared to the sequences disclosed herein.

[0080] The systems and methods may further include an engineered DNA mismatch repair (MMR)-inhibitor (e.g., protein or silencing RNA), or a nucleic acid encoding thereof. See Chen, et al., Cell. 2021 Oct 28;184(22):5635-5652.e29, incorporated herein by reference.

[0081] Nucleic acids of the present disclosure can comprise any of a number of promoters known to the art, wherein the promoter is constitutive, regulatable or inducible, cell type specific, tissue-specific, or species specific. In addition to the sequence sufficient to direct transcription, a promoter sequence of the invention can also include sequences of other regulatory elements that are involved in modulating transcription (e.g., enhancers, Kozak sequences and introns). Many promoter/regulatory sequences useful for driving constitutive expression of a gene are available in the art and include, but are not limited to, for example, CMV (cytomegalovirus promoter), EF1a (human elongation factor 1 alpha promoter), SV40 (simian vacuolating virus 40 promoter), PGK (mammalian phosphoglycerate kinase promoter), Ubc (human ubiquitin C promoter), human beta-actin promoter, rodent beta-actin promoter, CBh (chicken beta-actin promoter), CAG (hybrid promoter contains CMV enhancer, chicken beta actin promoter, and rabbit beta-globin splice acceptor), TRE (Tetracycline response element promoter), H1 (human polymerase III RNA promoter), U6 (human U6 small nuclear promoter), and the like. Additional promoters

that can be used for expression of the components of the present system, include, without limitation, cytomegalovirus (CMV) intermediate early promoter, a viral LTR such as the Rous sarcoma virus LTR, HIV-LTR, HTLV-1 LTR, Maloney murine leukemia virus (MMLV) LTR, myeoloproliferative sarcoma virus (MPSV) LTR, spleen focus-forming virus (SFFV) LTR, the simian virus 40 (SV40) early promoter, herpes simplex tk virus promoter, elongation factor 1-alpha (EF1-α) promoter with or without the EF1-α intron. Additional promoters include any constitutively active promoter. Alternatively, any regulatable promoter may be used, such that its expression can be modulated within a cell.

[0082] Moreover, inducible expression can be accomplished by placing the nucleic acid encoding such a molecule under the control of an inducible promoter/regulatory sequence. Promoters that are well known in the art can be induced in response to inducing agents such as metals, glucocorticoids, tetracycline, hormones, and the like, are also contemplated for use with the invention. Thus, it will be appreciated that the present disclosure includes the use of any promoter/regulatory sequence known in the art that is capable of driving expression of the desired protein operably linked thereto.

[0083] The present disclosure also provides for vectors containing the nucleic acids and cells containing the nucleic acids or vectors, thereof. The vectors may be used to propagate the nucleic acid in an appropriate cell and/or to allow expression from the nucleic acid (e.g., an expression vector). The person of ordinary skill in the art would be aware of the various vectors available for propagation and expression of a nucleic acid sequence.

[0084] In certain embodiments, vectors of the present disclosure can drive the expression of one or more sequences in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, Nature (1987) 329:840, incorporated herein by reference) and pMT2PC (Kaufman, et al., EMBO J. (1987) 6:187, incorporated herein by reference). When used in mammalian cells, the expression vector's control functions are typically provided by one or more regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, simian virus 40, and others disclosed herein and known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells see, e.g., Chapters 16 and 17 of Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL. 2nd eds., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989, incorporated herein by reference.

[0085] The vectors of the present disclosure may direct the expression of the nucleic acid in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Such regulatory elements include promoters that may be tissue specific or cell specific. The term "tissue specific" as it applies to a promoter refers to a promoter that is capable of directing selective expression of a nucleotide sequence of interest to a specific type of tissue (e.g., seeds) in the relative absence of expression of the same nucleotide sequence of interest in a different type of tissue. The term "cell type specific" as applied to a promoter refers to a promoter that is capable of directing selective expression of a nucleotide sequence of interest in a specific type of cell in the relative absence of expression of the same nucleotide sequence of interest in a different type of cell within the same tissue. The term "cell type specific" when applied to a promoter also means a promoter capable of promoting selective expression of a nucleotide sequence of interest in a region within a single tissue. Cell type specificity of a promoter may be assessed using methods well known in the art, e.g., immunohistochemical staining. In some embodiments, the promoter is specific to retinal cells. In some embodiments, the promoter directs expression in rod and/or cone photoreceptor cells.

[0086] Additionally, the vector may contain, for example, some or all of the following: a selectable marker gene for selection of stable or transient transfectants in host cells; transcription termination and RNA processing signals; 5'-and 3'-untranslated regions; internal ribosome binding sites (IRESes), versatile multiple cloning sites; and reporter gene for assessing expression of the chimeric receptor. Suitable vectors and methods for producing vectors containing transgenes are well known and available in the art. Selectable markers include chloramphenicol resistance, tetracycline resistance, spectinomycin resistance, neomycin, streptomycin resistance, erythromycin resistance, rifampicin resistance, bleomycin resistance, thermally adapted kanamycin resistance, gentamycin resistance, hygromycin resistance, trimethoprim resistance, dihydrofolate reductase (DHFR), GPT; the URA3, HIS4, LEU2, and TRP1 genes of S. cerevisiae.

[0087] When introduced into a cell, the vectors may be maintained as an autonomously replicating sequence or extrachromosomal element or may be integrated into host DNA.

[0088] Thus, the disclosure further provides for cells comprising a system for modifying the crumbs homologue-1 gene, or one or more nucleic acids or vectors encoding thereof, as disclosed herein.

[0089] Viral and non-viral based gene transfer methods can be used to introduce the nucleic acids into cells, tissues, or a subject. Such methods can be used to administer the nucleic acids to cells in culture, or in a host organism. Non-viral vector delivery systems include DNA plasmids, cosmids, RNA (e.g., a transcript of a vector described herein), a nucleic acid, and a nucleic acid complexed with a delivery vehicle.

[0090] Viral vector delivery systems include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell. A variety of viral constructs may be used to deliver the present nucleic acids to the cells, tissues and/or a subject. Viral vectors include, for example, retroviral, lentiviral, adenoviral, adeno-associated, baculoviral, and herpes simplex viral vectors. Nonlimiting examples of such recombinant viruses include recombinant adeno-associated virus (AAV), recombinant adenoviruses, recombinant lentiviruses, recombinant retroviruses, recombinant herpes simplex viruses, recombinant baculoviruses, recombinant poxviruses, phages, etc. The present disclosure provides vectors capable of integration in the host genome, such as retrovirus or lentivirus. See, e.g., Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989; Kay, M. A., et al., 2001 Nat. Medic. 7(1):33-40; and Walther W. and Stein U., 2000 Drugs, 60(2): 249-71, incorporated herein by reference.

[0091] Vectors according to the present disclosure can be transformed, transfected, or otherwise introduced into a wide variety of host cells. Transfection refers to the taking up of a vector by a cell whether or not any coding sequences are in fact expressed. Numerous methods of transfection are known to the ordinarily skilled artisan, for example, lipofectamine, calcium phosphate co-precipitation, electroporation, DEAE-dextran treatment, microinjection, viral infection, and other methods known in the art. Transduction refers to entry of a virus into the cell and expression (e.g., transcription and/or translation) of sequences delivered by the viral vector genome. In the case of a recombinant vector, "transduction" generally refers to entry of the recombinant viral vector into the cell and expression of a nucleic acid of interest delivered by the vector genome.

[0092] Methods of delivering vectors to cells may include DNA or RNA electroporation, transfection reagents such as liposomes or nanoparticles to delivery DNA or RNA; delivery of DNA, RNA, or protein by mechanical deformation (see, e.g., Sharei et al. Proc. Natl. Acad. Sci. USA (2013) 110(6): 2082-2087, incorporated herein by reference); or viral transduction. In some embodiments, the vectors are delivered to host cells by viral transduction. Nucleic acids can be

delivered as part of a larger construct, such as a plasmid or viral vector, or directly, e.g., by electroporation, lipid vesicles, viral transporters, microinjection, and biolistics (high-speed particle bombardment). Similarly, the construct containing the one or more transgenes can be delivered by any method appropriate for introducing nucleic acids into a cell. In some embodiments, the construct or the nucleic acid encoding the components of the present system is a DNA molecule. In some embodiments, the nucleic acid encoding the components of the present system is a DNA vector and may be electroporated to cells. In some embodiments, the nucleic acid encoding the components of the present system is an RNA molecule, which may be electroporated to cells.

[0093] Additionally, delivery vehicles such as nanoparticle- and lipid-based delivery systems can be used. Further examples of delivery vehicles include lentiviral vectors, ribonucleoprotein (RNP) complexes, lipid-based delivery system, gene gun, hydrodynamic, electroporation or nucleofection microinjection, and biolistics. Various gene delivery methods are discussed in detail by Nayerossadat et al. (Adv Biomed Res. 2012; 1: 27) and Ibraheem et al. (Int J Pharm. 2014 Jan 1;459(1-2):70-83), incorporated herein by reference.

[0094] As such, the disclosure provides an isolated cell comprising the vector(s) or nucleic acid(s) disclosed herein. Preferred cells are those that can be easily and reliably grown, have reasonably fast growth rates, have well characterized expression systems, and can be transformed or transfected easily and efficiently. Examples of suitable prokaryotic cells include, but are not limited to, cells from the genera *Bacillus* (such as *Bacillus subtilis* and *Bacillus brevis*), *Escherichia* (such as *E. coli*), *Pseudomonas*, *Streptomyces*, *Salmonella*, and *Envinia*. Suitable eukaryotic cells are known in the art and include, for example, yeast cells, insect cells, and mammalian cells. Examples of suitable yeast cells include those from the genera *Kluyveromyces*, *Pichia*, *Rhino-sporidium*, *Saccharomyces*, and *Schizosaccharomyces*. Exemplary insect cells include Sf-9 and HIS (Invitrogen, Carlsbad, Calif.) and are described in, for example, Kitts et al., *Biotechniques*, *14*: 810-817 (1993); Lucklow, *Curr. Opin. Biotechnol.*, *4*: 564-572 (1993); and Lucklow et al., *J. Virol.*, *67*: 4566-4579 (1993), incorporated herein by reference. A number of suitable mammalian and human host cells are known in the art, and many are available from the American Type Culture Collection (ATCC, Manassas, Va.). Examples of suitable mammalian cells include, but are not limited to, Chinese hamster ovary cells (CHO) (ATCC No. CCL61), CHO DHFR-cells (Urlaub et al., Proc. Natl. Acad. Sci. USA, 97: 4216-4220 (1980)), human

embryonic kidney (HEK) 293 or 293T cells (ATCC No. CRL1573), and 3T3 cells (ATCC No. CCL92). Other suitable mammalian cell lines are the monkey COS-1 (ATCC No. CRL1650) and COS-7 cell lines (ATCC No. CRL1651), as well as the CV-1 cell line (ATCC No. CCL70). Further exemplary mammalian host cells include primate, rodent, and human cell lines, including transformed cell lines. Normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Other suitable mammalian cell lines include, but are not limited to, mouse neuroblastoma N2A cells, HeLa, HEK, A549, HepG2, mouse L-929 cells, and BHK or HaK hamster cell lines. Methods for selecting suitable mammalian cells and methods for transformation, culture, amplification, screening, and purification of cells are known in the art.

[0095] In some embodiments, the cell is a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is a human cell. In some embodiments, the cell is in vitro. In some embodiments, the cell is ex vivo. In some embodiments, the cell is in vivo and delivery to the cell comprises administration to a subject.

Methods of Treating a Disease or Disorder

[0096] Also disclosed herein are methods for treating or preventing a disease or disorder in a subject caused or mitigated by mutations in the crumbs homologue-1 (CRB1) gene. In some embodiments, the subject has a mutation in in the CRB1 gene. In some embodiments, the subject has a CRB1gene with any of the mutations shown in FIG. 6.

[0097] The methods comprise administering to a subject: a sequence-specific nuclease, or a nucleic acid encoding thereof; an RNA-dependent DNA polymerase, or a nucleic acid encoding thereof; one or more RNA polynucleotides comprising a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence, or one or more nucleic acids encoding thereof; and optionally, a nicking guide RNA (ngRNA), or a nucleic acid encoding thereof. The systems and components disclosed herein are applicable to the methods for treatment and prevention of a disease or disorder.

