

ABCA4 TRANS-SPLICING MOLECULES**CROSS REFERENCE**

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 63/341,665, filed on May 13, 2022 and U.S. Provisional Patent Application No. 63/478,472, filed on January 4, 2023, each of which is incorporated herein by reference in its entirety.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been filed electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on May 10, 2023, is named 61313-703_601_SL.xml and is 1,315,456 bytes in size.

FIELD OF THE INVENTION

[0003] Compositions, methods, and uses involving trans-splicing molecules for correcting mutations in ABCA4 (ATP-binding cassette, Subfamily A, Member 4) are described herein.

BACKGROUND

[0004] Recessively inherited mutations in ABCA4 are causal in the development of progressive forms of blindness including Stargardt disease and cone-rod dystrophy. Localized along the rim region of photoreceptor cell outer segment disc membranes, ABCA4 functions as a transporter (or “flippase”) of N-retinylidene-phosphatidylethanolamine (N-Ret-PE) from the lumen to the cytoplasmic leaflet of disc membranes. Defective ABCA4 protein leads to the accumulation of toxic retinoid compounds and ultimately to photoreceptor and retinal pigment epithelial cell death. The 6882-bp coding sequence of ABCA4 is too large to be delivered in its entirety by a single AAV vector. Moreover, with hundreds of disease-causing mutations found throughout the gene, a single base-editing approach would not address a significant number of patients. Therefore, there is a need in the field for alternative strategies for correcting mutations in the ABCA4 gene.

SUMMARY

[0005] ABCA4-related retinopathies, such as Stargardt disease, are severe rare genetic retinal diseases affecting both adults and children, caused by biallelic pathogenic mutations in the ABCA4 gene. The estimated prevalence of Stargardt disease is 1 in 8–10,000 individuals worldwide and accounts for roughly 7% of retinal dystrophies. ABCA4 was characterized as the gene responsible for autosomal recessive Stargardt disease, however, several other retinal phenotypes are also associated with mutations in ABCA4. Indeed, ABCA4-related retinopathies are considered collectively to be the most frequent cause of retinal degeneration due to Mendelian inheritance. Individuals with ABCA4-related retinopathies are typically diagnosed in the first two decades of life and depending on the severity of specific mutations may completely lose their central field of vision by adulthood.

[0006] Presently, there are no approved treatments indicated for *ABCA4*-related retinopathies. Therefore, there is a significant unmet medical need to develop disease modifying therapies for both adults and children impacted by loss-of-function *ABCA4* mutations.

[0007] Through extensive research, the present inventors have identified nucleic acid elements that confer upon nucleic acid trans-splicing molecules in which they are incorporated properties that address technical challenges associated with, for example, increasing trans-splicing efficiency and the level of trans-spliced protein produced thereby. Inclusion of the identified elements in nucleic acid trans-splicing molecules (e.g., RNA trans-splicing molecules), therefore, presents a variety of solutions to the problems associated with implementation of nucleic acid trans-splicing molecules in the past, including problems relating to poor trans-splicing efficiency. Accordingly, the present inventors have designed nucleic acid trans-splicing molecules that achieve the long-sought objective of providing options for therapeutic intervention in diseases associated with genetic mutations, whereby correction of those genetic mutations at the RNA level could provide enormous benefit to patients. Many patients with diseases associated with genetic mutations have few, if any, options for treatment and therefore, have largely unmet needs. Patients with progressive forms of blindness associated with mutations in *ABCA4*, for example, represent a patient population having unmet needs.

[0008] To address these unmet needs, a trans-splicing solution, also referred to as exon editing, is described herein which comprises delivering a therapeutic nucleic acid trans-splicing construct that encodes an RNA exon editor that is designed to trans-splice into endogenous *ABCA4* pre-mRNA that comprises at least one mutation, whereby trans-splicing replaces exons in the endogenous *ABCA4* pre-mRNA (at least one exon of which comprises at least one mutation) with exons from the therapeutic RNA exon editor that encode functional (e.g., wildtype) amino acid sequences and thus, corrects the at least one mutation in the endogenous *ABCA4* pre-mRNA. In some embodiments, RNA exon editors may be delivered directly to cells. The terms “nucleic acid trans-splicing molecule”, “RNA exon editing molecule”, “RNA exon editor”, “Exon Editor”, and “RNA trans-splicing construct” are used interchangeably herein.

[0009] In some embodiments, a nucleic acid trans-splicing molecule is described herein comprising: (a) a sequence (CDS) comprising *ABCA4* exons; (b) a linker domain comprising SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27; and (c) a binding domain that anneals to a binding site within endogenous *ABCA4* pre-mRNA, wherein the CDS, the linker domain, and the binding domain are operatively linked in a 5'-to-3' direction. Use of same for treating retinopathies associated with mutations in *ABCA4* and in the preparation of medicaments for treating retinopathies associated with mutations in *ABCA4*, as well as methods for treating retinopathies associated with mutations in *ABCA4* are also described herein.

[0010] In some embodiments, a nucleic acid trans-splicing molecule is described herein comprising: (a) a CDS comprising, consisting essentially of, or consisting of a variant of any one of SEQ ID NOs: 53-55, wherein the variant of any one of SEQ ID NOs: 53-55 comprises at least one

nucleotide variation in at least one cryptic splice site listed in Table 3, wherein the at least one nucleotide variation reduces cryptic splice site use at each of the at least one cryptic splice sites comprising at least one nucleotide variation; and (b) optionally, a linker domain; and (c) optionally, a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the CDS; the linker domain, when present; and the binding domain, when present; are operatively linked in a 5'-to-3' direction. In some embodiments, the variant of SEQ ID NO: 55 comprises any one of SEQ ID NOs: 56-59. In some embodiments, the variant of SEQ ID NO: 55 comprises SEQ ID NO: 56. Use of same for treating retinopathies associated with mutations in *ABCA4* and in the preparation of medicaments for treating retinopathies associated with mutations in *ABCA4*, as well as methods for treating retinopathies associated with mutations in *ABCA4* are also described herein.

[0011] In one aspect, provided herein is a nucleic acid trans-splicing molecule (or a vector containing or encoding a nucleic acid trans-splicing molecule), wherein the nucleic acid trans-splicing molecule contains, operatively linked in a 5'-to-3' direction: (a) a 5' regulatory domain comprising, e.g., a native 5' ABCA4 untranslated region; (b) a CDS comprising a functional sequence of 5' ABCA4 exons; (c) a splicing domain configured to mediate trans-splicing; and (d) a binding domain configured to bind a binding site within endogenous ABCA4 pre-mRNA, wherein the nucleic acid trans-splicing molecule is configured to trans-splice the CDS to endogenous ABCA4, thereby replacing an endogenous 5' ABCA4 exon sequence with the functional sequence of 5' ABCA4 exons. In some embodiments, the 5' regulatory domain further includes a constitutive promoter, such as, e.g., a CMV/CMV promoter (CMV enhancer and promoter), or a variant thereof. In some embodiments, the native 5' ABCA4 untranslated region comprises a nucleic acid sequence having at least 80% sequence identity to SEQ ID NO: 12 or 13 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to SEQ ID NO: 12 or 13. In some embodiments, the 5' regulatory domain comprises a nucleic acid sequence having at least 80% sequence identity to SEQ ID NO: 14 or 15 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to SEQ ID NO: 14 or 15.

[0012] In some embodiments, the binding site is within ABCA4 intron 22. In some embodiments, the binding site comprises any one or more of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16 (e.g., any six or more, any eight or more, any 10 or more, any 12 or more, any 18 or more, any 20 or more, any 24 or more, any 50 or more, any 100 or more, or any 150 or more (e.g., 150)) of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16. In some embodiments, the binding site comprises any one or more of nucleotides 1160 to 1309 of SEQ ID NO: 16 (e.g., any six or more,

any eight or more, any 10 or more, any 12 or more, any 18 or more, any 20 or more, any 24 or more, any 50 or more, any 100 or more, or any 150 or more (e.g., 150)) of nucleotides 1160 to 1309 of SEQ ID NO: 16. In some embodiments, a binding domain that binds to a binding site has least 80% sequence identity to a binding site within SEQ ID NO: 16 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity to a binding site within SEQ ID NO: 16. In some embodiments, the binding domain comprises SEQ ID NO: 17 or 18.

[0013] In some embodiments, the splicing domain comprises GTAAGT, GUAAGT, GTAAGG, GUAAGG, GTAAGC, GUAAGC, GTAAGT, or GUAACU.

[0014] In some embodiments, the nucleic acid trans-splicing molecule further comprises a linker domain. In some embodiments, the linker domain is longer than 25 nucleotides in length (e.g., between 25 and 50 nucleotides in length, between 35 and 45 nucleotides in length, or about 40 nucleotides in length (e.g., a 40-mer linker)). In some instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with

CTGGTGCCCGCGGGCCGCGGAACCGGTTGGGGGCATGTAC

(SEQ ID NO: 27). In other instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CCGAATACGACACGTACAAGATCT (SEQ ID NO: 29). In other instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with GCTGCTGCTCAGTCTCCTGGGCTGG (SEQ ID NO: 31).

[0015] In another aspect, provided is a nucleic acid trans-splicing molecule comprising, operatively linked in a 5'-to-3' direction: (a) a CDS comprising a functional sequence of 5' ABCA4 exons; (b) a splicing domain configured to mediate trans-splicing, wherein the splicing domain comprises GTAAGT, GUAAGT, GTAAGG, GUAAGG, GTAAGC, GUAAGC, GTAAGT, or GUAACU; and (c) a binding domain configured to bind a binding site within endogenous ABCA4 intron 22, wherein

the nucleic acid trans-splicing molecule is configured to trans-splice the CDS to endogenous ABCA4 exon 23, thereby replacing an endogenous 5' ABCA4 exon sequence with the functional sequence of 5' ABCA4 exons. In some embodiments, the nucleic acid trans-splicing molecule comprises a 5' regulatory domain operatively linked 5' to the CDS. In some embodiments, the 5' regulatory domain comprises a native 5' ABCA4 untranslated region, e.g., a native 5' ABCA4 untranslated region having at least 80% sequence identity to SEQ ID NO: 12 or 13 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to SEQ ID NO: 12 or 13. In some embodiments, the 5' regulatory domain further comprises a constitutive promoter, e.g., a CMV promoter.

[0016] In some embodiments, the nucleic acid trans-splicing molecule further comprises a linker domain. In some embodiments, the linker domain is longer than 25 nucleotides in length (e.g., between 25 and 50 nucleotides in length, between 35 and 45 nucleotides in length, or about 40 nucleotides in length (e.g., a 40-mer linker)). In some instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CTGGTGCCCGCGGGCCGCGGAACCGGTTGGGGGCATGTAC (SEQ ID NO: 27). In other instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CCGAATACGACACGTACAAGATCT (SEQ ID NO: 29). In other instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with GCTGCTGCTCAGTCTCCTGGGCTGG (SEQ ID NO: 31).

[0017] In some embodiments, the binding site comprises any one or more of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16 (e.g., any six or more, any eight or more, any 10 or more, any 12 or more, any 18 or more, any 20 or more, any 24 or more, any 50 or more, any 100 or more, or any 150 or more (e.g., 150)) of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16. In some embodiments, the binding site comprises any one or more of nucleotides 1160 to 1309 of SEQ ID NO:

16 (e.g., any six or more, any eight or more, any 10 or more, any 12 or more, any 18 or more, any 20 or more, any 24 or more, any 50 or more, any 100 or more, or any 150 or more (e.g., 150)) of nucleotides 1160 to 1309 of SEQ ID NO: 16. In some embodiments, the binding domain that binds to a binding site has least 80% sequence identity to SEQ ID NO: 17 or 18 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity to SEQ ID NO: 17 or 18). In some embodiments, the binding domain comprises SEQ ID NO: 17 or 18.

[0018] In another aspect, provided is a nucleic acid trans-splicing molecule comprising, operatively linked in a 5'-to-3' direction: (a) a CDS comprising a functional sequence of 5' ABCA4 exons; (b) a splicing domain configured to mediate trans-splicing; (c) a linker domain between 25 and 50 nucleotides in length; and (d) a binding domain configured to bind a binding site within endogenous ABCA4 intron 22, wherein the nucleic acid trans-splicing molecule is configured to trans-splice the CDS to endogenous ABCA4 exon 23, thereby replacing an endogenous 5' ABCA4 exon sequence with the functional sequence of 5' ABCA4 exons. In some embodiments, the linker domain is between 35 and 45 nucleotides in length, or about 40 nucleotides in length (e.g., a 40-mer linker). In some instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CTGGTGCCCGCGGCCGCGGAACCGGTTGGGGGCATGTAC (SEQ ID NO: 27). In other instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CCGAATACGACACGTACAAGATCT (SEQ ID NO: 29). In other instances, the linker domain is, or comprises, a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with GCTGCTGCTCAGTCTCCTGGGCTGG (SEQ ID NO: 31).

[0019] In some embodiments, the nucleic acid trans-splicing molecule comprises a 5' regulatory domain operatively linked 5' to the CDS. In some embodiments, the 5' regulatory domain comprises a native 5' ABCA4 untranslated region, e.g., a native 5' ABCA4 untranslated region having

at least 80% sequence identity to SEQ ID NO: 12 or 13 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to SEQ ID NO: 12 or 13. In some embodiments, the 5' regulatory domain further comprises a constitutive promoter, e.g., a CMV/CMV promoter.

[0020] In some embodiments, the splicing domain comprises GTAAGT, GUAAGT, GTAAGG, GUAAGG, GTAAGC, GUAAGC, GTAAC, or GUAACU.

[0021] In some embodiments, the binding site comprises any one or more of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16 (e.g., any six or more, any eight or more, any 10 or more, any 12 or more, any 18 or more, any 20 or more, any 24 or more, any 50 or more, any 100 or more, or any 150 or more (e.g., 150)) of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16. In some embodiments, the binding site comprises any one or more of nucleotides 1160 to 1309 of SEQ ID NO: 16 (e.g., any six or more, any eight or more, any 10 or more, any 12 or more, any 18 or more, any 20 or more, any 24 or more, any 50 or more, any 100 or more, or any 150 or more (e.g., 150)) of nucleotides 1160 to 1309 of SEQ ID NO: 16. In some embodiments, the binding domain that binds to a binding site has least 80% sequence identity to SEQ ID NO: 17 or 18 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity to SEQ ID NO: 17 or 18). In some embodiments, the binding domain comprises SEQ ID NO: 17 or 18.

[0022] In some embodiments of any of the previous aspects, the nucleic acid trans-splicing molecule comprises a 3' transcription terminator domain. In some embodiments, the 3' transcription terminator domain forms a triple helical structure that effectively caps the 3' end of the trans-splicing molecule. In some embodiments, the 3' transcriptional terminator domain comprises a wildtype MALAT1 sequence. In some embodiments, the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 80% sequence identity to SEQ ID NO: 32 or 33 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity to SEQ ID NO: 32 or 33).

[0023] In some embodiments of any of the preceding aspects, the CDS comprises ABCA4 exons 1-22 or variants thereof (e.g., codon-optimized variants thereof).

[0024] In some embodiments of any of the preceding aspects, the nucleic acid trans-splicing molecule comprises a nucleic acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 40-50 (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91%

sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity to any one of SEQ ID NOs: 40-50).

[0025] In some embodiments of any of the preceding aspects, the endogenous ABCA4 pre-mRNA has at least one mutation associated with an ABCA4-associated retinal dystrophy (e.g., a recessively inherited ABCA4-associated retinal dystrophy, e.g., Stargardt disease or cone-rod dystrophy). In some embodiments, the mutation is expressed in a photoreceptor cell and/or a retinal pigment epithelial cell.

[0026] In another aspect, a nucleic acid trans-splicing molecule is described comprising, operatively linked in a 5'-to-3' direction: (a) a CDS comprising a functional sequence of 5' ABCA4 exons; (b) a splicing domain configured to mediate trans-splicing; and (c) a binding domain having at least 80% sequence identity (e.g., at least 85% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to any one of SEQ ID NOs: 17-24, wherein the nucleic acid trans-splicing molecule is configured to trans-splice the CDS to endogenous ABCA4 exon 23, thereby replacing an endogenous 5' ABCA4 exon sequence with the functional sequence of 5' ABCA4 exons.

[0027] In another aspect, provided herein is an isolated DNA molecule, or vector thereof, comprising SEQ ID NO: 18.

[0028] In another aspect, provided herein is an isolated RNA molecule, or vector thereof, comprising SEQ ID NO: 17.

[0029] In another aspect, provided herein is an isolated nucleic acid (e.g., DNA or RNA) encoding the nucleic acid trans-splicing molecule of any of the preceding embodiments or any of the preceding aspects, e.g., a vector encoding the nucleic acid trans-splicing molecule of any of the preceding embodiments or any of the preceding aspects.

[0030] In another aspect, provided herein is a proviral plasmid comprising the nucleic acid trans-splicing molecule of any of the preceding embodiments or any of the preceding aspects.

[0031] In another aspect, provided herein is an adeno-associated virus (AAV) comprising a sequence encoding the nucleic acid trans-splicing molecule of any of the preceding embodiments or any of the preceding aspects. In some embodiments, the AAV preferentially targets a photoreceptor cell and/or a retinal pigment epithelial cell. In some embodiments, the AAV is AAV8, AAV5, or AAV2.

[0032] In another aspect, provided herein is a composition comprising the nucleic acid trans-splicing molecule, the vector, the proviral plasmid, or the AAV of any of the preceding embodiments or any of the preceding aspects. In some embodiments, the composition includes a pharmaceutically acceptable excipient.

[0033] In another aspect, provided herein is a method of expressing functional ABCA4 in a target cell (e.g., *in vitro* or *in vivo*). In some embodiments, the method includes transducing the target cell with the nucleic acid trans-splicing molecule, the vector, the AAV, or the composition of any one of the previous aspects. In some embodiments, the method replaces 20% or more of the target ABCA4 mRNA in the target cell (e.g., in cultured human cells, e.g., as measured by trans-spliced ABCA4 RNA or corresponding protein resulting from translation of the trans-spliced RNA). In some embodiments, the method replaces 38% or more of the target ABCA4 mRNA in the target cell (e.g., in cultured human cells e.g., as measured by trans-spliced ABCA4 RNA or corresponding protein resulting from translation of the trans-spliced RNA). In some embodiments, the functional ABCA4 (corresponding protein resulting from translation of the trans-spliced RNA) is full-length ABCA4 protein.

[0034] In another aspect, provided herein is a method of reducing ABCA4 retinopathy-associated lipofuscin and/or A2E accumulation in a subject (e.g., a mammal, e.g., a primate, e.g., a human). In some embodiments, the method includes transducing a target retinal cell in the subject with the nucleic acid trans-splicing molecule, the vector, the AAV, or the composition of any of the preceding embodiments or any of the preceding aspects. In some embodiments, the subject is a non-human primate or a human (e.g., a human who has been diagnosed with an ABCA4-associated retinal dystrophy). In some embodiments, the nucleic acid trans-splicing molecule, the vector, or the AAV is administered to the subject by ocular administration (e.g., by subretinal administration).

[0035] In another aspect, provided herein is a method of correcting at least one mutation in a 5' ABCA4 exon sequence in a target cell of a subject (e.g., a mammal, e.g., a primate, e.g., a human). In some embodiments, the method includes administering to the subject the nucleic acid trans-splicing molecule, the vector, the composition, or the AAV of any of the preceding embodiments or any of the preceding aspects. In some embodiments, the subject is a non-human primate or a human (e.g., a human who has been diagnosed with an ABCA4-associated retinal dystrophy). In some embodiments, the nucleic acid trans-splicing molecule, the vector, or the AAV is administered to the subject by ocular administration (e.g., by subretinal administration).

[0036] In another aspect, provided herein is a method of treating an ABCA4-associated retinal dystrophy in a subject (e.g., a mammal, e.g., a primate, e.g., a human). In some embodiments, the method includes administering to the subject the nucleic acid trans-splicing molecule, the vector, the AAV, or the composition of any of the preceding embodiments or any of the preceding aspects in a therapeutically effective amount. In some embodiments, the ABCA4-associated retinal dystrophy is associated with a mutation in a 5' ABCA4 exon sequence (e.g., a mutation in any one or more of ABCA4 exons 1-22). In some embodiments, the ABCA4-associated retinal dystrophy is Stargardt disease or cone-rod dystrophy.

BRIEF DESCRIPTION OF THE DRAWINGS

[0037] FIG. 1 is a graph showing trans-splicing efficiency across ABCA4 trans-splicing molecules having various permutations of splice domains (SS), linkers, and binding domains (BD).

[0038] FIG. 2 is a graph showing trans-splicing efficiency across ABCA4 trans-splicing molecules having various splice domain sequences.

[0039] FIGS. 3A and 3B show trans-splicing efficiency of various ABCA4 trans-splicing molecules transduced by AAV in cultured human cells. FIG. 3A shows results by RT-qPCR. FIGS. 3B and 3C show results by Western blot. FIG. 3D is a graph quantifying the results of FIG. 3C.

[0040] FIG. 4 shows a Western blot of AAV8 preparations of trans-splicing molecules, showing VP1, VP2, and VP3 bands.

[0041] FIGS. 5A-5C show GFP protein expression by GFP trans-splicing molecules *in vivo* in African Green Monkey (AGM) non-human primates (NHP) following subretinal injection. FIG. 5A is a live image showing GFP protein. FIG. 5B is a photomicrograph showing anti-GFP immunohistochemistry (IHC) staining of retinal sections (photoreceptors and choroid are labeled) in a negative control at day 28. FIG. 5C is a photomicrograph showing anti-GFP IHC staining of retinal sections (photoreceptors and choroid are labeled) in a GFP trans-splicing molecule-treated animal at day 28.

[0042] FIG. 6 is a set of cSLO and OCT images of baseline, post-subretinal injections at days 14 and 28. Blebs are labeled in the "post injection" panels.

[0043] FIG. 7 is a set of color fundus images (anterior, posterior, superior, and inferior views) of NHP eyes treated with vehicle, exon editor 443, and exon editor 274 at day 28.

[0044] FIG. 8A is a schematic diagram showing RT-qPCR testing methodology for trans-splicing molecules. A trans-splicing molecule (Exon Editor RNA) with a 5' tag trans-splices into AGM pre-mRNA, and the edited (e.g., trans-spliced), full-length mRNA is quantitated by its resulting unique sequence. FIG. 8B shows quantitative RT-qPCR results for exon editors 443, 274, 497, and 505. As shown, each of Exon Editors 443, 274, 497, and 505 trans-splices with the endogenous AGM pre-mRNA *in vivo* in AGM retina at a statistically significant frequency as determined by quantitating the number of edited copies/10ng RNA.

[0045] FIG. 9 is a Western blot showing full-length chimeric (human/AGM) ABCA4 protein produced as a result of Exon Editors 443, 274, 497, and 505 trans-splicing into AGM pre-mRNA in AGM retina. The protein tag was detected in the membrane extracts of the retina.

[0046] FIGS. 10A and 10B are mass spectrometry results showing detection of unique peptides of human and AGM ABCA4 protein. FIG. 10A shows results for a negative control (vehicle-treated) retinal tissue. The top row shows the control sample lacking a spike where the human protein spike should appear; the second row shows the control sample spiked with standard to show where the human protein spike would appear; the third row shows the control sample containing a spike representing AGM protein; and the bottom row shows the sample spiked with standard, which contains the AGM protein spike. FIG. 10B shows results for the exon editor-treated retinal tissue, showing detection of both human and AGM peptides. The top row shows the control sample, which exhibits a human peptide spike, which reflects the presence of a human/AGM chimeric protein (in contrast to the negative control shown in FIG. 10A); the second row shows the control sample spiked

with standard; the third row shows the control sample containing a spike representing AGM protein; and the bottom row shows the sample spiked with standard, which contains the AGM protein spike.

[0047] FIGS. 11A-11F are bar graphs showing ocular health in AGM after treatment with AAV8 exon editors or vehicles administered subretinally. Data is shown as mean of vehicle and mean of exon editors 443, 274, 497, and 505 +/- standard deviation. FIG. 11A shows retinal thickness at baseline, day 14, and day 28. FIG. 11B shows intraocular pressure by tonometry at baseline, day 7, day 14, day 21, and day 28. FIGS. 11C-11F show results for metrics of ocular inflammation. FIG. 11C shows total clinical scores at baseline, day 7, day 14, day 21, and day 28; FIG. 11D shows vitreous cell score at baseline, day 7, day 14, day 21, and day 28; FIG. 11E shows aqueous cell score at baseline, day 7, day 14, day 21, and day 28; and FIG. 11F shows inflammatory keratic particles at baseline, day 7, day 14, day 21, and day 28.

[0048] FIG. 12A and 12B show results for general health metrics of AGM after treatment with AAV8 exon editors or vehicles administered subretinally. FIG. 12A shows body weight measurements at baseline, day 7, day 14, day 21, and day 28. FIG. 12B shows heart rate (beats per minute (bpm)) at baseline and day 28.

[0049] FIG. 13 depicts elements of an exemplary RNA exon editor that is designed to trans-splice exons 1-22 of the human ABCA4 gene of the exon editor to exons 23-50 of the endogenous ABCA4 pre-mRNA transcript. The exemplary exon editor is designed to treat subjects carrying mutations that span exons 1-22 of ABCA4. Numerical tags as follows: 1: AAV2-based Inverted Terminal Repeat Sequences (ITR) for AAV packaging of the ABCA4 RNA Exon Editor; 2: Cytomegalovirus (CMV) promoter (CMV immediate early enhancer and CMV immediate early promoter) drives ABCA4 RNA Exon Editor expression in photoreceptor cells; 3: Native 5' Untranslated Region (UTR) from Human ABCA4 mRNA maintains native 5' regulation; 4: ABCA4 exons 1-21, which comprise synonymous codon changes, replace corresponding endogenous coding sequence via trans-splicing. Synonymous codon changes may be used to, e.g., remove cryptic splice acceptor sites to mitigate potential alternative splicing reactions; 5: ABCA4 exon 22 comprising synonymous codon changes that enable sequence-based quantification; 6: Splice donor (also referred to herein as a splicing domain) trans-splices into endogenous ABCA4 intron 22 splice acceptor site leading to replacement of the endogenous ABCA4 pre-mRNA upstream sequence with the ABCA4 RNA Exon Editor Coding Domain Sequence (CDS) including Exon 22; 7: STOP codon (double stop) reduces/prevents translation of amino acids from the ABCA4 Exon Editor downstream of Exon 22 (e.g. Linker, 3x U1, ARE, BD domain regions); 8: Linker provides structural flexibility and accessibility to splice domain site; 9: Three tandem U1 binding sites (3x U1), which are predicted to associate with U1 snRNP and are demonstrated herein to potentiate ABCA4 RNA Exon Editor-mediated activity; 10: AU-rich element (ARE) which is demonstrated herein to decrease stability of any non-spliced ABCA4 RNA Exon Editor entering the cytoplasm; 11: Binding domain (BD) mediates targeting of ABCA4 RNA Exon Editor to endogenous ABCA4 pre-mRNA Intron 22 via sequence complementarity; 12: MALAT1 terminator promotes efficient trans-splicing; 13: AAV2-based ITR for AAV packaging of the Exon Editor

cassette. An exemplary sequence corresponding to the exemplary RNA exon editor depicted is presented in SEQ ID NO: 90 (ABCA4-01). SEQ ID NO: 90 corresponds to nucleotides 4284-9073 of SEQ ID NO: 82. SEQ ID NO: 90 comprises SEQ ID NO: 78 (spanning from the 5' UTR to the end of the MALAT1-Masc RNA comprising anti-mut1) and SEQ ID NO: 69 (spanning from the ATG within the 5' UTR to the end of the MALAT1-Masc RNA comprising anti-mut1). SEQ ID NO: 82 encodes SEQ ID NO: 90.

[0050] FIGs. 14A and 14B depict RNA trans-splicing and cis-splicing reactions. (A) Illustrates an RNA trans-splicing reaction schematic wherein the general mechanism of Exon Editing via trans-splicing is presented with respect to an ABCA4 RNA Exon Editor. An exemplary ABCA4 Exon Editor RNA is depicted above the targeted ABCA4 pre-mRNA. A binding domain (BD) complementary to intron 22 associates the Exon Editor with the pre-mRNA and enables the spliceosome-mediated trans-splicing reaction to occur between the Exon Editor's splice donor (SD) site (or splicing domain) and the native intron 22 splice acceptor (SA) site - leading to replacement of the upstream region of the targeted pre-mRNA with the coding sequence (CDS) provided by the Exon Editor. The pre-mRNA molecules progressively proceed through the common cis-splicing process that removes any remaining downstream introns, resulting in mature mRNA that can then undergo translation to form a functional (biologically active) ABCA4 protein. The red "x" mutations in the endogenous pre-mRNA illustration depict representative individual allelic mutations present in the ABCA4 gene on a single pre-mRNA transcript. The SD and SA sites are indicated with red arrows. The exon-exon junction formed via trans-splicing is encircled in red. (B) Depicts RNA cis- and trans-splicing mechanisms. Both RNA cis- and trans-splicing rely on the endogenous spliceosome to mediate splicing and begin with covalent bond formation between the Branch Point (BP), which is upstream of a polypyrimidine track (Py), and the first nucleotide of the SD site. During cis-splicing a lariat is formed, and during trans-splicing a Y-branch is formed. Next, the spliceosome mediates the ligation of the upstream exon to the downstream exon while the lariat and Y-branch molecules are rapidly degraded.

[0051] FIG. 15 presents a schematic of RT-qPCR analysis of individual RNA exon editor activity. Following RNA extraction from cells or tissues treated with an RNA exon editor, RNA is reverse-transcribed and subjected to qPCR or ddPCR. In the schematic, RNA exon editor RNA copy number is measured using the purple forward primer (labeled 1) and green reverse primer (labeled 2). Native ABCA4 RNA can be specifically quantified using the blue forward (labeled 3) and reverse (labeled 4) primer pair. Owing to synonymous codon changes in the ABCA4 RNA exon editor's terminal exon 22, the blue forward primer (labeled 3) has low complementarity to this position and is replaced with the purple forward primer (labeled 1) in order to quantify the exon-edited ONT RNA product instead. The purple forward (labeled 1) and reverse (labeled 5) primer pair is used to quantify the RNA exon editor, ONT, and OFT RNA molecules together.

[0052] FIG. 16 depicts screening and identification of exemplary ABCA4 Exon Editors. Both an NGS library-based approach (top) and a single Exon Editor bioluminescence-based approach

(bottom) were applied to identify BD117 as an exemplary position in the ABCA4 gene intron 22 to direct the Exon Editor.

[0053] FIGs. 17A and 17B show that the 40mer linker dramatically outperforms the LinkD linker. (A) Presents a histogram depicting percent replacement of ABCA4 RNA for each construct indicated or an untransfected negative control. An exon editor that comprises the 40mer linker exhibits a dramatic and statistically significant increased percent replacement of ABCA4 RNA compared to an exon editor that contains a progenitor linker (LinkD). (B) Depicts western blots showing that the 40mer linker promotes higher trans-splicing activity relative to the LinkD linker as reflected by dramatically higher levels of protein translated from on-target (ONT) trans-spliced RNA. In transfection experiments using HEK293T 17+06 cells engineered to overexpress mutant ABCA4 mRNA, an exon editor that contains the 40mer linker has dramatically higher trans-spliced ABCA4 protein (ONT) levels and significantly lower non-spliced protein (NSP) levels compared to an exon editor that contains the LinkD linker. The western blot was probed with anti-ABCA4 antibodies specific for the C-terminus of ABCA4 (top panel), anti-V5 epitope tag antibodies specific for the V5 epitope tag which is incorporated into the N-terminus of the trans-spliced protein and NSP (middle panel), and antibodies specific for sodium/potassium ATPase (Na⁺/K⁺ ATPase) which is a protein loading control.

[0054] FIG. 18 Cartoon illustrating full length proteins and NSP. C terminal Ab immunoprecipitates only full length ABCA4 protein.