[0098] In some embodiments, the disease or disorder is an ocular disease or disorder. In some embodiments, the disease or disorder is a retinal disease or disorder. In some embodiments, the disease or disorder comprises retinal degeneration, blindness or severe visual impairment, night blindness, loss of peripheral vision, progressive photoreceptor degeneration, nonrecordable

electroretinogram (ERG), nystagmus, hypermetropia, sluggish or absent pupillary responses and oculodigital reflexes, macular dystrophy, foveal schisis, or a combination thereof.

[0099] A wide range of retinal dystrophies can be caused by mutations in the CRB1 gene. In some embodiments, disease or disorder characterized by CRB1 mutations includes, but is not limited to, autosomal recessive retinitis pigmentosa (RP) and Leber congenital amaurosis (LCA).

[0100] In some embodiments, the systems or components thereof are configured for delivery to retinal cells. In some embodiments, the system is configured for delivery to rod and cone photoreceptor cells and/or Müller glial cells. For example, in some embodiments, the nucleic acids encoding the components may comprise a retinal cell (e.g., rod and/or cone photoreceptor cell or Müller glial cell) promoter which directs expression of the components in the retinal cells. In some embodiments, mini promoters, minimal promoter element(s) designed for expression in specific types (e.g., rod and/or cone photoreceptor cell or Müller glial cell) are used.

[0101] Suitable retinal, rod, and/or cone photoreceptor cell promoters include, but are not limited to: 770En_454P(h*GRM6*), a human *GRM6* gene-derived, short promoter; promoters based on the 2.1-kb human L-opsin promoter (pR2.1); promoter derived from the rhodopsin kinase (RK) gene; promoter derived from the rhodopsin gene; a promoter derived from the *Nrl* gene; murine rhodopsin promoter (mOP); G-protein-coupled receptor protein kinase 1 (GRK1) promoter; retinol-binding protein 3, interstitial (RBP3) promoter; RPE65 promoter; human inter-photoreceptor retinoid binding protein/retinol-binding protein 3 (IRBP) promoter; and retinaldehyde binding protein 1 (*RLBP1)* promoter. Suitable Müller glial cell promoters include but, are not limited to: RLBP1 (Retinaldehyde Binding Protein 1), GFAP (Glial fibrillary acidic protein), GfaABC1D (a truncated GFAP promoter), and synthetic promoters ProB2 and PROC17. Suitable photoreceptor cell promoters include, but are not limited to: interphotoreceptor retinoid-binding protein (IRBP), cone arrestin (CAR), rhodopsin (RHO), PR1.7 (a truncated version of version of the L-opsin promoter), synthetic promoters: ProA1, ProA6, ProC1, ProA14, and ProA36, and G protein-coupled receptor kinase 1 (GRK1).

[0102] Additionally, or alternatively, the systems or components are configured for administration to the eye and/or retina, rather than systemic administration.

[0103] Administration may be through any suitable mode of administration, including but not limited to: intravenous, intra-arterial, intramuscular, intracardiac, intrathecal, subventricular, epidural, intracerebral, intracerebroventricular, sub-retinal, intravitreal, intraarticular, intraocular,

intraperitoneal, intrauterine, intradermal, subcutaneous, transdermal, transmucosal, topical, and inhalation. In some embodiments, the systems or components are delivered to the tissue(s) of interest. Such delivery may be either via a single dose, or multiple doses.

[0104] In some embodiments, an effective amount of the components of the systems, methods or compositions as described can be administered. As used herein the term "effective amount" may be used interchangeably with the term "therapeutically effective amount" and refers to that quantity that is sufficient to result in a desired activity upon administration to a subject in need thereof. Within the context of the present disclosure, the term "effective amount" refers to that quantity of the components of the system such that successful modification of the CRB1 gene is achieved.

[0105] When utilized as a method of treatment, the effective amount may depend on the particular condition being treated, the severity of the condition, the individual patient parameters including age, physical condition, size, gender and weight, the duration of the treatment, the nature of concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. In some embodiments, the effective amount alleviates, relieves, ameliorates, improves, reduces the symptoms, or delays the progression of any disease or disorder in the subject. In some embodiments, the subject is a human.

Examples

[0106] The following are examples of the present invention and are not to be construed as limiting.

Materials and Methods

[0107] *Plasmid Constructs and Cloning.* The following plasmids were used: PEmax NGG and NGA (Chen et al., 2021 Cell), pU6-ePegRNA-GG-acceptor (Nelson et al., 2022 Nat Biotechnol) and pU6- spacer-acceptor (Tsai et al., 2022 Methods Mol Biol). The pegRNA was cloned into the pU6-ePegRNA-GG-acceptor and nicking guide RNA (ngRNA) was cloned into the pU6- spacer-acceptor, both using BsaI Golden Gate assembly (NEB), as described previously (Anzalone et al., 2019. Nature; Nelson et al., 2022 Nat Biotechnol; Tsai et al., 2022 Methods Mol Biol). All oligos were ordered from IDT with relevant overhangs for above described

cloning methodology. A "G" was added to all spacer sequences that did not start with a G, for high expression the human U6 promoter prefers a "G" at transcription start.

[0108] *HEK293 Transfection.* To test editing efficiency of the prime editing, the plasmids were transfected into HEK293 cells. The HEK293 cells were seeded one day before at 50000 cells/well in 24 well plates. On the next day, the medium was refreshed with 500 ul complete medium. For the transfection, the plasmid constitution is 1050ng:393.75ng:78.75ng of CMV-PEmax NGG or CMV-PEmax NGA: U6-pegRNA with append: U6-ngRNA. Lipofectamine 2000 (Thermo Fisher) was mixed at 1:1 mass ratio with plasmid DNA. The cells were then collected after 72 hours post transfection for DNA extraction and analysis.

[0109] *iPSC Transfection.* To test prime editing efficiency in *CRB1* patient or control iPSC lines the plasmids were nucleofected into cells using the Lonza P3 Primary Cell Nucleofector Kit and the Amaxa 4D-Nucleofector (Lonza Biosciences). Successful editing was observed with both the DS-150 and CA-137 nucleofection programs. iPSCs were pre-treated for 2 hours with mTeSR plus and rock inhibitor or Clone R2 (Stem Cell Technologies). Cells were lifted using either ReLeSR (Stem Cell Technologies) or TrypLE (Gibco) and counted. 2 x10^5 cells were used and nucleofected with plasmid at the following ratio: 700ng:262.5ng:52.5ng of CMV-PEmax NGG or CMV-PEmax NGA: U6-pegRNA with append: U6-ngRNA. Post nucleofection cells were resuspended in 150ul of mTeSR plus and rock inhibitor or Clone R2 and seeded in matrigel-covered wells in 96 well plate. Medium was refreshed at 48h post nucleofection. The cells were then collected after 72 hours post transfection for DNA extraction and analysis.

[0110] *DNA extraction.* The cells were detached from the wells by trypsin for HEK293 and ReLeSR (Stem Cell Technologies) for iPSCs. The cells from each well were washed with DPBS (without Ca2+ and Mg2+) and re-suspended with 50ul DPBS. The cells were then incubated at 95°C for 20 mins. Subsequently, after cooling, 4ul of 20mg/ml Proteinase K (Promega) was added to each sample. The samples were then incubated at 56°C for one hour followed by a 30 min incubation at 95°C to stop the proteinase K digestion. This crude extract of DNA is then ready for PCR purpose.

[0111] *Analysis of Prime Editing Efficiency.* For the determination of editing efficiency, the relevant CRB1 locus was amplified using primers with Illumina adaptors. The amplicon was submitted to Genewiz for the Amplicon EZ service. The analysis of the sequencing data was determined using CRISPResso2 (crispresso.pinellolab.partners.org/submission). The reads of the

amplicons less than 0.1% of the total frequency were excluded for analysis. In some cases, the CRB1 locus was amplified with primers and editing efficiency was evaluated using ICE analysis (Synthego).

[0112] *Correcting Human CRB1 c.2843G>A pathogenic mutation, NGA PAM*

pegRNA_name	protospacer	SEQ ID NO	RT_temp	SEQ ID NO	PBS	SEQ ID NO
NGA_PBS11_RTT16	TATTATCACC TTCTCTCATT	1	ATTTGCAAT ACCTAAT	2	GAGAG AAGGT G	4
NGA_PBS13_RTT16	TATTATCACC TTCTCTCATT	1	ATTTGCAAT ACCTAAT	2	GAGAG AAGGT GAT	5
NGA_PBS15_RTT16	TATTATCACC TTCTCTCATT	1	ATTTGCAAT ACCTAAT	2	GAGAG AAGGT GATAA	6
NGA_PBS11_RTT26	TATTATCACC TTCTCTCATT	1	TAAAAACA GCATTTGCA ATACCTAAT	3	GAGAG AAGGT G	4
NGA_PBS13_RTT26	TATTATCACC TTCTCTCATT	1	TAAAAACA GCATTTGCA ATACCTAAT	3	GAGAG AAGGT GAT	5
NGA_PBS15_RTT26	TATTATCACC TTCTCTCATT	1	TAAAAACA GCATTTGCA ATACCTAAT	3	GAGAG AAGGT GATAA	6

Nicking sgRNA name	protospacer	SEQ ID NO
2843G>A nsgRNA 1	CATTGCTTCTGAATAATATT	7
2843G>A nsgRNA 2	GGTAATATTCCCATTGCTTC	8
2843G>A nsgRNA 3	GAAACCAAATGTGATATTGG	9
2843G>A nsgRNA 4	CCTTGTTCTGAAACCAAATG	10
2843G>A nsgRNA 5	ATTACCGTTATTAATAGTAA	11

[0113] *Correcting Human CRB1 c.2843G>A pathogenic mutation with silent mutation, NGA PAM*

pegRNA_name	protospacer	SEQ ID NO	RT_temp	SEQ ID NO	PBS	SEQ ID NO
SM-948-PBS11_RTT16	TATTATCACC TTCTCTCATT	1	ATTTGCgATg cCTAAT	12	GAGAGAA GGTG	4
SM-948-PBS13_RTT16	TATTATCACC TTCTCTCATT	1	ATTTGCgATg cCTAAT	12	GAGAGAA GGTGAT	5
SM-948-PBS15_RTT16	TATTATCACC TTCTCTCATT	1	ATTTGCgATg cCTAAT	12	GAGAGAA GGTGATAA	6
SM-948-PBS11_RTT26	TATTATCACC TTCTCTCATT	1	TAAAAACAG CATTTGCgA TgcCTAAT	13	GAGAGAA GGTG	4

SM-948-PBS13_RTT26	TATTATCACC TTCTCTCATT	1	TAAAAACAG CATTTGCgA TgcCTAAT	13	GAGAGAA GGTGAT	5
SM-948-PBS15_RTT26	TATTATCACC TTCTCTCATT	1	TAAAAACAG CATTTGCgA TgcCTAAT	13	GAGAGAA GGTGATAA	6

[0114] *Installing Human CRB1 c.2843G>A pathogenic mutation, NGG PAM*

pegRNA_name	protospacer	SEQ ID NO	RT_temp	SEQ ID NO	PBS	SEQ ID NO
2843G>A Knockin_Proto_1_NGG_ PBS11_RTT26	TATTATC ACCTTCT CTCATT	1	TAAAAACAG CATTTGCAA TATCTAAT	14	GAGAG AAGGTG	4

Nicking sgRNA name	protospacer	SEQ ID NO
2843G>A Knockin-Nick1	GTGTGATATTGGTGAGTTCTC	15
2843G>A Knockin-Nick2	GTCTGAAACCAAATGTGATAT	16

[0115] *Correcting Human CRB1 c.3307G>A pathogenic mutation, NGG PAM*

pegRNA_name	protospacer	SEQ ID NO	RT_temp	SEQ ID NO	PBS	SEQ ID NO
hCRB1-3307G_Proto_1 _NGG_PBS11_ RTT15	CTAAGTACA ATAGAAATC AG	17	AGATAAATG CCTCCG	18	ATTTCT ATTGT	20
hCRB1-3307G_Proto_1 _NGG_PBS13_ RTT15	CTAAGTACA ATAGAAATC AG	17	AGATAAATG CCTCCG	18	ATTTCT ATTGTA C	21
hCRB1-3307G_Proto_1 _NGG_PBS15_ RTT15	CTAAGTACA ATAGAAATC AG	17	AGATAAATG CCTCCG	18	ATTTCT ATTGTA CTT	22
hCRB1-3307G_Proto_1 _NGG_PBS11_ RTT25	CTAAGTACA ATAGAAATC AG	17	AAAGTAAGA GAGATAAAT GCCTCCG	19	ATTTCT ATTGT	20
hCRB1-3307G_Proto_1 _NGG_PBS13_ RTT25	CTAAGTACA ATAGAAATC AG	17	AAAGTAAGA GAGATAAAT GCCTCCG	19	ATTTCT ATTGTA C	21
hCRB1-3307G_Proto_1 _NGG_PBS15_ RTT25	CTAAGTACA ATAGAAATC AG	17	AAAGTAAGA GAGATAAAT GCCTCCG	19	ATTTCT ATTGTA CTT	22

Nicking sgRNA name	protospacer	SEQ ID NO
3307G>A nsgRNA 1	TTGAGAAATTGCTCTTCCTG	23
3307G>A nsgRNA 2	ACAGCCAGTGACCACTGAAT	24
3307G>A nsgRNA 3	TCCTCCATGCAAACAGGGGT	25
3307G>A nsgRNA 4	TATTATCCTTCAAAAAGTTG	26
3307G>A nsgRNA 5	TTGTACTTAGACACCCTTGC	27

[0116] *Correcting CRB1 c.3307G>A pathogenic mutation with silent mutations, NGG PAM*

pegRNA_name	protospacer	SEQ ID NO	RT_temp	SEQ ID NO	PBS	SEQ ID NO
SM-1103-RTT15_PBS11	CTAAGTACAATAGAAATCAG	17	AGATAAATtCCgCca	28	ATTTCTATTGT	20
SM-1103-RTT15_PBS13	CTAAGTACAATAGAAATCAG	17	AGATAAATtCCgCca	28	ATTTCTATTGTAC	21
SM-1103-RTT15_PBS15	CTAAGTACAATAGAAATCAG	17	AGATAAATtCCgCca	28	ATTTCTATTGTACTT	22
SM-1103-RTT25_PBS11	CTAAGTACAATAGAAATCAG	17	AAAGTAAGAGAGATAAATtCCgCca	29	ATTTCTATTGT	20
SM-1103-RTT25_PBS13	CTAAGTACAATAGAAATCAG	17	AAAGTAAGAGAGATAAATtCCgCca	29	ATTTCTATTGTAC	21
SM-1103-RTT25_PBS15	CTAAGTACAATAGAAATCAG	17	AAAGTAAGAGAGATAAATtCCgCca	29	ATTTCTATTGTACTT	22

[0117] *Installing Human CRB1 c.3307G>A pathogenic mutation, NGG PAM*

pegRNA_name	protospacer	SEQ ID NO	RT_temp	SEQ ID NO	PBS	SEQ ID NO
3307G>A Knockin_NGG_RTT25_PBS15	CTAAGTACAATAGAAATCGG	23	AAAGTAAGAGAGATAAATGCCTCTG	54	ATTTCTATTGTACTT	22
3307G>A Knockin_NGG_RTT15_PBS11	CTAAGTACAATAGAAATCGG	23	AGATAAATGCCTCTG	55	ATTTCTATTGT	20
3307G>A Knockin-PBS12-RTT15 - Design2	TGTCTAAGTACAATAGAAAT	24	TAAATGCCTCtGATT	56	TCTATTGTACTT	30
3307G>A Knockin-PBS12-RTT15-Design 3	GTAAGAGAGATAAATGCCTC	25	CAATAGAAATCaGAG	57	GCATTTATCTCT	31

Nicking sgRNA name	protospacer	SEQ ID NO
3307G>A NGA nsgRNA 1	TATGTGGGAGACAGAGCTAT	32
3307G>A NGA nsgRNA 2	TGACCAGCACAATTGCTACT	33
3307G>A NGA nsgRNA 3	GGTTTCATTAATAAACCTCA	34
3307G>A NGA nsgRNA 4	ACTCCAACCCCTGTTTGCAT	35

[0118] *Prime Editing Analysis.* The top ten most prevalent CRB1 mutations described in the LOVD database were evaluated for correction using prime editing. Only the use of both the canonical NGG PAM prime editor and the NGA PAM prime editor were considered. For those mutations with multiple possible designs, the NGG or NGA design that led to a nick closest to the mutation was chosen. Designs in the top DNA strand are labeled as positive (+) and those in the bottom strand are labeled as negative (-). Designs up to 10 nucleotides from the nick were included in the evaluation. Designs were made using only pegRNA; no nicking sgRNA designs were considered. Specific protospacers were identified for each mutation from which 3'-extensions were designed. Prime editing efficiencies are highly associated with the composition of the pegRNAs, as such optimizations of primer binding sequences (PBS) and reverse transcriptase (RT) template are necessary to achieve the best prime editing activities. Here, a PBS of 13 nucleotides in length and a RT template of 16 nucleotides in length were chosen, not accounting for insertions and deletions that require a longer RT.

Example 1

[0119] Thirty combinations of prime editing guide RNA (pegRNA) and nicking single guide RNA (sgRNA) were tested per mutation, c.2843G>A (p.(Cys948Tyr)) and c.3307G>A (p.(Gly1103Arg)), on the corresponding CRB1 patient iPSC lines using nucleofection or Stem Cell lipofectamine. DNA was extracted 3 days after transfection/nucleofection. The target loci were amplified by PCR and editing efficiency was evaluated by Next Generation Sequencing (NGS) or Sanger sequencing followed by ICE analysis.

[0120] Contrary to conventional editors and base editing techniques, p.(Gly1103Arg) and p.(Cys948Tyr) patients' iPSC lines were amenable to prime editing, with editing efficiencies as high as 72% dependent on the combination of pegRNA (primer binding site (*PBS*) and reverse transcription template (*RTT*) length) and nicking sgRNA for a particular mutation. See FIGS. 3, 4, 5, and 7. As shown in FIG. 10, a split intein system may be utilized to deliver prime editing components to facilitate successful editing.

[0121] Prime editing was also used to insert the c.2843G>A (p.(Cys948Tyr)) and c.3307G>A (p.(Gly1103Arg)) mutations into control iPSCs and HEK293 cells (FIGS. 8A-8B).

Example 2

[0122] To assess the prevalence and spectrum of *CRB1* associated pathogenic variants amenable to base and prime editing, analysis of the Leiden Open Variation Database was carried out. Editable variants accounted for 54.5% for base editing and 99.8% for prime editing of all *CRB1* pathogenic variants in the Leiden Open Variation Database. The 10 most common editable pathogenic variants for *CRB1* accounted for 34.95% of all pathogenic variants, with the c.2843G>A, p.(Cys948Tyr) being the most common editable *CRB1* variant.

[0123] Editable variants G>A (28.4%), C>T (12.6%), T>C (11.5%) and deletions (12.2%) were the most prevalent types (FIG. 11A) Missense variants were the predominant editable variants observed in the *CRB1* gene (64.8%) followed by termination (define as SNVs that are associated with in-frame codons being altered to stop codons) (16.5%) and deletion variants (12.2%) (FIG. 11B).

[0124] Of the top ten *CRB1* mutations, 6 are transitions, 2 are transversions, and 2 are deletions. All of those have suitable prime editing designs using either NGG or NGA PAM designs (Table 1). Mutations positioned far away from the PAM (at position ±33, for example) are still targetable by prime editing and can give rise to efficient editing as prime editing has been shown to have a low dependency on PAM placement. Here, example prime editing strategies with the edit positions close to the PAM are highlighted. All the top 10 *CRB1* mutations show the ability to be prime editable (See FIG. 12 for a schematics of exemplary prime editing designs).

Table 1

Mutation	PAM	Edit position	Spacer	SEQ ID NO	3' Extension	SEQ ID NO
c.2843G>A	AGA	+6	TATTATC ACCTTCTC TCATT	1	ATTTGCAATACCTAATG AGAGAAGGTGAT	44
c.2234C>T	CGG	-1	AGCCTGA TGGTTGA AGCATT	36	CCATGTTTGTCCGAACG CTTCAACCATCA	45
c.2401A>T	AGA	-4	CAGTTCA ATTTTATA TGGCT	37	TCTTTGAAAATCAAGCC ATATAAAATTGA	46

c.2290C>T	AGA	-4	GCCGCGC TCTAGCC AGACAC	38	TATCAATATATCCGTGTC TGGCTAGAGCG	47
c.2688T>A	GGG	-1	CAAACAC CTCCATT GTGACA	39	TCAGTCCAACCCCTGTC ACAATGGAGGTG	48
c.613_619del	AGA	+4	GAGGCTA CATGCCT CAATGA	40	TTGTGGGGACAGATACA AGTATATCTTCCTATTTC ATTGAGGCATGTAG	49
c.498_506del	TGG	+6	AATGGGG CCGTGTG CCAGGA	41	GGACACAGAAGCAGGA GTAACCATCAATTCCAT CCTGGCACACGGCCC	50
c.1576C>T	TGA	+4	CAATGGC TCTTCTAC TTTTC	42	CCCTGTTGCTTCGGAAA AGTAGAAGAGCC	51
c.3307G>A	AGG	+2	CTAAGTA CAATAGA AATCAG	17	GAGATAAATGCCTCCGA TTTCTATTGTAC	52
c.614T>C	AGG	+3	GCTACAT GCCTCAA TGAAAC	43	AGTATATCTTCCTTTTTC ATTGAGGCATG	53

[0125] The scope of the present invention is not limited by what has been specifically shown and described hereinabove. Those skilled in the art will recognize that there are suitable alternatives to the depicted examples of materials, configurations, constructions, and dimensions. Variations, modifications, and other implementations of what is described herein will occur to those of ordinary skill in the art without departing from the spirit and scope of the invention.

[0126] Numerous references, including patents and various publications, are cited and discussed in the description of this invention. The citation and discussion of such references is provided merely to clarify the description of the present invention and is not an admission that any reference is prior art to the invention described herein. All references cited and discussed in this specification are incorporated herein by reference in their entirety.

CLAIMS

What is claimed is:

1. A system for modifying the crumbs homologue-1 (CRB1) gene comprising:

a Cas protein, or a nucleic acid encoding thereof;

a reverse transcriptase, or a nucleic acid encoding thereof;

one or more RNA polynucleotides comprising a spacer sequence and an extension sequence comprising a primer binding sequence (PBS) and a reverse transcriptase template (RTT) sequence; or one or more nucleic acids encoding thereof; and

optionally, a nicking guide RNA (ngRNA), or a nucleic acid encoding thereof,

wherein the RTT sequence encodes one or more base substitutions, deletions, or additions to modify the CRB1 gene sequence.

2. The system of claim 1, wherein the spacer sequence and the extension sequence are contained within a single RNA polynucleotide.

3. The system of claim 1 or 2, wherein the spacer sequence comprises any of SEQ ID NOs: 1, 17, 23-25, and 36-43, the extension sequence comprises any of SEQ ID NOs: 44-56, the PBS comprises any of SEQ ID NOs: 4-6, 20-22, 30-31, the RTT sequence comprises any of SEQ ID NOs: 2-3, 12-13, 14, 18-19, 28-29, and 54-57, and the ngRNA comprises any of SEQ ID NOs: 7-11, 15-16, 23-27, and 32-35.

4. The system of any of claims 1-3, wherein the Cas protein is Cas9 or a variant or fragment thereof.

5. The system of any of claims 1-4, wherein the Cas protein is a Cas9 nickase.

6. The system of any of claims 1-5, wherein the Cas protein comprises a Cas protein variant configured to target an expanded range of PAM sequences.

7. The system of any of claims 1-6, wherein the Cas protein and the reverse transcriptase are contained within a single fusion protein.

8. The system of any of claims 1-7, wherein the CRB1 gene is a mutant CRB1 gene comprising one or more disease-causing mutations and wherein the RTT sequence corrects at least one disease-causing mutation in the CRB1 gene or wherein the CRB1 gene is a wild-type CRB1 gene and wherein the RTT sequence inserts at least one mutation in the CRB1 gene.