[0055] FIGs. 19A and 19B illustrate the activity of each of the indicated RTM elements, alone and in combination, in the context of an exemplary ABCA4 5' RTM. (A) Presents a histogram depicting percent replacement of ABCA4 RNA for each construct indicated. The percent replacement ranges from around 20-30% for each construct. Intriguingly, the combination of 40mer + 3x repeat of a U1 binding site (3XUBS) shows a trend toward increasing percent replacement of ABCA4 RNA relative to 40mer alone. (B) Depicts western blots showing that the indicated constructs 40mer, 40mer + 3XUBS, 40mer + AU-rich elements (ARE), and 40mer + 3XUBS + ARE express trans-spliced ABCA4 protein (ONT) and in combination, these RTM elements exhibit a trend toward increased ONT levels relative to 40mer alone. The western blots also show that each of the combinations of 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE exhibits a combinatorial effect on non-spliced protein (NSP) reduction relative to 40mer alone. The combinatorial effect on NSP reduction is particularly pronounced for 40mer + 3XUBS + ARE relative to 40mer alone. In a minimal NSP context (contains stop codons immediately following the splice donor sequence in the exon editor), the combination of two strategies that independently reduced NSP levels –the inclusion of 3x UBS and an AU-rich element – showed concerted effect on NSP level reduction. This combination reduced the NSP level by ~70% compared to 40mer only, down to the level seen in a non-stop exon editor. Exemplary constructs comprise SEQ ID NO: 84 (40mer lane); SEQ ID NO: 85 (40mer + 3XUBS lane); SEQ ID NO: 86 (40mer + ARE); SEQ ID NO: 87 (40mer + 3XUBS + ARE).

[0056] FIGs. 20A and 20B show that the MALAT1 terminator (triple helix + mascRNA) confers significant improvement in ONT trans-splicing in an exemplary ABCA4 5' exon editor and the

anti-mut1 MALAT1 terminator exhibits similar activity to that of the MALAT1 terminator with respect to ONT protein and NSP expression. (A) Presents a histogram depicting relative trans-splicing conferred by the MALAT1 terminator as compared to a progenitor terminator in the context of the exemplary ABCA4 5' exon editor. The MALAT1 terminator confers a 23-fold increase in trans-splicing relative to the progenitor terminator. (B) Depicts western blots showing that for three different ABCA4 5' exon editors, MALAT1 and anti-mut1 MALAT1, exhibit similar activity with respect to expression levels of trans-spliced ABCA4 protein (ONT) and non-spliced protein (NSP). The western blot was probed with anti-ABCA4 antibodies specific for the N-terminus of ABCA4 (top panel), anti-ABCA4 antibodies specific for the C-terminal of ABCA4 (middle panel), and antibodies specific for sodium/potassium ATPase (Na^+/K^+ ATPase) which is a protein loading control.

[0057] FIGs. 21A and 21B show a dose response of an exemplary AAV2-ABCA4 RNA exon editor construct transduced into HEK293T 17+06 cells engineered to overexpress mutant ABCA4 mRNA, wherein the ABCA4 RNA exon editor is depicted in FIG. 13 (spanning from the 5' UTR to the end of the anti-mut1 MALAT1 terminator) ; SEQ ID NO: 78) and comprises the cassette of the 40mer linker + 3XUBS + ARE (SEQ ID NO: 73). The AAV2-ABCA4 exon editor was transduced into HEK293T 17+06 cells in biological duplicate at a series of multiplicity of infection (MOI) as follows: 0, 1k, 10k (1e4), 100k (1e5), 1M (1e6) followed by (A) RNA extraction for ddRT-PCR % RNA replacement or (B) membrane protein extraction and Western blotting probed with an ABCA4-specific antibody. The western blot probed with anti-ABCA4 antibodies specific for the N-terminus of ABCA4 (top panel) shows that ONT protein and NSP are detectable at MOI 1e4 and increase through MOI 1e6. The western blot was probed with anti-ABCA4 antibodies specific for the N-terminus of ABCA4 (top panel; 20 second exposure) and antibodies specific for sodium/potassium ATPase (Na^+/K^+ ATPase; 5 second exposure) which is a protein loading control. These results demonstrate the expected dose-response effect of increasing MOIs in terms of % ABCA4 RNA and protein replacement levels.

[0058] FIG. 22 shows *in vivo* ABCA4 RNA Exon Editing. Exon Editing in the AGM Retina Measured by RT-qPCR. Neural retina punches of the treated area were analyzed by RT-qPCR. (A) The % RNA replacement is shown as % of the total ABCA4 (native and edited). (B) The Exon Editor copy number per 10ng RNA was normalized to RNF20. The exemplary AAV8 RNA exon editor construct (AAV8-REEC) used in this *in vivo* study comprises an RNA exon editor comprising SEQ ID NO: 77. Results are shown for 1-month (4 weeks) and 3-months (12 weeks) time points.

[0059] FIG. 23 presents a schematic representation of an AAV vector plasmid (SEQ ID NO: 82) that encodes ABCA4-01 (SEQ ID NO: 90).

[0060] FIG. 24 depicts functional activity of protein translated from exon-edited ABCA4 RNA. These results provide evidence confirming the functional activity of ABCA4 protein rescued via ABCA4 RNA exon editing. ATPase activity of ABCA4 protein that was enriched by immunoprecipitation from ABCA4 KO cells following no treatment (negative control) or transfection with a V5 epitope-tagged version of the Exon Editor comprising SEQ ID NO: 78 (V5 tagged REEC),

wherein the activity of the V5 epitope-tagged version (SEQ ID NO: 81) is shown. ABCA4 ATPase activity measurements with no ATR (baseline) or with the addition of 40 μ M ATR (+ ATR) are shown. Data are single biological replicates with technical duplicate or triplicate measurements.

[0061] FIG. 25 depicts RNA species captured by OFT-seq. GSP targets codon optimized (CO) sequences in Exon 22 (Ex22) of the Exon Editor. Red dotted regions are sequenced by Illumina sequencing platform. The purple region illustrates the ABCA4 CDS provided by the Exon Editor. Three putative OFT junctions (A, B, C) and two putative self-spliced Exon Editor sequences are provided for illustrative purposes.

[0062] FIG. 26 presents results depicting cryptic splice sites identified and frequency of use thereof in HEK 17+06 cells for exemplary ABCA4 constructs comprising the indicated elements. A similar panel of cryptic splice sites were identified and characterized in vivo in NHP retina. Experimentally determined cryptic splice site frequency for an ABCA4 construct comprising a precursor sequence before introduction of synonymous mutations in identified cryptic splice sites (SEQ ID NO: 46) (top panel) and after introduction of synonymous mutations in identified cryptic splice sites therein (bottom panel), introduction of which generated a splice site resistant sequence (SEQ ID NO: 79) as evidenced by a significant reduction in splicing frequency.

[0063] FIGs. 27A and 27B illustrate activity for each of the indicated RTM elements, alone and in combination, in the context of a 5' RTM for an exemplary intron (Intron A) of a target pre-mRNA examined in vitro. **(A)** Presents a histogram depicting percent replacement of the target RNA for each construct indicated. Intriguingly, the combination of 40mer + 3XUBS, the combination of 40mer + ARE, or the combination of 40mer + 3XUBS + ARE shows a trend toward decreasing percent replacement of the target RNA relative to 40mer alone. **(B)** Depicts western blots showing that the indicated constructs 40mer, 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE express trans-spliced on-target protein (ONT) and that each of the combinations exhibits a combinatorial effect on NSP reduction relative to 40mer alone. More particularly, the combination of 40mer + 3XUBS reduced NSP levels by about 75%; the combination of 40mer + ARE reduced NSP levels by about 40%; and the combination of 40mer + 3XUBS + ARE reduced NSP levels by about 88% relative to 40mer alone.

[0064] FIGs. 28A and 28B illustrate the activity of each of the indicated RTM elements, alone and in combination, in the context of a 5' RTM for an exemplary intron (Intron B) of a target pre-mRNA examined in vitro. **(A)** Presents a histogram depicting percent replacement of the target RNA for each construct indicated. Intriguingly, the combination of 40mer + 3XUBS or the combination of 40mer + ARE shows a trend toward increasing percent replacement of the target RNA relative to 40mer alone. **(B)** Depicts western blots showing that the indicated constructs 40mer, 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE express trans-spliced target protein (ONT) and that each of the combinations of 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE exhibits a combinatorial effect on NSP reduction relative to 40mer alone. The combinatorial effect on NSP reduction is particularly pronounced for 40mer + 3XUBS + ARE relative to 40mer alone. More

particularly, the combination of 40mer + 3XUBS reduced NSP levels by about 38%; the combination of 40mer + ARE reduced NSP levels by about 33%; and the combination of 40mer + 3XUBS + ARE reduced NSP levels by about 66% relative to 40mer alone.

[0065] FIG. 29 depicts the activity of each of the indicated RTM elements, alone and in combination, in the context of a 5' RTM for an exemplary intron (Intron C) of a target pre-mRNA examined *in vitro*. Briefly, HEK293T cells were transfected with RTMs containing Non-Binding Domain (NBD) + 40mer; BD34 (which binds specifically to Intron C) + 40mer; BD34 + 40mer + 3XUBS; or no RTM (Transfection control). Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The percent replacement for BD34 + 40mer was about 1% replacement and that increased to about 3% replacement for an RTM comprising BD34 + 40mer + 3XUBS. Accordingly, the 3x UBS element increased percent replacement by about 3-fold in the exemplary RTM targeting Intron C shown.

[0066] FIGs. 30A-30C present results demonstrating robust ABCA4 RNA and Protein Replacement In Vivo in NHP Six (6) Months After Treatment with an exemplary AAV8 RNA exon editor construct. (A) Viral genome copy (VGC) (x-axis) and exon editor RNA copies numbers (y-axis); (B) % ABCA4 RNA replacement; and (C) Protein replacement in n=3 cynomolgus animals at 6-months following treatment with ABCA4-01 (3.5E11 vg/eye). In A and B, DNA and RNA levels were evaluated using Droplet Digital Polymerase Chain Reaction (ddPCR) based assays. In C, percent total protein (edited ABCA4 relative to the sum of edited and NHP native protein) was quantified using a qualified IA MS assay. The exemplary AAV8 RNA exon editor construct used in this in vivo study comprises an RNA exon editor comprising SEQ ID NO: 90. Results for four (4) NHPs are shown and designated in (B) and (C) as 1-4: 1 (vehicle treated); 2-4 (treated with 3.5E11 vg/eye).

[0067] FIG. 31 presents results demonstrating that an exemplary AAV8 RNA exon editor construct edits ABCA4 RNA in human retinas ex vivo. The exemplary AAV8 RNA exon editor construct used in this ex vivo study comprises an RNA exon editor comprising SEQ ID NO: 90. The graph depicts percent ABCA4 RNA replacement in human retinal explants from 3 independent donors, 21 days after treatment with vehicle or exemplary AAV8 RNA exon editor construct comprising SEQ ID NO: 90.

DETAILED DESCRIPTION

[0068] The compositions and methods described herein involve trans-splicing molecules (e.g., pre-mRNA trans-splicing molecules) for treating diseases or disorders caused by a mutation in the ABCA4 gene. The compositions and methods described herein employ pre-mRNA trans-splicing molecules for gene therapy (e.g., *in vivo* gene therapy, e.g., delivered by adeno-associated virus) to treat diseases caused by at least one mutation in ABCA4, such as Stargardt disease or cone-rod dystrophy.

[0069] ABCA4 is approximately 6.8 kb in size, contains 50 exons, and encodes a 250 kDa ATP-binding cassette transporter protein that is expressed in photoreceptor cells and RPE cells.

ABCA4 is a transmembrane protein that is specifically localized at the rims and incisures of outer segment discs of rod and cone photoreceptors. In the retina, ABCA4 transports N-retinylidene-phosphatidylethanolamine from the lumen to the cytoplasmic leaflet of photoreceptor disc membranes, a process that in combination with chemical isomerization facilitates the removal of potentially toxic retinoid compounds. Conversely, impaired ABCA4 function results in an accumulation of bisretinoids such as the di-retinal conjugate A2E in the retinal pigment epithelium (RPE). This debris, also known as lipofuscin, is thought to disturb RPE metabolism and ultimately contributes to cell death. Current genotype–phenotype models correlate the degree of residual ABCA4 activity with the severity of the retinal disease phenotype. Since *ABCA4*-related retinopathy is typically inherited in an autosomal recessive manner, individuals with 50% *ABCA4* levels in photoreceptors are expected to have normal retinal function. Data from *Abca4* knock-out mice have shown that a range of 10-25% correction of ABCA4 levels is sufficient to achieve relevant therapeutic effects, including reduction of toxic lipid accumulation in the retina.

[0070] Significantly, the present inventors have demonstrated complete rescue of ABCA4 protein levels in an *in vitro* engineered human *ABCA4* protein knockout (KO) cell line that carries disease mimicking premature stop codon mutations and achieved therapeutically relevant levels of RNA exon editing *in vivo* in NHP retinas and in human retinal explants. More specifically, the present inventors have observed that AAV-*ABCA4* Exon Editor treatment resulting in >20% ABCA4 RNA exon editing efficiency (% RNA replacement) is sufficient to achieve 100% rescue of ABCA4 protein levels *in vitro* in the *ABCA4* protein KO cell line. See, e.g., FIG. 3C. The parental ABCA4 knock-in (KI) cell line, which contains a strong CAG promoter knocked in upstream of the endogenous ABCA4 locus was used as a reference to compare ABCA4 protein rescue efficiency. As described herein, the *in vitro* engineered human *ABCA4* protein KO cell line was derived from the parental ABCA4 KI cell line by introducing disease mimicking premature stop codon mutations into the ABCA4 gene. Accordingly, the *in vitro* engineered human *ABCA4* protein knockout cell line serves as an *in vitro* model system for predicting therapeutic efficacy of ABCA4 RNA exon editors (nucleic acid trans-splicing molecules) described herein.

[0071] Moreover, the present inventors have shown that a single subretinal administration of an exemplary AAV8-based ABCA4 exon editing molecule resulted in up to 37% RNA replacement *in vivo* at a 3-month (12 week) timepoint in NHP eyes. See, e.g., FIG. 22. As used herein, 37% replacement means that 37% of the total ABCA4 mRNA quantified from the treated cells or tissue is present in an exon-edited form. The exemplary AAV8-based ABCA4 exon editing molecule used in this study comprises SEQ ID NO: 77, which shares the same regulatory elements and the double stop codon present in SEQ ID NO: 90. Translation of full-length human/NHP ABCA4 protein expression resulting from RNA trans-splicing has also been demonstrated in this study and expression levels trend with those of % RNA replacement. Additional NHP studies using a variety of other AAV8-based ABCA4 exon editing molecules examined at different time points post treatment have demonstrated durable RNA exon editing and human/NHP ABCA4 protein expression resulting from RNA trans-splicing. See, e.g., FIGs. 9 and 10B. Accordingly, the present inventors have achieved therapeutically relevant RNA exon editing via spliceosome-mediated trans-splicing and resultant full-length ABCA4

protein expression mediated by a single AAV vector in NHP. These studies demonstrated that the selected regulatory elements and 40mer linker and BD shared in common with ABCA4-01 (SEQ ID NO: 90) confer persistent expression and exon editing activity out to a 3-month time-point in NHP. See, e.g., FIGs. 22A and 22B.

[0072] Additional evidence presented herein demonstrates that ABCA4-01 (SEQ ID NO: 90), which is an exemplary AAV8-packaged RNA exon editor encoded by an AAV vector plasmid comprising SEQ ID NO: 82, can mediate trans-splicing of endogenous pre-RNA to a degree sufficient to at least partially restore ABCA4 biological activity in cells in vitro and in vivo. ABCA4-01 (SEQ ID NO: 90) is depicted in FIG. 13 and spans the length of RNA exon editor from 5' ITR to 3' ITR. In some embodiments, SEQ ID NO: 78 [a portion of SEQ ID NO: 90 spanning from the 5' UTR to the end of the MALAT1-MascRNA (anti-mut1); See FIG. 13] may be used as an RNA exon editor. In some embodiments, a portion of SEQ ID NO: 78 spanning from Exons 1-22 CDS to the end of the MALAT1-MascRNA comprising anti-mut1 depicted in FIG. 13 (e.g., SEQ ID NO: 69) can be combined with a different 5' UTR in an ABCA4 RNA exon editor with a reasonable expectation of success in achieving comparable restoration of ABCA4 biological activity.

[0073] Further to the above, FIGs. 30A-30C present results demonstrating robust ABCA4 RNA and protein replacement in vivo in NHP six (6) months after treatment with an exemplary AAV8 RNA exon editor construct (ABCA4-01; SEQ ID NO: 90). In brief, the present inventors treated wild-type cynomolgus macaques with a single subretinal dose of ABCA4-01 (3.5E11 vg/eye). AAV vector genome DNA copy (VGC) biodistribution, Exon Editor RNA expression, and the percent of edited ABCA4 RNA and protein at a 6-month timepoint were then evaluated. Exon Editor RNA expression correlated well with VGC (FIG. 30A) and resulted in robust % RNA replacement (conversion of native ABCA4 RNA copies to edited copies), calculated as the percent of edited ABCA4 RNA out of the total (edited + native) ABCA4 RNA population present in the neural retina samples (FIG. 30B). As shown in FIG. 30B, the percent RNA replacement ranged from about 40%-60% for this dose of ABCA4-01. This resulted in therapeutic levels of the resultant human-NHP chimeric ABCA4 protein as measured by a qualified IA-MS assay (FIG. 30C). More specifically, human NHP chimeric ABCA4 protein, which results from ABCA4 RNA editing in these animals, was found to be present at 20-40% of the total (human-NHP + NHP) ABCA4 protein in the tested samples. The present inventors have, moreover, demonstrated up to 66% RNA replacement and up to 45% human/NHP chimeric ABCA4 protein in wild-type cynomolgus macaques treated with a higher dose of ABCA4-01 delivered via single subretinal injection. These results demonstrate that ABCA4-01 achieves therapeutically relevant levels of edited ABCA4 protein expression in NHP, which levels exceed the levels of rescued ABCA4 protein previously shown to confer therapeutic benefit in an Abca4 KO mouse model.

[0074] The present inventors also evaluated ABCA4 exon editing in human photoreceptors by treating human donor retinal explants with ABCA4-01, an exemplary AAV8 RNA exon editor construct comprising SEQ ID NO: 90. As described in detail herein, the present inventors treated explants from multiple donors with ABCA4-01 (3.9E11 vg/retinal punch), incubated the cells for 21

days, and then extracted RNA. Analysis of the extracted RNA demonstrated approximately 20-30% RNA replacement of the endogenous exons 1-22 with the same corresponding exons encoded by ABCA4-01 (FIG. 31).

[0075] The present inventors have also demonstrated that ABCA4 protein rescued in ABCA4 protein KO cells transfected with an overexpression plasmid encoding an N-terminal V5-epitope tagged version of the RNA exon editor of ABCA4-01 (SEQ ID NO: 81) exhibited ATPase activity. As described herein, the activity of ABCA4 protein rescued following trans-splicing mediated by the V5-epitope tagged version of the RNA exon editor of ABCA4-01 (SEQ ID NO: 81) was assessed in the absence or presence of 40 μ M all-trans retinal (ATR) following ABCA4 protein immunoprecipitation. The results show that ABCA4 protein rescued in ABCA4 protein KO cells exhibited ATPase activity as reflected by an increase in basal level of ATPase activity in response to the addition of 40 μ M all-trans retinal. See FIG. 24. These results provide evidence that the exon-edited ABCA4 protein generated from ABCA4-01 (SEQ ID NO: 90) treatment exhibited the functional ATPase activity required for clearance of ATR and 11-cis-retinal from photoreceptors. These results depict restoration of ABCA4 biological activity following expression of an exemplary ABCA4 RNA exon editor.

[0076] Accordingly, results presented herein demonstrate the therapeutic potential for ABCA4 RNA exon editors described herein (e.g., RNA exon editors comprising any one of SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69) and compositions comprising same for use in treating ABCA4-related retinopathies and in the preparation of medicaments for treating ABCA4-related retinopathies, as well as in methods for treating subjects having ABCA4-related retinopathies. ABCA4-related retinopathies that may be treated using ABCA4 RNA exon editors described herein (e.g., RNA exon editors comprising any one of SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69) and compositions comprising same include Stargardt Disease-1 (STGD1); Fundus flavimaculatus, Retinitis Pigmentosa-19 (RP19); Cone-Rod Dystrophy-3 (CORD3); and Age-Related Macular Degeneration-2 (ARMD2). In some embodiments, ABCA4-related retinopathies that may be treated using ABCA4 RNA exon editors described herein (e.g., RNA exon editors comprising any one of SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69) and compositions comprising same comprise inherited retinal diseases with two allelic pathogenic mutations spanning Exons 1-22 in the ABCA4 gene that present with a phenotype consistent with cone-rod dystrophy, Stargardt macular dystrophy, or fundus flavimaculatus. Based on available data, approximately 60% of the known mutations in ABCA4 are correctable using ABCA4 RNA exon editors described herein (e.g., RNA exon editors encoded by nucleic acid trans-splicing molecules comprising any one of SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69) and 70-85% of patients diagnosed with ABCA4-related retinopathies would benefit from treatment with ABCA4 RNA exon editors described herein (e.g., RNA exon editors comprising any one of SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69).

[0077] Further to the above, persons skilled in the art appreciate that new mutations are regularly identified in the ABCA4 gene and have not necessarily been ascribed to specific phenotype/s. Skilled persons also appreciate that CORD3→STGD1→RP19 appear to represent a

spectrum of phenotype based on severity of mutation. Accordingly, RNA exon editors and constructs encoding same, as well as compositions comprising same may be used to advantage in the treatment of conditions/diseases associated with any and all mutations identified in exons 1-22 of the ABCA4 gene, with the understanding that subsequently ascribed phenotypes may be characterized as knowledge about ABCA4 retinopathies evolves.

[0078] In some embodiments, subjects/patients are identified as having at least one mutation in exons 1-22 of the ABCA4 gene. In some embodiments, the subjects/patients are genotyped to identify those having at least one mutation in exons 1-22 of the ABCA4 gene. In some embodiments, subjects/patients are obtained who have already been genotyped to identify those having at least one mutation in exons 1-22 of the ABCA4 gene. Under any model, subjects/patients are selected based on having at least one mutation in exons 1-22 of the ABCA4 gene, irrespective of the timing of the genotyping on which basis the selection process is made relative to a proposed treatment of such subjects/patients with RNA exon editors described herein. Such subjects/patients will benefit from use of ABCA4 RNA exon editors described herein for treating ABCA4-related retinopathies, medicaments comprising ABCA4 RNA exon editors described herein for treating ABCA4-related retinopathies, as well as methods for treating ABCA4-related retinopathies that call for administering ABCA4 RNA exon editors described herein or compositions comprising same.

[0079] Stargardt Disease-1 (STGD1) is caused by homozygous or compound heterozygous mutations in the ABCA4 gene (OMIM #601691) on chromosome 1p22. More than 900 different mutations including missense, splicing, truncating, and frameshift alterations have been reported in the ABCA4 gene and shown to be associated with retinal degeneration. Many of these mutations change single amino acids in the ABCA4 protein and result in the absence of, or decreased, ABCA4 activity in photoreceptor cells. Reduced ABCA4 activity leads to the accumulation of vitamin A derivatives and accumulation of lipofuscin in retinal cells, which is toxic to retinal pigment epithelium and photoreceptors. Stargardt macular degeneration is a genetic eye disorder that affects the retina and more specifically, affects an area called the macula that is responsible for sharp central vision, which is important for a variety of visual tasks including reading, driving, and facial recognition. In addition to central visual loss, subjects with Stargardt macular degeneration have impaired night vision and some affected individuals also have impaired color vision. Late in the disease course the peripheral visual fields may show moderate to extensive restriction as well. The signs and symptoms of Stargardt macular degeneration typically present in childhood to early adulthood and patients exhibit progressive bilateral visual loss over time, often reaching 20/200 or worse.

[0080] Fundus flavimaculatus (FFM) is an allelic subtype of Stargardt disease that has been associated with mutation in the ABCA4 gene and the peripherin-2 (PRPH2) gene. Fundus flavimaculatus, which is a form of fleck fundus disease, derives its name from the occurrence of many white-yellow spots due to deposition of lipofuscin, which are widely distributed over the fundus. FFM is characterized by network atrophy of the retinal pigment epithelium and choroidal vascular atrophy. Central visual loss, loss of color vision, photophobia, paracentral scotoma, and slow dark adaptation

are characteristic features of FFM. FFM presents later in life than other Stargardt subtypes and tends to run a slower course of progression. If loss of visual acuity begins in the first 2 decades of life, a patient is typically diagnosed with Stargardt, whereas if disease presents later in life and has a more slowly progressive course, a patient is typically diagnosed as having FFM.

[0081] Retinitis Pigmentosa-19 (RP19) can be caused by homozygous or compound heterozygous mutations in the ABCA4 gene on chromosome 1p22. Retinitis pigmentosa is an inherited eye disorder characterized by progressive loss of peripheral vision and impaired night vision. Many genes have been linked to retinitis pigmentosa. Retinitis Pigmentosa-19 is characterized by night-blindness, peripheral loss of vision, progressive retinal degeneration, tunnel vision, progressive vision loss, decreased vision at night or in low light, loss of central vision in advanced phases, and retinal pigment epithelium (RPE) mottling. Clinically, the fundus demonstrates widespread outer retinal degeneration, bone-spicule pigment deposition, vessel attenuation, waxy disc pallor, and severe atrophic changes as well as choroidal scarring of the macula.

[0082] Cone-Rod Dystrophy-3 (CORD3) is caused by homozygous or compound heterozygous mutation in the ABCA4 gene on chromosome 1p22. ABCA4 is one of the four major causative genes involved in the pathogenesis of cone rod dystrophies and ABCA4 mutations are associated with 30 to 60% of autosomal recessive cone rod dystrophies. CORD3 is an autosomal recessive, clinically heterogeneous retinal disorder typically characterized by symptomatic rapidly progressing visual loss combined with photophobia and loss of color discrimination noted initially, followed later by nyctalopia, and progressive visual field loss. Cone degeneration appears early in life with central involvement of the retina, followed by degeneration of rods several years later. The diagnosis of cone rod dystrophies is based on clinical history, fundus examination and electroretinogram. Cone-rod dystrophies typically present earlier and progress much more rapidly than the other ABCA4-retinopathies, with extensive macular atrophy occurring several years after onset. Currently, there is no therapy that stops the evolution of the disease or restores vision, and the visual prognosis is poor. Management aims at slowing down the degenerative process, treating the complications, and preparing patients for the challenges associated with the social and psychological impact of blindness.

[0083] Susceptibility to Age-Related Macular Degeneration-2 (ARMD2) is conferred by variation in the ABCA4 gene on chromosome 1p22. ARMD2 is a complex disorder characterized by the accumulation of drusen (a buildup of yellowish deposits) in and under the RPE and the progressive atrophy of the macular RPE. These changes result in loss of photoreceptor function and vision impairment. Age-related macular degeneration is a leading cause of vision loss in older people in developed countries. The vision loss usually becomes noticeable in a person's sixties or seventies and tends to worsen over time. Age-related macular degeneration mainly affects central vision, which is important for detailed tasks such as reading, driving, and facial recognition. Two major types of age-related macular degeneration have been described: the dry form and the wet form. The dry form is much more common, accounting for 85 to 90 percent of all cases of age-related macular

degeneration. It is characterized by a buildup of drusen beneath the retina, atrophy in the macula, and slowly progressive vision loss. The condition typically affects both eyes. The wet form of age-related macular degeneration is associated with sudden-onset severe vision loss that can worsen rapidly. This form of the condition is characterized by the growth of abnormal, fragile blood vessels underneath the macula due to over-expression of VEGF. These vessels leak blood and fluid, which damages the macula, thereby blurring and distorting central vision. Wet macular degeneration can be treated using a variety of therapies. Dry AMD has a single available treatment, a C3 inhibitor only for those who have geographic atrophy. No therapies currently exist for earlier forms of ARMD, where only drusen deposition has occurred, without atrophy or fluid leakage.

I. Definitions

[0084] As used herein, “trans-splicing” means joining a first RNA molecule containing one or more exons (e.g., exogenous exons or exons that are part of a CDS of a trans-splicing molecule) to a second RNA molecule (e.g., a pre-mRNA molecule, e.g., an endogenous pre-mRNA molecule) and replacing a portion of the second RNA molecule with a portion of the first RNA molecule through a spliceosome-mediated mechanism. The general mechanism for an RNA trans-splicing reaction is illustrated in FIG. 14B.

[0085] A “nucleic acid trans-splicing molecule” or “trans-splicing molecule” has three main elements: (a) a binding domain that confers specificity by tethering the trans-splicing molecule to its target gene (e.g., pre-mRNA); (b) a splicing domain (e.g., a splicing domain having a 3' or 5' splice site); and (c) a CDS configured to be trans-spliced onto the target nucleic acid, which can replace one or more exons in the target nucleic acid (e.g., one or more mutated exons). A “pre-mRNA trans-splicing molecule” or “RTM” refers to a nucleic acid trans-splicing molecule that targets pre-mRNA. The terms “nucleic acid trans-splicing molecule” and “trans-splicing molecule” refers to both (1) DNA that encodes RNA, wherein the RNA transcript is the effector molecule that physically binds the target pre-mRNA; and (2) the RNA transcript itself. For clarity, the term “-encoding sequence” (e.g., trans-splicing molecule-encoding sequence) is used herein to specify that the subject encodes the effector (e.g., the encoding sequence is DNA and the effector is RNA). In some embodiments, a trans-splicing molecule-encoding sequence can include cDNA, e.g., as part of a functional exon (e.g., a functional ABCA4 exon) for replacement of a mutated ABCA4 exon.

[0086] As used herein, the term “exon editor” may be used to refer to a trans-splicing molecule or a vector comprising same (e.g., an AAV vector comprising DNA encoding an RNA transcript that is a trans-splicing molecule).

[0087] As used herein, “trans-splicing efficiency” refers to a ratio of detected expression level of the desired trans-spliced RNA product (i.e., a chimeric RNA molecule that includes the functional exons of the trans-splicing molecule operably linked to endogenous target pre-mRNA generated by an RNA trans-splicing reaction) to the amount of DNA or RNA introduced for the trans-splicing molecule (or reference molecule). In some instances, expression level of a trans-spliced RNA product is detected from RNA that is isolated from cells or tissues using RNA-seq.

[0088] As used herein, “% RNA replacement” refers to the portion of the total target mRNA population that has undergone successful trans splicing (TS), and is calculated via the following equation: % on-target (ONT) TS = $100 * (\text{ONT copy number} / (\text{ONT copy number} + \text{Native copy number}))$.

[0089] As used herein, “relative trans-splicing efficiency” refers to a ratio of a test trans-splicing efficiency to a reference trans-splicing efficiency, wherein the test trans-splicing efficiency is the trans-splicing efficiency of a trans-splicing molecule (e.g., a nucleic acid trans-splicing molecule described herein), and the reference trans-splicing efficiency is the trans-splicing efficiency of a reference molecule (e.g., a reference molecule having the same elements as the nucleic acid trans-splicing molecule except that the binding domain is replaced with a scrambled binding domain (e.g., a binding domain comprising, or consisting of, SEQ ID NO: 1, e.g., a binding domain comprising, or consisting of, SEQ ID NO: 2). Relative trans-splicing efficiency of a trans-splicing molecule may be given as a ratio (a.k.a. fold increase) of the test trans-splicing RNA efficiency over the reference trans-splicing efficiency tested under similar conditions.

[0090] As used herein, the term “operably linked” or “operatively linked” refers to an arrangement of elements, wherein the components so described are configured so as to perform their usual function. A nucleic acid is “operably linked” to another nucleic acid sequence when it is placed into a functional relationship with the other nucleic acid sequence. Elements need not be contiguous with to be operably linked. Thus, for example, intervening sequences can be present between operably linked sequences (e.g., a binding domain and a coding sequence can be separated by intervening sequences and the binding domain is still considered to be “operably linked” to the coding sequence).

[0091] As used herein, the term “CDS” refers to a nucleic acid sequence (e.g., an RNA sequence, a DNA sequence, or combination of RNA and DNA) that encodes a portion of a protein (e.g., a target protein in which a mutation is being corrected). Thus, a CDS may include one or more functional exons (e.g., a sequence of functional exons). In some instances, one or more functional exons of a CDS are not separated by introns (e.g., as in endogenous pre-mRNA) but adjacent to one another (e.g., as cDNA). In some instances, a CDS can include one or more introns (e.g., native introns) or untranslated regions (UTRs, e.g., native UTRs) between or otherwise adjacent to (e.g., upstream or downstream of) exons.

[0092] As used herein, a “native 5' ABCA4 untranslated region” or “native 5' ABCA4 UTR” refers to a sequence greater than 20 nucleotides in length that has at least 90% sequence identity with a region of a native ABCA4 gene (e.g., a human ABCA4 gene) that is 5' to the ATG start codon. An example of a native 5' ABCA4 untranslated region is given by the DNA sequence of SEQ ID NO: 13 or the RNA sequence of SEQ ID NO: 12. Also encompassed herein are variants of a native 5' ABCA4 untranslated region such as, for example, the DNA sequence of SEQ ID NO: 64 or the RNA sequence of SEQ ID NO: 65.