9. A method for modifying the crumbs homologue-1 (CRB1) gene comprising contacting a DNA encoding the CRB1 gene with a system of any of claims 1-8.

10. The method of claim 9, wherein the DNA encoding the CRB1 gene is in a cell and the contacting comprises introducing the system into the cell.

11. The method of claim 10, wherein the cell is in vitro, ex vivo, or in vivo.

12. The method of claim 10 or 11, wherein the cell is a human cell.

13. A method of treating or preventing a disease or disorder in a subject in need thereof comprising administering of a system of any of claims 1-8 to the subject,

wherein the disease or disorder is caused or mitigated by mutations in CRB1 gene.

14. The method of claim 13, wherein the disease or disorder comprises autosomal recessive retinitis pigmentosa (RP) or Leber congenital amaurosis (LCA).

15. The method of claim 13 or 14, wherein the system is configured for delivery to photoreceptor cells and/or Müller glial cells.

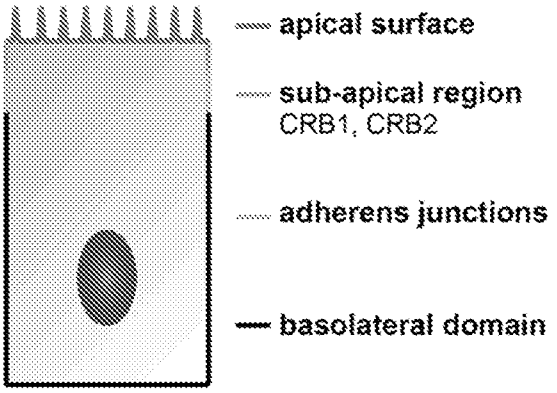

FIG. 1A

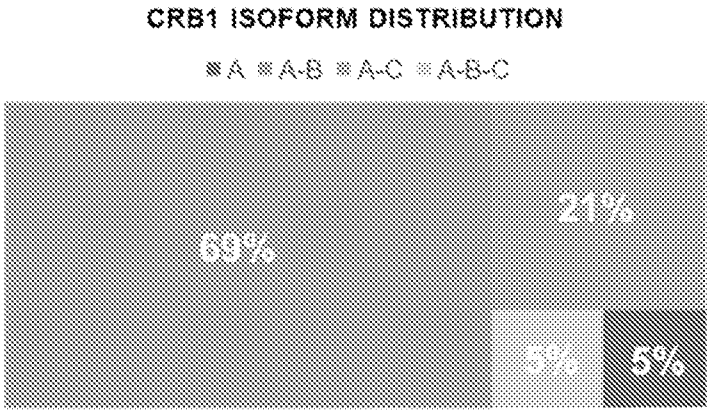

FIG. 1B

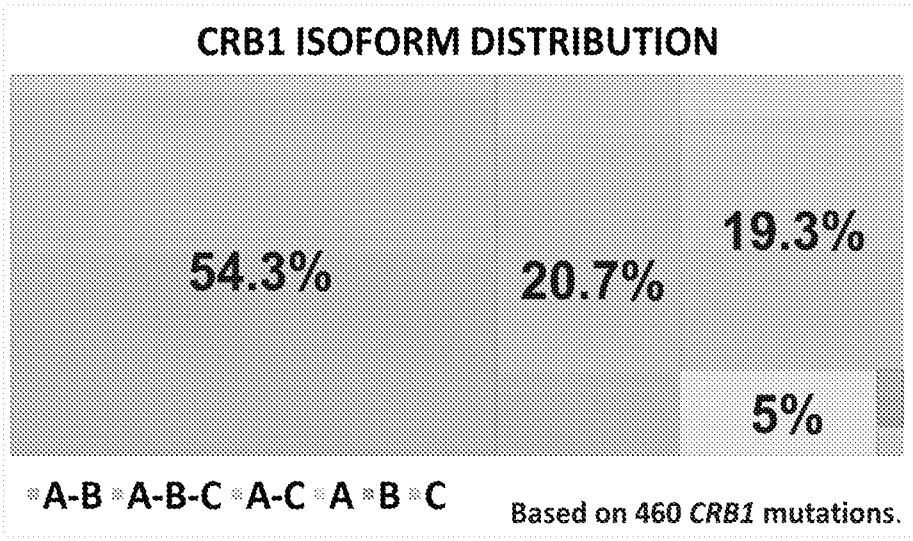

FIG. 1C

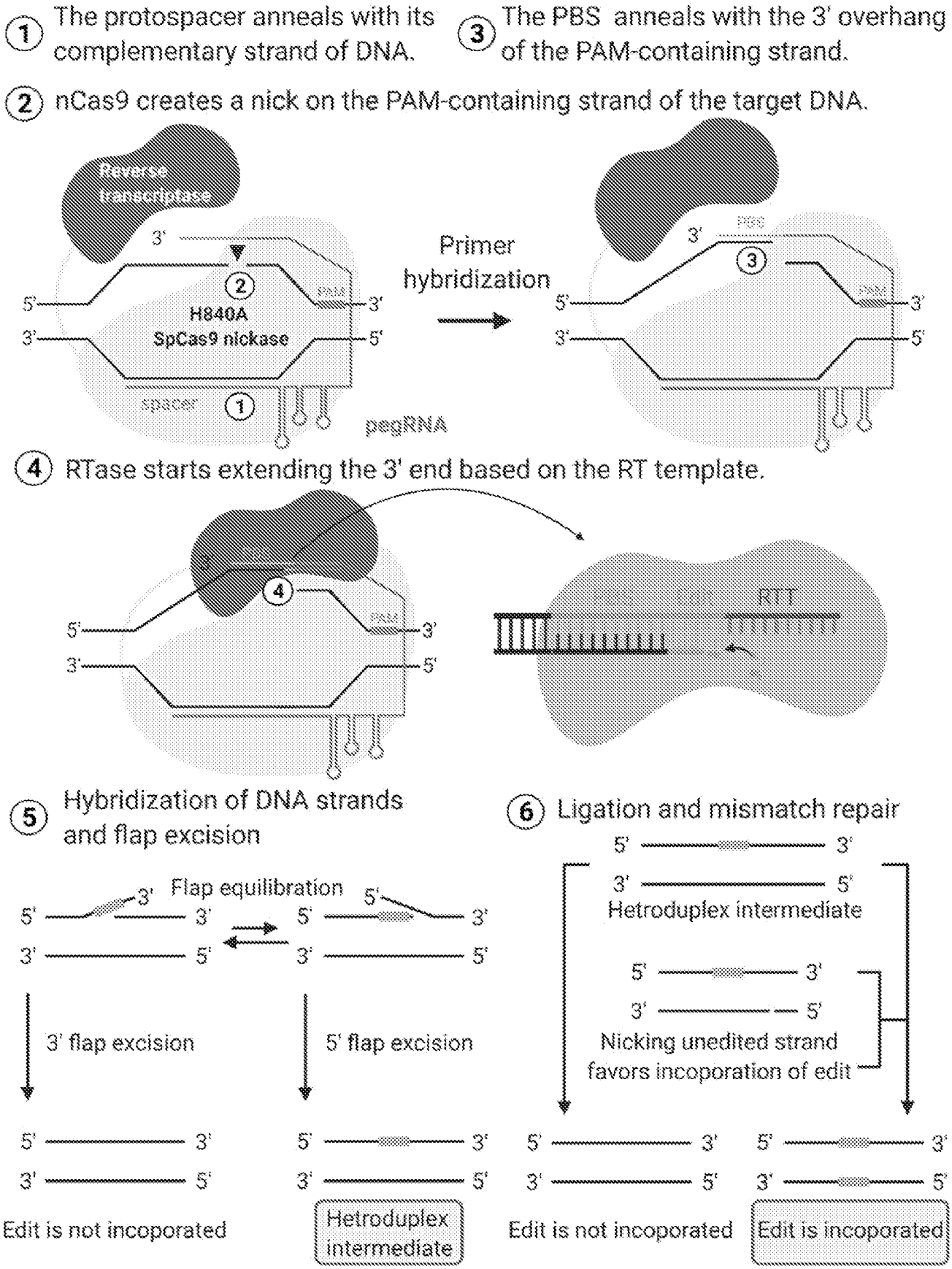
1 The protospacer anneals with its complementary strand of DNA.
3 The PBS anneals with the 3' overhang of the PAM-containing strand.
2 nCas9 creates a nick on the PAM-containing strand of the target DNA.
Reverse transcriptase
3'
2
5'
3'
H840A
SpCas9 nickase
PAM
3'
5'
spacer
1
pegRNA
Primer hybridization
PBS
3
4 RTase starts extending the 3' end based on the RT template.
4
RTT
5 Hybridization of DNA strands and flap excision
Flap equilibration
3' flap excision
5' flap excision
Edit is not incoporated
Hetroduplex intermediate
6 Ligation and mismatch repair
Hetroduplex intermediate
Nicking unedited strand favors incoporation of edit
Edit is not incoporated
Edit is incoporated

FIG. 2

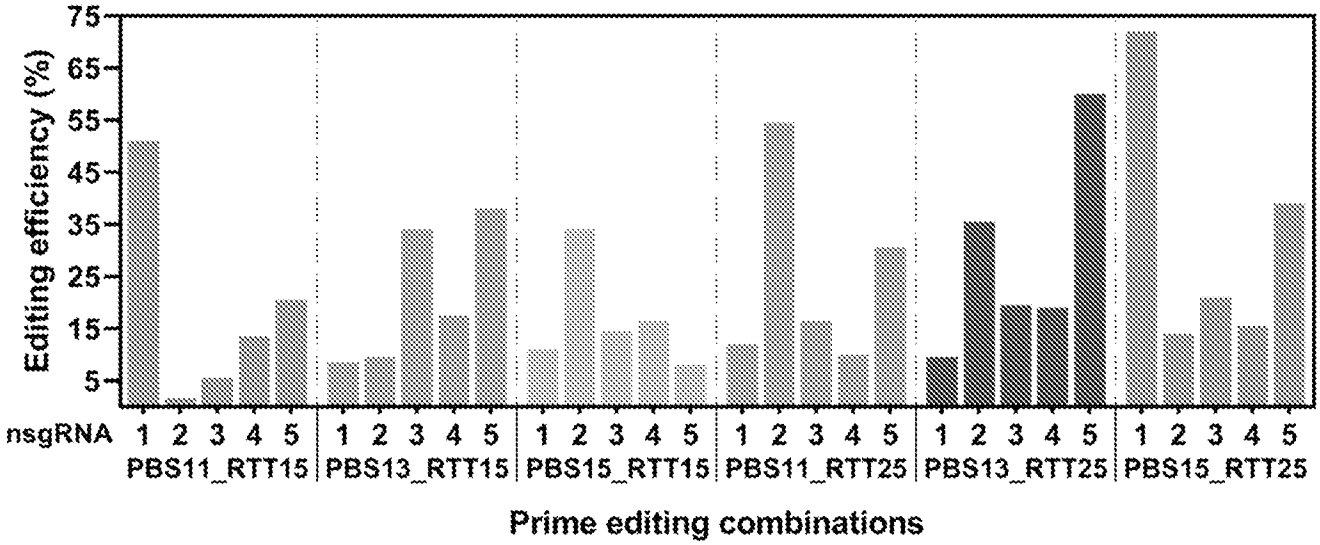

FIG. 3

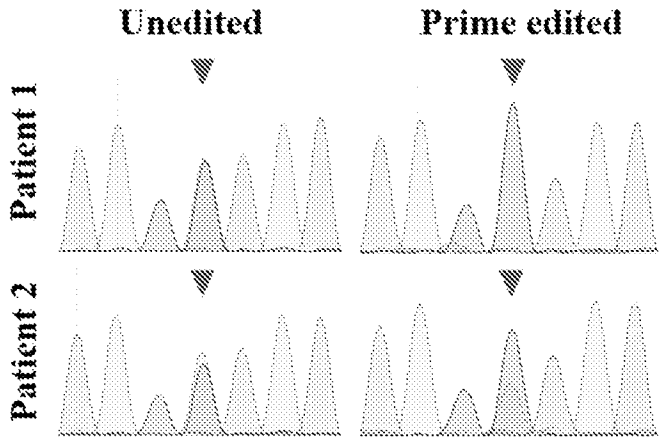

FIG. 4A

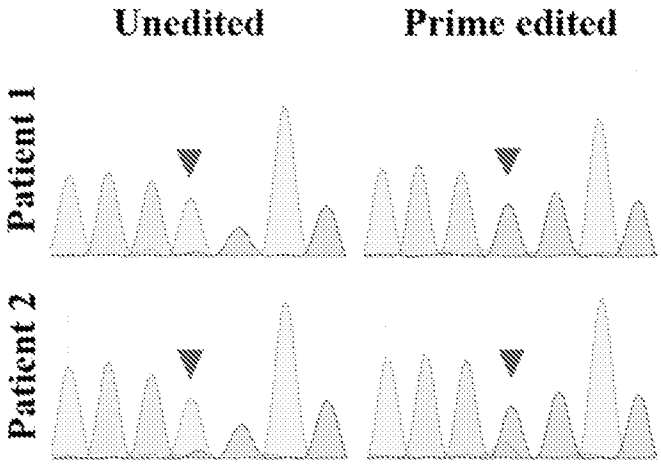

FIG. 4B

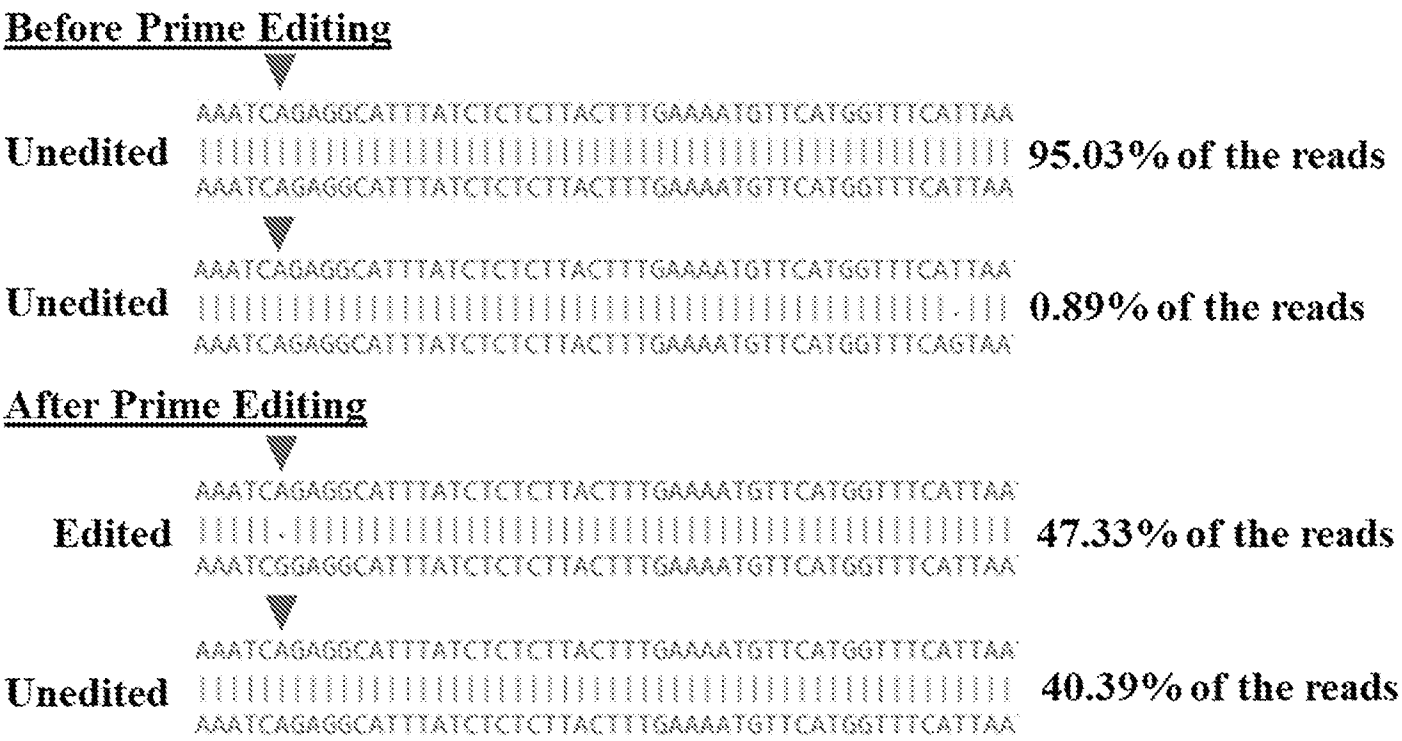
Before Prime Editing
Unedited
95.03% of the reads
Unedited
0.89% of the reads
After Prime Editing
Edited
47.33% of the reads
Unedited
40.39% of the reads

FIG. 5A

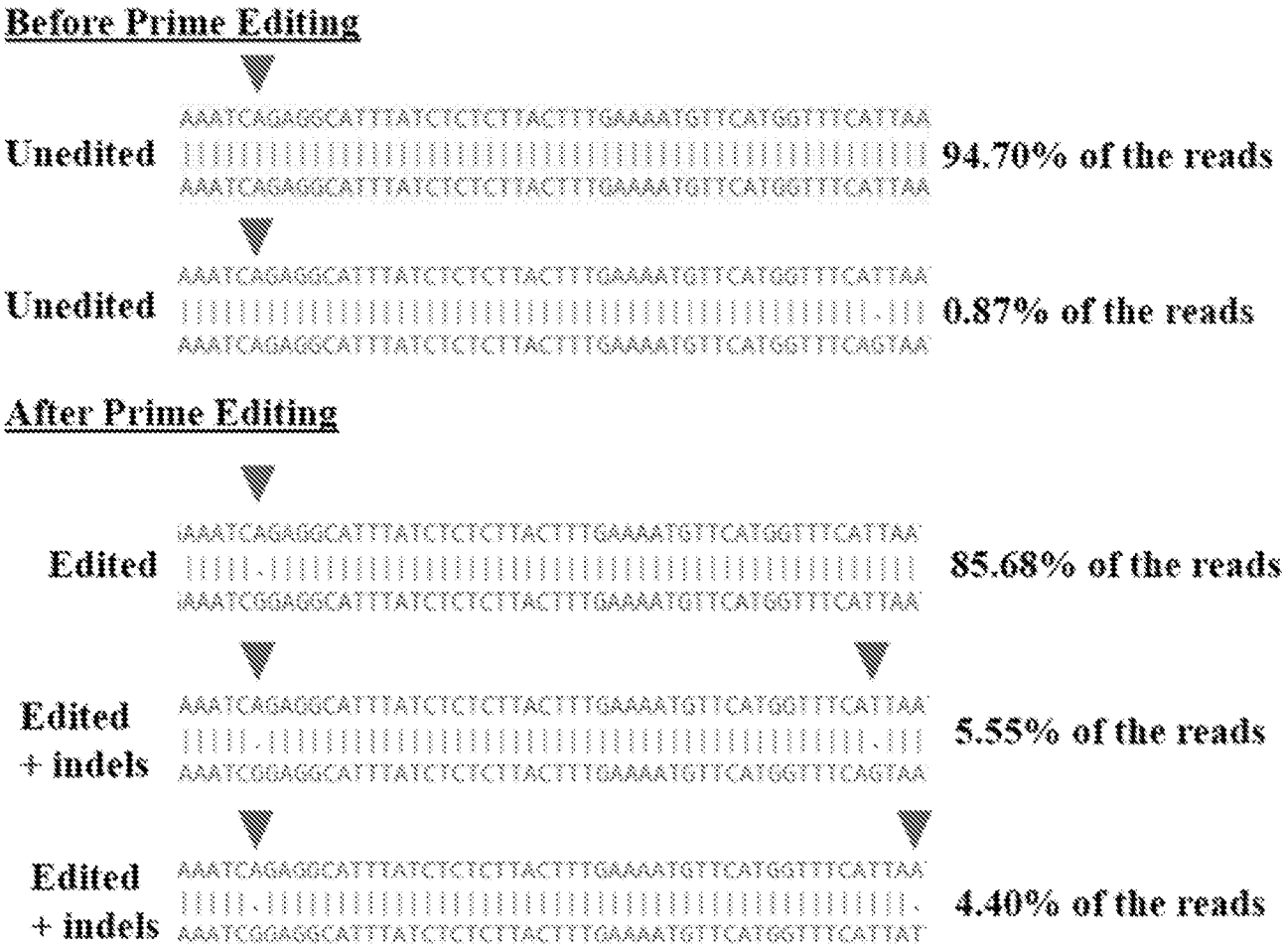
Before Prime Editing
Unedited
94.70% of the reads
Unedited
0.87% of the reads
After Prime Editing
Edited
85.68% of the reads
Edited
+ indels
5.55% of the reads
Edited
+ indels
4.40% of the reads

FIG. 5B

		cDNA Change	Protein Change	Alleles (n)	Proportion of pathogenic alleles (%)
LOVD	1	c.2843G>A	p.(Cys948Tyr)	172	13.66
	2	c.2401A>T	p.(Lys801*)	50	3.97
	3	c.2234C>T	p.(Thr745Met)	47	3.73
	4	c.2290C>T	p.(Arg764Cys)	39	3.10
	5	c.2688T>A	p.(Cys896*)	29	2.30
	6	c.613_619del	p.(Ile205Aspfs*13)	27	2.14
	7	c.498_506del	p.(Ile167_Gly169del)	22	1.75
	8	c.3307G>A	p.(Gly1103Arg)	20	1.59
	9	c.4121_4130del	p.(Ala1374Glufs*20)	17	1.35
	10	c.1148G>A	p.(Cys383Tyr)	17	1.35
	Total			440	34.95