[0093] As used herein, a “functional sequence of 5' ABCA4 exons” refers to a nucleic acid sequence comprising one or more of ABCA4 exons 1-22 that encode a functional (biologically active) portion of ABCA4 protein. When trans-spliced to an endogenous ABCA4 exon 3' to the binding site, a

functional sequence of 5' ABCA4 exons provides expression of functional ABCA4 protein (e.g., non-mutated ABCA4 protein). In some instances, the functional sequence of 5' ABCA4 exons includes a sequence of exons abutting the exon to which the trans-splicing molecule is being trans-spliced (e.g., a trans-splicing molecule that binds ABCA4 intron 22 and trans-splices with endogenous ABCA4 exon 23 can include a functional sequence of 5' ABCA4 exons that includes exons 22, 21, 20, 19, etc.).

[0094] As used herein, the term "functional", when used in the context of a protein, refers to a biologically active protein. The term "functional" may also be used to refer to the amount of activity of a protein that is necessary to support normal cellular functions. With respect to ABCA4, the term "functional" may be used to refer to the amount of ABCA4 protein activity that is necessary to restore ABCA4 activity levels to support normal cellular functions within the context of, for example, photoreceptor and/or retinal epithelial cells. Such levels are sufficient to reduce or prevent the accumulation of toxic levels of bisretinoid compounds in photoreceptor and/or retinal epithelial cells. More particularly, defective (non-functional) ABCA4 protein leads to the accumulation of 11-cis and all-trans-retinal in photoreceptors and lipofuscin in the retinal pigment epithelium. In the context of treating a condition associated with reduced or negligible ABCA4 activity or use of a therapeutic agent comprising a nucleic acid trans-splicing molecule described herein, functional refers to restoring an amount of ABCA4 protein sufficient to eliminate one or more symptoms of a condition associated with reduced levels of ABCA4 activity, e.g., an ABCA4-associated retinal dystrophy. ABCA4-associated retinal dystrophies include, for example, Stargardt macular dystrophy (Stargardt disease), fundus flavimaculatus, and ABCA4-related cone-rod dystrophy. In some embodiments, such methods or uses lead to an increase in ABCA4 protein activity (functional biologically active ABCA4 protein activity). In some embodiments, such an increase in ABCA4 protein activity restores ABCA4 activity levels to at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% of ABCA4 activity (e.g., 96%, 97%, 98%, 99%, or 100%) relative to that of a cell in which ABCA4 is present at normal, wildtype levels such as those present in cells in which non-mutated ABCA4 is expressed (e.g., photoreceptors and or retinal epithelial cells).

[0095] As used herein, "cryptic splice site corrected", "cryptic splice site mitigated", or "cryptic splice site resistant" refers to a nucleic acid trans-splicing molecule or a portion thereof (e.g., a coding domain sequence therein) that has been modified to change individual nucleotides therein to reduce the frequency of splicing that occurs at a cryptic splice site identified in the context of the nucleic acid trans-splicing molecule. In some embodiments, the modifications do not result in any changes in amino acid sequences encoded thereby. In some embodiments, the modifications result in changes in amino acid sequences encoded thereby. In some embodiments, the cryptic splice site resistant nucleic acid sequence within a nucleic acid trans-splicing molecule is a coding domain sequence (CDS). In some embodiments, the cryptic splice site resistant ABCA4 CDS comprises, consists essentially of, or consists of exons 1-22 of the ABCA4 gene, wherein cryptic splice sites have been identified in the context of the nucleic acid trans-splicing molecule and wherein at least one of the cryptic splice sites has been modified to reduce the frequency of splicing at the at least one site,

while not altering amino acids encoded thereby. Exemplary cryptic splice sites identified in ABCA4 exons 1-22 and frequencies of splicing at these sites are presented in Table 3. The nucleotide position indicated in the column having a header of rtm_pos in Table refers to the position of the last nucleotide indicated in column 3 or column 4 relative to the ATG codon (the translational initiation site), with the A in this codon being designated as position 1. Exemplary cryptic splice resistant ABCA4 exons 1-22 sequences comprise, consist essentially of, or consist of any one of SEQ ID NOs: 56-59. In some embodiments, an exemplary cryptic splice resistant ABCA4 exons 1-22 sequence comprises, consists essentially of, or consists of SEQ ID NO: 56.

[0096] A “splicing domain,” as used herein, refers to a nucleic acid sequence having motifs that are recognized by the spliceosome and mediate trans-splicing. A splicing domain includes a splice site (e.g., a single splice site, i.e., one and only one splice site), which can be a 3' splice site or a 5' splice site. A splicing domain may include other regulatory elements. For example, in some embodiments, a splicing domain includes splicing enhancers (e.g., exonic splicing enhancers (ESE) and intronic splicing enhancers (ISE)). In some embodiments, the splicing domain comprises GUAAGT or GTAAGT. In some embodiments, the splice site consists essentially of GUAAGT or GTAAGT. In some embodiments, the splice site consists of GUAAGT or GTAAGT.

[0097] As used herein, the “binding domain” of a trans-splicing molecule is a polynucleotide sequence that binds a target gene at a binding site via hybridization (i.e., full or partial complementarity to the binding site).

[0098] As used herein, the term “binding site” refers to an endogenous polynucleotide sequence in the target mRNA (e.g., an mRNA of an endogenous gene, e.g., *ABCA4*) that is bound by the binding domain of a nucleic acid trans-splicing molecule. The binding site extends from the 5'-most nucleotide bound by the binding domain to the 3'-most nucleotide bound by the binding domain. In some embodiments, the binding site is the same length as the binding domain. In other embodiments, the binding site is within 1-10 residues longer or shorter than the binding domain (i.e., some of the residues of either the binding site or the binding domain are unhybridized). In embodiments involving binding domains having at least two non-overlapping sequences with at least 80% complementarity to the binding site, the binding site may be substantially shorter than the binding domain.

[0099] As used herein, “complementarity,” and grammatical variations thereof, refers to the percentage of nucleotide bases of a given sequence that pairs through hydrogen bonding with a reference sequence. In the absence of a given percentage of complementarity, the terms “complement” and “complementary” refer to 100% complementarity.

[0100] As used herein, a given sequence (e.g., a binding domain sequence) is “100% complementary to,” or has “100% complementarity” with a reference sequence (e.g., an endogenous pre-mRNA binding site) if each of the nucleotide bases of the given sequence pairs through hydrogen bonding with the reference sequence, thereby hybridizing to form a double stranded sequence (e.g., through Watson-Crick base-pairing, e.g., each A pairs with a T or U, and each C pairs with a G). For instance, a binding domain that is in an antisense orientation to a binding site is complementary to the

binding site. RNA pairing includes G pairing with U (wobble base pairing); therefore, an RNA binding domain having G-U pairing with its binding site can be 100% complementary with the binding site. Accordingly, a binding domain that is exactly the reverse complement of its binding site (i.e., A's of the binding domain are paired with U's of the binding site) can be modified to replace any one or more of the A's with G's without substantially affecting binding.

[0101] As used herein, a given sequence (e.g., a binding domain sequence) is "at least X% complementary to," or has "X% complementarity" with a reference sequence (e.g., an endogenous pre-mRNA binding site) if X% of the nucleotide bases of the given sequence pairs through hydrogen bonding with the reference sequence, e.g., hybridizing to form a double stranded sequence (e.g., through Watson-Crick base-pairing, e.g., A pairs with T or U, and C pairs with G). For instance, a binding domain sequence having a length of 150 bases is at least 90% complementary to a binding site having a length of 150 bases if at least 135 of its 150 residues pair through hydrogen bonding with the binding site through Watson-Crick base pairing, leaving 15 or fewer mismatched nucleotides.

[0102] "Binding" between a binding domain and an intron, as used herein, refers to hydrogen bonding (e.g., double helix formation, or Watson Crick pairing) between the binding domain and the target intron in a degree sufficient to mediate trans-splicing by bringing the trans-splicing molecule into association with the target (e.g., pre-mRNA). In some embodiments, the hydrogen bonds between the binding domain and the target intron are between nucleotide bases that are complementary to and in antisense orientation from one another (e.g., hybridized to one another).

[0103] The sequence of a binding domain may be at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.5% complementary to a target sequence of a target sequence (e.g., a pre-mRNA). In some embodiments, the sequence of a binding domain may be 100% complementary to a target sequence of a pre-mRNA transcript.

[0104] In some embodiments, the sequence of a binding domain may have 4 or fewer mismatches with respect to a target sequence of a pre-mRNA transcript. In some embodiments, the sequence of a binding domain may have 3 or fewer mismatches with respect to a target sequence of a pre-mRNA transcript. In some embodiments, the sequence of a binding domain may have 2 or fewer mismatches with respect to a target sequence of a pre-mRNA transcript. In some embodiments, the sequence of a binding domain may have 1 or fewer mismatches with respect to a target sequence of a pre-mRNA transcript. In some embodiments, the sequence of a binding domain may have no mismatches with respect to a target sequence of a pre-mRNA transcript.

[0105] A binding domain may specifically hybridize to a target sequence of a pre-mRNA transcript. For example, the binding domain may have 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or 100% sequence complementarity to a target sequence of the pre-mRNA transcript. In some embodiments, the hybridization may be under physiological hybridization conditions (in a cell *in vitro* or *in vivo*). In some embodiments, the hybridization may be under high stringent hybridization conditions.

[0106] In some embodiments, binding domains "specifically hybridize" to or are "specific" for a target sequence of a pre-mRNA. Such hybridization typically occurs with a T_m substantially greater than 37° C., more particularly, at least 50° C., and typically between 60° C. to approximately 90° C.

Such hybridization more particularly corresponds to stringent hybridization conditions. At a given ionic strength and pH, the T_m is the temperature at which 50% of a target sequence hybridizes to a complementary oligonucleotide.

[0107] As used herein, an “artificial intron” refers to a noncoding nucleic acid sequence that links (directly or indirectly) a binding domain to a CDS. An artificial intron includes a splicing domain and may further include one or more spacer sequences and/or other regulatory elements.

[0108] As used herein, the term “mutation” may be used to refer to any aberrant nucleic acid sequence that causes a defective protein product (e.g., a non-functional protein product, a non-biologically active protein, a protein product having reduced function, a protein product having aberrant function, and/or a protein product that is produced in less than normal or greater than normal quantities). Mutations include base pair mutations (e.g., single nucleotide polymorphisms), missense mutations, frameshift mutations, deletions, insertions, and splice mutations. In some embodiments, a mutation refers to a nucleic acid sequence that is different in one or more portions of its sequence than a corresponding wildtype nucleic acid sequence or functional variant thereof. In some embodiments, a mutation refers to a nucleic acid sequence that encodes a protein having an amino acid sequence that is different from a corresponding wildtype protein or functional variant thereof. A “mutated exon” (e.g., a mutated ABCA4 exon) refers to an exon containing a mutation or an exon sequence that reflects a mutation in a different region, such as a cryptic exon resulting from a mutation in an intron.

[0109] The term “ABCA4” (ATP Binding Cassette Subfamily A Member 4) refers to any native ABCA4 (also known as ARMD2, CORD3, ABCR, FFM, Retinal-specific phospholipid-transporting ATPase ABCA4, Stargardt Disease Protein, RIM ABC Transporter, RIM Protein, STGD1, RP19, or STGD) from any vertebrate source, including mammals such as primates (e.g., human, African green monkeys, and cynomolgus monkeys) and rodents (e.g., mice and rats), unless otherwise indicated, as well as functionally equivalent or improved variants (e.g., natural or synthetic variants), mutants, muteins, analogs, subunits, receptor complexes, isotypes, splice variants, and fragments thereof. Functionally equivalent and improved variants can be determined on the basis of known ABCA4 signaling. ABCA4 encompasses full-length, unprocessed ABCA4, as well as any form of ABCA4 that results from native processing in the cell. An exemplary human ABCA4 sequence is provided as National Center for Biotechnology Information (NCBI) Reference Sequence: NG_009073.1. In some instances, the ABCA4 is encoded by a therapeutic gene having at least 95% sequence identity to SEQ ID NO: 11 (e.g., at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to SEQ ID NO: 11, a functional portion thereof, and/or a codon-optimized variant thereof).

[0110] As used herein, a “variant” refers to a polynucleotide that differs in at least one nucleic acid residue from the reference polynucleotide sequence, such as a naturally occurring polynucleotide sequence, or a polypeptide (e.g., an AAV capsid sequence) that differs in at least one amino acid residue from the reference polypeptide sequence, such as a naturally occurring polypeptide sequence or, e.g., any of the rAAV sequences described herein. In this context, the difference in at least one residue may include, for example, a substitution of a nucleic acid residue to

another nucleic acid, a deletion, or an insertion, or a substitution of an amino acid residue to another amino acid. A variant may be a homolog, isoform, or transcript variant of a polynucleotide as defined herein, wherein the homolog, isoform or transcript variant is characterized by a degree of identity or homology, respectively, as defined herein.

[0111] In some instances, a variant of a polynucleotide or polypeptide includes at least one nucleic acid substitution (e.g., 1-100 nucleic acid or amino acid substitutions, 1-50 nucleic acid or amino acid substitutions, 1-20 nucleic acid or amino acid substitutions, 1-10 nucleic acid or amino acid substitutions, e.g., 1 nucleic acid or amino acid substitution, 2 nucleic acid or amino acid substitutions, 3 nucleic acid or amino acid substitutions, 4 nucleic acid or amino acid substitutions, 5 nucleic acid or amino acid substitutions, 6 nucleic acid or amino acid substitutions, 7 nucleic acid or amino acid substitutions, 8 nucleic acid or amino acid substitutions, 9 nucleic acid or amino acid substitutions, or 10 nucleic acid or amino acid substitutions). Nucleic acid substitutions that result in the expressed polypeptide having an exchanged amino acid from the same class are referred to herein as conservative substitutions. In particular, these are amino acids having aliphatic side chains, positively or negatively charged side chains, aromatic groups in the side chains or amino acids, the side chains of which can form hydrogen bridges, e.g., side chains which have a hydroxyl function. By conservative constitution, e.g., an amino acid having a polar side chain may be replaced by another amino acid having a corresponding polar side chain, or, for example, an amino acid characterized by a hydrophobic side chain may be substituted by another amino acid having a corresponding hydrophobic side chain (e.g., serine (threonine) by threonine (serine) or leucine (isoleucine) by isoleucine (leucine)).

[0112] In some instances, insertions, deletions, and/or non-conservative substitutions are also encompassed by the term variant, e.g., at those positions that do not cause a substantial modification of the three-dimensional structure of the protein. Modifications to a three-dimensional structure by insertion(s) or deletion(s) can readily be determined by a person of skill in the art, e.g., using CD spectra (circular dichroism spectra).

[0113] The term "homologous" refers to the degree of identity between sequences of two nucleic acid sequences. The homology of sequences is determined by comparing two sequences aligned under standard conditions over the sequence length to be compared. The sequences to be compared herein may have an addition or deletion (for example, gap and the like) in the optimum alignment of the two sequences. In some embodiments, sequence homology is calculated by creating an alignment using, for example, the ClustalW algorithm (*Nucleic Acid Res.*, 1994, 22(22): 4673-4680). Commonly available sequence analysis software, such as, Vector NTI, GENETYX, BLAST, or analysis tools provided by public databases may also be used.

[0114] The term "AAV" or "AAV serotype" as used herein refers to the dozens of naturally occurring and available adeno-associated viruses, as well as artificial AAVs. Among the AAVs isolated or engineered from human or non-human primates (NHP) and well characterized, human AAV2 is the first AAV that was developed as a gene transfer vector; it has been widely used for efficient gene transfer experiments in different target tissues and animal models.

[0115] As used herein, relating to AAV, the term variant means any AAV sequence which is derived from a known AAV sequence, including those sharing at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% or greater sequence identity over the amino acid or nucleic acid sequence. In another embodiment, the AAV capsid includes variants which may include up to about 10% variation from any described or known AAV capsid sequence. That is, the AAV capsid shares about 90% identity to about 99.9% identity, about 95% to about 99% identity or about 97% to about 98% identity to an AAV capsid provided herein and/or known in the art. In one embodiment, the AAV capsid shares at least 95% identity with an AAV capsid. When determining the percent identity of an AAV capsid, the comparison may be made over any of the variable proteins (e.g., vp1, vp2, or vp3).

[0116] The ITRs or other AAV components may be readily isolated or engineered using techniques available to those of skill in the art from an AAV. Such AAV may be isolated, engineered, or obtained from academic, commercial, or public sources (e.g., the American Type Culture Collection, Manassas, Va.). Alternatively, the AAV sequences may be engineered through synthetic or other suitable means by reference to published sequences such as are available in the literature or in databases such as, e.g., GenBank, PubMed, or the like. AAV viruses may be engineered by conventional molecular biology techniques, making it possible to optimize these particles for cell specific delivery of nucleic acid sequences, for minimizing immunogenicity, for tuning stability and particle lifetime, for efficient degradation, for accurate delivery to the nucleus, etc.

[0117] As used herein, the term "subject," "individual," or "patient" includes any mammal in need of these methods of treatment or prophylaxis, including primates, such as humans. Other mammals in need of such treatment or prophylaxis include non-human primates (NHP; e.g., cynomolgus monkeys and African green monkeys), dogs, cats, or other domesticated animals, horses, livestock, laboratory animals (e.g., mice, rats, or rabbits), etc. The individual may be male or female. In one embodiment, the individual has a disease or disorder caused by a mutation in the ABCA4 gene (e.g., an ABCA4-associated retinal dystrophy). In another embodiment, the individual is at risk for developing a disease or disorder caused by a mutation in the ABCA4 gene. In another embodiment, the individual has shown clinical signs of a disease or disorder caused by a mutation in the ABCA4 gene, such as an ABCA4-associated retinal dystrophy. The individual may be any age during which treatment or prophylactic therapy may be beneficial. For example, in some embodiments, the individual is 0-5 years of age, 5-10 years of age, 10-20 years of age, 20-30 years of age, 30-50 years of age, 50-70 years of age, or more than 70 years of age. In another embodiment, the individual is 12 months of age or older, 18 months of age or older, 2 years of age or older, 3 years of age or older, 4 years of age or older, 5 years of age or older, 6 years of age or older, 7 years of age or older, 8 years of age or older, 9 years of age or older, or 10 years of age or older. In another embodiment, the individual has viable retinal cells (e.g., viable retinal cells sufficient to maintain partial visual function or full visual function).

[0118] As used herein, the terms "disorder associated with a mutation" or "mutation associated with a disorder" refer to a correlation between a disorder and a mutation. In some embodiments, a disorder associated with a mutation is known or suspected to be wholly or partially,

or directly or indirectly, caused by the mutation. For example, an individual having the mutation may be at risk of developing the disorder, and the risk may additionally depend on other factors, such as other (e.g., independent) mutations (e.g., in the same or a different gene), or environmental factors.

[0119] As used herein, the term “treatment,” or a grammatical derivation thereof, is defined as reducing the progression of a disease, reducing the severity of a disease symptom, retarding progression of a disease symptom, removing a disease symptom, or delaying onset of a disease. In some embodiments, the term “treatment” is used to refer to a persistent or durable effect of a therapeutic agent such as an RNA exon editor described herein. Evidence demonstrating the durable therapeutic effects of RNA exon editor described herein are presented in, for example, FIGs. 22 and 31. In some embodiments, the effect of a treatment provided herein persists for at least 1, 2, 3, 4, 5, 6 or more months. In some embodiments, the nucleic acid trans-splicing molecules are administered only once over the course of a lifetime.

[0120] As used herein, the term “prevention” of a disorder, or a grammatical derivation thereof, is defined as reducing the risk of onset of a disease, e.g., as a prophylactic therapy for an individual who is at risk for developing a disorder associated with a mutation. An individual can be characterized as “at risk” for developing a disorder by identifying a mutation associated with the disorder, according to any suitable method known in the art or described herein. In some embodiments, an individual who is at risk for developing a disorder has one or more ABCA4 mutations associated with the disorder. Additionally, or alternatively, an individual can be characterized as “at risk” for developing a disorder if the individual has a family history of the disorder.

[0121] As used herein, the term “ABCA4-related retinopathies” comprises diseases/conditions associated with mutations in ABCA4, which comprises, for example, Stargardt Disease-1 (STGD1); Fundus flavimaculatus, Retinitis Pigmentosa-19 (RP19); Cone-Rod Dystrophy-3 (CORD3); and Age-Related Macular Degeneration-2 (ARMD2).

[0122] Treating or preventing a disorder in an individual can be performed by directly administering the trans-splicing molecule or RNA exon editor (e.g., within a vector, e.g., an AAV vector or AAV particle) to the individual. Alternatively, host cells containing the trans-splicing molecule may be administered to the individual.

[0123] The term “administering” or a grammatical derivation thereof, as used in the methods described herein, refers to delivering a trans-splicing molecule or RNA exon editor (e.g., within a vector, e.g., an AAV vector or AAV particle) or a composition thereof, or an *ex vivo*-treated cell, to the individual in need thereof, e.g., an individual having a mutation or defect in ABCA4. For example, in one embodiment in which ocular cells (e.g., photoreceptors) are targeted, the method involves delivering a trans-splicing molecule or RNA exon editor (e.g., within a vector, e.g., an AAV vector or AAV particle) or a composition thereof to the individual by subretinal injection. In another embodiment, intravitreal injection or injection via the palpebral vein may be performed as administration. In another embodiment, the composition is administered systemically (e.g., intravenously). Still other methods of administration may be selected by one of skill in the art, in view of this disclosure.

[0124] As used herein, “modulating expression of ABCA4” refers to decreasing the expression of endogenous mutated (non-functional) ABCA4 and/or increasing the expression of trans-spliced ABCA4. Modulating expression of ABCA4 may be used to refer, e.g., to decreasing the expression of endogenous (e.g., mutated) ABCA4 and/or increasing the expression of trans-spliced ABCA4 (e.g., ABCA4 transcript or protein product that has a trans-splicing molecule-mediated corrected mutation site relative to its endogenous mutated transcript or protein product. Upon replacement of the endogenous ABCA4 exon comprising the mutation site via trans-splicing, a functional ABCA4 protein is expressed.

[0125] As used herein, “codon optimization” refers to modifying a nucleic acid sequence to change individual nucleic acids without any resulting change in the encoded amino acid. Sequences modified in this way are referred to herein as “codon-optimized.” This process may be performed on any of the sequences described in this specification to enhance expression or stability. Codon optimization may be performed in a manner such as that described in, e.g., U.S. Patent Nos. 7,561,972, 7,561,973, and 7,888,112, each of which is incorporated herein by reference in its entirety. The sequence surrounding the translational start site can be converted to a consensus Kozak sequence according to known methods. See, e.g., Kozak et al, 1987. *Nucleic Acids Res.* 15 (20): 8125-8148, which is incorporated herein by reference in its entirety.

[0126] The term “pharmaceutically acceptable” means safe for administration to a mammal, such as a human. In some embodiments, a pharmaceutically acceptable composition is approved by a regulatory agency of the Federal or a state government or listed in the U. S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans.

[0127] The term “carrier” refers to a diluent, adjuvant, excipient, or vehicle with which a therapeutic molecule (e.g., a trans-splicing molecule or a trans-splicing molecule including a vector or cell of the present invention) is administered. Examples of suitable pharmaceutical carriers are described in “Remington's Pharmaceutical Sciences,” Mack Publishing Co., Easton, PA., 18th edition.

[0128] The terms “a” and “an” mean “one or more of.” For example, “a gene” is understood to represent one or more such genes. As such, the terms “a” and “an,” “one or more of a (or an),” and “at least one of a (or an)” are used interchangeably herein.

[0129] Unless specified otherwise, the left-hand end of single-stranded nucleic acid (e.g., pre-mRNA transcript, oligonucleotide, etc.) sequences is the 5' end and the left-hand direction of single or double-stranded nucleic acid sequences is referred to as the 5' direction. Similarly, the right-hand end or direction of a nucleic acid sequence (single or double stranded) is the 3' end or direction. Generally, a region or sequence that is 5' to a reference point in a nucleic acid is referred to as “upstream,” and a region or sequence that is 3' to a reference point in a nucleic acid is referred to as “downstream.” Generally, the 5' direction or end of an mRNA is where the initiation or start codon is located, while the 3' end or direction is where the termination codon is located.

[0130] As used herein, the term “about” refers to a value within $\pm 10\%$ variability from the reference value, unless otherwise specified.

II. Trans-Splicing Molecules

[0131] Exon Editor design and discovery employs a combined approach that includes Next Generation Sequencing (NGS)-based screening of Exon Editor libraries followed by computational analysis, as well as transfection of individual Exon Editors followed by RT-qPCR analysis to assess RNA replacement efficiency and Western blot analysis to assess protein rescue. The former approach utilizes rational design and cumulative knowledge gained from past screens along with the power of multi-parallelized throughput, and the latter approach is used to both test new hypotheses and validate conclusions resulting from the NGS-based method.

[0132] Provided herein are nucleic acid trans-splicing molecules useful for correcting mutations in ABCA4 by replacing at least one mutated ABCA4 exon with a functional ABCA4 exon (e.g., a ABCA4 exon 5' to the binding site, e.g., exons 1-22 of ABCA4). In some embodiments, the nucleic acid trans-splicing molecule is a pre-RNA trans-splicing molecule (RTM). The design of the trans-splicing molecule permits replacement of the defective or mutated portion of the pre-mRNA exon(s) with a nucleic acid sequence, e.g., the exon(s) having a functional (e.g., normal) sequence without the mutation. The functional sequence can be a wildtype, naturally occurring sequence or a corrected sequence with some other modification, e.g., codon optimization.

[0133] Trans-splicing molecules comprise a binding domain, a splicing domain, and a CDS. In some embodiments, the nucleic acid trans-splicing molecule has a 5' regulatory domain having a native 5' ABCA4 untranslated region (e.g., a sequence having at least 80% sequence identity with any one of SEQ ID NOS: 12-15). In some embodiments, the nucleic acid trans-splicing molecule has a splice site of GUAAGT or GTAAGT. In some embodiments, the nucleic acid trans-splicing molecule has a linker domain that is longer than 25 nucleotides in length. In some embodiments, the nucleic acid trans-splicing molecule has a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27.

[0134] In some embodiments, a trans-splicing molecule described herein includes, operatively linked in a 5' to 3' direction: a 5' untranslated region, a coding domain sequence (e.g., a CDS, e.g., a sequence encoding a functional sequence of ABCA4 exons, e.g., a functional sequence of ABCA4 exons 5' to the binding site), a splicing domain (e.g., a splice site), a linker domain, a binding domain, a 3' downstream region, and a terminator domain.

[0135] In some embodiments, nucleic acid trans-splicing molecules described herein are configured to correct at least one mutation (e.g., two different mutations, wherein each mutation is present on a different allele of the ABCA4 gene in a subject) located in a 5' region of the ABCA4 gene (e.g., a region 5' to intron 22) by binding to target intron 22 and mediating trans-splicing of a CDS having a functional sequence of 5' ABCA4 exons to an endogenous ABCA4 exon 3' to the target intron. Such trans-splicing thereby repairs the defective ABCA4 gene in the target cell of an individual by replacing the defective exon/s and removing the defective portion of the target pre-mRNA, yielding a functional ABCA4 mRNA capable of transcribing a functional ABCA4 protein in the cell.

[0136] In some embodiments, the trans-splicing molecule is up to 5,000 nucleotide bases in length, e.g., up to 4,900 nucleotide bases in length, up to 4,800 nucleotide bases in length, or up to 4,700 nucleotide bases in length (e.g., from 3,000 to 4,000 nucleotide bases in length, from 3,100 to 3,800 nucleotide bases in length, from 3,200 to 3,700 nucleotide bases in length, or from 3,300 to 3,500 nucleotide bases in length, e.g., from 3,000 to 3,100 nucleotide bases in length, from 3,100 to 3,200 nucleotide bases in length, from 3,200 to 3,300 nucleotide bases in length, from 3,300 to 3,400 nucleotide bases in length, from 3,400 to 3,500 nucleotide bases in length, from 3,500 to 3,600 nucleotide bases in length, from 3,600 to 3,700 nucleotide bases in length, from 3,700 to 3,800 nucleotide bases in length, from 3,800 to 3,900 nucleotide bases in length, from 3,900 to 4,000 nucleotide bases in length, from 4,000 to 4,100 nucleotide bases in length, from 4,100 to 4,200 nucleotide bases in length, from 4,200 to 4,300 nucleotide bases in length, from 4,300 to 4,400 nucleotide bases in length, from 4,400 to 4,500 nucleotide bases in length, from 4,500 to 4,600 nucleotide bases in length, from 4,600 to 4,700 nucleotide bases in length, from 4,700 to 4,800 nucleotide bases in length, from 4,800 to 4,900 nucleotide bases in length, or from 4,900 to 5,000 nucleotide bases in length).

[0137] Due to the large size of the ABCA4 gene and the size constraints of AAV delivery, a single trans-splicing molecule configured for packaging within an AAV vector (e.g., as a trans-splicing molecule-encoding sequence) may not span all mutations in a ABCA4 gene that may be associated with a disorder and therefore, may not correct mutations along the length of the entire ABCA4 gene. Accordingly, the trans-splicing molecules described herein can be adapted as part of methods described below to correct multiple mutations spanning the entire length of the ABCA4 gene.

ABCA4

[0138] An ABCA4 gene targeted by a trans-splicing molecule described herein can contain one or multiple mutations that are associated with ABCA4-associated retinal dystrophies (e.g., cause, or are correlated with) a disease, such as Stargardt Disease or cone-rod dystrophy. An exemplary human ABCA4 sequence is provided as National Center for Biotechnology Information (NCBI) Reference Sequence: NG_009073. In addition to published sequences, all corrections later obtained or naturally occurring conservative and non-disease-causing variant sequences that occur in the human or other mammalian population are also included. Additional conservative nucleotide replacements or those causing codon optimizations are also included. The sequences as provided by the database accession numbers may also be used to search for homologous sequences in the same or another mammalian organism.

[0139] It is anticipated that the ABCA4 nucleic acid sequences and resulting proteins expressed may tolerate certain minor modifications at the nucleic acid level to include, for example, modifications to the nucleotide bases which are silent, e.g., preference codons. In other embodiments, nucleic acid base modifications which change the amino acids, e.g., to improve expression of the resulting peptide/protein (for example, codon optimization) are envisioned. In some embodiments, modification of allelic variations, caused by the natural degeneracy of the genetic code are envisioned.

[0140] Also included as modifications of ABCA4 genes are analogs, or modified versions of the encoded amino acid sequences. Typically, such analogs differ from the specifically identified proteins by only one to four codon changes. Conservative replacements are those that take place within a family of amino acids that are related in their side chains and chemical properties.

[0141] The nucleic acid sequence of a functional ABCA4 gene may be derived from any mammal which natively expresses functional ABCA4 or a homolog thereof. In other embodiments, certain modifications are made to the ABCA4 gene sequence in order to enhance expression in the target cell. Such modifications include codon optimization.

[0142] Mutations in particular ABCA4 exons are listed, for example, in International Patent Publication No. WO 2017/087900, incorporated herein by reference.

[0143] As described herein above, ABCA4 retinopathies are caused by pathogenic variants in the ABCA4 gene, which are inherited in an autosomal recessive manner. Such pathogenic variants (mutations) include missense, nonsense, splicing, structural, and deep intronic variants. Missense variants account for most variants associated with ABCA4 retinopathies. In an individual having an ABCA4 retinopathy, the individual has a mutation in ABCA4 on each allele. Compositions comprising trans-splicing molecules can correct the mutations on both alleles, regardless of the location of the mutation within the ABCA4 gene. For instance, for an individual having a mutated ABCA4 exon 2 on a first allele and a mutated ABCA4 exon 9 on a second allele, a single 5' trans-splicing molecule (5' RNA editing molecule) encoding exons 1-22 may be used to replace the mutated ABCA4 exon 2 and the mutated ABCA4 exon 9. See, e.g., FIG. 14A. Accordingly, in some embodiments, wherein two or more mutations are located on a portion of the ABCA4 gene that can be replaced by the same trans-splicing molecule that spans the two or more mutations, a single trans-splicing molecule having a coding region containing a functional ABCA4 exon/s can replace the one or more exons containing the mutations.

[0144] Alternatively, in some embodiments, an individual has ABCA4 mutations that involve mutations at distant exons within the ABCA4 gene. For instance, wherein an individual has a mutated ABCA4 exon 2 on a first allele and a mutated ABCA4 exon 49 on a second allele, two RNA editing molecules may be used in conjunction (e.g., in a composition) to replace each of the mutated exons. Such an approach would involve use of a 5' trans-splicing molecule to replace the mutated ABCA4 exon 2 and a 3' trans-splicing molecule to replace the mutated ABCA4 exon 49. In such embodiments, the two trans-splicing molecules can be co-delivered as part of the same AAV vector or delivered in separate AAV vectors (e.g., in the case in which both trans-splicing molecules exceed the packaging limit of AAV).

[0145] Coding Domain Sequences

[0146] In some embodiments, the CDS of a 5' trans-splicing molecule includes all ABCA4 exons (e.g., functional ABCA4 exons) that are 5' to the target ABCA4 intron (e.g., ABCA4 intron 22). For example, in embodiments in which a 5' trans-splicing molecule targets ABCA4 intron 22, the CDS may include functional ABCA4 exons 1-22. In such embodiments, the CDS may be from 2,000 to 4,500 bp in length. In some instances, both allelic mutations occur in the 5' portion of the target gene,

and a 5' trans-splicing molecule is selected to correct both mutations. In one embodiment, the binding domain binds to intron 22, and the CDS includes functional ABCA4 exons 1-22.

[0147] In some embodiments, a CDS- (e.g., of a transgene encoding an RTM) includes cDNA of ABCA4 exons (e.g., functional ABCA4 exons) for replacement of mutated ABCA4 exon/s. For example, one or more functional ABCA4 exons within the CDS can be a cDNA sequence. In some embodiments, the entire CDS is a cDNA sequence. Additionally, or alternatively, all or a portion of the CDS, or one or more functional ABCA4 exons thereof, can be a naturally occurring sequence (e.g., a sequence having 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity with an endogenous ABCA4 exon).

[0148] In some embodiments, all or a portion of the CDS, or one or more functional ABCA4 exons thereof, is a codon-optimized sequence in which a nucleic acid sequence has been modified, e.g., to enhance expression or stability, without resulting in a change in the encoded amino acid. Codon optimization may be performed in a manner such as that described in, e.g., U.S. Patent Nos. 7,561,972, 7,561,973, and 7,888,112, each of which is incorporated herein by reference in its entirety. For delivery via a recombinant AAV, as described herein, in one embodiment, the CDS can be a nucleic acid sequence of up to 4,000 nucleotide bases in length (e.g., from 3,000 to 4,000 nucleotide bases in length, from 3,100 to 3,800 nucleotide bases in length, from 3,200 to 3,700 nucleotide bases in length, or from 3,300 to 3,500 nucleotide bases in length, e.g., from 3,000 to 3,100 nucleotide bases in length, from 3,100 to 3,200 nucleotide bases in length, from 3,200 to 3,300 nucleotide bases in length, from 3,300 to 3,400 nucleotide bases in length, from 3,400 to 3,500 nucleotide bases in length, from 3,500 to 3,600 nucleotide bases in length, from 3,600 to 3,700 nucleotide bases in length, from 3,700 to 3,800 nucleotide bases in length, from 3,800 to 3,900 nucleotide bases in length, or from 3,900 to 4,000 nucleotide bases in length, e.g., about 3,108 nucleotide bases in length, about 3,285 nucleotide bases in length, about 3,375 nucleotide bases in length, about 3,503 nucleotide bases in length, about 3,630 nucleotide bases in length, about 3,540 nucleotide bases in length, about 3,363 nucleotide bases in length, about 3,273 nucleotide bases in length, about 3,145 nucleotide bases in length, or about 3,018 nucleotide bases in length).

[0149] In some embodiments, all or a portion of the CDS, or one or more functional ABCA4 exons thereof, is modified to mutate cryptic splice sites identified therein. The present inventors discovered through extensive experimental and computational analyses of results that cryptic splice sites are utilized in the context of some embodiments of the ABCA4 exon editors described herein. In short, the present inventors performed extensive experimentation and analyses, including the generation of an off-target (OFT) library and bioinformatic analyses thereof, and surprisingly discovered numerous locations within the ABCA4 coding sequence (CDS) comprising exons 1-22 wherein cryptic splicing occurs in the context of trans-splicing molecules described herein. This discovery revealed a previously unappreciated problem that RNA exon editors (also referred to herein as RNA trans-splicing molecules or nucleic acid trans-splicing molecules) can undergo splicing at cryptic splice sites which can lead to unwanted species of splicing variants having potential downstream effects. To improve trans-splicing efficiency and fidelity of trans-splicing molecules described herein and thereby provide a solution to the problem of cryptic splice site utilization, the

sequences of a plurality of the cryptic splice sites were modified to impair and/or disable the cryptic splice sites by introduction of synonymous mutations. In some embodiments, cryptic splice sites that were modified were identified based on the number of times the cryptic splice site was utilized as detected by bioinformatic analyses of the OFT library. In some embodiments, the cryptic splice sites modified accounted for a plurality of cryptic splice site events. Such an approach takes advantage of the degeneracy of the genetic code, which permits introduction of the desired modifications to mitigate cryptic splice site use without resulting in a change in the encoded amino acid. See, e.g., Table 6 and FIG. 26.

[0150] Using a Next Generation Sequencing (NGS) approach, the present inventors were able to identify cryptic splice sites and quantitate the frequency of cryptic splice use at each of the different locations within the ABCA4 CDS comprising exons 1-22, thereby establishing a protocol for data-driven, prioritized mitigation of cryptic splice use therein. See Table 3 and FIG. 26. As shown in Table 3, the cryptic splice sites identified are arranged in order of frequency of use as determined by the average counts per million (CPM) reads mapped/incorporated in a sequence that does not include synonymous mutations that disrupt the cryptic splice site. Accordingly, the results and bioinformatic analyses thereof establish a prioritized list of cryptic splice sites in ABCA4 exons 1-22 that may be modified to generate an ABCA4 CDS having improved properties and functionality in the context of a trans-splicing molecule.

[0151] In accordance with results presented herein, the present inventors established a frequency of use threshold above which analysis and design for cryptic splice site mitigation may advantageously be implemented where possible via introduction of synonymous mutations. See Table 6. In some embodiments, the threshold based on results presented in Table 3 is equal to or greater than 10 CPMs.

[0152] An NGS approach was also implemented to assess the frequency of cryptic splice usage following introduction of synonymous mutations. See Table 3 and FIG. 26. Confirmation of cryptic splice site mitigation by such an approach linked the structural change of the sequence modification (synonymous mutation) to the function conferred thereby, namely generation of cryptic splice site resistant ABCA4 CDSs.

[0153] In some embodiments, an RNA exon editor molecule, e.g., an ABCA4 exon editor, has a cryptic splice site-mitigating nucleotide change at one or more sites associated with off-target splicing within the exon editor molecule. Any or all of the changed nucleotide positions may be at locations identified in Table 3. In some embodiments, an ABCA4 exon editor has a nucleotide change at one or more of the following positions (numbering according to, e.g., SEQ ID NO: 56), wherein the nucleotide positions indicated below refer to the 5' and 3' ends of the cryptic splice site nucleotide sequence mitigated relative to the G of the ATG codon (the translational initiation site), or any combination thereof. The nucleotide change made is underlined in the sequences listed below. See also Table 3:

Sequence spanning nucleotide positions 3066-3086:

ACATGCTGTTCTATGCCCAG (SEQ ID NO: 91)

ACATGCTGTTCTATGCACAA (SEQ ID NO: 92)

Sequence spanning nucleotide positions 2911-2931:

CACCTTGTCCATCCTCACAG (SEQ ID NO: 93)

CACCTTGTCCATCCTCACTG (SEQ ID NO: 94)

Sequence spanning nucleotide positions 2556-2576:

TCGCTTGGTACCTTGATCAG (SEQ ID NO: 95)

TCGCTTGGTACCTTGATCAA (SEQ ID NO: 96)

Sequence spanning nucleotide positions 3009-3029 (splice site corrected):

GCCTTGGCATGTGTCCACAG (SEQ ID NO: 97)

CCCTTGGCATGTGTCCACAA (SEQ ID NO: 98)

Sequence spanning nucleotide positions 107-127:

CTGGTCTTGATCTGGTTAAG (SEQ ID NO: 99)

CTGGTCTTGATCTGGTTACG (SEQ ID NO: 100)

Sequence spanning nucleotide positions 2328-2348:

TCCTGTGCTTCGCCTGGCAG (SEQ ID NO: 101)

TCCTGTGCTTCGCCTGGCAA (SEQ ID NO: 102)

Sequence spanning nucleotide positions 1815-1835:

GCGGGTTTGCCTATCTGCAG (SEQ ID NO: 103)

GCGGGTTTGCATATCTGCAA (SEQ ID NO: 104)

Sequence spanning nucleotide positions 1198-1218:

CACTCCTGATTACCTGCAG (SEQ ID NO: 105)

CACTCCTGATTACCTGCTG (SEQ ID NO: 106)

Sequence spanning nucleotide positions 3063-3083 [the nucleotide change indicated below is a result of changes made in the context of position 3086 (proximal to position 3083) and did not correct a splice site, polypyrimidine tract, and/or branchpoint for the 3083 position]:

AGCACATGCTGTTCTATGCCC (SEQ ID NO: 107)

AGCACATGCTGTTCTATGCAA (SEQ ID NO: 108)

Sequence spanning nucleotide positions 2090-2110:

TGTACCTGGTTCCTGGACAG (SEQ ID NO: 109)

TGTACCTGGTTCCTGGACTC (SEQ ID NO: 110)

Last nucleotide listed below corresponds to position 2608-2628 (branchpoint corrected):

TTGGTACTTTCTTCTACAAAG (SEQ ID NO: 111)

TTGGTACTTTCTTCTGCAAG (SEQ ID NO: 112)

Sequence spanning nucleotide positions 3083-3103 [the nucleotide change indicated below is a result of changes made in the context of position 3086 (proximal to position 3103) and did not correct a splice site, polypyrimidine tract, and/or branchpoint for the 3103 position]:

CAGCTGAAAGGAAAGTCCCA (SEQ ID NO: 113)

CAACTGAAAGGAAAGTCCCA (SEQ ID NO: 114)

Sequence spanning nucleotide positions 3073-3093 [the nucleotide change indicated below is a result of changes made in the context of position 3086 (proximal to position 3093) and did not correct a splice site, polypyrimidine tract, and/or branchpoint for the 3093 position]:

GTTCTATGCCCAGCTGAAAG (SEQ ID NO: 115)
 GTTCTATGCACAACTGAAAG (SEQ ID NO: 116)
 Sequence spanning nucleotide positions 681-701 (splice site corrected):
 CCCTGTGCTCCCTCTCCCAG (SEQ ID NO: 117)
 CCCTGTGCTCCCTCTCCAA (SEQ ID NO: 118)
 Sequence spanning nucleotide positions 508-528:
 CATCGGCCTGTCTGACTCAG (SEQ ID NO: 119)
 CATCGGCCTGTCTGACTCTG (SEQ ID NO: 120)
 Sequence spanning nucleotide positions 1431-1451 (branchpoint corrected):
 TCCTAAACTTCCTCTACAAG (SEQ ID NO: 121)
 TCCTGAACTTCCTCTACAAG (SEQ ID NO: 122)
 Sequence spanning nucleotide positions 1995-2015 (splice site corrected):
 CTGTCTCCATGACTGTGAAG (SEQ ID NO: 123)
 CTGTCTCCATGACTGTAAAA (SEQ ID NO: 124)
 Sequence spanning nucleotide positions 3159-3179 (could be corrected with non-synonymous mutation):
 AGCGGAATGAAGAGGCTCAG (SEQ ID NO: 125)
 AGCGGAATGAAGAGGCTCAG (SEQ ID NO: 125)
 Sequence spanning nucleotide positions 2234-2254 (splice site corrected):
 ATGCTGTGCTTTCTGCTCAG (SEQ ID NO: 126)
 ATGCTGTGCTTTCTGCTCTC (SEQ ID NO: 127)
 Sequence spanning nucleotide positions 1104-1124 (splice site, branchpoint corrected):
 TTTGTAATGCATTGATCCAG (SEQ ID NO: 128)
 TTTGTAATGCCTTGATACAA (SEQ ID NO: 129)
 Sequence spanning nucleotide positions 1887-1907 (splice site corrected):
 TTGGAATCTACCTCCAGCAG (SEQ ID NO: 130)
 TTGGAATCTACCTCCAACAA (SEQ ID NO: 131)

[0154] A skilled practitioner would, however, appreciate that other nucleotide changes could be made to disrupt the cryptic splice sites identified and would be able to envision and design such changes based on results presented herein. In some embodiments, the ABCA4 exon editor has one or more of the nucleotide changes at the above positions. See also Table 3. In some embodiments, the ABCA4 exon editor has a CDS sequence having at least 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 99.9% sequence identity or having 100% sequence identity to any one of the sequences set forth SEQ ID NOs: 56-59. In some embodiments, the ABCA4 exon editor has a cryptic splice site-mitigating nucleotide change at one or more of the above positions (numbering according to, e.g., SEQ ID NO: 56) and has a CDS sequence having at least 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 99.9% sequence identity or having 100% sequence identity to any one of the sequences set forth SEQ ID NOs: 56-58. In some embodiments, the ABCA4 exon editor has a cryptic splice site-mitigating nucleotide change at one or more of the above positions (numbering according to, e.g.,

SEQ ID NO: 56) and has a CDS sequence having at least 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 99.9% sequence identity or having 100% sequence identity to the sequence set forth SEQ ID NO: 56. In some embodiments, an ABCA4 exon editor has cryptic splice site-mitigating nucleotide changes at least one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, or 45 of the positions set forth in Table 3.

[0155] In some embodiments, an ABCA4 exon editor has cryptic splice site-mitigating nucleotide changes at at least one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 of the positions set forth in Table 3 that have a frequency of splice site usage of at least 10 counts per million (also referred to herein as CPMs), with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093). In some embodiments, the ABCA4 exon editor molecule has a cryptic splice site-mitigating nucleotide change at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or all cryptic splice sites having a frequency of usage of at least 10 cpms [with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093)], as measured by an assay described in Example 7 below.

[0156] In some embodiments, an ABCA4 exon editor has cryptic splice site-mitigating nucleotide changes at at least one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 of the positions set forth in Table 3 that have a frequency of splice site usage of at least 5 counts per million (also referred to herein as CPMs), with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093). In some embodiments, the ABCA4 exon editor molecule has a cryptic splice site-mitigating nucleotide change at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or all cryptic splice sites having a frequency of usage of at least 5 cpms [with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093)], as measured by an assay described in Example 7 below.

[0157] In some embodiments, an ABCA4 exon editor has cryptic splice site-mitigating nucleotide changes at at least one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 of the positions set forth in Table 3 that have a frequency of splice site usage of at least 3 counts per million (also referred to herein as CPMs), with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093). In some embodiments, the ABCA4 exon editor molecule has a cryptic splice site-mitigating nucleotide change at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or all cryptic splice sites having a frequency of usage of at least 3 cpms [with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093)], as measured by an assay described in Example 7 below.

[0158] In some embodiments, an ABCA4 exon editor has cryptic splice site-mitigating nucleotide changes at at least one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 of the positions set forth in Table 3 that have a frequency of splice site usage of at least 2 counts per million (also referred to herein as CPMs), with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093). In some embodiments, the ABCA4 exon editor molecule has a cryptic splice site-mitigating nucleotide change at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or all cryptic splice sites having a frequency of usage of at

least 2 cpms [with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093)], as measured by an assay described in Example 7 below.

[0159] In some embodiments, an ABCA4 exon editor has cryptic splice site-mitigating nucleotide changes at at least one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 of the positions set forth in Table 3 that have a frequency of splice site usage of at least 1 count per million (also referred to herein as CPMs), with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093). In some embodiments, the ABCA4 exon editor molecule has a cryptic splice site-mitigating nucleotide change at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 or all cryptic splice sites having a frequency of usage of at least 1 cpms [with the exception of those sites listed in Table 3 that do not include a splice site (e.g., positions 3083, 3103, or 3093)], as measured by an assay described in Example 7 below.

[0160] In addition to ABCA4 exon editors, embodiments described herein include exon editors specific for other gene targets with cryptic splice site-mitigating nucleotide changes. Cryptic splice site-mitigating nucleotide changes can include changes that eliminate or reduce the ability of a cryptic splice site to be used in a splicing reaction. For example, a cryptic splice site identified in the context of an RNA exon editor typically comprises a splice site, a polypyrimidine tract, and a branchpoint. In some embodiments, one or more nucleotide changes may be introduced into at least one of a splice site, a polypyrimidine tract, or a branchpoint, or any combination thereof of a cryptic splice site identified in the context of an RNA exon editor. In some embodiments, the nucleotide change is determined so as to minimize the potential impact on a protein encoded thereby. A person of skill in the art would appreciate that if a nucleotide change made to reduce the frequency of cryptic splice site usage also altered the amino acid encoded by a trans-spliced RNA, conservative amino acid changes would be preferred over non-conservative amino acid changes. Moreover, such a skilled person could readily analyze the protein sequence and structure with an eye toward functional domains and significant sequences therein to evaluate whether such changes could reasonably be expected to alter function of a protein encoded by a trans-spliced protein. A skilled person could also test a protein comprising such an amino acid change to determine if biological activity is altered using assays known in the art. In some embodiments, more than one nucleotide is changed within a cryptic splice site identified. Under some circumstances, a determination of how many nucleotides should be changed is made empirically based on in silico predictions and/or experimental results. In some embodiments, one or more (also referred to herein as at least one) synonymous mutations may be introduced into at least one of a splice site, a polypyrimidine tract, or a branchpoint, or any combination thereof of a cryptic splice site identified in the context of an RNA exon editor. Synonymous mutations do not alter the amino acid sequence of a protein encoded by a trans-spliced RNA. In some embodiments, more than one synonymous mutation may be introduced into at least one of a splice site, a polypyrimidine tract, or a branchpoint, or any combination thereof of a cryptic splice site identified in the context of an RNA exon editor.

[0161] Further to the above, the present inventors analyzed experimental results and sequence information generally as follows. Changes to remove cryptic splice sites identified experimentally were made by searching for and replacing certain elements of splice donor sites. AG

sites (and more strongly CAG sites) at the end of a splice donor site were prioritized for introduction of nucleotide changes. If an AG site was not found or could not be changed without introducing a non-synonymous mutation, then the sequence 42 - 4 basepairs upstream of the splice site was scanned for branch points (sequences matching YNAH). Any such branch point sequences identified were then analyzed and considered for introduction of a nucleotide mutation/s to reduce cryptic splice site usage at the experimentally identified cryptic splice site. In addition, the sequences were also scanned for the presence of polypyrimidine tracts (multiple Ys immediately upstream of the terminal AG). Typically, such polypyrimidine tracts comprise at least 5 pyrimidines within 10 base pairs upstream of the splice site. Once identified, such polypyrimidine tracts were then analyzed and considered for introduction of a nucleotide mutation/s to reduce cryptic splice site usage at the experimentally identified cryptic splice site.

[0162] In some embodiments, the cryptic splice site, or off-target splice site, that is changed to mitigate off-target splicing is a site that has been identified empirically as a site of off-target splicing. Such sites can be identified, for example, using techniques described in Examples 6 and 7 below. In some embodiments, all cryptic splice sites that have a frequency of usage above a predetermined threshold are changed by cryptic splice site-mitigating nucleotide changes.

[0163] In some embodiments, the cryptic splice site, or off-target splice site, that is changed to mitigate off-target splicing is a site that has been predicted to be a site of off-target splicing. Such predictions can be made based on sequence analysis to identify a canonical splice site, polypyrimidine tract, and/or branchpoint of a putative cryptic splice site therein. Cryptic splice sites outnumber authentic splice sites in the human genome by an order of magnitude and may have the same splicing recognition sequences as authentic splice sites but are normally repressed by thus far poorly understood molecular mechanisms. Cryptic 5' splice sites frequently have a consensus NNN/GUNNNN or NNN/GCNNNN where N is any nucleotide and "/" is the exon-intron boundary. Cryptic 3' splice sites frequently have a consensus NAG/N. Cryptic splice site activation is positively influenced by surrounding nucleotides that may serve to make cryptic splice sites appear more like those of a typical consensus of authentic splice sites, namely MAG/GURAGU and YAG/G, respectively, where M is C or A, R is G or A, and Y is C or U.

[0164] In some embodiments, a cryptic splice site-mitigating nucleotide change causes a nucleotide sequence that matches a canonical splice site consensus sequence to no longer match the canonical sequence. In some embodiments, a cryptic splice site-mitigating nucleotide change eliminates a potential splice site nucleotide. In some embodiments, a cryptic splice site-mitigating nucleotide change eliminates a potential polypyrimidine tract nucleotide. In some embodiments, a cryptic splice site-mitigating nucleotide change eliminates a potential branch point nucleotide. In some embodiments, a cryptic splice site-mitigating nucleotide change is a synonymous nucleotide change. In some embodiments, a cryptic splice site-mitigating nucleotide change causes a change in an amino acid encoded by the exon editor. In some embodiments, the amino acid change is a conservative amino acid substitution.

[0165] As described in, for example, Table 3 and Examples 6 and 7, the present inventors experimentally observed cryptic splice site usage in RNA exon editors described herein. Further

analysis of the cryptic splice sites localized/identified led to the identification of putative splice sites, polypyrimidine tracts and branchpoints within the experimentally identified cryptic splice sites, which in turn provided guidance regarding what sequences should be modified to achieve mitigation of cryptic splice site engagement. These results are broadly applicable to RNA exon editors in general.

[0166] In some embodiments, an RNA molecule is described herein comprising, in a 5' to 3' direction: (a) a cDNA coding domain sequence; (b) a splice donor sequence; and (c) a binding domain sequence configured to bind to an intron of an endogenous RNA molecule; wherein the coding domain sequence comprises a nucleotide mutation relative to the endogenous RNA molecule sequence, wherein the nucleotide mutation disrupts a cryptic splice site within the coding domain sequence. In some embodiments, the nucleotide mutation is a synonymous nucleotide mutation. In some embodiments, the cryptic splice site is identified experimentally. In some embodiments, the cryptic splice site is predicted based on in silico analysis.

[0167] Also encompassed herein is a method for modifying an RNA molecule in a cell, the method comprising providing to the cell an exogenous RNA molecule comprising, in a 5' to 3' direction: (a) a cDNA coding domain sequence comprising a nucleotide mutation that differs from that of an endogenous target RNA molecule in the cell; (b) a splice donor sequence configured to splice to a splice acceptor sequence of the endogenous target RNA molecule; and (c) a binding domain sequence configured to bind to an intron of the endogenous target RNA molecule; wherein the nucleotide mutation disrupts a cryptic splice site within the coding domain sequence of the exogenous RNA molecule. In some embodiments, the nucleotide mutation is a synonymous nucleotide mutation. In some embodiments, the cryptic splice site is identified experimentally. In some embodiments, the cryptic splice site is predicted based on in silico analysis.

[0168] Also encompassed herein is a method of increasing trans-splicing efficiency of an RNA exon editor comprising introducing a mutation into a coding domain sequence of the RNA exon editor, wherein the mutation disrupts a cryptic splice site in the coding domain sequence of the RNA exon editor. In some embodiments, the nucleotide mutation is a synonymous nucleotide mutation. See, e.g., Table 6. In some embodiments, the cryptic splice site is identified experimentally. In some embodiments, the cryptic splice site is predicted based on in silico analysis.

Binding Domains

[0169] ABCA4 trans-splicing molecules described herein feature a binding domain (BD) configured to bind/anneal a target ABCA4 intron and/or exon. In some instances, the target ABCA4 intron is ABCA4 intron 22. In one embodiment, the binding domain is a nucleic acid sequence that is at least 80% complementary to (e.g., at least 85% complementary to, at least 90% complementary to, at least 91% complementary to, at least 92% complementary to, at least 93% complementary to, at least 94% complementary to, at least 95% complementary to, at least 96% complementary to, at least 97% complementary to, at least 98% complementary to, at least 99% complementary to, or 100% complementary to) a sequence of the target ABCA4 intron pre-mRNA (e.g., a target ABCA4 intron), which may suppress endogenous target cis-splicing while enhancing trans-splicing between the trans-splicing molecule and the target ABCA4 pre-mRNA (e.g., by creating a chimeric molecule having a portion of endogenous ABCA4 mRNA and a coding domain sequence having one or more functional

ABCA4 exons that encode wildtype ABCA4 amino acid sequences). In one embodiment involving trans-splicing molecule-encoding sequences (e.g., vectors encoding trans-splicing molecules), the binding domain-encoding sequence encodes a nucleic acid sequence that is at least 80% complementary to (e.g., at least 85% complementary to, at least 86% complementary to, at least 87% complementary to, at least 88% complementary to, at least 89% complementary to, at least 90% complementary to, at least 91% complementary to, at least 92% complementary to, at least 93% complementary to, at least 94% complementary to, at least 95% complementary to, at least 96% complementary to, at least 97% complementary to, at least 98% complementary to, at least 99% complementary to, or 100% complementary to) a sequence of the target ABCA4 intron pre-mRNA.

[0170] In some instances, the present invention provides trans-splicing molecules (or vectors thereof) that bind ABCA4 at intron 22, e.g., wherein the nucleic acid trans-splicing molecule is configured to trans-splice a CDS to endogenous ABCA4 exon 23. In particular, trans-splicing molecules described herein include those in which the binding domain binds a binding site having any one or more (e.g., six or more, eight or more, ten or more, or twelve or more) of nucleotides 1 to 510 or 880 to 1,350 of SEQ ID NO: 16.

[0171] In some instances, the binding domain includes any six or more consecutive nucleotides within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22 (e.g., any eight or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any ten or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 12 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 20 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 30 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 40 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 50 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 100 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 150 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, any 200 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, or any 250 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22).

[0172] In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 880 to 1,350 of ABCA4 intron 22 (e.g., any eight or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any ten or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 12 or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 20 or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 30 or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 40 or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 50 or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 100 or more consecutive nucleic acids within

nucleotides 880 to 1,350 of ABCA4 intron 22, any 150 or more consecutive nucleic acids within nucleotides 880 to 1,350 of ABCA4 intron 22, any 200 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22, or any 250 or more consecutive nucleic acids within nucleotides 1 to 510 or 880 to 1,350 of ABCA4 intron 22).

[0173] In some instances, a binding domain has at least two non-overlapping sequences with at least 80% complementarity to the binding site.

[0174] In some instances, the binding domain includes a nucleic acid sequence having at least 80% identity (e.g., at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to
 UUGUUGGUAAGGGUGUACAGCAUUUCCUGGCUAUAGUUGCAAAAUCAGGAGUUAACUAA
 AAAAAAACCCAAGGGAACUAAUUCAGCAGCAAAUCCAGCAUUAUUGGACAAUUAUACCAAC
 AUUUCUAGCUUCCUACAUAC (SEQ ID NO: 17).

[0175] In some embodiments, the binding domain is a DNA sequence having at least 80% identity (e.g., at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 18
 TTGTTGGTAAAGGGTGTACAGCAATTTCTGGCTATATAGTTGCAAAAATCAGGAGTAACTAAAAA
 AAAAAACCCAAGGGAACATAATTCAGCAGCAAATTCAGCATATTGGGACAATAATAACCAACATTTC
 ATAGCTTCCTACATAC (SEQ ID NO: 18).

[0176] In some instances, the binding domain includes a nucleic acid sequence having at least 80% identity (at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 19.

[0177] In some embodiments, the binding domain is a DNA sequence having at least 80% identity (e.g., at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 20.

[0178] In some instances, the binding domain includes a nucleic acid sequence having at least 80% identity (at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 21.

[0179] In some embodiments, the binding domain is a DNA sequence having at least 80% identity (e.g., at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least

87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 22.

[0180] In some instances, the binding domain includes a nucleic acid sequence having at least 80% identity (e.g., at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 23.

[0181] In some embodiments, the binding domain is a DNA sequence having at least 80% identity (e.g., at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identity; e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to SEQ ID NO: 24.

[0182] A binding domain can be operably linked 3' to a splicing domain (e.g., directly connected to a splicing domain or have intervening sequences connecting the 3' end of the splicing domain and the 5' end of the binding domain).

[0183] As detailed herein, the first stage of ABCA4 Exon Editor design began with the screening and selection of a highly efficient BD sequence, which is complementary to the targeted pre-mRNA intron. To achieve a high CDS replacement potential and provide an Exon Editor useful for correcting mutations in a high percentage of the ABCA4-related retinopathy patient population within the limits of AAV packaging capacity (~4.7 Kb), intron 22 of the ABCA4 gene was selected for a BD screen. The present inventors applied both an NGS-based BD library approach and a bioluminescence-based approach to scan along the intron 22 sequence with exon editor variants that differed in their BD sequence. Results from both methods aligned and successfully identified regions within the intron that were amenable to RNA exon editing, leading to the selection of BD117 in some embodiments. See FIG. 16.

Splicing Domains

[0184] The splicing domain can include a splice site, a branch point, and/or a polypyrimidine tract (PPT) tract to mediate trans-splicing. In some embodiments, a splicing domain has a single splice site, which denotes that the splice site is designed for preferential trans-splicing, but not cis-splicing, due to the lack of a corresponding splice site.

[0185] Alternative splicing domains may be selected by one of skill in the art according to known methods and principles. In one embodiment, the 5' splice site consensus sequence is the nucleic acid sequence AG/GURAGU (where / indicates the splice site). In another embodiment, the endogenous splice sites that correspond to the exon and intron proximal to the splice site can be employed to maintain any splicing regulatory signals.

[0186] In one embodiment, a suitable 5' splice site with spacer is: 5'- GTA AGA GAG CTC GTT GCG ATA TTA T-3' (SEQ ID NO: 25). In one embodiment, a suitable 5' splice site is AGGT. In some embodiments, the splice site (e.g., 5' splice site) is, or comprises, GTAAGT or GUAAGT. In

some embodiments, the splice site (e.g., 5' splice site) is, or comprises, GTAAGG, GUAAGG, GTAAGC, GUAAGC, GTAACU, GUAACU, CAAAGT, or CAAAGU.

[0187] A splicing domain can be operably linked 5' to a terminator domain (e.g., directly connected to a terminator domain or have intervening sequences connecting the 3' end of the splicing domain and the 5' end of the terminator domain, e.g., a linker domain and/or a 3' downstream sequence).

[0188] 5' Untranslated Region

[0189] In some instances, the nucleic acid trans-splicing molecule includes a 5' untranslated region. In some embodiments, the 5' untranslated region comprises, consists essentially of, or consists of a native 5' ABCA4 untranslated region. In some embodiments, the 5' untranslated region comprises a sequence having at least 80% sequence identity (e.g., at least 85% sequence identity, at least 86% sequence identity, at least 87% sequence identity, at least 88% sequence identity, at least 89% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to any one of SEQ ID NOS: 12-13.

[0190] In some embodiments, a 5' untranslated region can be operably linked 5' to a CDS (e.g., directly connected to a CDS or have intervening sequences connecting the 3' end of the 5' untranslated region and the 5' end of the CDS).

5' Regulatory Domains

[0191] In some instances, the nucleic acid trans-splicing molecule is operatively linked to a 5' regulatory domain operatively linked 5' to the CDS (e.g., directly linked to the CDS, or linked through an intermediate domain, e.g., an untranslated region). A 5' regulatory domain can include a promoter (e.g., a constitutive promoter, e.g., CMV promoter or an EF1-alpha promoter). In some instances, the 5' regulatory domain includes a promoter (e.g., a constitutive promoter, e.g., CMV promoter) operatively linked to a native 5' ABCA4 untranslated region. In some embodiments, the 5' regulatory domain operatively linked to a native 5' ABCA4 untranslated region comprises a sequence having at least 80% sequence identity (e.g., at least 85% sequence identity, at least 86% sequence identity, at least 87% sequence identity, at least 88% sequence identity, at least 89% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) to any one of SEQ ID Nos: 14-15.

[0192] In some embodiments, a 5' regulatory domain can be operably linked 5' to a CDS (e.g., directly connected to a CDS or have intervening sequences connecting the 3' end of the 5' regulatory domain and the 5' end of the CDS).