FIG. 6

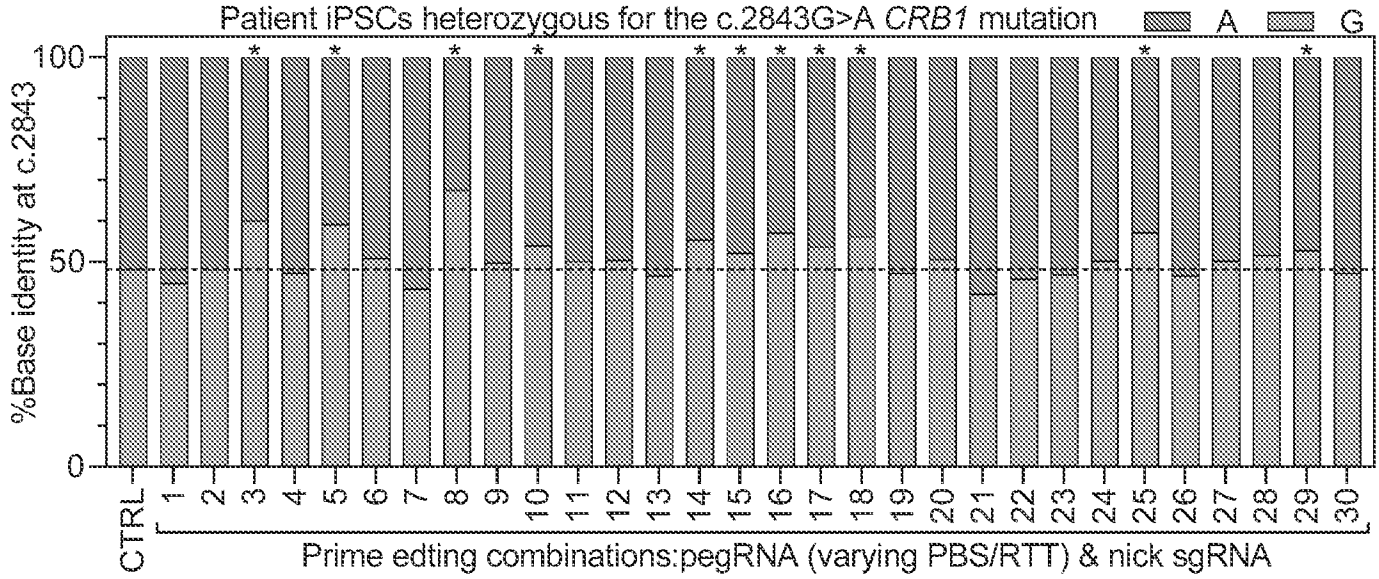

FIG. 7

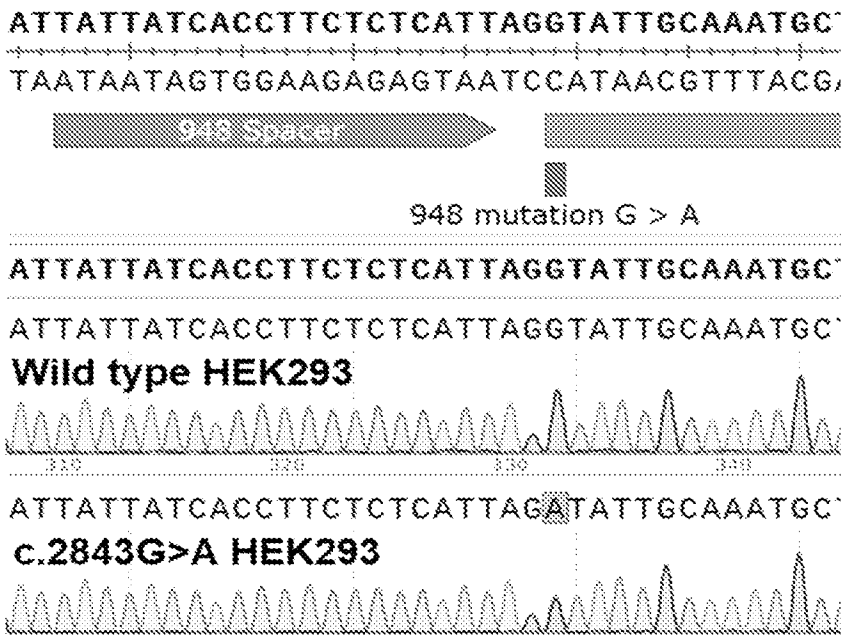

FIG. 8A

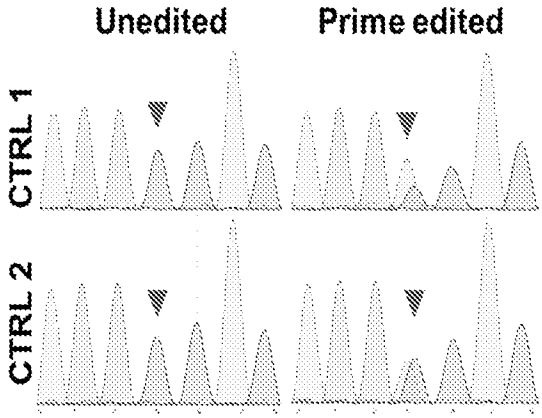

FIG. 8B

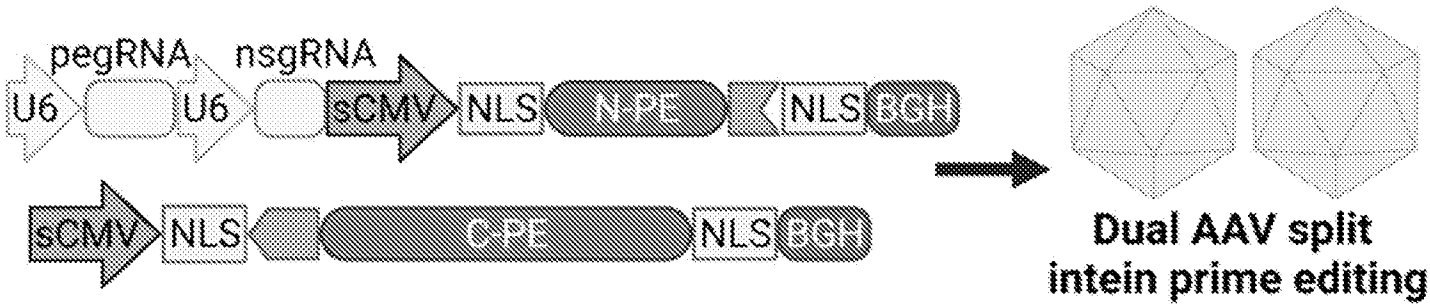

FIG. 9A

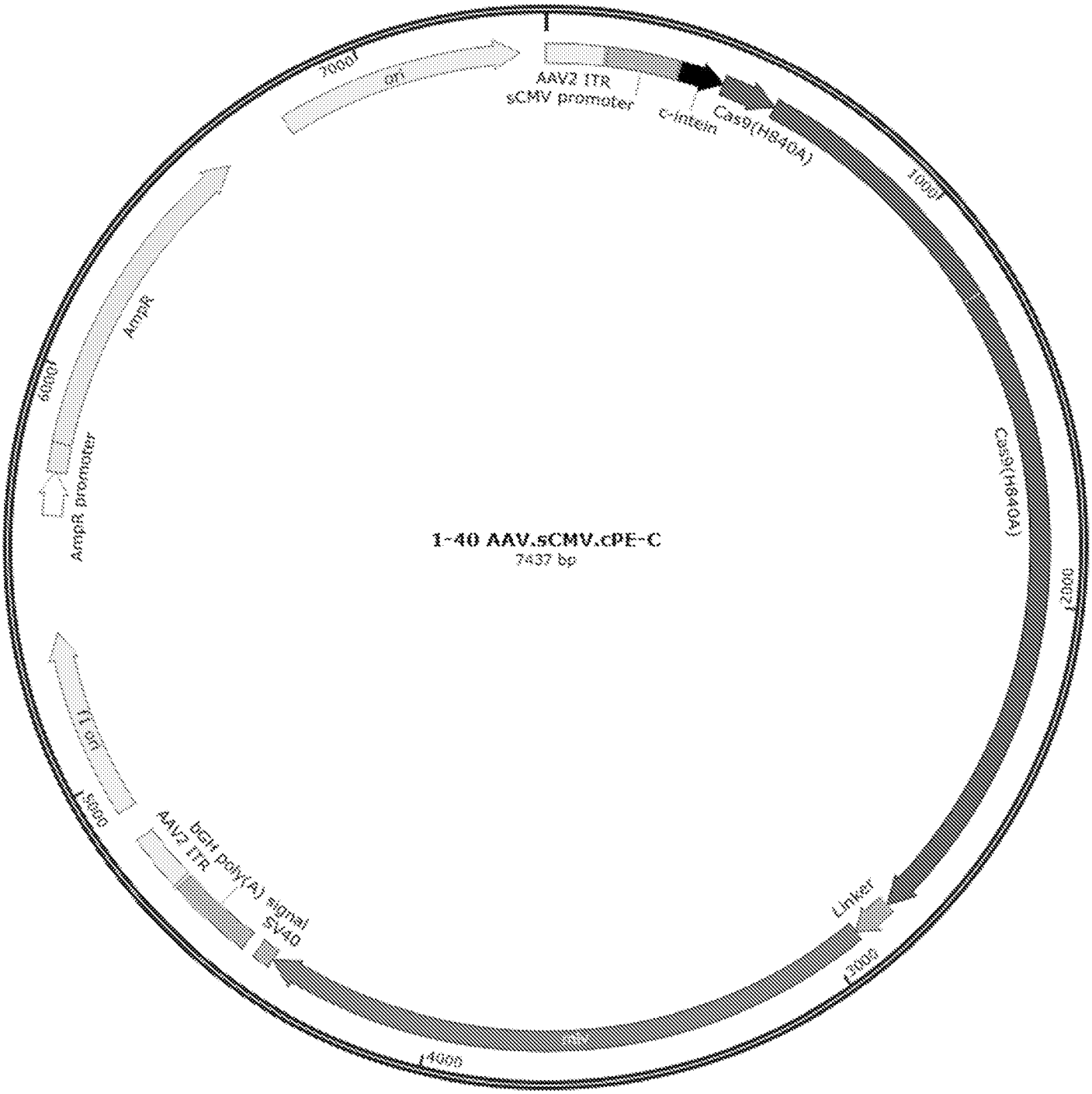
1000
AAV2 ITR
sCMV promoter
c-intein
Cas9(H840A)
2000
Cas9(H840A)
3000
Linker
3500
4000
bGH poly(A) signal
SV40
AAV2 ITR
4500
f1 ori
5000
AmpR promoter
5500
AmpR
6000
ori
7000
1-40 AAV.sCMV.cPE-C
7437 bp

FIG. 9B


1-36 AAV.sCMV.cPE-N insertion construct
5384 bp
AAV2 ITR
MSC
c-myc NLS
sCMV promoter
CMV d1 promoter
1000
CMV d1
2000
N-INTEIN
AAV2 ITR
bGH poly(A) signal
3000
f1 ori
4000
AmpR promoter
AmpR
5000
ori

FIG. 9C

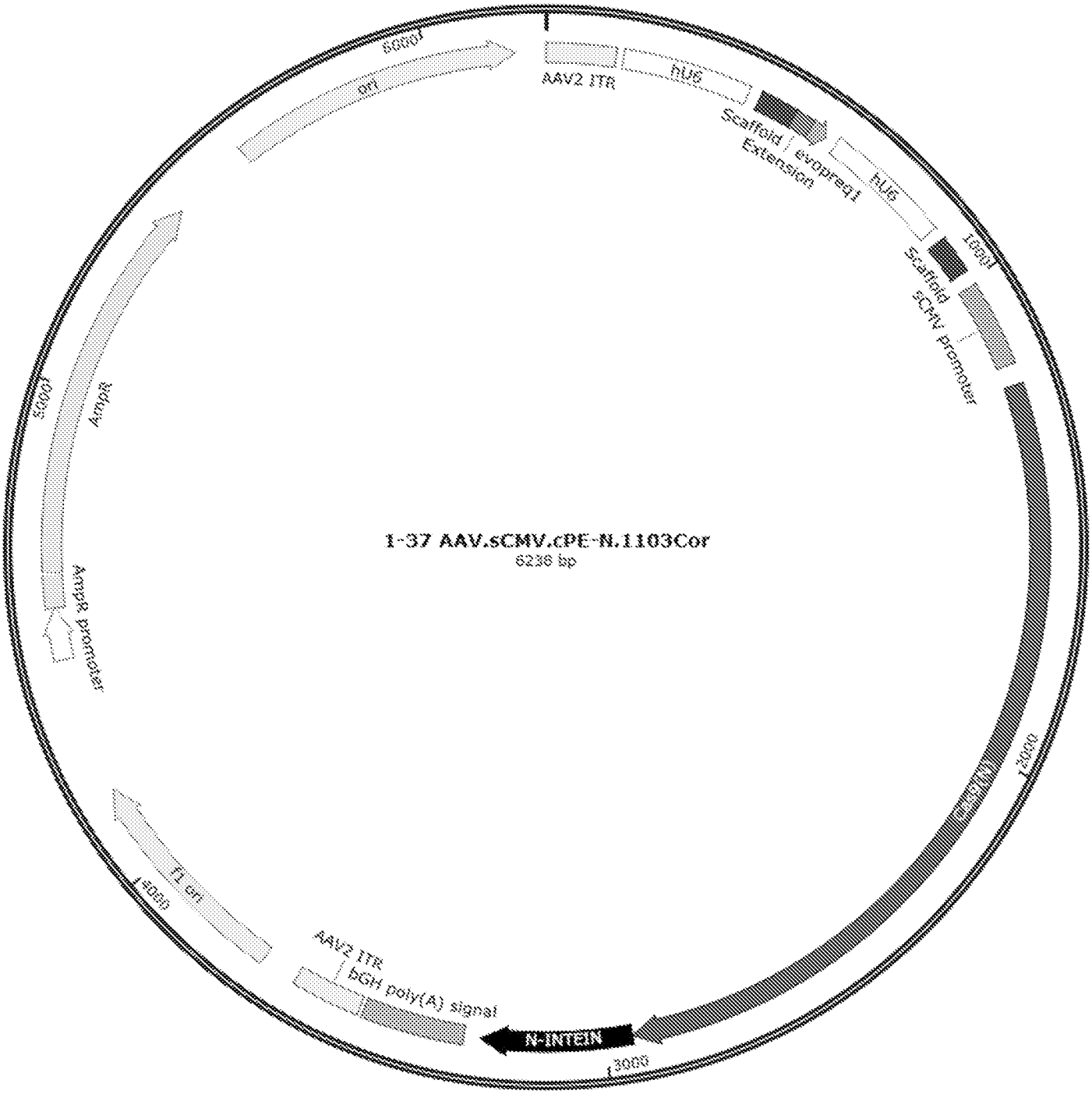
1-37 AAV.sCMV.cPE-N.1103Cor
6238 bp
6000
ori
AAV2 ITR
hU6
Scaffold / evopreq1 Extension
hU6
Scaffold
sCMV promoter
1000
3000
N-INTEIN
3000
AAV2 ITR
bGH poly(A) signal
f1 ori
5000
AmpR promoter
AmpR
5000

FIG. 9D

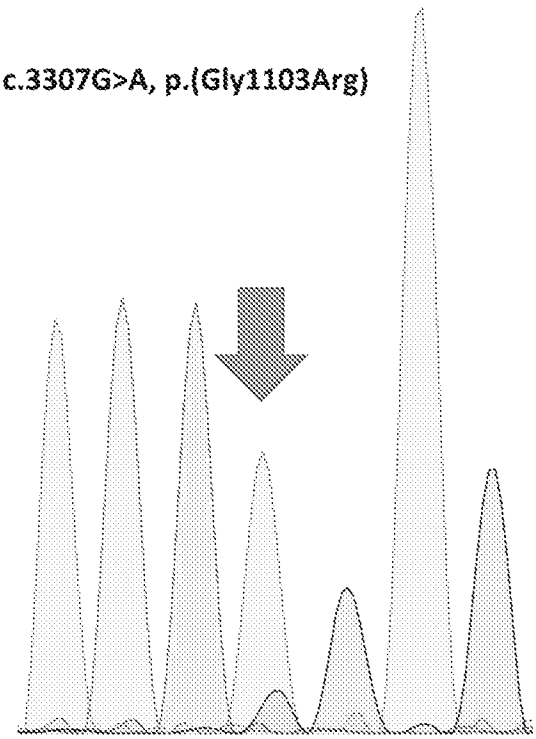

FIG. 10

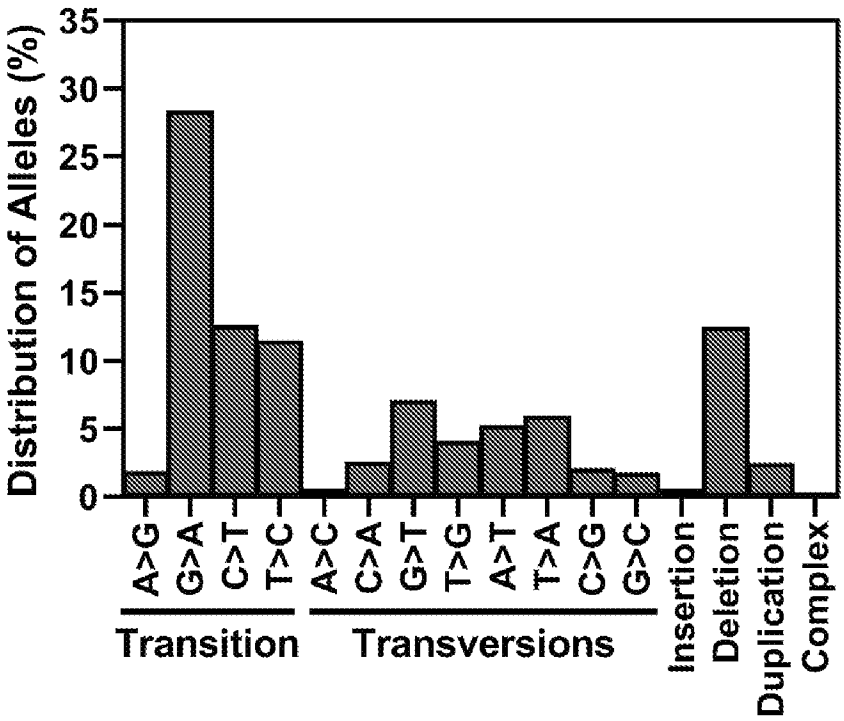

FIG. 11A

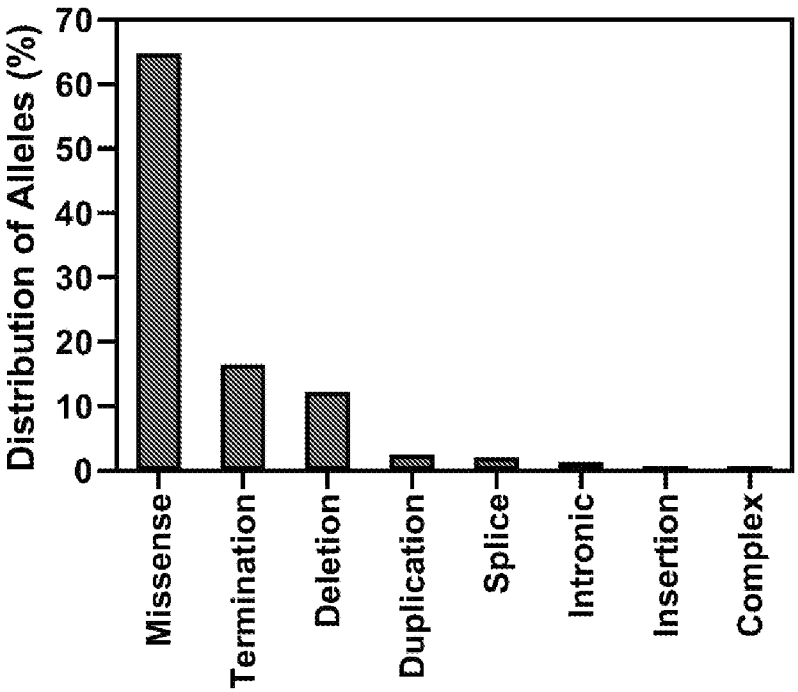
Distribution of Alleles (%)
70
60
50
40
30
20
10
0
Missense
Termination
Deletion
Duplication
Splice
Intronic
Insertion
Complex

FIG. 11B

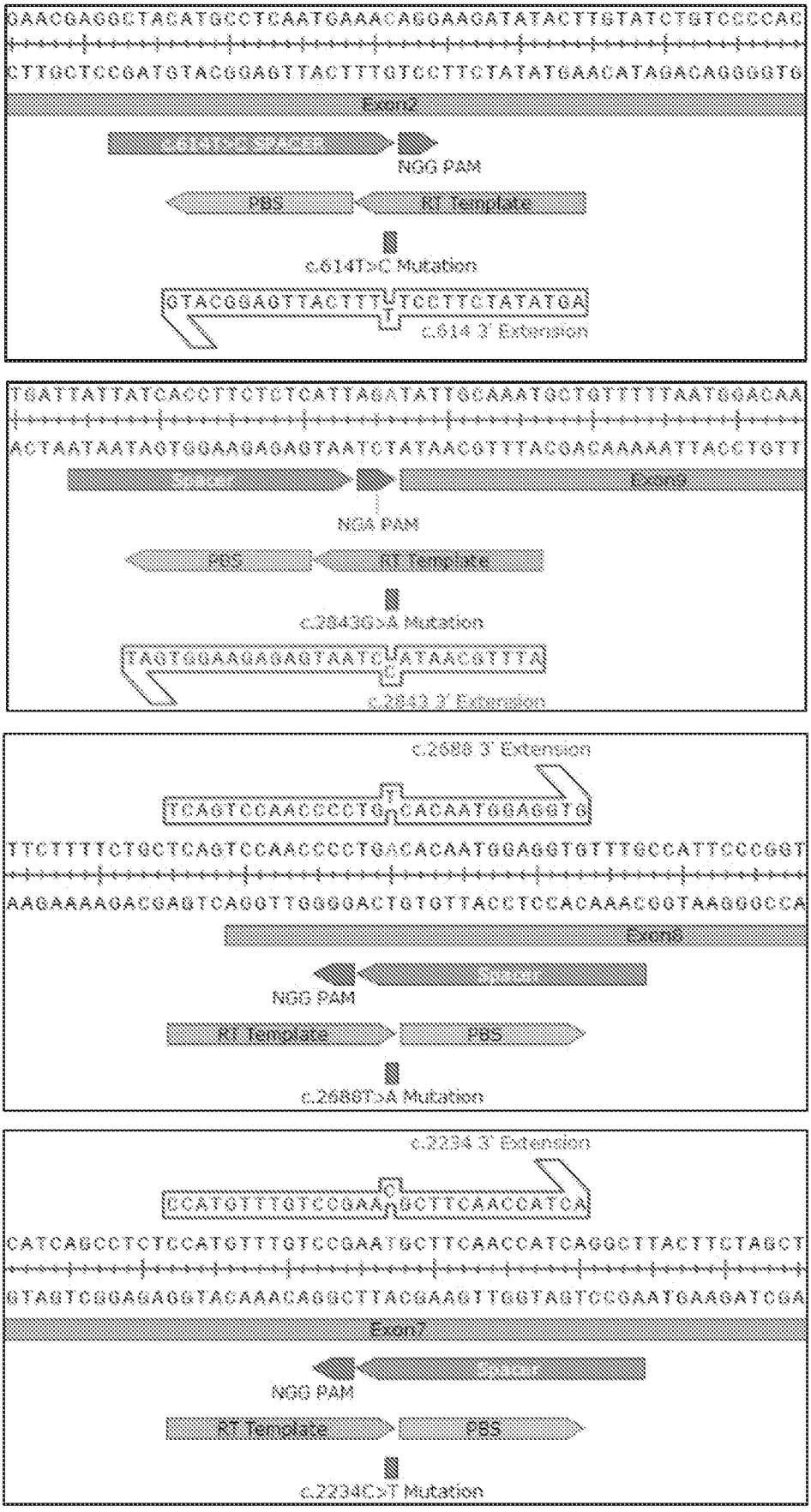
GAACGAGGCTACATGCCTCAATGAAACAGGAAGATATACTTGTATCTGTCCCCAC
CTTGCTCCGATGTACGGAGTTACTTTGTCCTTCTATATGAACATAGACAGGGGTG
Exon2
NGG PAM
PBS
RT Template
c.614T>C Mutation
GTACGGAGTTACTTT TCCTTCTATATGA
c.614 3' Extension
TGATTATTATCACCTTCTCTCATTAGATATTGCAAATGCTGTTTTTAATGGACAA
ACTAATAATAGTGGAAGAGAGTAATCTATAACGTTTACGACAAAAATTACCTGTT
Exon7
NGA PAM
PBS
RT Template
c.2843G>A Mutation
TAGTGGAAGAGAGTAATC ATAACGTTTA
c.2843 3' Extension
c.2688 3' Extension
TCAGTCCAACCCCTG CACAATGGAGGTG
TTCTTTTCTGCTCAGTCCAACCCCTGACACAATGGAGGTGTTTGCCATTCCCGGT
AAGAAAAGACGAGTCAGGTTGGGGACTGTGTTACCTCCACAAACGGTAAGGGCCA
Exon8
NGG PAM
RT Template
PBS
c.2680T>A Mutation
c.2234 3' Extension
CCATGTTTGTCCGAA GCTTCAACCATCA
CATCAGCCTCTCCATGTTTGTCCGAATGCTTCAACCATCAGGCTTACTTCTAGCT
GTAGTCGGAGAGGTACAAACAGGCTTACGAAGTTGGTAGTCCGAATGAAGATCGA
Exon7
NGG PAM
RT Template
PBS
c.2234C>T Mutation