Linker Domains

[0193] As discussed herein, increasing trans-splicing efficiency remains an important goal for implementation of nucleic acid trans-splicing molecules as therapeutic agents. In an effort to engineer nucleic acid trans-splicing molecules having trans-splicing efficiencies that could satisfy the requirements for their use in therapeutic intervention, the present inventors tested thousands of linker sequences between the splice domain (SD) and binding domain (BD) and identified an exemplary linker, termed 40mer, that improved performance relative to that of a progenitor linker (LinkD) to a statistically significant degree. See, for example, FIGs. 17A and 17B.

[0194] Further to the above, nucleic acid trans-splicing molecules may include a linker domain at one or more positions with the molecule. In some embodiments, the linker domain is operatively linked 3' to the splicing domain or splice site (e.g., directly connected to the splicing domain or splice site). The linker domain may be any suitable size. In some embodiments, the linker domain is longer than 25 nucleotides in length (e.g., between 25 and 50 nucleotides in length, between 35 and 45 nucleotides in length, or about 40 nucleotides in length (e.g., a 40-mer linker)). In some instances, the linker domain comprises, consists essentially of, or consists of a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 86% sequence identity, at least 87% sequence identity, at least 88% sequence identity, at least 89% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CUGGUGCCCGCGGGCCGCGGAACCGGUUGGGGGCAUGUAC (SEQ ID NO: 26) or CTGGTGCCCGCGGGCCGCGGAACCGGTTGGGGGCATGTAC (SEQ ID NO: 27).

[0195] As described above, the 40mer linker, comprising SEQ ID NO: 26 (RNA) or SEQ ID NO: 27 (DNA), was selected from a library screen wherein thousands of different elements were assessed for function in the context of trans-splicing molecules. The superiority of the 40mer linker is, for example, evident in FIG. 1, which compares three different linkers in the context of different combinations of elements. In each context, the presence of the 40mer linker (indicated as the Linker corresponding to the lightest grayscale tone; the third linker indicated from left to right in the context of each ss) was associated with the highest relative trans-splicing efficiency. A linker sequence is frequently included in trans-splicing molecules, wherein it is positioned between the splice donor and the binding domain to offer flexibility and accessibility to each element. As described herein, functional contributions of different elements were assessed in different combinations for activity conferred thereby in the context of trans-splicing molecules. The 40mer linker emerged from this screen as a lead candidate element that conferred improved trans-splicing activity in the context of trans-splicing molecules, as exemplified by a statistically significant increase in on-target (ONT) % replacement of ABCA4 RNA (FIG. 17A) and a significant increase in trans-spliced ABCA4 protein (ONT) levels and a significant decrease in non-spliced protein (NSP) levels relative to LinkD (FIG. 17B).

[0196] In other instances, the linker domain comprises, consists essentially of, or consists of a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least

86% sequence identity, at least 87% sequence identity, at least 88% sequence identity, at least 89% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with CCGAAUACGACACGUACAAGAUCU (SEQ ID NO: 28) or CCGAATACGACACGTACAAGATCT (SEQ ID NO: 29). In other instances, the linker domain comprises, consists essentially of, or consists of a nucleic acid sequence having at least 80% identity (e.g., at least 85% sequence identity, at least 86% sequence identity, at least 87% sequence identity, at least 88% sequence identity, at least 89% sequence identity, at least 90% sequence identity, at least 91% sequence identity, at least 92% sequence identity, at least 93% sequence identity, at least 94% sequence identity, at least 95% sequence identity, at least 96% sequence identity, at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with GCUGCUGCUCAGUCUCCUGGGCUGG (SEQ ID NO: 30) or GCTGCTGCTCAGTCTCCTGGGCTGG (SEQ ID NO: 31).

Potentiators of RNA Exon Editor Activity

[0197] The present inventors used a multi-faceted experimental approach to engineer nucleic acid trans-splicing molecules having properties advantageous for their use in therapeutic intervention. Accordingly, the present inventors sought to identify additional elements that could enhance RNA exon editor activity by, for example, increasing on target (ONT) trans-splicing, increasing ONT ABCA4 trans-spliced protein levels, or decreasing non-spliced protein (NSP) levels, or any combination thereof. Accordingly, a variety of elements were tested to determine if the presence of a tested element potentiated the activity of an RNA exon editor relative to that observed for the RNA exon editor without the tested element. Through extensive trial and error procedures whereby scores of different elements and different combinations thereof were assessed in the context of individual trans-splicing molecules, the present inventors identified a sequence element, a U1 binding site (UBS), that conferred both an increase in ONT trans-splicing and a decrease in NSP levels to RNA exon editors comprising a UBS as an additional element relative to an RNA exon editor lacking the UBS. See, for example, FIG. 13 (3xU1) and FIGs. 19A and 19B (3xUBS). In some embodiments, 3xUBS (also referred to as 3xU1) comprises SEQ ID NO: 62. In some embodiments, 3xUBS (also referred to as 3xU1) comprises at least one sequence comprising GGTAAGT or GGTAAGTAGAGTG (SEQ ID NO: 61) or GGTGAGTAGAGTG (SEQ ID NO: 83). In a particular embodiment thereof, 3xUBS is positioned 3' of the linker domain. In a particular embodiment thereof, 3xUBS is positioned 3' of the linker domain and 5' to the binding domain. In a particular embodiment thereof, 3xUBS comprises at least one sequence comprising GGTAAGT, wherein each of the at least one sequences comprising GGTAAGT further comprises flanking sequences. In embodiments thereof, 3xUBS comprises at least one of GGTAAGTAGAGTG (SEQ ID NO: 61) or GGTGAGTAGAGTG (SEQ ID NO: 83). In some embodiments, 3xUBS comprises at least one of

CGT GGTAAGT AGAGTG (SEQ ID NO: 74; UBS1); ATG GGTAAGT AGAGTG (SEQ ID NO: 75; UBS2); or TGG GGTAAGT AGAGTG (SEQ ID NO: 76; UBS3), or any combination thereof.

[0198] Accordingly in some embodiments, an RNA exon editor comprising at least one UBS comprises at least one sequence comprising GGTAAGT or GGTAAGTAGAGTG (SEQ ID NO: 61). In a particular embodiment thereof, the 3xUBS is positioned 3' of the linker domain in the RNA exon editor. In a particular embodiment thereof, the 3xUBS is positioned 3' of the linker domain and 5' to the binding domain in the RNA exon editor. In some embodiments, an RNA exon editor comprising at least one UBS comprises at least one of CGT GGTAAGT AGAGTG (UBS1; SEQ ID NO: 74); ATG GGTAAGT AGAGTG (UBS2; SEQ ID NO: 75); or TGG GGTAAGT AGAGTG (UBS3; SEQ ID NO: 76); or any combination thereof. In some embodiments, an RNA exon editor comprising at least one UBS comprises at least one CGT GGTAAGT AGAGTG (UBS1; SEQ ID NO: 74). In some embodiments, an RNA exon editor comprising at least one UBS comprises CGT GGTAAGT AGAGTG (SEQ ID NO: 74; UBS1), ATG GGTAAGT AGAGTG (SEQ ID NO: 75; UBS2), and TGG GGTAAGT AGAGTG (SEQ ID NO: 76; UBS3).

[0199] In accordance with the results presented herein, at least one UBS present in the context of a trans-splicing molecule is referred to herein as a potentiator. An exemplary potentiator described herein comprises at least one of CGT GGTAAGT AGAGTG (UBS1; SEQ ID NO: 74), ATG GGTAAGT AGAGTG (UBS2; SEQ ID NO: 75), or TGG GGTAAGT AGAGTG (UBS3; SEQ ID NO: 76), or any combination thereof. In some embodiments, an exemplary potentiator described herein comprises at least one UBS comprising at least one CGT GGTAAGT AGAGTG (UBS1; SEQ ID NO: 74). In some embodiments, an exemplary potentiator described herein comprises a UBS comprising CGT GGTAAGT AGAGTG (UBS1; SEQ ID NO: 74), ATG GGTAAGT AGAGTG (UBS2; SEQ ID NO: 75), and TGG GGTAAGT AGAGTG (UBS3; SEQ ID NO: 76).

[0200] The ability of a UBS (included as an additional element in an RNA exon editor and positioned therein 3' relative to the linker and 5' relative to the BD; See FIG. 13) to enhance RNA exon editor activity was surprising at least because it was reasonable to predict that the presence of at least one additional element that could potentially act as an alternative 5' splice site might reduce ONT trans-splicing by potentially competing with and/or interfering with the splice domain (See, e.g., FIG. 13; SD; existing *bona fide* 5' splice site) of the RNA exon editor, which competition/interference would reduce ONT trans-splicing. Accordingly, the increase in ONT trans-splicing observed as conferred by the presence of a potentiator (at least one of a UBS, e.g., UBS1, UBS2, or UBS3, or any combination thereof) in an RNA exon editor was unexpected and surprising. The potentiator also conferred a statistically significant decrease in NSP, which observation further underscored the surprising functional properties of this element in the context of an RNA exon editor.

[0201] The functional properties of potentiators described herein have also been demonstrated in the context of RNA exon editors comprising different binding domains that bind/target different introns in different target pre-mRNAs other than ABCA4 pre-mRNA. Results from exemplary RNA exon editors demonstrate that a potentiator comprising 3xUBS confers surprising beneficial properties in the context of RNA exon editors targeting Intron A and Intron B in non-ABCA4

pre-mRNAs. More particularly, RNA exon editors targeting non-ABCA4 pre-mRNA that comprise a combination of 40mer + 3xUBS (SEQ ID NO: 62) exhibited significantly reduced levels of NSP relative to RNA exon editors comprising only 40mer (SEQ ID NO: 27). See, e.g., FIGs. 27B and 28B. While reducing NSP, the combination of 40mer + 3xUBS (SEQ ID NO: 62) maintained similar levels of ONT trans-spliced protein relative to that of 40mer alone. By way of comparison, each of the RNA exon editors in which 3xUBS was shown to reduce NSP [Intron 22 targeting ABCA4 pre-mRNA (FIG. 19B); Intron A (FIG. 27B); and Intron B (FIG. 28B)] included the same cassette of 40mer + 3xUBS sequence, which supports a conclusion that the function conferred by the presence of the 3xUBS sequence (SEQ ID NO: 62) is consistent irrespective of the target intron and/or target pre-mRNA of the RNA exon editor.

[0202] Functionality conferred by the presence of the 3xUBS sequence (SEQ ID NO: 62) in nucleic acid trans-splicing molecules was also evident on the level of % RNA trans-splicing. FIG. 19A (Intron 22 targeting ABCA4 pre-mRNA), FIG. 28A (Intron B), and FIG. 29 (Intron C) demonstrate that the presence of the 3xUBS sequence (SEQ ID NO: 62) can confer an increase in % RNA trans-splicing as reflected by a trend toward increased % replacement of RNA. Indeed, in the exemplary trans-splicing molecule targeting Intron C, the presence of the 3xUBS sequence (SEQ ID NO: 62) conferred a 3-fold increase in the level of % replacement. These results suggest that this additional functional property of the 3xUBS sequence (SEQ ID NO: 62) to promote a trend toward increased % replacement is consistent irrespective of the target intron and/or target pre-mRNA of the RNA exon editor. It is noteworthy that the RNA exon editor comprising a binding domain (BD34) that targets Intron C included an alternative 5' splice site that differs from the 5' splice site used in each of the RNA exon editors comprising binding domains that target Intron 22 of ABCA4 pre-mRNA, Intron A, and Intron B. These results demonstrate that the 3xUBS potentiator can confer functionality in the context of RNA exon editors that differ with respect to elements in addition to binding domain and pre-mRNA target specificity.

[0203] Accordingly, the present inventors have presented evidence that potentiators such as those described herein (e.g., SEQ ID NO: 62) function in a binding domain- and target-independent manner and therefore, may be used in the context of RNA exon editors in general with an expectation that the presence of a potentiator will confer advantageous properties, including reduced NSP levels and potentially increased trans-splicing activity as reflected in % RNA replacement.

Translational Potentiator of RNA Exon Editor Activity

[0204] The search for additional elements that could enhance RNA exon editor activity by increasing on target (ONT) trans-splicing, increasing ONT ABCA4 trans-spliced protein levels, or decreasing non-spliced protein (NSP) levels, or any combination thereof also led the present inventors to the discovery that AU-rich elements (ARE) exhibit surprising functional properties in the context of RNA exon editors described herein. The surprising properties of AREs were discovered via extensive trial and error procedures whereby scores of different elements and different combinations thereof were assessed in the context of individual trans-splicing molecules. The present inventors showed that the presence of an ARE conferred a decrease in NSP levels to RNA exon editors

comprising an ARE as an additional element (40mer + ARE) relative to those lacking an ARE (e.g., comprising the 40mer alone). In some experiments, the presence of an ARE also conferred an increase in ONT trans-spliced ABCA4 protein to RNA exon editors comprising an ARE as an additional element (40mer + ARE) relative to those lacking an ARE (e.g., comprising the 40mer alone). See, for example, FIG. 13 (ARE) and FIG. 19B (40mer + ARE). Since AREs conferred a decrease in NSP levels and an increase in ONT trans-spliced ABCA4 protein levels, they are referred to herein as “translational potentiators”.

[0205] In some embodiments, an exemplary ARE is a granulocyte-macrophage colony stimulating factor (GM-CSF) ARE comprising a sequence comprising ATTTATATATTTATATTTTAAATATTTATTTATTTATTTATTTA (SEQ ID NO: 63) or a sequence having at least 90% identity to SEQ ID NO: 63. In some embodiments, the translational potentiator has at least 91% identity to SEQ ID NO: 63, has at least 92% identity to SEQ ID NO: 63, has at least 93% identity to SEQ ID NO: 63, has at least 94% identity to SEQ ID NO: 63, has at least 95% identity to SEQ ID NO: 63, has at least 96% identity to SEQ ID NO: 63, has at least 97% identity to SEQ ID NO: 63, has at least 98% identity to SEQ ID NO: 63, or has at least 99% identity to SEQ ID NO: 63. In some embodiments, SEQ ID NO: 63 or the sequence having at least 90% identity to SEQ ID NO: 63 is operably linked 3' to the potentiator (e.g., SEQ ID NO: 62). The core sequence is typically AUUUA (RNA; ATTTA in the DNA vector encoding same) and repeats of this core AUUUA element are often required for function. An ARE element is typically about equal to or greater than 50 bases. Accordingly, in some embodiments the ARE comprises repeats of the core AUUUA element about equal to or greater than 50 bases.

[0206] The ability of an ARE (included as an additional element in an RNA exon editor; See FIG. 13) to enhance RNA exon editor activity was surprising at least because it increased ONT trans-spliced ABCA4 protein levels, while decreasing NSP levels. Accordingly, ARE-mediated effects on protein levels encoded thereby promote the therapeutic potential of the RNA exon editor by preferentially increasing therapeutic protein levels (ONT trans-spliced ABCA4 protein levels) and decreasing NSP levels.

[0207] It is also noteworthy that the cassette of an exemplary potentiator in combination with an exemplary translational potentiator (40mer + 3XUBS + ARE; SEQ ID NO: 73) confers an even more significant decrease in NSP levels relative to that of RNA exon editors comprising 40mer only, 40mer + 3XUBS, or 40mer + ARE. This combination of elements (40mer + 3XUBS + ARE; SEQ ID NO: 73), therefore, exhibits particularly advantageous properties in the context of trans-splicing molecules such as those described herein.

[0208] The functional properties of translational potentiators described herein have also been demonstrated in the context of RNA exon editors comprising different binding domains that bind/target different introns in different target pre-mRNAs other than ABCA4 pre-mRNA. Results from exemplary RNA exon editors demonstrate that a translational potentiator comprising ARE (SEQ ID NO: 63) confers surprising beneficial properties in the context of RNA exon editors targeting Introns A and B in non-ABCA4 pre-mRNAs. More particularly, RNA exon editors targeting non-ABCA4 pre-

mRNA that comprise a combination of 40mer + ARE (SEQ ID NO: 63) exhibited significantly reduced levels of NSP relative to RNA exon editors comprising only 40mer (SEQ ID NO: 27). See, e.g., FIGs. 27B and 28B. While reducing NSP, the combination of 40mer + ARE (SEQ ID NO: 63) maintained similar levels of ONT trans-spliced protein relative to that of 40mer alone. By way of comparison, each of the RNA exon editors in which ARE was shown to reduce NSP [Intron 22 targeting ABCA4 pre-mRNA (FIG. 19B); Intron A (FIG. 27B); and Intron B (FIG. 28B)] included the same cassette of 40mer + ARE sequence (sequence 3444-3529 of SEQ ID NO: 86), which supports a conclusion that the function conferred by the presence of the ARE sequence (SEQ ID NO: 63) is consistent irrespective of the target intron and/or target pre-mRNA of the RNA exon editor. The reduction in NSP levels was even more pronounced with the cassette of 40mer + 3xUBS + ARE (SEQ ID NO: 73), indicating a combinatorial effect of these elements. These results demonstrate that ARE translational potentiators can confer functionality in the context of RNA exon editors that differ with respect to binding domain and pre-mRNA target specificity and can act in a combinatorial fashion with other elements to promote RNA exon editor efficacy.

[0209] Accordingly, the present inventors have presented evidence that translational potentiators such as those described herein (e.g., SEQ ID NO: 63) function in a binding domain- and target-independent manner and therefore, may be used in the context of RNA exon editors in general with an expectation that the presence of a translational potentiator will confer advantageous properties, including reduced NSP levels.

Additional Components or Modifications

[0210] In some instances, the trans-splicing molecule includes a 3' transcription terminator domain. In some embodiments, such 3' transcription terminator domains form a triple helical structure that effectively caps the 3' end of the trans-splicing molecule. In some instances, the 3' transcription terminator domain is from the human long non-coding RNA MALAT1 (e.g., wildtype MALAT1). In some embodiments, the 3' transcription terminator domain includes a tRNA-like domain. 3' transcription terminator domains useful as part of the present ABCA4 trans-splicing molecules are described in International Patent Publication No. WO 2020/214973, which is herein incorporated by reference in its entirety. For example, in some embodiments, the region of an RTM operably linked to the 3' end of the binding domain includes a terminator domain that comprises, consists essentially of, or consists of a wildtype MALAT1+mascRNA domain, such as SEQ ID NO: 32 or SEQ ID NO: 33. In some embodiments, the region of an RNA exon editor operably linked to the 3' end of the binding domain includes a terminator domain that comprises, consists essentially of, or consists of a mutated MALAT1+masc RNA (anti-Mut1 masc RNA) domain, such as SEQ ID NO: 66 (DNA) or SEQ ID NO: 67 (RNA).

[0211] In some instances, the trans-splicing molecule includes a 3' downstream (DS) region downstream of (i.e., operatively linked 3' to) the binding domain and/or upstream of (i.e., operatively linked 5' to) the terminator domain. In some instances, the 3' DS region is part of the terminator domain. The 3' DS region may directly connect the binding domain to the terminator region. In some embodiments, the 3' DS region is 6-10 nucleotides in length (e.g., 8 nucleotides in length). In some

embodiments, the 3' DS region comprises any one of 3' DS- MALAT1 flanking 8-mer RNA, AGGGUCAU; 3' DS- MALAT1 flanking 8-mer DNA, AGGGTCAT; 3' DS- random 8-mer RNA, CGAGCCUC; 3' DS- random 8-mer DNA, CGAGCCTC; 3' DS- MALAT1 flanking first 6 of 8-mer RNA, AGGGU; and 3' DS- MALAT1 flanking first 6 of 8-mer DNA, AGGGT. In some embodiments, the 3' DS region consists of any one of 3' DS- MALAT1 flanking 8-mer RNA, AGGGUCAU; 3' DS- MALAT1 flanking 8-mer DNA, AGGGTCAT; 3' DS- random 8-mer RNA, CGAGCCUC; 3' DS- random 8-mer DNA, CGAGCCTC; 3' DS- MALAT1 flanking first 6 of 8-mer RNA, AGGGU; and 3' DS- MALAT1 flanking first 6 of 8-mer DNA, AGGGT.

[0212] In some embodiments, binding of a trans-splicing molecule to the target pre-mRNA is mediated by percent complementarity (i.e., based on base-pairing characteristics of nucleic acids), triple helix formation, or protein-nucleic acid interaction (as described in documents cited herein). In one embodiment, the nucleic acid trans-splicing molecule includes DNA, RNA, or DNA/RNA hybrid molecules, wherein the DNA or RNA is either single or double stranded. Also included herein are RNAs or DNAs, which can hybridize to one of the aforementioned RNAs or DNAs, preferably under stringent conditions, for example, at 60°C in 2.5x SSC buffer and several washes at 37°C at a lower buffer concentration, for example, 0.5x SSC buffer. These nucleic acids can encode proteins exhibiting lipid phosphate phosphatase activity and/or association with plasma membranes. When trans-splicing molecules are synthesized in vitro, such trans-splicing molecules can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization to the target mRNA, transport into the cell, stability in the cells to enzymatic cleavage, etc. For example, modification of a trans-splicing molecule to reduce the overall charge can enhance the cellular uptake of the molecule. In addition, modifications can be made to reduce susceptibility to nuclease or chemical degradation. The nucleic acid molecules may be synthesized in such a way as to be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

[0213] Various other well-known modifications to the nucleic acid molecules can be introduced as a means of increasing intracellular stability and half-life (see also above for oligonucleotides). Possible modifications are known to the art. Modifications, which may be made to the structure of synthetic trans-splicing molecules include backbone modifications.

Exemplary ABCA4 RNA Exon Editor Cassette Map and Overview

[0214] In general terms, an ABCA4 RNA exon editor cassette (ABCA4 REEC) comprises a regulatory component that drives expression of the RNA exon editor, a 5' UTR and CDS that replaces the corresponding native sequence and any mutations that may exist within the target native sequence, a splice donor (SD) that interacts with the native splice acceptor (SA), a linker that may enable flexibility and accessibility for this interaction to occur, regulatory elements that enhance trans-splicing (TS), a BD that targets the correct position in the pre-mRNA, and a terminator sequence that promotes TS and stability. These features are indicated and briefly described in Figure 13 and in Table 1 below which present an exemplary ABCA4 REEC:

Table 1:

1	AAV2-based Inverted Terminal Repeat Sequences (ITR) essential for AAV packaging of the <i>ABCA4</i> Exon Editor cassette.
2	Cytomegalovirus (CMV) promoter drives <i>ABCA4</i> Exon Editor expression.
3	Native 5' Untranslated Region (UTR) from Human <i>ABCA4</i> mRNA maintains native 5' regulation.
4	<i>ABCA4</i> exons 1-21 comprising synonymous codon changes designed to replace corresponding endogenous coding sequence via RNA exon editing. Synonymous codon changes remove cryptic splice acceptor sites to mitigate potential alternative splicing reactions.
5	<i>ABCA4</i> exon 22 comprising synonymous codon changes that enable sequence-based quantification.
6	Splice donor (SD) trans-splices into endogenous <i>ABCA4</i> intron 22 splice acceptor site leading to replacement of the endogenous <i>ABCA4</i> pre-mRNA upstream sequence with the <i>ABCA4</i> Exon Editor Coding Domain Sequence (CDS) including Exon 22.
7	Double STOP codon prevents translation of amino acids from the <i>ABCA4</i> Exon Editor downstream of Exon 22 (e.g., Linker, 3x U1, AU-Rich Element [ARE], BD domain regions).
8	Linker provides structural flexibility and accessibility to splice donor site.
9	Three tandem U1 binding sites (3x U1) hypothesized to associate with U1 snRNP; demonstrated to increase RNA exon editing efficiency.
10	ARE hypothesized to decrease the stability of any non-spliced <i>ABCA4</i> Exon Editor entering the cytoplasm.
11	Binding domain (BD) mediates targeting of <i>ABCA4</i> Exon Editor to endogenous <i>ABCA4</i> pre-mRNA Intron 22 via sequence complementarity.
12	MALAT1 terminator (e.g., MALAT1-MascRNA anti- mut1) promotes efficient RNA exon editing.
13	AAV2-based ITR essential for AAV packaging of the Exon Editor cassette.

ABCA4 REEC Features

[0215] Details and background regarding the features of the exemplary *ABCA4* REEC embodiment presented in Figure 13 are outlined below.

[0216] Inverted terminal repeat (ITR) sequence: 5' and 3' ITR sequences flank the exemplary RNA Exon Editor sequence and are essential for Adeno-associated virus (AAV) packaging of the transgene. An exemplary *ABCA4* REEC contains “flop orientation” 130 nucleotide ITR sequences derived from AAV2, which are commonly used in gene therapies.

[0217] CMV/CMV promoter: This regulatory sequence drives the expression of the exemplary Exon Editor transcript. It comprises the Cytomegalovirus (CMV) immediate-early enhancer and CMV immediate-early enhancer, which has been shown to drive high expression in non-human primate (NHP) photoreceptors and is shown herein to drive high expression in NHP retina and human retina. See, e.g., FIGs. 8, 9, 31, and 32.

[0218] 5' UTR: The 5' end of the exemplary Exon Editor RNA molecule begins with the native *ABCA4* 5' UTR sequence, which maintains the native RNA sequence in the final ONT mRNA molecule resulting from TS. The UTR sequence has been modified with five point mutations that remove suspected or data-driven cryptic splice acceptor sites to mitigate the potential occurrence of cis-splicing or TS between Exon Editor transcripts.

[0219] Coding sequence (CDS): An exemplary ABCA4 REEC encodes exons 1 to 22 of the ABCA4 mRNA sequence to replace the corresponding native sequence following TS, (additional details regarding Exon 22 are presented below). Exons 1 to 21 have been modified with over 40 synonymous point mutations that remove suspected or data-driven cryptic splice acceptor sites to mitigate the potential occurrence of cis-splicing or TS between Exon Editor transcripts.

[0220] Terminal Exon 22: The encoded CDS in the exemplary ABCA4 REEC includes Exon 22, referred to as the Terminal Exon, as it immediately precedes the splice donor (SD) site in the Exon Editor sequence and is later, via TS, ligated to the native Exon 23 forming a correct Exon 22-23 junction. To facilitate the identification and quantification of the ONT molecule formed by TS, Exon 22 has been modified with 29 synonymous point mutations. These mutations provide opportunities to differentiate trans-spliced transcripts from endogenous transcripts by facilitating selective primer/probe design for RT-qPCR analysis (quantification), as well as next-generation sequencing (NGS) – based analysis (identification and quantification) without eliciting amino acid changes to the protein sequence.

[0221] Splice donor site (splicing domain; SD): The exemplary ABCA4 REEC comprises a consensus U1 splice donor site (SD) located immediately downstream of Terminal Exon 22. This SD interacts with the endogenous Intron 22 splice acceptor (SA) site during the TS reaction mediated by the spliceosome, in place of the natural cis-splicing reaction that would otherwise occur between the endogenous Intron 22 splice donor site and SA site. The result of this TS event is a precisely formed Exon 22-23 junction in which the upstream region originates from the Exon Editor and the downstream region comprises the remainder of the native ABCA4 pre-mRNA, which proceeds through its natural cis-splicing events toward the formation of the mature mRNA molecule.

[0222] STOP codon: The exemplary ABCA4 REEC includes a double STOP codon immediately following the SD site. In fact, the last nucleotide of the GTAAGT SD site serves as the first nucleotide in the TAGTGA double STOP codon. One purpose of this intentional STOP codon is to eliminate the addition of foreign amino acids at the C-terminus under circumstances in which the pre-spliced Exon Editor transcript exits the nucleus, enters the cytoplasm and undergoes translation that results in a truncated, non-spliced protein (NSP). Any such foreign amino acids at the C-terminus of the NSP could pose a potential immunogenicity risk that is avoided via the inclusion of the STOP codon. The STOP codon can also contribute to AU rich element (ARE) mediated decay, which the present inventors hypothesized would serve to degrade pre-spliced Exon Editor RNA that escapes into the cytoplasm, as discussed hereinbelow.

[0223] 40mer Linker: A linker sequence positioned between the SD site and downstream elements may provide flexibility to the RNA structure and accessibility to the SD, which can be advantageous for effective TS. The 40mer Linker emerged as a highly effective sequence from comprehensive screens of thousands of linker candidates. See, for example, FIG. 1. The predicted secondary structure of the relevant region of an Exon Editor containing the 40mer Linker suggests that the presence of the 40mer Linker results in an open and accessible SD.

[0224] 3X U1 sites: An exemplary ABCA4 REEC comprises a potentiator comprising three tandem U1 sites downstream of the 40mer Linker. U1 sites are bound by U1 snRNP during the

process of spliceosome mediated splicing. Inclusion of three tandem U1 sites in this Exon Editor downstream of the primary SD site led to an increase in RNA TS, as well as a decrease in observed NSP levels. The present inventors hypothesized that these results could potentially have been due to increased association with the spliceosome and retention in the nucleus where TS occurs.

[0225] AU rich element (ARE): An exemplary ABCA4 REEC comprises an AU rich element (ARE) derived from the mRNA of the granulocyte-macrophage colony-stimulating factor (GM-CSF) gene. The presence of this ARE along with the upstream STOP codon was designed to mimic a 3' UTR-like sequence. The present inventors hypothesized that in cases in which pre-spliced Exon Editor RNA enters the cytoplasm, ARE RNA binding proteins might recognize the ARE sequence and lead to rapid degradation of the molecule, thereby reducing levels of the NSP.

[0226] Binding domain (BD): The exemplary ABCA4 REEC 's encoded BD directly associates the Exon Editor RNA with the native ABCA4 pre-mRNA. Its 150 nucleotide (nt) sequence is complementary to a region of the same length within the pre-mRNA's intron 22 (chr1:94041452-94041601). The BD's intronic position was selected through an iterative process of scanning intron 22 with an array of Exon Editors which diverge with respect to their BD sequences at distinct intronic positions. This scanning process entailed both a massively parallel library NGS-based approach that assessed thousands of BD's targeting intron 22 and assessments at the individual Exon Editor level using bioluminescence or qPCR as a readout. See, e.g., FIG. 16.

[0227] MALAT1 triple-helix terminator: The exemplary ABCA4 REEC also includes a terminator sequence derived from the MALAT1 long non-coding RNA (lncRNA). This portion of the MALAT1 lncRNA contains a triple helix structure that may confer stability and nuclear retention to the Exon Editor RNA. The MALAT1 terminator leads to TS efficiency that is far beyond that which is achieved when applying alternative terminators, such as self-cleaving ribozymes. See, e.g., FIG. 20A. The triple helix is followed by an Rnase P cleavage site and a short sequence comprising the MALAT1-associated small cytoplasmic RNA (mascRNA), discussed below. The Rnase P cleavage downstream of the triple helix defines the 3' end of the Exon Editor transcript.

[0228] MascRNA: The mascRNA sequence in the exemplary ABCA4 REEC is flanked by an upstream Rnase P cleavage site and a downstream Rnase Z cleavage site, and once processed, is thought to exit the nucleus while the Exon Editor transcript remains and engages in TS. While the mascRNA is no longer present within the Exon Editor transcript following Rnase P cleavage, including its sequence downstream of the MALAT1 triple helix as it naturally occurs may contribute to achieving a highly efficient TS outcome. In the cytoplasm, mascRNA is suspected to promote global protein translation and cell proliferation via its interaction with and upregulation of the multi-tRNA synthetase complex component glutamyl-tRNA synthetase (QARS). To mitigate this potential risk, the mascRNA sequence encoded in the exemplary ABCA4 REEC has been mutated to include a point mutation within its anticodon stem loop, termed anti-Mut1, which abolishes its interaction with QARS and the potentially ensuing increase in global protein synthesis rate. A second point mutation helps maintain the mascRNA's natural cloverleaf structure by base-pairing with the anti-Mut1 mutation. Significantly, introduction of the anti-Mut1 into the mascRNA sequence does not impair TS efficiency. See, e.g., FIG. 20B.

[0229] As described herein, an exemplary ABCA4 REEC encodes an exemplary ABCA4 5' Exon Editor RNA molecule that functions in a 5' trans-splicing reaction mediated by the endogenous spliceosome. As shown in FIG. 14A, following transcription of the Exon Editor RNA, an RNA sequence (Binding domain; BD) within the Exon Editor that is complementary to Intron 22 of the native ABCA4 pre-mRNA localizes the Exon Editor RNA to the pre-mRNA. Next, a trans-splicing (TS) reaction occurs between the Exon Editor's splice donor site (SD) and the pre-mRNA's intron 22 splice acceptor site (SA). This reaction results in the replacement of the native endogenous pre-mRNA's sequence from exon 1 to exon 22 with the corresponding coding sequence (CDS) provided by the Exon-Editor RNA, thereby correcting any mutations that might exist within that region of the native endogenous pre-mRNA. The resulting on-target (ONT) molecule undergoes translation to form a biologically active ABCA4 protein. See, e.g., FIG. 14A.