FIG. 12

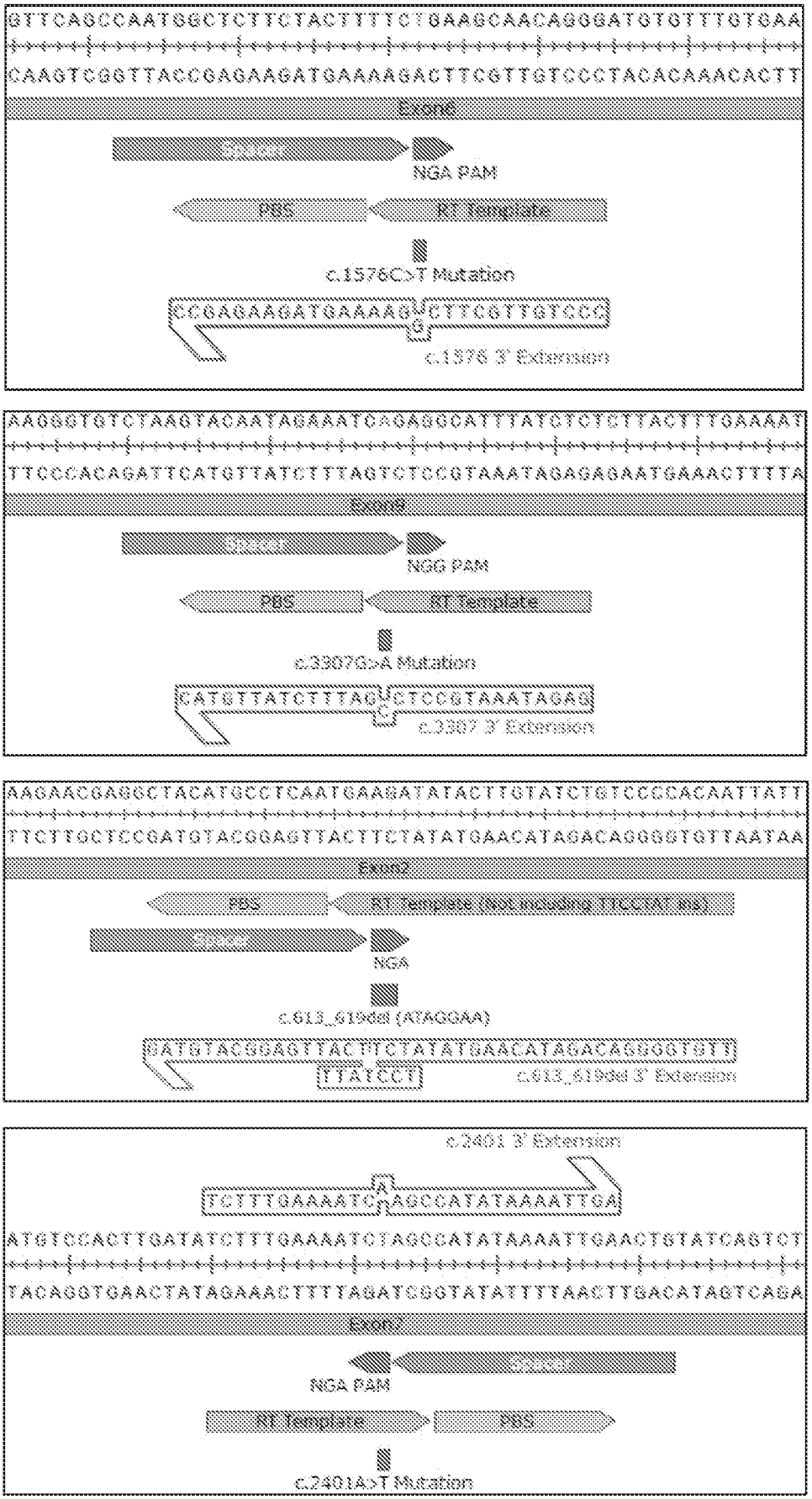
GTTCAGCCAATGGCTCTTCTACTTTTCTGAAGCAACAGGGATGTGTTTGTGAA
CAAGTCGGTTACCGAGAAGATGAAAAGACTTCGTTGTCCCTACACAAACACTT
Exon6
NGA PAM
PBS
RT Template
c.1576C>T Mutation
CCGAGAAGATGAAAAG CTTCGTTGTCCC
G
c.1576 3' Extension
AAGGGTGTCTAAGTACAATAGAAATCAGAGGCATTTATCTCTCTTACTTTGAAAAT
TTCCCACAGATTCATGTTATCTTTAGTCTCCGTAAATAGAGAGAATGAAACTTTTA
Exon9
NGG PAM
PBS
RT Template
c.3307G>A Mutation
CATGTTATCTTTAG CTCCGTAAATAGAG
C
c.3307 3' Extension
AAGAACGAGGCTACATGCCTCAATGAAGATATACTTGTATCTGTCCCCACAATTATT
TTCTTGCTCCGATGTACGGAGTTACTTCTATATGAACATAGACAGGGGTGTTAATAA
Exon2
PBS
RT Template (Not including TTCCTAT nts)
NGA
c.613_619del (ATAGGAA)
GATGTACGGAGTTACT TCTATATGAACATAGACAGGGGTGTT
TTATCCT
c.613_619del 3' Extension
c.2401 3' Extension
A
TCTTTGAAAATC AGCCATATAAAATTGA
ATGTCCACTTGATATCTTTGAAAATCTAGCCATATAAAATTGAACTGTATCAGTCT
TACAGGTGAACTATAGAAACTTTTAGATCGGTATATTTTAACTTGACATAGTCAGA
Exon7
NGA PAM
RT Template
PBS
c.2401A>T Mutation

FIG. 12, cont'd

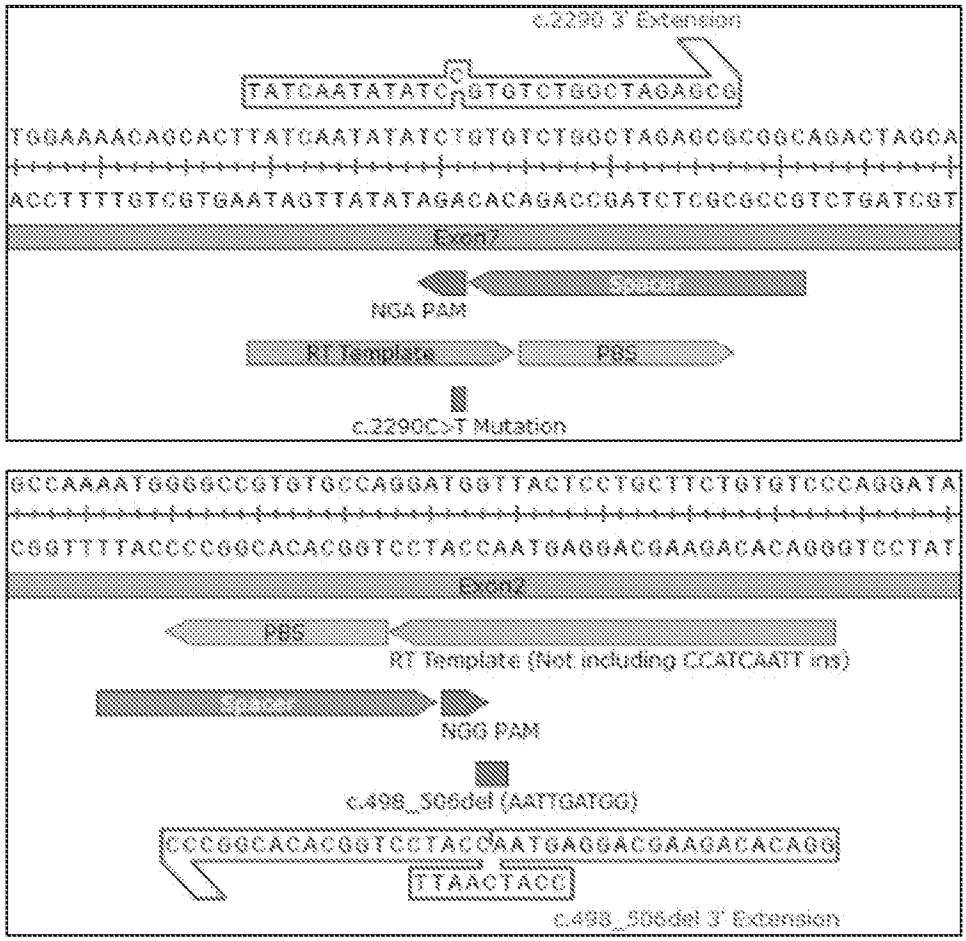
c.2290 3' Extension
TATCAATATATC GTGTCTGGCTAGAGCG
TGGAAAACAGCACTTATCAATATATCTGTGTCTGGCTAGAGCGCGGCAGACTAGCA
ACCTTTTGTCGTGAATAGTTATATAGACACAGACCGATCTCGCGCCGTCTGATCGT
Exon7
NGA PAM
RT Template
PBS
c.2290C>T Mutation
GCCAAAATGGGGCCGTGTGCCAGGATGGTTACTCCTGCTTCTGTGTCCCAGGATA
CGGTTTTACCCCGGCACACGGTCCTACCAATGAGGACGAAGACACAGGGTCCTAT
Exon7
PBS
RT Template (Not including CCATCAATT ins)
NGG PAM
c.498_506del (AATTGATGG)
CCCGGCACACGGTCCTACCAATGAGGACGAAGACACAGG
TTAACTACC
c.498_506del 3' Extension

FIG. 12, cont'd.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/77100

A. CLASSIFICATION OF SUBJECT MATTER

IPC - INV. A61P 27/02; A61K 9/00; C12N 15/113; C12N 15/86 (2022.01)
ADD.

CPC - INV. A61P 27/02; A61K 9/0048; C12N 15/113; C12N 15/86
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

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

Electronic database consulted during the international search (name of database and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2020/191249 A1 (THE BROAD INSTITUTE, INC.) 24 September 2020; Title; abstract; paragraphs [0009], [0045]-[0046], [0077]; claim 1; claim 11	1-3
Y	US 2016/0194374 A1 (ACADEMISCH ZIEKENHUIS LEIDEN H.O.D.N. LEIDS UNIVERSITAIR MEDISCH CENTRUM) 07 July 2016; Abstract; paragraph [0007]	1-3
Y	WO 2014/165753 A1 (THE WISTAR INSTITUTE OF ANATOMY AND BIOLOGY) 09 October 2014; Page 16	3
Y	W. JIANG ET AL. "Artesunate has its enhancement on antibacterial activity" 339–345. The Journal of Antibiotics (2013) 66. Web. 03 April 2013; Page 344, second column, third paragraph; DOI: 10.1038/ja.2013.22	3
Y	US 2013/0011881 A1 (LEAMON JOHN) 10 January 2013; SEQ ID 32,492	3
Y	WO 2014/140051 A1 (FONDAZIONE TELETHON) 18 September 2014; SEQ ID: 160	3

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 November 2022 (19.11.2022)	FEB 08 2023
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Shane Thomas Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/77100

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

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

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

   b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)),

      ☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/77100

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 4-15
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

This International Searching Authority found multiple inventions in this international application, as follows:
-***-Please See Supplemental Page-***-

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-3, SEQ ID NO: 1 (spacer sequence), SEQ ID NO: 44 (extension sequence), SEQ ID NO: 4 (primer binding sequence), SEQ ID NO: 2 (RTT sequence);

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/77100

-***-Continued From Box No. III: Observations where unity of invention is lacking-***-

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+, claims 1-3, SEQ ID NO: 1 (spacer sequence), SEQ ID NO: 44 (extension sequence), SEQ ID NO: 4 (primer binding sequence), SEQ ID NO: 2 (RTT sequence), are directed to systems for modifying the crumbs homologue-1 (CRB 1) gene.

The systems of Claims 1-2 and 3 (in-part) are believed to encompass the first named invention of Groups I+ and are the claims that will be searched to the extent that they comprise a RNA polynucleotide encompassing SEQ ID NO: 1 (first exemplary spacer sequence), SEQ ID NO: 44 (first exemplary extension sequence), SEQ ID NO: 4 (first exemplary primer binding sequence), and SEQ ID NO: 2 (first exemplary RTT sequence). This first named invention of Group I+ has been selected to encompass the first species of each of the genera found in claim 3 based on the guidance set forth in section 10.54 of the PCT International Search and Preliminary Examination Guidelines.

Applicant is invited to elect additional spacer, extension, PBS, and RTT sequences, with specified SEQ ID NO: for each, or with specified substitution(s) at specified site(s) of a SEQ ID NO:, such that the sequence of each elected species is fully specified (i.e. no optional or variable residues or substituents), and where available as an option within at least one searchable claim, to be searched. Additional sequence(s) will be searched upon the payment of additional fees. Applicants must specify the searchable claims that encompass any additionally elected sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election would be SEQ ID NO:17 (spacer sequence), SEQ ID NO: 45 (extension sequence), SEQ ID NO: 5 (primer binding sequence), and SEQ ID NO: 3 (RTT sequence).

No technical features are shared between the RNA polynucleotide sequences of Groups I+ and, accordingly, these groups lack unity a priori.