[0230] Further to the above, without wishing to be bound by theory, RNA exon editors described herein are thought to operate via spliceosome-mediated trans-splicing, which proceeds via a mechanism similar to that of cis-splicing. Following transcription, the exon editor RNA diffuses within the nucleus with U1 snRNP deposited on its unpaired SD. Also in the nucleus, the endogenous (native) pre-mRNA target is transcribed from its genomic locus, and the removal of its introns via cis-splicing between their flanking SD and splice acceptor (SA) sites occurs within seconds of the emergence of each intron's SA. The trans-splicing process is initiated by the base-pairing of an exon editor's binding domain (BD) to its complementary sequence located within intron 22 of the endogenous ABCA4 pre-mRNA, which must occur before the removal of the intron through cis-splicing. At this point, the spliceosome may either proceed with (1) lariat formation toward cis-splicing using the endogenous SD and SA sites or (2) "Y branch" formation toward trans-splicing using the exon editor's SD paired with the endogenous SA. In the latter case, the first nucleotide of the exon editor's SD (Guanine), forms a covalent bond with the intron's branchpoint, which is typically 15-50 bp upstream of the SA. Formation of this bond results in the detachment of the downstream portion of the exon editor from its CDS portion ending in exon 22, which remains in interaction with the spliceosome. Next, the spliceosome mediates the ligation of the exon editor's exon 22 to the endogenous pre mRNA's exon 23 while the intron is detached, resulting in a correct exon 22 - 23 junction. Finally, the Y branch RNA, consisting of the downstream portion of the Exon Editor RNA and the detached intronic sequence, undergoes rapid degradation, similar to the fate of the lariat RNA formed during cis splicing. See, e.g., FIG. 14B.

[0231] ABCA4 exon editors can include or exclude one or more of the components set forth in Table 1 above. Individual functional components can be advantageously included in an ABCA4 exon editor without including all of the components set forth in Table 1. For example, while elements 3, 7, 8, 9, 10, and 12 provide advantages to an ABCA4 exon editor, an effective ABCA4 exon editor can be made without one or more of these elements.

Cell Line Assays

[0232] In some instances, trans-splicing molecules described herein are tested in cultured cell lines. To screen, select, and improve functionality of RNA exon editors, cultured cell lines may be

acquired or engineered to express the targeted ABCA4 pre-mRNA at a sufficient level. Since ABCA4 expression is extremely low in a common, robust and readily transfectable/transducible transformed cell line, a HEK293T-based cell line, termed B6, was engineered that expresses both high ABCA4 RNA and protein levels. This was achieved via the knock-in (KI) of a constitutive CAGGS promoter between the endogenous ABCA4 promoter and exon 1 via CRISPR-mediated homology-directed repair (HDR). PCR on genomic DNA revealed insertion of the CAG promoter in at least one allele in the B6 cells. ABCA4 RNA expression was then confirmed in B6 cells using four different qPCR assays. This stands in contrast to parent HEK293T cells, which do not express ABCA4. Protein translation was confirmed in the B6 cells via Western blot. See FIG. 3C.

[0233] To characterize an ABCA4 RNA exon editor's ability to effectively correct a targeted native pre-mRNA, the resulting trans-spliced on-target (ONT) RNA's ability to undergo translation and yield a full-length ABCA4 protein was assessed. To this end, a knock-out (KO) cell line (termed 17+06) was engineered based on the B6 cell line. 17+06 cells continued to produce high levels of ABCA4 RNA, but in contrast to the B6 cell line, 17+06 cells do not produce ABCA4 protein. Engineering 17+06 cells was achieved by introducing mutations into the ABCA4 gene coding sequence using CRISPR and guide RNAs (gRNA) targeting exons 3 and 4. B6 cells were subjected to two sequential transfection rounds to achieve high mutational efficiency and disrupt the ABCA4 gene sequence – first by applying a gRNA targeting exon 3 and then by applying a mixture of gRNA targeting exon 3 and gRNA targeting exon 4. Mutations were confirmed by Sanger sequencing applied to PCR amplicons produced from 17+06 genomic DNA, as well as cDNA, with the latter serving as confirmation that the mutations occurred within alleles that had undergone successful CAG insertion during the process of generating the parental cell line, B6. RT-qPCR confirmed the persistence of ABCA4 RNA expression in the 17+06 cell line, and Western blot analysis indicated that ABCA4 protein translation was abolished. See FIG. 3C.

RNA Exon Editor Screening Platform

[0234] As described in Figure 13, 5' RNA exon editors comprise several functional sequence elements, such as a binding domain (BD) for pre-mRNA targeting and a linker that enables access to the splice donor (SD) site. When engineering an RNA exon editor for a given gene target, a variety of sequence options for each of these elements is tested for its ability to contribute to high trans-splicing (TS) efficiency. Such testing may be accomplished via (A) cloning and transfecting individual RNA exon editor variants and analyzing efficiency via RT-qPCR/ddPCR and Western blot or (B) cloning and pooling RNA exon editors in a high-throughput (HT) library-based approach that relies on next generation sequencing (NGS) and computational analysis to assess efficiency. ABCA4 REEC, an exemplary ABCA4 RNA exon editor construct, was developed via a combination of these approaches. Both approaches are described below.

RNA exon editor screening in individual format:

[0235] This approach can be applied to test a small-scale number of variable elements within an RNA exon editor sequence prior to initiating a library-based multiplexed screen, to validate the

performance of RNA exon editors identified in a multiplexed screen, and/or to improve performance of lead candidates. Evaluation of TS efficiency occurs at the RNA and protein levels.

[0236] At the RNA level, TS activity is evaluated via isolation of total RNA from cells followed by reverse transcription and real-time quantitative PCR (RT-qPCR) measuring RNA copy numbers of, e.g., the following targets: RNF20 (housekeeping gene for normalization); Native (ABCA4) mRNA; Exon Editor RNA; On-target, exon-edited RNA (ONT), which is the product of positive TS; ONT+Exon Editor+OFT (off-target) – a single assay that captures all three of these targets. OFT represents incorrect RNA molecules to which the RNA exon editor may trans-splice.

[0237] ONT TS efficiency, also referred to as percent replacement, represents the portion of the total ABCA4 mRNA population that has undergone successful TS, and is calculated via the following equation: $\% \text{ ONT TS} = 100 * (\text{ONT copy number} / (\text{ONT copy number} + \text{Native copy number}))$.

[0238] RNA exon editor TS efficiency is the portion of the RNA exon editor transcript population that has been correctly trans-spliced into the ABCA4 RNA and is calculated via the following equation: $\% \text{ Exon Editor TS} = 100 * (\text{ONT copy number} / (\text{ONT copy number} + \text{Exon Editor copy number} + \text{OFT copy number}))$.

[0239] At the protein level, TS activity is measured via Western blot analysis applied to protein extracted from cell or tissue samples via a protocol that enriches membrane proteins in the case of ABCA4. Na⁺/K⁺ ATPase, which is a membrane protein, is used as a loading control. The level of ONT protein is measured via ABCA4-specific antibodies (Ab). For those constructs that include a tag, e.g., a V5 tag at their N-terminus, an Ab specific for the tag (e.g., a V5 specific Ab) can be used to probe Western blots to assess ONT protein levels.

RNA exon editor screening in library format:

[0240] Further to the above, RNA exon editors can be screened in a high-throughput library format that relies on synthesis of DNA oligo libraries that may focus on a particular feature of an Exon Editor, such as a BD or linker.

Confirmation of the Full-length Exon-edited ABCA4 mRNA Sequence:

[0241] To confirm that the ABCA4 mature mRNA generated from exon editing has the sequence predicted to result from a precise trans-splicing reaction, including the correct Exon 22-Exon23 junction, the present inventors isolated and sequenced the edited ABCA4 RT-PCR product from ABCA4 KO cells that were transiently transfected with a plasmid that encodes for the ABCA4-01 Exon Editor. Sanger sequencing confirmed the correct mature mRNA sequence in multiple, randomly selected clones. In addition to confirming the correct Exon 22-Exon 23 junction, these sequencing results confirmed that the full-length exon-edited ABCA4 mature mRNA molecule has the expected sequence over its entire length, including the Exon Editor-derived CDS spanning Exons 1 to 22 and the endogenous sequence spanning Exons 23 to 50. These results provide additional confidence that

the ABCA4 exon editing mediated by treatment with ABCA4-01 results in the intended, correct ABCA4 mature mRNA sequence.

ATPase Functional Activity Assessment of Exon edited ABCA4 Protein:

[0242] The ABCA4 protein is a member of the superfamily of ATP-binding cassette (ABC) transporters and is preferentially localized along the rim region of photoreceptor outer segment disc membranes. ABC transporters utilize the energy of ATP hydrolysis to unidirectionally translocate a diverse set of substrates, ranging from ions to lipids and peptides, across cellular membranes. The ABCA4 protein becomes active once light entering the eye is converted into an electrical signal (phototransduction). Potentially toxic vitamin A derivatives are formed during phototransduction, and the ABCA4 protein helps remove at least two of these substances (all trans retinal and 11-cis-retinal) from photoreceptor outer segments. This photoreceptor clearance mechanism helps protect both photoreceptors and retinal pigment epithelial cells from the buildup and secondary damage that can be induced by these molecules. As described above, recessively inherited loss-of-function ABCA4 mutations result in the accumulation of these vitamin A derivatives causing a buildup of lipofuscin, including A2E (a bis retinoid pyridinium salt N-retinylidene-N-retinylethanolamine fluorophore found in lipofuscin), resulting in cellular toxicity in the retina, and culminating in a progressive loss of vision.

[0243] ATPase activity, including an increase in activity due to substrate binding, is critical for ABCA4 protein function. Accordingly, assessing the ATPase activity of ABCA4 protein that was translated from an ABCA4 mRNA generated via exon editing serves as an overall indicator of ABCA4 biological activity. Previous reports have shown that incubation of 40 μ M all trans retinal (ATR) with purified wild-type (WT) ABCA4 protein results in a 1.8- to 2.5-fold increase in ATPase activity. With this information in mind, experiments were designed to assess the biological activity of exon-edited ABCA4 protein.

[0244] ATPase activity of the ABCA4 protein rescued in ABCA4 KO cells transfected with an overexpression plasmid encoding an N-terminal V5-epitope tagged version of the RNA exon editor of ABCA4-01 (SEQ ID NO: 81) was assessed in the absence or presence of 40 μ M all-trans retinal (ATR) following ABCA4 protein immunoprecipitation utilizing established methods. Initial data demonstrates a basal level of ATPase activity that increases in response to the addition of 40 μ M all-trans retinal. See FIG. 24. Minimal background ATPase activity was detected in an untransfected sample that lacks ABCA4. These results provide evidence that the exon-edited ABCA4 protein generated from ABCA4-01 (SEQ ID NO: 90) treatment will exhibit the functional ATPase activity required for transporting all-trans-retinal and excess 11-cis-retinal from photoreceptor outer segments. In some embodiments, a fragment of SEQ ID NO: 90 (e.g., SEQ ID NO: 69) can be combined with a different 5' UTR with equivalent results. These results depict restoration of ABCA4 biological activity following expression of an exemplary ABCA4 RNA exon editor.

Summary of in vivo NHP studies supporting use an ABCA4 Exon Editor therapeutic for patients with ABCA4-related retinopathies:

[0245] Example 4 presents evidence demonstrating that AAV8-based V5 epitope-tagged ABCA4 Exon Editors result in therapeutically relevant exon editing activity in African Green Monkey (AGM) 1-month after a single subretinal administration of 1×10^{11} vg/injection. Of note, an Early Lead vector comprising (SEQ ID NO: 42) tested in this study was used as the basis for further optimization and ultimate selection of elements present in ABCA4-01 (SEQ ID NO: 90), which comprises an RNA exon editor comprising SEQ ID NO: 78. See, for example, FIGs. 4-12; Tables 1 and 2.

[0246] Example 8 presents evidence demonstrating that the selected regulatory elements in ABCA4-01 (SEQ ID NO: 90) confer persistent expression and exon editing activity out to a 3-month time-point in AGM. The AAV8 construct tested in this study comprises an RNA Exon Editor comprising SEQ ID NO: 77 (of which the first 104 nucleotides comprise the native ABCA4 5' UTR followed by the ATG start codon for translation initiation) that closely resembles Early Lead vector comprising (SEQ ID NO: 42), but differs from it in that it lacks the V5 epitope tag, contains a double stop-codon following the splice domain, and produces higher levels of a truncated ABCA4 non-spliced protein (NSP) *in vitro* when compared to ABCA4-01 (SEQ ID NO: 90). Notably, the AAV8 construct tested in this study comprises an Exon Editor (SEQ ID NO: 77) that closely resembles Early Lead vector comprising (SEQ ID NO: 42), which was well tolerated at 1- and 3-months in all animals following a single subretinal dose of 4.3×10^{10} vg/eye, further supporting a clinical approach using the related construct ABCA4-01 (SEQ ID NO: 90). See FIG. 22. The ABCA4-01 RNA Exon Editor comprising SEQ ID NO: 90 (which in turn comprises SEQ ID NO: 78) differs from the RNA exon editor comprising SEQ ID NO: 77 used in this study in that SEQ ID NO: 77 does not include a potentiator (e.g., SEQ ID NO: 62) or a translational potentiator (e.g., SEQ ID NO: 63). Each of SEQ ID NO: 77 and SEQ ID NO: 78, however, share the same regulatory elements and double STOP codon.

[0247] Example 10 presents evidence demonstrating that ABCA4-01 (SEQ ID NO: 90) confers persistent expression and exon editing activity out to a 6-month time-point in NHP. More particularly, FIGs. 30A-30C present results demonstrating robust ABCA4 RNA and protein replacement *in vivo* in NHP six (6) months after treatment with an exemplary AAV8 RNA exon editor construct (ABCA4-01; SEQ ID NO: 90). Briefly, wild-type cynomolgus macaques were treated with a single subretinal dose of ABCA4-01 (results for an exemplary dose of 3.5×10^{11} vg/eye are presented). The percent RNA replacement ranged from about 40%-60% for this dose of ABCA4-01. This resulted in therapeutic levels of the resultant human-NHP chimeric ABCA4 protein as measured by a qualified IA-MS assay (FIG. 30C). More specifically, human NHP chimeric ABCA4 protein, which results from ABCA4 RNA editing in these animals, was found to be present at 20-40% of the total (human-NHP + NHP) ABCA4 protein in the tested samples. Up to 66% RNA replacement and up to 45% human/NHP chimeric ABCA4 protein were detected in wild-type cynomolgus macaques treated with a higher dose of ABCA4-01 (1×10^{12} vg/eye) delivered via single subretinal injection. These results demonstrate that ABCA4-01 achieves therapeutically relevant levels of edited ABCA4 protein expression in NHP, which

exceed the levels of rescued ABCA4 protein previously shown to confer therapeutic benefit in an *Abca4* KO mouse model.

[0248] Example 12 presents evidence that ABCA4-01, an exemplary AAV8 RNA exon editor construct comprising SEQ ID NO: 90, functions in the context of human photoreceptors. Briefly, donor-derived retinal explants from multiple human donors were treated with ABCA4-01 (3.9E11 vg) to evaluate ABCA4-01 activity in human photoreceptors. ABCA4-01 treatment resulted in approximately 20-30% RNA replacement of the endogenous exons 1-22 with the same corresponding exons encoded by ABCA4-01 (FIG. 31). Of note, the same lot of ABCA4-01 was also used in the investigative NHP study referred to in Example 10, further evidencing the translatability of ABCA4-01 results from NHP to human subjects.

III. Vectors

[0249] Trans-splicing molecules can be delivered to target cells of an individual using various techniques, e.g., using recombinant adeno-associated virus (AAV) vectors or other vector modalities, such as non-viral vectors. Thus, provided herein are vectors comprising/encoding trans-splicing molecules (e.g., viral or non-viral vectors comprising/encoding trans-splicing molecules, e.g., DNA vectors comprising/encoding trans-splicing molecules). Any suitable nucleic acid vector may be used in conjunction with the present compositions and methods to design and assemble the components of the trans-splicing molecule and a recombinant AAV. In one embodiment, the vector is a recombinant AAV carrying the trans-splicing molecule driven by a promoter that expresses a trans-splicing molecule in selected cells of an individual. Methods for assembly of the recombinant vectors are known in the art. See, e.g., Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1989; Kay, M. A. et al., *Nat. Medic*, 2001, 7(1):33-40; and Walther W. and Stein U., *Drugs* 2000, 60(2):249-71.

[0250] In certain embodiments described herein, the trans-splicing molecule is delivered to the selected cells, e.g., photoreceptor cells, in need of treatment by means of an AAV vector. A variety of naturally occurring serotypes of AAV are available. Many natural variants in the AAV capsid exist, allowing identification and use of an AAV with properties specifically suited for ocular or cochlear cells. Artificial AAV vectors may be engineered by conventional molecular biology techniques, making it possible to optimize these particles for cell specific delivery of the trans-splicing molecule nucleic acid sequences, for minimizing immunogenicity, for tuning stability and particle lifetime, for efficient degradation, for accurate delivery to the nucleus, etc. For example, such artificial capsids may be generated by any suitable technique, using a selected AAV sequence (e.g., a fragment of a vp1 capsid protein) in combination with heterologous sequences which may be obtained from a different selected AAV, non-contiguous portions of the same AAV, from a non-AAV viral source, or from a non-viral source. An artificial AAV may be, without limitation, a pseudotyped AAV, a chimeric AAV capsid, a recombinant AAV capsid, or a "humanized" AAV capsid. Pseudotyped vectors, wherein the capsid of one AAV is replaced with a heterologous capsid protein, are useful for delivering trans-splicing molecules described herein.

[0251] The expression of trans-splicing molecules described herein can be achieved in the selected cells through delivery by recombinantly engineered AAVs or artificial AAVs that contain sequences comprising/encoding the desired trans-splicing molecule. The use of AAVs is a common mode of exogenous delivery of DNA as it is relatively non-toxic, provides efficient gene transfer, and can be easily optimized for specific purposes. Among the well-characterized serotypes of AAVs isolated from human or non-human primates, human serotype 2 has been widely used for efficient gene transfer experiments in different target tissues and animal models.

[0252] In some embodiments, the AAV is AAV1 or a variant thereof (e.g., SEQ ID NO: 6 or 64 of US20030138772 or SEQ ID NO: 11 or 27 of US20150159173), AAV2 or a variant thereof (e.g., SEQ ID NO: 7 or 70 of US20030138772, SEQ ID NO: 7 or 23 of US20150159173, or SEQ ID NO: 7 of US20150159173), AAV2G9 or a variant thereof, AAV3 or a variant thereof (e.g., SEQ ID NO: 8 or 71 of US20030138772), AAV3a or a variant thereof, AAV3b or a variant thereof (e.g., SEQ ID NO: 1 and 10 of U.S. Pat. No. 6,156,303), AAV3-3 or a variant thereof (e.g., SEQ ID NO: 200 and 217 of WO2005033321), AAV4 or a variant thereof (e.g., SEQ ID NO: 63 of US20030138772), AAV4-4 or a variant thereof (e.g., SEQ ID NO: 201 or 218 of WO2005033321), AAV5 or a variant thereof (e.g., SEQ ID NO: 114 of US20030138772), AAV6 or a variant thereof (e.g., SEQ ID NO: 65 of US20030138772), AAV6.1 or a variant thereof (e.g., SEQ ID NO: 29 of US20150159173), AAV6.2 or a variant thereof, AAV6.1.2 or a variant thereof, AAV7 or a variant thereof (e.g., SEQ ID NO: 1-3 of US20030138772), AAV7.2 or a variant thereof, AAV8 or a variant thereof (e.g., SEQ ID NO: 4 and 95 of US20030138772 or AAV8(b) (having the amino acid sequence of Pro-Glu-Arg-Thr-Ala-Met-Ser-Leu-Pro (SEQ ID NO: 132) at amino acid positions 587-595 as compared to wildtype AAV8, as described in U.S. Patent No. 9,567,376, which is incorporated herein by reference in its entirety)), AAV9 or a variant thereof (e.g., SEQ ID NO: 5 and 100 of US20030138772), AAV9.9 or a variant thereof, AAV9.11 or a variant thereof, AAV9.13 or a variant thereof, AAV9.16 or a variant thereof, AAV9.24 or a variant thereof, AAV9.45 or a variant thereof, AAV9.47 or a variant thereof, AAV9.61 or a variant thereof, AAV9.68 or a variant thereof, AAV9.84 or a variant thereof (see, e.g., N. Pulicherla et al. *Molecular Therapy* 19(6):1070-1078 (2011), herein incorporated by reference in its entirety), AAV10 or a variant thereof (e.g., SEQ ID NO: 117 of US20030138772), AAV11 or a variant thereof (e.g., SEQ ID NO: 118 of US20030138772), AAV12 or a variant thereof (e.g., SEQ ID NO: 119 of US20030138772), AAV16.3 or a variant thereof, AAV24.1 or a variant thereof, AAV27.3 or a variant thereof, AAV42.12 or a variant thereof, AAV42-1b or a variant thereof, AAV42-2 or a variant thereof, AAV42-3a or a variant thereof, AAV42-3b or a variant thereof, AAV42-4 or a variant thereof, AAV42-5a or a variant thereof, AAV42-5b or a variant thereof, AAV42-6b or a variant thereof, AAV42-8 or a variant thereof, AAV42-10 or a variant thereof, AAV42-11 or a variant thereof, AAV42-12 or a variant thereof, AAV42-13 or a variant thereof, AAV42-15 or a variant thereof, AAV42-aa or a variant thereof, AAV43-1 or a variant thereof, AAV43-12 or a variant thereof, AAV43-20 or a variant thereof, AAV43-21 or a variant thereof, AAV43-23 or a variant thereof, AAV43-25 or a variant thereof, AAV43-5 or a variant thereof, AAV44.1 or a variant thereof, AAV44.2 or a variant thereof, AAV44.5 or a variant thereof, AAV223.1 or a variant thereof, AAV223.2 or a variant thereof, AAV223.4 or a variant thereof, AAV223.5 or a variant thereof, AAV223.6 or a variant thereof, AAV223.7 or a variant thereof, AAV1-

7/rh.48 or a variant thereof, AAV1-8/rh.49 or a variant thereof, AAV2-15/rh.62 or a variant thereof, AAV2-3/rh.61 or a variant thereof, AAV2-4/rh.50 or a variant thereof, AAV2-5/rh.51 or a variant thereof, AAV3.1/hu.6 or a variant thereof, AAV3.1/hu.9 or a variant thereof, AAV3-9/rh.52 or a variant thereof, AAV3-11/rh.53 or a variant thereof, AAV4-8/rh.64 or a variant thereof, AAV4-9/rh.54 or a variant thereof (e.g., SEQ ID NO: 116 of WO2005033321), AAV4-19/rh.55 or a variant thereof (e.g., SEQ ID NO: 117 of WO2005033321), AAV5-3/rh.57 or a variant thereof, AAV5-22/rh.58 or a variant thereof, AAV7.3/hu.7 or a variant thereof, AAV16.8/hu.10 or a variant thereof, AAV16.12/hu.11 or a variant thereof, AAV29.3/bb.1 or a variant thereof, AAV29.5/bb.2 or a variant thereof, AAV106.1/hu.37 or a variant thereof, AAV114.3/hu.40 or a variant thereof, AAV127.2/hu.41 or a variant thereof, AAV127.5/hu.42 or a variant thereof, AAV128.3/hu.44 or a variant thereof, AAV130.4/hu.48 or a variant thereof, AAV145.1/hu.53 or a variant thereof, AAV145.5/hu.54 or a variant thereof, AAV145.6/hu.55 or a variant thereof, AAV161.10/hu.60 or a variant thereof, AAV161.6/hu.61 or a variant thereof, AAV33.12/hu.17 or a variant thereof, AAV33.4/hu.15 or a variant thereof, AAV33.8/hu.16 or a variant thereof, AAV52/hu.19 or a variant thereof, AAV52.1/hu.20 or a variant thereof, AAV58.2/hu.25 or a variant thereof, AAVA3.3 or a variant thereof, AAVA3.4 or a variant thereof, AAVA3.5 or a variant thereof, AAVA3.7 or a variant thereof, AAVC1 or a variant thereof, AAVC2 or a variant thereof, AAVC5 or a variant thereof, AAV-DJ or a variant thereof (e.g., SEQ ID NO: 2 or 3 of US20140359799), AAV-DJ8 or a variant thereof, AAVF3 or a variant thereof, AAVF5 or a variant thereof, AAVH2 or a variant thereof, AAVH6 or a variant thereof, AAVLK03 or a variant thereof, AAVH-1/hu.1 or a variant thereof, AAVH-5/hu.3 or a variant thereof, AAVLG-10/rh.40 or a variant thereof, AAVLG-4/rh.38 or a variant thereof, AAVLG-9/hu.39 or a variant thereof, AAVN721-8/rh.43 or a variant thereof, AAVCh.5 or a variant thereof (e.g., SEQ ID NO 46 of US20150159173), AAVCh.5R1 or a variant thereof, AAVcy.2 or a variant thereof, AAVcy.3 or a variant thereof, AAVcy.4 or a variant thereof, AAVcy.5 or a variant thereof (e.g., SEQ ID NO: 8 and 24 of US20150159173), AAVCy.5R1 or a variant thereof, AAVCy.5R2 or a variant thereof, AAVCy.5R3 or a variant thereof, AAVCy.5R4 or a variant thereof, AAVcy.6 or a variant thereof, AAVhu.1 or a variant thereof (e.g., SEQ ID NO: 144 of WO2005033321), AAVhu.2 or a variant thereof (e.g., SEQ ID NO: 143 of WO2005033321), AAVhu.3 or a variant thereof (e.g., SEQ ID NO: 145 of WO2005033321), AAVhu.4 or a variant thereof (e.g., SEQ ID NO: 141 of WO2005033321), AAVhu.5 or a variant thereof, AAVhu.6 or a variant thereof (e.g., SEQ ID NO: 84 of WO2005033321), AAVhu.7 or a variant thereof (e.g., SEQ ID NO: 150 of WO2005033321), AAVhu.9 or a variant thereof (e.g., SEQ ID NO: 155 of WO2005033321), AAVhu.10 or a variant thereof (e.g., SEQ ID NO: 156 of WO2005033321), AAVhu.11 or a variant thereof (e.g., SEQ ID NO: 153 of WO2005033321), AAVhu.13 or a variant thereof (SEQ ID NO: 16 and 32 of US20150159173), AAVhu.15 or a variant thereof (e.g., SEQ ID NO: 147 of WO2005033321), AAVhu.16 or a variant thereof (e.g., SEQ ID NO: 148 of WO2005033321), AAVhu.17 or a variant thereof (e.g., SEQ ID NO: 83 of WO2005033321), AAVhu.18 or a variant thereof (e.g., SEQ ID NO: 149 of WO2005033321), AAVhu.19 or a variant thereof (e.g., SEQ ID NO: 133 of WO2005033321), AAVhu.20 or a variant thereof (e.g., SEQ ID NO: 134 of WO2005033321), AAVhu.21 or a variant thereof (e.g., SEQ ID NO: 135 of WO2005033321), AAVhu.22 or a variant thereof (e.g., SEQ ID NO: 138 of WO2005033321), AAVhu.23.2 or a variant

thereof (e.g., SEQ ID NO: 137 of WO2005033321), AAVhu.24 or a variant thereof (e.g., SEQ ID NO: 136 of WO2005033321), AAVhu.25 or a variant thereof (e.g., SEQ ID NO: 146 of WO2005033321), AAVhu.26 or a variant thereof (e.g., SEQ ID NO: 17 and 33 of US20150159173), AAVhu.27 or a variant thereof (e.g., SEQ ID NO: 140 of WO2005033321), AAVhu.28 or a variant thereof (e.g., SEQ ID NO: 42 of US20150159173), AAVhu.29 or a variant thereof (e.g., SEQ ID NO: 132 of WO2005033321), AAVhu.29R or a variant thereof, AAVhu.31 or a variant thereof (e.g., SEQ ID NO: 121 of WO2005033321), AAVhu.32 or a variant thereof (SEQ ID NO: 122 of WO2005033321), AAVhu.34 or a variant thereof (e.g., SEQ ID NO: 125 of WO2005033321), AAVhu.35 or a variant thereof (e.g., SEQ ID NO: 164 of WO2005033321), AAVhu.37 or a variant thereof (e.g., SEQ ID NO: 18 and 34 of US20150159173), AAVhu.39 or a variant thereof (e.g., SEQ ID NO: 102 of WO2005033321), AAVhu.40 or a variant thereof (e.g., SEQ ID NO: 87 of WO2005033321), AAVhu.41 or a variant thereof (e.g., SEQ ID NO: 91 of WO2005033321), AAVhu.42 or a variant thereof (e.g., SEQ ID NO: 85 of WO2005033321), AAVhu.43 or a variant thereof (e.g., SEQ ID NO: 160 of WO2005033321), AAVhu.44 or a variant thereof (e.g., SEQ ID NO: 45 of US20150159173), AAVhu.44R1 or a variant thereof, AAVhu.44R2 or a variant thereof, AAVhu.44R3 or a variant thereof, AAVhu.45 or a variant thereof (e.g., SEQ ID NO: 127 of WO2005033321), AAVhu.46 or a variant thereof (e.g., SEQ ID NO: 159 of WO2005033321), AAVhu.47 or a variant thereof (e.g., SEQ ID NO: 128 of WO2005033321), AAVhu.48 or a variant thereof (e.g., SEQ ID NO: 38 of US20150159173), AAVhu.48R1 or a variant thereof, AAVhu.48R2 or a variant thereof, AAVhu.48R3 or a variant thereof, AAVhu.49 or a variant thereof (e.g., SEQ ID NO: 189 of WO2005033321), AAVhu.51 or a variant thereof (e.g., SEQ ID NO: 190 of WO2005033321), AAVhu.52 or a variant thereof (e.g., SEQ ID NO: 191 of WO2005033321), AAVhu.53 or a variant thereof (e.g., SEQ ID NO: 19 and 35 of US20150159173), AAVhu.54 or a variant thereof (e.g., SEQ ID NO: 188 of WO2005033321), AAVhu.55 or a variant thereof (e.g., SEQ ID NO: 187 of WO2005033321), AAVhu.56 or a variant thereof (e.g., SEQ ID NO: 192 of WO2005033321), AAVhu.57 or a variant thereof (e.g., SEQ ID NO: 193 of WO2005033321), AAVhu.58 or a variant thereof (e.g., SEQ ID NO: 194 of WO2005033321), AAVhu.60 or a variant thereof (e.g., SEQ ID NO: 184 of WO2005033321), AAVhu.61 or a variant thereof (e.g., SEQ ID NO: 185 of WO2005033321), AAVhu.63 or a variant thereof (e.g., SEQ ID NO: 195 of WO2005033321), AAVhu.64 or a variant thereof (e.g., SEQ ID NO: 196 of WO2005033321), AAVhu.66 or a variant thereof (e.g., SEQ ID NO: 197 of WO2005033321), AAVhu.67 or a variant thereof (e.g., SEQ ID NO: 198 of WO2005033321), AAVhu.14/9 or a variant thereof, AAVhu.t 19 or a variant thereof, AAVrh.2 or a variant thereof (e.g., SEQ ID NO: 39 of US20150159173), AAVrh.2R or a variant thereof, AAVrh.8 or a variant thereof (e.g., SEQ ID NO: 41 of US20150159173), AAVrh.8R or a variant thereof, AAVrh.10 or a variant thereof (e.g., SEQ ID NO: 9 and 25 of US20150159173), AAVrh.12 or a variant thereof, AAVrh.13 or a variant thereof (e.g., SEQ ID NO: 10 and 26 of US20150159173), AAVrh.13R or a variant thereof, AAVrh.14 or a variant thereof, AAVrh.17 or a variant thereof, AAVrh.18 or a variant thereof, AAVrh.19 or a variant thereof, AAVrh.20 or a variant thereof (e.g., SEQ ID NO: 1 of US20150159173), AAVrh.21 or a variant thereof, AAVrh.22 or a variant thereof, AAVrh.23 or a variant thereof, AAVrh.24 or a variant thereof, AAVrh.25 or a variant thereof, AAVrh.31 or a variant thereof, AAVrh.32 or a variant thereof, AAVrh.33 or a variant thereof, AAVrh.34

or a variant thereof, AAVrh.35 or a variant thereof, AAVrh.36 or a variant thereof, AAVrh.37 or a variant thereof (e.g., SEQ ID NO: 40 of US20150159173), AAVrh.37R2 or a variant thereof, AAVrh.38 or a variant thereof (e.g., SEQ ID NO: 86 of WO2005033321), AAVrh.39 or a variant thereof (e.g., SEQ ID NO: 3, 20, or 36 of US20150159173), AAVrh.40 or a variant thereof (e.g., SEQ ID NO: 92 of WO2005033321), AAVrh.43 or a variant thereof (e.g., SEQ ID NO: 21 and 37 of US20150159173), AAVrh.46 or a variant thereof (e.g., SEQ ID NO: 4 and 22 of US20150159173), AAVrh.48 or a variant thereof (e.g., SEQ ID NO: 44 of US20150159173), AAVrh.48.1 or a variant thereof (e.g., SEQ ID NO: 44 of US20150159173), AAVrh.48.1.2 or a variant thereof, AAVrh.48.2 or a variant thereof, AAVrh.49 or a variant thereof (e.g., SEQ ID NO: 103 of WO2005033321), AAVrh.50 or a variant thereof (e.g., SEQ ID NO: 108 of WO2005033321), AAVrh.51 or a variant thereof (e.g., SEQ ID NO: 104 of WO2005033321), AAVrh.52 or a variant thereof (e.g., SEQ ID NO: 96 of WO2005033321), AAVrh.53 or a variant thereof (e.g., SEQ ID NO: 97 of WO2005033321), AAVrh.54 or a variant thereof (e.g., SEQ ID NO: 49 of US20150159173), AAVrh.56 or a variant thereof (e.g., SEQ ID NO: 152 of WO2005033321), AAVrh.57 or a variant thereof (e.g., SEQ ID NO: 105 of WO2005033321), AAVrh.58 or a variant thereof (e.g., SEQ ID NO: 48 of US20150159173), AAVrh.61 or a variant thereof (e.g., SEQ ID NO: 107 of WO2005033321), AAVrh.62 or a variant thereof (e.g., SEQ ID NO: 114 of WO2005033321), AAVrh.64 or a variant thereof (e.g., SEQ ID NO: 43 of US20150159173), AAVrh.64R1 or a variant thereof, AAVrh.64R2 or a variant thereof, AAVrh.67 or a variant thereof (e.g., SEQ ID NO: 47 of US20150159173), AAVrh.73 or a variant thereof (e.g., SEQ ID NO: 5 of US20150159173), or AAVrh.74 or a variant thereof (e.g., SEQ ID NO: 6 of US2015015917). Non-limiting examples of variants include SEQ ID NOs: 9, 27-45, 47-62, 66-69, 73-81, 84-94, 96, 97, 99, and 101-113 of US20030138772, the contents of which are herein incorporated by reference in its entirety, and SEQ ID NOs: 1, 2, 4-82, 89, 90, 93-95, 98, 100, 101, 109-113, 118-120, 124, 126, 131, 139, 142, 151, 154, 158, 161, 162, 165-183, 202, 204-212, 215, 219, and 224-236 of WO2005033321, the contents of which are herein incorporated by reference in its entirety. In one embodiment, the AAV serotype is any of those described in U.S. 2021/0189430, the contents of which are herein incorporated by reference in its entirety. The amino acid sequence of the AAV may include one or more amino acid substitutions in an AAV capsid protein at one or more positions that interacts with a heparan sulfate proteoglycan or at one or more positions corresponding to amino acids 484, 487, 527, 532, 585, or 588, numbering based on VP1 numbering of AAV2.

[0253] Unless otherwise specified, the AAV ITRs, and other selected AAV components described herein, may be readily selected from among any AAV serotype, including, without limitation, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 or other known and unknown AAV serotypes. In one embodiment, the ITRs are from AAV2. These ITRs or other AAV components may be readily isolated using techniques available to those of skill in the art from an AAV serotype. Such AAV may be isolated or obtained from academic, commercial, or public sources (e.g., the American Type Culture Collection, Manassas, VA). Alternatively, the AAV sequences may be obtained through synthetic or other suitable means by reference to published sequences such as are available in the literature or in databases such as, e.g., GenBank, PubMed, or the like.

[0254] Desirable AAV fragments for assembly into vectors include the cap proteins, including the vp1, vp2, vp3, and hypervariable regions, the rep proteins, including rep 78, rep 68, rep 52, and rep 40, and the sequences encoding these proteins. These fragments may be readily utilized in a variety of vector systems and host cells. Such fragments may be used alone, in combination with other AAV serotype sequences or fragments, or in combination with elements from other AAV or non-AAV viral sequences. As used herein, artificial AAV serotypes include, without limitation, AAV with a non-naturally occurring capsid protein. Such an artificial capsid may be generated by any suitable technique, using a selected AAV sequence (e.g., a fragment of a vp1 capsid protein) in combination with heterologous sequences which may be obtained from a different selected AAV serotype, non-contiguous portions of the same AAV serotype, from a non-AAV viral source, or from a non-viral source. An artificial AAV serotype may be, without limitation, a pseudotyped AAV, a chimeric AAV capsid, a recombinant AAV capsid, or a "humanized" AAV capsid. Pseudotyped vectors, wherein the capsid of one AAV is utilized with the ITRs from an AAV having a different capsid protein, are useful as described herein. In one embodiment, the AAV is AAV2/5 (i.e., an AAV having AAV2 ITRs and an AAV5 capsid). In another embodiment, the AAV is AAV2/8 (i.e., an AAV having AAV2 ITRs and an AAV8 capsid). In one embodiment, the AAV includes an AAV8 capsid. Such AAV8 capsid includes the amino acid sequence found under NCBI Reference Sequence: YP_077180.1. In another embodiment, the AAV8 capsid includes a capsid encoded by nt 2121 to 4337 of GenBank accession: AF513852.1.

[0255] In one embodiment, the vectors useful in compositions and methods described herein contain, at a minimum, sequences encoding a selected AAV serotype capsid, e.g., an AAV2 capsid, or a fragment thereof. In another embodiment, useful vectors contain, at a minimum, sequences encoding a selected AAV serotype rep protein, e.g., AAV2 rep protein, or a fragment thereof. Optionally, such vectors may contain both AAV cap and rep proteins. In vectors in which both AAV rep and cap are provided, the AAV rep and AAV cap sequences can both be of one serotype origin, e.g., an AAV2 origin.

[0256] Alternatively, vectors may be used in which the rep sequences are from an AAV serotype which differs from that which is providing the cap sequences. In one embodiment, the rep and cap sequences are expressed from separate sources (e.g., separate vectors, or a host cell and a vector). In another embodiment, these rep sequences are fused in frame to cap sequences of a different AAV serotype to form a chimeric AAV vector, such as those described in U.S. Patent No. 7,282,199, which is incorporated by reference herein.

[0257] A suitable recombinant AAV (rAAV) is generated by culturing a host cell which contains a nucleic acid sequence encoding an AAV serotype capsid protein, or fragment thereof, as defined herein; a functional rep gene; a minigene composed of, e.g., AAV ITRs and a trans-splicing molecule nucleic acid sequence; and sufficient helper functions to permit packaging of the minigene into the AAV capsid protein. The components required to be cultured in the host cell to package an AAV minigene in an AAV capsid may be provided to the host cell in trans. Alternatively, any one or more of the required components (e.g., minigene, rep sequences, cap sequences, and/or helper

functions) may be provided by a stable host cell which has been engineered to contain one or more of the required components using methods known to those of skill in the art.

[0258] In one embodiment, the AAV includes a promoter (or a functional fragment of a promoter). The selection of the promoter to be employed in the rAAV may be made from among a wide number of constitutive or inducible promoters that can express the selected transgene in the desired target cell. See, e.g., the list of promoters identified in International Patent Publication No. WO 2014/012482, incorporated by reference herein. In one embodiment, the promoter is cell-specific. The term "cell-specific" means that the particular promoter selected for the recombinant vector can direct expression of the selected transgene in a particular cell type. In one embodiment, the promoter is specific for expression of the transgene in photoreceptor cells. In another embodiment, the promoter is specific for expression in the rods and/or cones. In another embodiment, the promoter is specific for expression of the transgene in retinal pigment epithelium (RPE) cells. In another embodiment, the promoter is specific for expression of the transgene in ganglion cells. In another embodiment, the promoter is specific for expression of the transgene in Mueller cells. In another embodiment, the promoter is specific for expression of the transgene in bipolar cells. In another embodiment, the promoter is specific for expression of the transgene in horizontal cells. In another embodiment, the promoter is specific for expression of the transgene in amacrine cells. In another embodiment, the transgene is expressed in any of the above noted cells.

[0259] In another embodiment, the promoter is the native promoter for the target gene to be expressed. Useful promoters include, without limitation, a rod opsin promoter, a red-green opsin promoter, a blue opsin promoter, a cGMP-phosphodiesterase promoter, a mouse opsin promoter, a rhodopsin promoter, an alpha-subunit of cone transducing, a beta phosphodiesterase (PDE) promoter, a retinitis pigmentosa promoter, a NXNL2/NXNL 1 promoter, the RPE65 promoter, the retinal degeneration slow/peripherin 2 (Rds/perph2) promoter, and the VMD2 promoter.

[0260] Other conventional regulatory sequences contained in the mini-gene or rAAV are also disclosed in documents such as WO 2014/124282 and others cited and incorporated by reference herein. One of skill in the art may make a selection among these, and other, expression control sequences without departing from the scope described herein.

[0261] The selected genetic element may be delivered by any suitable method, including those described herein. The methods used to construct any embodiment described herein are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY. Similarly, methods of generating rAAV virions are well known and the selection of a suitable method is not a limitation on methods and constructs described herein. See, e.g., K. Fisher et al., *J. Virol.*, 1993 70: 520-532 and U.S. Patent 5,478,745, each of which is incorporated by reference herein.

[0262] In another embodiment, the trans-splicing molecule is included in a proviral plasmid, such as those disclosed in International Patent Publication No. WO 2012/158757, incorporated herein by reference. Such a proviral plasmid contains a modular recombinant AAV genome comprising in operative association: a wildtype 5' AAV2 ITR sequence flanked by unique restriction sites that permit

ready removal or replacement of said ITR; a promoter comprising a 49-nucleic acid cytomegalovirus sequence upstream of a cytomegalovirus (CMV)-chicken beta actin sequence, or a photoreceptor-specific promoter/enhancer, the promoter flanked by unique restriction sites that permit ready removal or replacement of the entire promoter sequence, and the upstream sequence flanked by unique restriction sites that permit ready removal or replacement of only the upstream CMV or enhancer sequence, from the promoter sequence. The trans-splicing molecule described herein can be inserted into the site of a multi-cloning poly linker, wherein the trans-splicing molecule is operably linked to, and under the regulatory control of, the promoter. A bovine growth hormone polyadenylation sequence flanked by unique restriction sites that permit ready removal or replacement of said poly A sequence; and a wildtype 3' AAV2 ITR sequence flanked by unique restriction sites that permit ready removal or replacement of the 3' ITR; are also part of such a plasmid. The plasmid backbone comprises the elements necessary for replication in bacterial cells, e.g., a kanamycin resistance gene, and is itself flanked by transcriptional terminator/insulator sequences.

[0263] In one embodiment, a proviral plasmid comprises: (a) a modular recombinant AAV genome comprising in operative association: (i) a wildtype 5' AAV2 ITR sequence flanked by unique restriction sites that permit ready removal or replacement of said ITR; (ii) a promoter comprising (A) a 49-nucleic acid CMV sequence upstream of a CMV-chicken beta actin sequence; (B) a photoreceptor-specific promoter/enhancer; or (C) a neuronal cell-specific promoter/enhancer. The promoter is flanked by unique restriction sites that permit ready removal or replacement of the entire promoter sequence, and the upstream sequence flanked by unique restriction sites that permit ready removal or replacement of only the upstream CMV or enhancer sequence, from the promoter sequence. Also part of this proviral plasmid is a multi-cloning polylinker sequence that permits insertion of a trans-splicing molecule sequence including any of those described herein, wherein the trans-splicing molecule is operably linked to, and under the regulatory control of, the promoter; a bovine growth hormone polyadenylation sequence flanked by unique restriction sites that permit ready removal or replacement of said poly A sequence; and a wildtype 3' AAV2 ITR sequence flanked by unique restriction sites that permit ready removal or replacement of the 3' ITR. The proviral plasmid also contains a plasmid backbone comprising the elements necessary for replication in bacterial cells, and further comprising a kanamycin resistance gene, said plasmid backbone flanked by transcriptional terminator/insulator sequences. The proviral plasmid described herein may also contain in the plasmid backbone a non-coding lambda phage 5.1 kb stuffer sequence to increase backbone length and prevent reverse packaging of non-functional AAV genomes.

[0264] In yet a further aspect, the promoter of the proviral plasmid is modified to reduce the size of the promoter to permit larger trans-splicing molecule sequences to be inserted in the rAAV. In one embodiment, the CMV/CBA hybrid promoter, which normally includes a non-coding exon and intron totaling about 1,000 base pairs, is replaced with a 130-base pair chimeric intron, as described in International Patent Publication No. WO 2017/087900, which is incorporated herein by reference in its entirety.

[0265] These proviral plasmids are then employed in currently conventional packaging methodologies to generate a recombinant virus expressing the trans-splicing molecule transgene carried by the proviral plasmids. Suitable production cell lines are readily selected by one of skill in the art. For example, a suitable host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Briefly, the proviral plasmid is transfected into a selected packaging cell, where it may exist transiently. Alternatively, the minigene or gene expression cassette with its flanking ITRs is stably integrated into the genome of the host cell, either chromosomally or as an episome. Suitable transfection techniques are known and may readily be utilized to deliver the recombinant AAV genome to the host cell. Typically, the proviral plasmids are cultured in the host cells which express the cap and/or rep proteins. In the host cells, the minigene consisting of the trans-splicing molecule with flanking AAV ITRs is rescued and packaged into the capsid protein or envelope protein to form an infectious viral particle. Thus, a recombinant AAV infectious particle is produced by culturing a packaging cell carrying the proviral plasmid in the presence of sufficient viral sequences to permit packaging of the gene expression cassette viral genome into an infectious AAV envelope or capsid.

[0266] Alternatively, trans-splicing molecules can be delivered using a non-AAV vector, e.g., a non-viral vector. Any suitable non-viral vector technology known in the art or described herein may be used. Such non-viral vectors amenable for delivery of trans-splicing molecules include liposomes (e.g., cationic liposomes, unilamellar liposomes, or multilamellar liposomes), nanoparticles (e.g., polymeric nanoparticles, lipid nanoparticles (LNPs), PEGylated nanoparticles (e.g., PEGylated LNPs), peptide nanoparticles, metal nanoparticles, and the like), dendrimers (e.g., cationic dendrimers, e.g., polypropylenimine dendrimers), exosomes (e.g., immunologically inert and/or targeted exosomes, e.g., made using techniques described in Alvarez-Erviti, et al., 2011, *Nat. Biotechnol.* 29:341), and microvesicles. In some instances, the trans-splicing molecules described herein may be delivered using cell penetrating peptides (CPPs), which can translocate the plasma membrane of a target cell and facilitate the delivery of a trans-splicing molecule to the interior of the target cell.

IV. Pharmaceutical Compositions and Kits

[0267] Provided herein are pharmaceutical compositions including a nucleic acid trans-splicing molecule, a proviral plasmid, or a rAAV comprising any of the ABCA4 nucleic acid trans-splicing molecules described herein. In some embodiments, the pharmaceutical composition includes any of the 5' trans-splicing molecules described herein. In some embodiments, the pharmaceutical composition includes a 5' trans-splicing molecule and a 3' trans-splicing molecule, e.g., wherein the 5' trans-splicing molecule and the 3' trans-splicing molecule together contain a functional sequence of ABCA4 exons and bind the same target ABCA4 intron (e.g., intron 22).

[0268] Such pharmaceutical compositions may be prepared so as to be pure of contamination and suitable for *in vivo* administration. The pharmaceutical compositions described herein may be assessed for contamination by conventional methods and then formulated into a pharmaceutical composition intended for a suitable route of administration. Still other compositions containing the trans-splicing molecule, e.g., naked DNA, may be formulated similarly with a suitable carrier. Such formulation involves the use of a pharmaceutically and/or physiologically acceptable

vehicle or carrier, particularly directed for administration to the target cell (e.g., an ocular cell). In one embodiment, carriers suitable for administration to the target cells include buffered saline, an isotonic sodium chloride solution, or other buffers, e.g., HEPES, to maintain pH at appropriate physiological levels, and, optionally, other medicinal agents, pharmaceutical agents, stabilizing agents, buffers, carriers, adjuvants, diluents, etc.

[0269] In some embodiments, the carrier is a liquid for injection. Exemplary physiologically acceptable carriers include sterile, pyrogen-free water and sterile, pyrogen-free, phosphate buffered saline. A variety of such known carriers are provided in U.S. Patent No. 7,629,322, incorporated herein by reference. In one embodiment, the carrier is an isotonic sodium chloride solution. In another embodiment, the carrier is balanced salt solution. In one embodiment, the carrier includes tween. If the virus is to be stored long-term, it may be frozen in the presence of glycerol or TWEEN 20.

[0270] In other embodiments, compositions containing trans-splicing molecules described herein include a surfactant. Useful surfactants, such as Pluronic F68 (Poloxamer 188, also known as LUTROL® F68) may be included as they prevent AAV from sticking to inert surfaces and thus ensure delivery of the desired dose. As an example, one illustrative composition designed for the treatment of ocular diseases described herein comprises a recombinant adeno-associated vector carrying a nucleic acid sequence encoding a 5' trans-splicing molecule or 3' trans-splicing molecule as described herein, under the control of regulatory sequences which express the trans-splicing molecule in an ocular cell of a mammalian individual, and a pharmaceutically acceptable carrier. The carrier is isotonic sodium chloride solution and includes a surfactant Pluronic F68. In one embodiment, the trans-splicing molecule is any of those described herein.

[0271] In yet another exemplary embodiment, the composition comprises a rAAV virus comprising any of the ABCA4 trans-splicing molecules described herein for ABCA4 gene correction, the nucleic acid sequence under the control of a promoter which directs expression of the trans-splicing molecule in photoreceptor cells, wherein the composition is formulated with a carrier and additional components suitable for subretinal injection. In still another embodiment, the composition or components for production or assembly of this composition, including carriers, rAAV particles, surfactants, and/or the components for generating the rAAV, as well as suitable laboratory hardware to prepare the composition, may be incorporated into a kit. Such kits may further include instructions for administering the composition to an individual, e.g., as a treatment for an ocular disease, such as Stargardt Disease or Usher Syndrome.

[0272] Additionally provided herein are kits containing a pharmaceutical composition comprising a 5' trans-splicing molecule (e.g., wherein the trans-splicing molecule is packaged in any AAV vector described herein). In some embodiments, the kit includes instructions for mixing the pharmaceutical composition prior to administration.

V. Methods and Uses

[0273] Patients with ABCA4-related retinopathy present with similar, but not identical clinical features. The vast majority have accumulation of lipofuscin and bilateral visual loss. ABCA4-related Stargardt disease typically shows a characteristic macular atrophy often with a beaten-bronze

appearance on funduscopy. Fundus flavimaculatus is similar but can reveal less macular atrophy and a greater presence of yellow-white flecks in the macula, extending to the mid-peripheral region. Subjects with ABCA4-related cone-rod dystrophy present with visual symptoms such as visual loss, impaired color vision, and a central scotoma.

[0274] The clinical presentation of ABCA4-related Stargardt disease can be divided into childhood-onset, early adult-onset and late adult-onset. Each of these groups may have their own variations in phenotype that correlate with specific ABCA4 variants.

[0275] Childhood-Onset: Children with biallelic severe ABCA4 variants present with more aggressive central visual loss attributed to macular dysfunction alone or in combination with cone-rod dystrophy, typically between the ages of five and eleven years.

[0276] Early Adult-Onset: These patients typically present with central visual loss due to foveal atrophy in their second or third decade of life. At presentation, funduscopy typically shows prominent macular atrophy with numerous paramacular and peripheral flecks. These patients often carry ABCA4 variants with an intermediate impact on ABCA4 function (e.g. c.6079C>T and c.[2588G>C;5603A>T]) either on both alleles or as compound heterozygous combinations with a null/severe ABCA4 variant.

[0277] Late Adult-Onset: The definition of a “late” onset disease has varied from those ≥ 35 to those ≥ 50 years of age. Thus, patients can present from the fourth decade onwards. Patients are often identified incidentally through retinal screening due to relatively well-preserved best corrected visual acuity (BCVA) thought to be explained by atrophy sparing the fovea.

[0278] The majority of ABCA4-related retinopathy subjects are diagnosed/symptomatic by the second decade of life, which can have a significant impact on their development, independence, education, and overall quality of life. In addition, some of the more severe forms, such as cone-rod dystrophy begin earlier in childhood, as young as 5 years of age. One study of 95 patients with Stargardt disease found that once visual acuity (VA) dropped below 20/40, there was rapid progression of visual loss to 20/200. Further, the probability of maintaining a VA of 20/40 or better in at least one eye was 52% at age 19 years but dropped to 22% by age 39. Earlier onset and greater severity cone-rod dystrophy, as well as childhood-onset Stargardt disease, generally speaking, are typically associated with biallelic null variants in ABCA4. These childhood/early-onset forms present with a symptomatic central scotoma and rapidly progressing atrophy, with the majority of pediatric patients demonstrating rod photoreceptor involvement at diagnosis, which is an indicator for worse prognosis. Another consideration is the development of amblyopia in younger patients with ABCA4 retinopathies. Amblyopia arises as a result of early deficiencies in visual development, and leads to impairment of visual acuity, reduced stereopsis and impaired fusion that is understood to be permanent if not corrected in the critical period of the first 7-9 years of life. There are some studies that indicate that adolescents may also retain some potential for neuroplasticity. This leads the present inventors to expect a greater efficacy in pediatric subjects that are treated with ABCA4 exon editors described herein for their retinopathy before the critical period completes.

[0279] The nucleic acid trans-splicing molecules (e.g., nucleic acid trans-splicing molecules and nucleic acid trans-splicing molecule-encoding vectors) and compositions described above are useful for expressing functional ABCA4, and/or modulating expression of ABCA4, in a target cell (e.g., an ocular cell, a retinal cell, a photoreceptor and/or an RPE cell) of an individual in, e.g., methods for treating diseases or disorders associated with mutations in the ABCA4 gene, such as ABCA4-associated retinal dystrophies, such as Stargardt Disease and cone-rod dystrophy, including delaying or ameliorating symptoms associated with the disease described herein. The nucleic acid trans-splicing molecules (e.g., nucleic acid trans-splicing molecules and nucleic acid trans-splicing molecule-encoding vectors) and compositions described above are furthermore useful for expressing functional ABCA4, and/or modulating expression of ABCA4, in a target cell (e.g., an ocular cell, a retinal cell, a photoreceptor and/or an RPE cell) of an individual as applied, e.g., to their use for treating diseases or disorders associated with mutations in the ABCA4 gene, such as ABCA4-associated retinal dystrophies, such as Stargardt Disease and cone-rod dystrophy, including delaying or ameliorating symptoms associated with the disease described herein or to their use in the preparation of a medicament for the treatment of diseases or disorders associated with mutations in the ABCA4 gene, such as ABCA4-associated retinal dystrophies, such as Stargardt Disease and cone-rod dystrophy, including delaying or ameliorating symptoms associated with the disease described herein. Such methods and uses involve contacting a target ABCA4 gene (e.g., ABCA4 pre-mRNA) with a trans-splicing molecule as described herein [e.g., a 5' trans-splicing molecule, or a mixture of both 3' and 5' trans-splicing molecules as described herein, a composition (e.g., a pharmaceutical composition) comprising same or a medicament comprising same], under conditions in which a CDS of the trans-splicing molecule is spliced to the target ABCA4 pre-mRNA to replace a part of the targeted pre-mRNA carrying one or more defects or mutations, with a functional (i.e., healthy), or normal or wildtype or corrected mRNA of the targeted gene, in order to correct expression of ABCA4 in the target cell. Thus, the methods and compositions are used to treat the ocular diseases/pathologies associated with the specific mutations and/or gene expression.

[0280] In some instances, provided herein are methods of expressing functional ABCA4 in a target cell, by contacting (e.g., transducing) the target cell with any of the nucleic acid trans-splicing molecules, vectors (e.g., AAV vectors), or compositions described herein. In one embodiment, the contacting involves direct administration of the composition (e.g., pharmaceutical composition) to the affected individual. In another embodiment, the contacting may occur ex vivo with the cultured cell and the treated ocular cell reimplanted in the individual. In another embodiment, the method involves administering an rAAV carrying any of the 5' ABCA4 trans-splicing molecules described herein. In still another embodiment, the method involves administering a mixture of rAAV carrying a 3' trans-splicing molecule and rAAV carrying a 5' trans-splicing molecule. These methods comprise administering to an individual in need thereof an effective concentration of a composition of any of those described herein.

[0281] In some embodiments, the methods include selecting one or more trans-splicing molecules for treating an individual having a disorder associated with mutation/s in ABCA4. In some embodiments, use of one or more trans-splicing molecules for treating an individual having a disorder

associated with mutation/s in ABCA4 or use of same in the preparation of a medicament for the treatment of an individual having a disorder associated with mutation/s in ABCA4 is encompassed herein. Such methods and uses include selecting one or more trans-splicing molecules for treating an individual having a disorder associated with a mutation in ABCA4 or for use of such selected one or more trans-splicing molecules in treating an individual having a disorder associated with a mutation in ABCA4 or for use of such selected one or more trans-splicing molecules in the preparation of a medicament for the treatment of an individual having a disorder associated with mutation/s in ABCA4. Such selection can be based on the genotype of the individual. In some embodiments, a disorder associated with ABCA4 may be an autosomal recessive disorder. In some instances, the individual is homozygous or compound heterozygous for mutation/s in ABCA4. Methods of screening for and identifying particular mutations in ABCA4 are known in the art.

[0282] Methods of the invention include selecting a single trans-splicing molecule based on the location of a single mutation in ABCA4 (e.g., a mutation of one allele of the individual). In some instances in the context of autosomal recessive mutations, correction of just one of two mutations can be sufficient to restore functional protein activity, for example, wherein the second allele has a mutation on the opposite portion of the ABCA4 gene, out of range of a single AAV-delivered trans-splicing molecule configured to correct the first mutation. Thus, in some embodiments, methods of the invention include administering a single trans-splicing molecule to correct a single mutation on the 5' portion of the target gene, e.g., without regard to the location of the mutation in the other allele.

[0283] Additionally, the invention provides methods involving selecting a single trans-splicing molecule to correct two or more mutations on the 5' portion of the ABCA4 gene, such that a single trans-splicing molecule capable of being packaged in an AAV vector is capable of spanning both or all of the mutations, thereby correcting both or all of the mutations.

[0284] In other embodiments, provided herein are methods for correcting multiple mutations within a ABCA4 gene using two trans-splicing molecules—a 5' trans-splicing molecule and a 3' trans-splicing molecule. In some embodiments, the entire ABCA4 gene is replaced upon binding and trans-splicing of both trans-splicing molecules, for example, where the 5' trans-splicing molecule and the 3' trans-splicing molecule bind the same target ABCA4 intron and replace the exons upstream and downstream, respectively, of the target intron(s).

[0285] In some embodiments, the compositions provided herein are administered to an individual as a method of treating a disease or disorder. In some embodiments, the individual has a genetic disease, such as an ABCA4-related retinopathy as described herein. In some embodiments, the individual is at risk of having a disease, such as an ABCA4-related retinopathy as described herein. In some embodiments, the individual is at increased risk of having a disease or disorder caused by insufficient amount of a protein (e.g., ABCA protein) or insufficient activity of a protein (e.g., ABCA protein). If an individual is "at an increased risk" of having a disease or disorder caused by insufficient amount of a protein or insufficient activity of a protein, the method involves preventative or prophylactic treatment. For example, an individual may be at an increased risk of having such a disease or disorder because of family history of the disease. Typically, individuals at an increased risk

of having such a disease or disorder benefit from prophylactic treatment (e.g., by preventing or delaying the onset or progression of the disease or disorder).

[0286] An effective concentration of a recombinant adeno-associated virus carrying a trans-splicing molecule as described herein ranges between about 10^8 and 10^{13} vector genomes per milliliter (vg/mL). The rAAV infectious units are measured as described in McLaughlin et al., *J. Virol.* 1988, 62: 1963. In another embodiment, the concentration ranges between 10^9 and 10^{13} vg/mL. In another embodiment, the effective concentration is about 1.5×10^{11} vg/mL. In another embodiment, the effective concentration is about 5×10^{11} vg/mL. In one embodiment, the effective concentration is about 1.5×10^{10} vg/mL. In another embodiment, the effective concentration is about 2.8×10^{11} vg/mL. In yet another embodiment, the effective concentration is about 1.5×10^{12} vg/mL. In another embodiment, the effective concentration is about 1.5×10^{13} vg/mL.

[0287] In some embodiments, the concentration ranges between 4.0×10^{10} and 1×10^{12} vg/eye. In some embodiments, the effective concentration is about 1.0×10^{11} vg/eye. In some embodiments, the effective concentration is about 4.3×10^{10} vg/eye.

[0288] It is desirable that the lowest effective dosage (total genome copies delivered) of virus be utilized in order to reduce the risk of undesirable effects, such as toxicity, and other issues related to administration to the eye, e.g., retinal dysplasia and detachment. An effective dosage of a recombinant adeno-associated virus carrying a trans-splicing molecule as described herein ranges between about 10^8 and 10^{13} vector genomes (vg) per dose (i.e., per injection). In one embodiment, the dosage ranges between 10^9 and 10^{13} vg. In another embodiment, the effective dosage is about 1.5×10^{11} vg. In another embodiment, the effective dosage is about 5×10^{11} vg. In one embodiment, the effective dosage is about 1.5×10^{10} vg. In another embodiment, the effective dosage is about 2.8×10^{11} vg. In yet another embodiment, the effective dosage is about 1.5×10^{12} vg. In another embodiment, the effective concentration is about 1.5×10^{13} vg. Still other dosages in these ranges or in other units may be selected by the attending physician, taking into account the physical state of the individual being treated, including the age of the individual; the composition being administered, and the particular disorder; the targeted cell and the degree to which the disorder, if progressive, has developed.

[0289] The composition may be delivered in a volume of from about 50 μ L to about 1 mL, including all numbers within the range, depending on the size of the area to be treated, the viral titer used, the route of administration, and the desired effect of the method. In one embodiment, the volume is about 50 μ L. In another embodiment, the volume is about 70 μ L. In another embodiment, the volume is about 100 μ L. In another embodiment, the volume is about 125 μ L. In another embodiment, the volume is about 150 μ L. In another embodiment, the volume is about 175 μ L. In yet another embodiment, the volume is about 200 μ L. In another embodiment, the volume is about 250 μ L. In another embodiment, the volume is about 300 μ L. In another embodiment, the volume is about 350 μ L. In another embodiment, the volume is about 400 μ L. In another embodiment, the volume is about 450 μ L. In another embodiment, the volume is about 500 μ L. In another embodiment, the volume is about 600 μ L. In another embodiment, the volume is about 750 μ L. In another embodiment, the volume is about 850 μ L. In another embodiment, the volume is about 1,000 μ L.

[0290] In some embodiments, a composition comprising an RNA exon editor described herein (e.g., an RNA exon editor comprising SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69) may be supplied as a suspension suitable for injection. In some embodiments, an RNA exon editor described herein (e.g., an RNA exon editor comprising SEQ ID NO: 90, SEQ ID NO: 78, or SEQ ID NO: 69) may be administered by subretinal injection following pars plana vitrectomy through a small gauge cannula in a fixed volume of 300 μ L.

[0291] In one embodiment, the volume and concentration of the rAAV composition are selected so that only certain anatomical regions having target cells are impacted. In another embodiment, the volume and/or concentration of the rAAV composition is a greater amount, in order to reach larger portions of the eye. Similarly, dosages are adjusted for administration to other organs.

[0292] In some embodiments, treatments and uses described herein replace 10% or more of the target ABCA4 mRNA in the target cell (e.g., 11% or more of the target ABCA4 mRNA in the target cell(s), 12% or more of the target ABCA4 mRNA in the target cell(s), 13% or more of the target ABCA4 mRNA in the target cell(s), 14% or more of the target ABCA4 mRNA in the target cell(s), 15% or more of the target ABCA4 mRNA in the target cell(s), 16% or more of the target ABCA4 mRNA in the target cell(s), 17% or more of the target ABCA4 mRNA in the target cell(s), 18% or more of the target ABCA4 mRNA in the target cell(s), or 19% or more of the target ABCA4 mRNA in the target cell(s). In some embodiments, treatments and uses described herein replace 20% or more of the target ABCA4 mRNA in the target cell [e.g., 21% or more of the target ABCA4 mRNA in the target cell(s), 22% or more of the target ABCA4 mRNA in the target cell(s), 23% or more of the target ABCA4 mRNA in the target cell(s), 24% or more of the target ABCA4 mRNA in the target cell(s), 25% or more of the target ABCA4 mRNA in the target cell(s), 26% or more of the target ABCA4 mRNA in the target cell(s), 27% or more of the target ABCA4 mRNA in the target cell(s), 28% or more of the target ABCA4 mRNA in the target cell(s), 29% or more of the target ABCA4 mRNA in the target cell(s), 30% or more of the target ABCA4 mRNA in the target cell(s), 31% or more of the target ABCA4 mRNA in the target cell(s), 32% or more of the target ABCA4 mRNA in the target cell(s), 33% or more of the target ABCA4 mRNA in the target cell(s), 34% or more of the target ABCA4 mRNA in the target cell(s), 35% or more of the target ABCA4 mRNA in the target cell(s), 36% or more of the target ABCA4 mRNA in the target cell(s), 37% or more of the target ABCA4 mRNA in the target cell(s), 38% or more of the target ABCA4 mRNA in the target cell(s), 39% or more of the target ABCA4 mRNA in the target cell(s), 40% or more of the target ABCA4 mRNA in the target cell(s), 41% or more of the target ABCA4 mRNA in the target cell(s), 42% or more of the target ABCA4 mRNA in the target cell(s), 43% or more of the target ABCA4 mRNA in the target cell(s), 44% or more of the target ABCA4 mRNA in the target cell(s), 45% or more of the target ABCA4 mRNA in the target cell(s), 46% or more of the target ABCA4 mRNA in the target cell(s), 47% or more of the target ABCA4 mRNA in the target cell(s), 48% or more of the target ABCA4 mRNA in the target cell(s), 49% or more of the target ABCA4 mRNA in the target cell(s), 50% or more of the target ABCA4 mRNA in the target cell(s), 51% or more of the target ABCA4 mRNA in the target cell(s), 52% or more of the target ABCA4 mRNA in the target cell(s), 53% or more of the target ABCA4 mRNA in the target cell(s), 54% or more of the target ABCA4 mRNA in the target cell(s), 55% or more of the target

ABCA4 protein in the target cell(s), 51% or more of the target ABCA4 protein in the target cell(s), 52% or more of the target ABCA4 protein in the target cell(s), 53% or more of the target ABCA4 protein in the target cell(s), 54% or more of the target ABCA4 protein in the target cell(s), 55% or more of the target ABCA4 protein in the target cell(s), 56% or more of the target ABCA4 protein in the target cell(s), 57% or more of the target ABCA4 protein in the target cell(s), 58% or more of the target ABCA4 protein in the target cell(s), 59% or more of the target ABCA4 protein in the target cell(s), 60% or more of the target ABCA4 protein in the target cell(s), 61% or more of the target ABCA4 protein in the target cell(s), 62% or more of the target ABCA4 protein in the target cell(s), 63% or more of the target ABCA4 protein in the target cell(s), 64% or more of the target ABCA4 protein in the target cell(s), 65% or more of the target ABCA4 protein in the target cell(s), 66% or more of the target ABCA4 protein in the target cell(s), 67% or more of the target ABCA4 protein in the target cell(s), 68% or more of the target ABCA4 protein in the target cell(s), 69% or more of the target ABCA4 protein in the target cell(s), 70% or more of the target ABCA4 protein in the target cell(s); 75% or more of the target ABCA4 protein in the target cell(s); 80% or more of the target ABCA4 protein in the target cell(s); 85% or more of the target ABCA4 protein in the target cell(s); 90% or more of the target ABCA4 protein in the target cell(s); 95% or more of the target ABCA4 protein in the target cell(s)].

[0294] Where reference is made to increasing functional ABCA4 protein levels, the increase may be clinically significant. The increase may be relative to the level of functional ABCA4 protein in the subject without treatment or relative to the amount of functional ABCA4 protein in a population of similar subjects. The increase may be at least 10% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 20% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 40% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 50% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 80% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 100% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 200% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects. The increase may be at least 500% more functional ABCA4 protein relative to the subject prior to treatment or to a population of similar subjects.

[0295] In some instances, methods and uses described herein reduce ABCA4-associated lipofuscin in a subject (e.g., a subject having an ABCA4-associated retinal dystrophy). In other instances, methods and uses described herein reduce A2E accumulation in a subject (e.g., a subject having an ABCA4-associated retinal dystrophy).

[0296] For each of the described methods and uses, the treatment or use may be used to prevent the occurrence of further damage or to rescue tissue having mild, moderate, or advanced

disease. As used herein, the term “rescue” means to prevent progression of the disease, prevent spread of damage to uninjured cells, and/or to improve damage in injured cells.

[0297] Thus, in one embodiment, the composition is administered before disease onset. In another embodiment, the composition is administered prior to the development of symptoms. In another embodiment, the composition is administered after development of symptoms. In yet another embodiment, the composition is administered when less than 90% of the target cells are functioning or remaining, e.g., as compared to a reference tissue. In yet another embodiment, the composition is administered when more than 10% of the target cells are functioning or remaining, e.g., as compared to a reference tissue. In yet another embodiment, the composition is administered when more than 20% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 30% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 40% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 50% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 60% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 70% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 80% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 90% of the target cells are functioning or remaining. In yet another embodiment, the composition is administered when more than 95% of the target cells are functioning or remaining.

[0298] In yet another embodiment, any of the above-described methods or uses is performed in combination with another, or secondary, therapy. The therapy may be any now known, or as yet unknown, therapy which helps prevent, arrest or ameliorate these mutations or defects or any of the effects associated therewith. The secondary therapy can be administered before, concurrent with, or after administration of a pharmaceutical composition described above. In one embodiment, a secondary therapy involves non-specific approaches for maintaining the health of the retinal cells, such as administration of neurotrophic factors, anti-oxidants, and/or anti-apoptotic agents. The non-specific approaches are achieved through injection of proteins, recombinant DNA, recombinant viral vectors, stem cells, fetal tissue, or genetically modified cells. The latter could include genetically modified cells that are encapsulated.

[0299] In another embodiment, the method includes performing functional and imaging studies to determine the efficacy of the treatment. These studies include ERG and *in vivo* retinal imaging, as described in U.S. Patent No. 8,147,823; in International Patent Publication Nos. WO 2014/011210 or WO 2014/124282, incorporated herein by reference. In addition, visual field studies, perimetry (e.g., kinetic perimetry or full field static perimetry) and microperimetry, slit lamp examination, intraocular pressure measurement, dilated fundus ophthalmoscopy, mobility testing, visual acuity, and/or color vision testing may be performed. Efficacy may be measured at time points following administration of the nucleic acid trans-splicing molecules and RNA exon editors encoded thereby as described herein or compositions comprising same (e.g., pharmaceutical compositions) to determine if symptoms of ABCA4-related retinopathies are, e.g., reduced in treated subjects versus

untreated subjects. In some embodiments, efficacy may be measured by, for example, an improvement in visual acuity and/or color vision or a decrease in lipofuscin, including A2E accumulation.

[0300] In certain embodiments, it is desirable to perform non-invasive retinal imaging and functional studies to identify areas of retained photoreceptors to be targeted for therapy. In these embodiments, clinical diagnostic tests are employed to determine the precise location(s) for one or more subretinal injection(s). These tests may include ERG, perimetry, topographical mapping of the layers of the retina and measurement of the thickness of its layers by means of confocal scanning laser ophthalmoscopy (cSLO) and optical coherence tomography (OCT), topographical mapping of cone density via adaptive optics (AO), functional eye exam, etc. In view of the imaging and functional studies, in some embodiments, one or more injections are performed in the same eye in order to target different areas of retained photoreceptors.

[0301] For use in these methods, the volume and viral titer of each injection is determined individually and may be the same or different from other injections performed in the same, or contralateral, eye. In another embodiment, a single, larger volume injection is made in order to treat the entire eye. The dosages, administrations, and regimens may be determined by the attending physician given the teachings of this disclosure.

[0302] The examples that follow do not limit the scope of the embodiments described herein. One skilled in the art will appreciate that modifications can be made in the following examples which are intended to be encompassed by the spirit and scope of the invention.

EXAMPLES

[0303] The following examples provide non-limiting methods for generating compositions described above and for using the compositions described above.

Example 1. Trans-splicing efficiency in cultured human cells

[0304] ABCA4 trans-splicing molecules were tested in cultured human cells using a high-throughput, NGS-based assay in which various permutations were tested in each trans-splicing molecule. Results are shown in FIG. 1. Each ABCA4 trans-splicing molecule in this example included a 5' regulatory domain including a CMV promoter and a native 5' UTR and a functional ABCA4 CDS. Domains that were varied across trans-splicing molecules were the linker domain, the binding domain (BD), the 3' DS, and the terminator. In this example, trans-splicing molecules were transfected into cells as plasmid and detected by unique molecule identifier (UMI) according to known RNA-seq methods. Trans-splicing efficiency readouts show that certain domains showed improved trans-splicing across many or all permutations. For instance, here, a 40-mer linker of SEQ ID NO: 27 unexpectedly exhibited the highest trans-splicing efficiency, relative to shorter linkers of SEQ ID NO: 29 and 31. Also, a binding domain having SEQ ID NO: 18 exhibited the highest trans-splicing efficiency among all binding domains tested. It is also noteworthy that certain combinations of linker domains, BDs, and splice sites (SSs) conferred surprising levels of relative trans-splicing efficiency.

Example 2. Splice site evaluation

[0305] Splice sites were tested using next generation sequencing across a library of ABCA4 nucleic acid trans-splicing molecules having a total of 550 linker sequences. The five splice sites tested were CAAAGT, GTAAGT, GTAAGC, GTAAGG, and GTAAGT. RNA sequences encoded thereby are referred to herein as GUAAGU, GUAAGG, GUAAGC, and GUAACU, respectively. FIG. 2 shows trans-splicing activity (exon editing activity) in arbitrary units. These results show that a splice site comprising GTAAGT conferred the highest trans-splicing activity out of the five splice sites tested. Accordingly, an exemplary splicing domain described herein comprises GUAAGT, which is encoded by GTAAGT.

Example 3. ABCA4 trans-splicing in cultured human cells

[0306] Cultured human cells (HEK-293T) cells were transduced with AAV encoding various ABCA4 trans-splicing molecules (RNA exons editors) at MOI = 100,000 and incubated for three days before cell harvest. AAV-56 is an AAV vector that encodes an ABCA4 trans-splicing molecule having SEQ ID NO: 40. AAV-274 is an AAV vector that encodes an ABCA4 trans-splicing molecule having SEQ ID NO: 41. AAV-443 is an AAV vector that encodes an ABCA4 trans-splicing molecule having SEQ ID NO: 42. AAV-497 is an AAV vector that encodes an ABCA4 trans-splicing molecule having SEQ ID NO: 43. AAV-505 is an AAV vector that encodes an ABCA4 trans-splicing molecule having SEQ ID NO: 44.

[0307] Membrane protein extraction was performed using conventional methods. Briefly, permeabilization buffer and protease/phosphatase inhibitor were added to cells, vortexed, incubated at 4°C for 10 minutes. Permeabilized cells were centrifuged, and the pellet was resuspended in solubilization buffer with protease/phosphatase inhibitor. Samples were incubated at 4°C for 30 minutes before centrifugation. Supernatants containing membrane protein were transferred and stored at -80°C.

[0308] RNA was extracted by conventional methods and quantified by RT-qPCR. Percent of on-target ABCA4 trans-splicing was normalized to trans-splicing molecule copy number. RNA results showing percentage of ABCA4 mRNA replaced by each trans-splicing molecule (exon editing) are shown in FIG. 3A. Western blots are shown in FIG. 3B, using an anti-ABCA4 C-terminal antibody and, for loading controls, an anti-NaK/ATPase antibody.

[0309] AAV-274, AAV-443, AAV-497, and AAV-505 each exhibited improved trans-splicing efficiency relative to AAV-56 (37.3%, 38.7%, 28.5%, and 28.5%, respectively, relative to 23.8%). AAV-274 and AAV-443 each exhibited significantly higher on-target copy number, compared to AAV-56.

[0310] An analogous assay was conducted to compare trans-splicing efficiencies between AAV-443-2 (SEQ ID NO: 45) and AAV-652 (SEQ ID NO: 46), using ABCA4 knock-in cells as a positive control and untreated HEK293T cells as a negative control. FIG. 3C shows a Western blot showing protein expression in replicates of each of AAV-443-2 and AAV-652. Results are given as fold change

in FIG. 3D. Remarkably, full restoration of ABCA4 protein levels was observed in ABCA4 knockout cells (FIG. 3D).

Example 4. Evaluation of ABCA4 trans-splicing molecules in non-human primates

[0311] Recessively inherited mutations in ABCA4 are causal in the development of progressive forms of blindness including Stargardt disease 1 and cone-rod dystrophy 3. The 6882-bp coding sequence of ABCA4 is too large to be delivered in its entirety by a single AAV vector. Moreover, with hundreds of disease-causing mutations found throughout the gene, a single base editing approach would not address a significant number of patients. A large-scale exon editing solution was designed to deliver a therapeutic RNA construct capable of trans-splicing into endogenous ABCA4 pre-mRNA, thereby introducing functional exons that correct a myriad of mutations localized in the functional exons delivered via trans-splicing. Indeed, RNA trans-splicing with a single AAV-based construct can address approximately 60% of all patient mutations. This example reports the editing efficiency and tolerability of AAV-ABCA4 exon editors *in vivo* in healthy African Green Monkey (AGM) non-human primates (NHP) following subretinal injection.

[0312] Exon editors 443 (SEQ ID NO: 47), 274 (SEQ ID NO: 48), 497 (SEQ ID NO: 49), and 505 (SEQ ID NO: 50) were packaged in AAV8 vectors. AAV8 vector preparations were characterized using qPCR of shared regulatory elements to titer, Stunner for measurement of full and empty capsids, and measurement of endotoxin levels. Results are shown in Table 2, below.

Table 2. AAV characterization

Exon Editor	Titer (vg/mL)	% full	Endotoxin (EU/mL)
443	1×10 ¹²	41.5	<0.5
274	1×10 ¹²	46.0	<0.5
497	1×10 ¹²	43.0	<0.5
505	1×10 ¹²	23.9	<0.5

[0313] VP1, VP2 and VP3 proteins for each preparation were detected and visualized on an SDS gel run at 5×10 vg/well (FIG. 4).

[0314] GFP-encoding AAV vectors containing a 5' regulatory domain of SEQ ID NO: 15 (i.e., containing a CMV promoter and a native 5' ABCA4 untranslated region) were administered subretinally in African Green Monkey (AGM) non-human primates (NHP). GFP protein expression is shown in FIGS. 5A-5C. FIG. 5A is a live image showing GFP protein. FIG. 5B is a photomicrograph showing anti-GFP immunohistochemistry (IHC) staining of retinal sections (photoreceptors and choroid are labeled) in a negative control at day 28. FIG. 5C is a photomicrograph showing anti-GFP IHC staining of retinal sections (photoreceptors and choroid are labeled) in a GFP exon editor-treated animal at day 28. These results show that the 5' regulatory domain confers robust protein expression in photoreceptors *in vivo*.

[0315] ABCA4 exon editors were formulated as AAV8 and were administered subretinally in AGM NHP as two blebs per eye. Each injection was 100 μ l containing 1×10^{11} vg. cSLO and OCT images of baseline, post subretinal injections, day 14 and day 28, are shown in FIG. 6. Superior and inferior blebs were observed post-injection. A slight 'shadow' of the bleb area was observed through day 28. At day 28, clinical scoring was conducted using color fundus imaging for vehicle, exon editor 443, and trans-splicing molecule 274 (FIG. 7). Total clinical scores for the vehicle, exon editor 443, and trans-splicing molecule 274 were 17, 0, and 1, respectively.

[0316] Exon editing by ABCA4 exon editors was assessed in AGM NHP retina using quantitative RT-PCR to quantify RNA and western blot to quantify ABCA4 protein. To measure trans-spliced ABCA4 mRNA, quantitative RT-PCR was used to detect the junction in chimeric mRNA between human and AGM sequences, as shown in FIG. 8A. For each ABCA4 exon editor, trans-spliced mRNA was quantified as edited copies per 10 ng RNA (FIG. 8B and Table 4).

Table 4. Edited copies / 10 ng RNA

Exon Editor	Mean	Standard Deviation
(Vehicle)	0	0
443	464880.29	464880.29
274	45380.29	45380.29
497	328749.88	328749.88
505	30594.16	30594.16

[0317] For each ABCA4 exon editor, tagged full-length trans-spliced ABCA4 protein was quantified in membrane extracts from retina by western blot of the tag. An exemplary western blot showing full-length ABCA4 expression for exon editors 443, 274, 497, and 505 is shown in FIG. 9.

[0318] To confirm that the protein expressed was chimeric human/AGM (and therefore trans-spliced product), a mass spectrometry experiment was conducted. Unique peptides distinguishing the fully AGM NHP ABCA4 and the chimeric, edited ABCA4 were identified by coupling immunoaffinity enrichment with MRM-MS. Identities of the peptides were confirmed by spiking in standards. In the vehicle-treated retina (FIG. 10A), only NHP peptide was observed. In the exon editor-treated retina (FIG. 10B), both human and NHP peptides were detected. This result confirms that the full-length exon edited mRNA is translated.

[0319] Ophthalmic health was monitored throughout the study. In general, ophthalmic health was maintained after treatment with AAV8 subretinally delivered exon editors, as measured by retinal thickness (FIG. 11A), intraocular pressure (FIG. 11B), total clinical score (FIG. 11C), vitreous cell score (FIG. 11D), aqueous cell score (FIG. 11E), and inflammatory keratic particles (FIG. 11F). Retinal thickness remained normal throughout the study (FIG. 11A). Tonometry showed intraocular pressure was reduced on day 7 in vehicle-treated eyes and stabilized at later timepoints (FIG. 11B). Intraocular pressure remained in a healthy range in exon editor-treated eyes. Ocular inflammation was assessed

by slit lamp biomicroscopy and fundoscopy (FIGS. 11C-11F). Individual eyes treated with exon editors showed some signs of mild-to-moderate inflammation, but not consistently for a certain editor or AAV preparation, suggesting that these observations were due to the procedure in an individual eye.

[0320] All animals on the study remained in good health. Body weight (FIG. 12A) and heart rate (FIG. 12B) measurements did not reveal any clinical signs indicative of adverse systemic effects caused by dosing of exon editors. Cage-side observations were performed daily and food consumption was monitored, and no changes were found.

[0321] The studies described herein constitute the first trans-splicing report at clinically relevant efficiencies in the NHP retina. Exon editing and replacement demonstrated efficiency and tolerability *in vivo* in African Green Monkey non-human primates following subretinal injection. All subretinal delivered AAV-mediated exon editing molecules were well-tolerated. These studies highlight the potential of RNA exon editing to treat ABCA4-related retinopathies.

Example 5. Evaluation of Exemplary ABCA4 trans-splicing molecules in vitro

Activity of an exemplary ABCA4 RNA Exon Editor following transfection into ABCA4 Protein KO Cell Line

[0322] As described hereinabove, a HEK293T-based cell line, termed B6, was engineered that expresses both high ABCA4 RNA and protein levels. This was achieved via the knock-in (KI) of a constitutive CAGGS promoter between the endogenous ABCA4 promoter and exon 1 via CRISPR-mediated homology-directed repair (HDR).

[0323] To create a model system in which to characterize an ABCA4 RNA exon editor's ability to correct a targeted native pre-mRNA, generate a trans-spliced on-target (ONT) RNA, and encode a full-length ABCA4 protein, a protein knock-out (KO) cell line (termed 17+06) was engineered based on the B6 cell line. 17+06 cells continued to produce high levels of ABCA4 RNA, but in contrast to the B6 cell line, 17+06 cells do not produce ABCA4 protein. Engineering 17+06 cells was achieved by introducing mutations into the ABCA4 gene coding sequence using CRISPR and guide RNAs (gRNA) targeting exons 3 and 4. RT-qPCR confirmed the persistence of ABCA4 RNA expression in the 17+06 cell line, and Western blot analysis indicated that ABCA4 protein translation was abolished. See FIG. 3C.

[0324] ABCA4 KO (17+06) cells were transfected with several ABCA4 RNA Exon Editor constructs to assess trans-splicing efficiency and resultant expression of ONT trans-spliced ABCA4 protein. The constructs were identical but for the elements indicated as distinguishing each as reflected in the labels for the 4 lanes in FIGs. 19A and 19B. An exemplary ABCA4 RNA Exon Editor Construct comprising an exemplary ABCA4 RNA Exon Editor depicted in FIG. 13 (5' UTR to the end of MALAT1-MascRNA (anti-mut1; SEQ ID NO: 78) emerged from this analysis as a lead candidate because it drives high levels of trans-splicing (~20% replacement of ABCA4 RNA) and high levels of ONT trans-spliced ABCA4 protein, while reducing levels of NSP relative to the other constructs tested. The activity of the exemplary ABCA4 RNA Exon Editor is shown in lane 4 of FIGs. 19A and

19B. As indicated in FIGs. 19A and 19B, the exemplary ABCA4 RNA Exon Editor comprises the 40mer linker + 3XUBS + ARE elements described in detail herein.

[0325] Further to the above, FIGs. 19A and 19B illustrate the activity of each of the indicated RNA exon editor (RTM) elements, alone and in combination, in the context of an exemplary ABCA4 5' construct. FIG. 19A shows a histogram depicting percent replacement of ABCA4 RNA for each construct indicated. The percent replacement ranges from around 20-30% for each construct. Intriguingly, the combination of 40mer + 3x repeat of a U1 binding site (3XUBS) (SEQ ID NO: 72) shows a trend toward increasing percent replacement of ABCA4 RNA relative to 40mer alone (SEQ ID NO: 27). FIG. 19B presents western blots showing that the indicated constructs comprising the 40mer (SEQ ID NO: 27), 40mer + 3XUBS (SEQ ID NO: 72), 40mer + AU-rich elements (sequences 3444-3529 of SEQ ID NO: 86), and 40mer + 3XUBS + ARE (SEQ ID NO: 73) express trans-spliced ABCA4 protein (ONT) and in combination, these RTM elements exhibit a trend toward increased ONT levels relative to 40mer alone. The western blots also show that each of the combinations of 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE exhibit a combinatorial effect on non-spliced protein (NSP) reduction relative to 40mer alone. The combinatorial effect on NSP reduction is particularly pronounced for 40mer + 3XUBS + ARE (SEQ ID NO: 73) relative to 40mer (SEQ ID NO: 27) alone. In a minimal NSP context (contains stop codons immediately following the splice donor sequence in the exon editor), the combination of two strategies that independently reduced NSP levels – the inclusion of 3x UBS and an AU-rich element – showed a concerted effect on NSP level reduction. This combination reduced the NSP level by ~70% compared to 40mer only, down to the level seen in a non-stop exon editor.

[0326] It is noteworthy that the cassette of the 40mer + 3XUBS + ARE (SEQ ID NO: 73) confers an even more significant decrease in NSP levels relative to that of RNA exon editors comprising 40mer only, 40mer + 3XUBS, or 40mer + ARE. This combination of elements (40mer + 3XUBS + ARE; SEQ ID NO: 73), therefore, exhibits particularly advantageous properties in the context of trans-splicing molecules such as those described herein.

[0327] Results presented in the context of ABCA4 RNA exon editors, were recapitulated in the context of other RNA exon editors targeting different introns in target pre-mRNAs of genes other than ABCA4. See, e.g. FIGs. 27 and 28, which are described in greater detail below, wherein the ability of each of the 40mer + 3XUBS (SEQ ID NO: 72), the 40mer + ARE (sequences 3444-3529 of SEQ ID NO: 86), and the 40mer + 3XUBS + ARE (SEQ ID NO: 73) to reduce NSP levels when compared to the 40mer alone (SEQ ID NO: 27) is depicted. The arrangement and positioning of the 40mer + 3XUBS (SEQ ID NO: 72), the 40mer + ARE (sequences 3444-3529 of SEQ ID NO: 86), and the 40mer + 3XUBS + ARE (SEQ ID NO: 73) in the context of different RNA exon editors is consistent with that of ABCA4 exon editors described herein and depicted in FIG. 13.

[0328] As shown in FIG. 27B, NSP was reduced in a combinatorial manner by the inclusion of 3X UBS and an ARE in the Intron A-targeting RNA Exon Editor for which results are depicted. The Intron A-targeting RNA Exon Editor targets an intron in a pre-mRNA that is different from ABCA4.

Briefly, HEK293 cells were transfected with Intron A-targeting RNA Exon Editors that varied at their linker region to include the indicated RNA exon editor elements (40mer, 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE). Cells were harvested 48 hours post-transfection, assayed for trans-splicing efficiencies by RT-qPCR (FIG. 27A) or subjected to Western Blot analysis (FIG. 28B). In Intron A-targeting Exon Editors, the present inventors show that 40mer + 3X UBS reduced NSP levels by ~ 75% relative to 40mer alone, 40mer + ARE reduced NSP levels by ~40% relative to 40mer alone, and 40mer + 3XUBS +ARE reduced NSP levels by 88% relative to the 40mer-only (FIG. 27B).

[0329] As shown in FIG. 28B, NSP was reduced in a combinatorial manner by the inclusion of 3X UBS and an ARE in the Intron B-targeting RNA Exon Editor for which results are depicted. The Intron B-targeting RNA Exon Editor targets an intron in a pre-mRNA that is different from ABCA4 and from the pre-mRNA targeted by the Intron-A targeting RNA Exon Editor. Briefly, HEK293 cells were transfected with Intron B-targeting RNA Exon Editors that varied at their linker region to include the indicated RNA exon editor elements (40mer, 40mer + 3XUBS, 40mer + ARE, and 40mer + 3XUBS + ARE). Cells were harvested 48 hours post-transfection, assayed for trans-splicing efficiencies by RT-qPCR (FIG. 28A) or subjected to Western Blot analysis (FIG. 28B). In Intron B-targeting Exon Editors, the present inventors show that 40mer + 3X UBS reduced NSP levels by ~ 38% relative to 40mer alone, 40mer + ARE reduced NSP levels by ~33% relative to 40mer alone, and 40mer + 3XUBS +ARE reduced NSP levels by 66% relative to the 40mer-only (FIG. 28B).

Dose-Response evaluation of exemplary ABCA4 RNA Exon Editor following transduction of ABCA4 KO Cell Line

[0330] To demonstrate confidence in the ABCA4 RNA Exon Editor Construct pharmacology and mechanism of action, the present inventors packaged an exemplary ABCA4 RNA Exon Editor Construct (ABCA4-01) comprising an exemplary ABCA4 RNA Exon Editor depicted in FIG. 13 (SEQ ID NO: 90, which comprises SEQ ID NO: 78) into a surrogate AAV2 capsid and evaluated a dose response in the ABCA4 KO cell line. ABCA4 KO cells were transduced with increasing MOIs 1×10^3 to 1×10^6 ($1e3$ - $1e6$) of AAV2-ABCA4 RNA Exon Editor Construct comprising an exemplary ABCA4 RNA Exon Editor depicted in FIG. 13 (SEQ ID NO: 90, which comprises EQ ID NO: 78), followed by membrane protein isolation and Western blotting with an ABCA4-specific Ab. Increasing MOI of the AAV2-ABCA4 RNA Exon Editor Construct comprising an exemplary ABCA4 RNA Exon Editor depicted in FIG. 13 (SEQ ID NO: 90) resulted in increasing levels of exon editing as reflected in on-target (ONT) ABCA4 protein expression and % RNA replacement. See FIGs. 21A and 21B.

Example 6. Off-target (OFT) discovery Next Generation Sequence (NGS) Method (OFT-seq)

[0331] OFT RNA sequencing (OFT-seq) is a global OFF-target discovery method developed by the present inventors. This method utilizes RNA extracted from cells, NHP neural retina samples, or human retinal explants that have either been transfected with plasmids or transduced with AAV encoding ABCA4 Exon Editors or relevant controls. The Exon Editor specific codon optimized CDS upstream of the 3' of Exon22 is used as a targeted primer (gene specific primer: GSP) region. Since

this method only targets upstream sequences of the Exon Editor, it can be used to capture unbiased global OFF-target non-ABCA4 RNA sequences. This method, therefore, allows capture of non-spliced Exon Editor, self-spliced Exon Editor, and intended ONT edited ABCA4 RNA and OFT (Figure 25). These sequences are then sequenced with an Illumina short read NGS platform. Output NGS reads are processed with proprietary computational bioinformatics analysis tools developed in house. This analysis provides overall sequencing quality and target junction sequences as well as quantitative information for each variant sequence.

[0332] In brief, RNA samples were first treated with DNase to eliminate residual genomic DNA, plasmid DNA/AAV DNA, and possible contaminant DNA. Then enrichment of RNA was performed with either ribosomal RNA (rRNA) depletion or PolyA enrichment. Enriched RNA samples were then fragmented. After Illumina adapter ligation and cDNA preparation, GSP-targeting ABCA4 codon optimized sequence and adapter targeting primer were used for PCR. This step amplifies target sequences bearing ABCA4 codon optimized sequences - including Exon Editor, ONT, OFT and self-spliced Exon Editor.

Example 7. Generating cryptic splice site resistant ABCA4 CDSs

[0333] The present inventors discovered through extensive experimental and computational analyses of results in vitro and in vivo (NHP) that cryptic splice sites are utilized in the context of the ABCA4 exon editors described herein. The present inventors performed extensive experimentation and analyses, including the generation of an off-target (OFT) library and bioinformatic analyses thereof, and surprisingly discovered numerous locations within the ABCA4 coding sequence (CDS) comprising exons 1-22 wherein cryptic splicing occurs in the context of trans-splicing molecules described herein. To improve trans-splicing efficiency and fidelity of trans-splicing molecules described herein, the sequences of a plurality of the cryptic splice sites were modified to impair and/or disable the cryptic splice sites by introduction of synonymous mutations therein. Such an approach takes advantage of the degeneracy of the genetic code, which permits introduction of the desired modifications to mitigate cryptic splice site use without resulting in a change in the encoded amino acid.

[0334] Further to the above, using a Next Generation Sequencing (NGS) approach, the present inventors identified cryptic splice sites and quantitated the frequency of cryptic splice use at each of the different locations within the ABCA4 CDS comprising exons 1-22, thereby establishing a protocol for data driven, prioritized mitigation of cryptic splice use therein. See Table 3 and FIG. 26. As shown in Table 3, the cryptic splice sites identified are arranged in order of frequency of use as determined by the average counts per million reads mapped (CPM) incorporated before introduction of synonymous mutations. Accordingly, the results and bioinformatic analyses thereof establish a prioritized list of cryptic splice sites in ABCA4 exons 1-22 that may be modified to generate an ABCA4 CDS having improved properties (e.g., cryptic splice site resistant ABCA4 CDS) and functionality in the context of a trans-splicing molecule.

[0335] An NGS approach was also implemented to assess the frequency of cryptic splice usage after introduction of synonymous mutations. See Table 3 and FIG. 26. Confirmation of cryptic

splice site mitigation by such an approach linked the structural change of the sequence modification (synonymous mutation) to the function conferred thereby, namely generation of cryptic splice site resistant ABCA4 CDSs. FIG. 26 illustrates the function conferred by introduction of synonymous mutations at identified cryptic splice sites in an exemplary ABCA4 exon editor by showing the frequency of cryptic splice site cleavage before introduction of synonymous mutations at identified cryptic splice sites (top panel) and after introduction of synonymous mutations at identified cryptic splice sites (bottom panel; splice site resistant sequence). FIG. 26 shows that a significant reduction in splicing frequency is observed following cryptic splice site mitigation.

Table 3: Lists cryptic splice sites identified in ABCA4 coding domain sequences (CDS) and frequency of splicing at each splice site identified before and after splice site mitigation. SEQ ID NO: 46 represents an exemplary sequence before introduction of synonymous mutations; SEQ ID NO: 79 represents an exemplary sequence after introduction of synonymous mutations. The nucleotide position indicated in the rtm_pos column refers to the position of the last nucleotide indicated in column 2 or column 3 relative to the G of the ATG codon (the translational initiation site). The abbreviation “not” is used to indicate that a nucleotide change was not made in the indicated sequence (SEQ ID NO: 79); the abbreviation for branchpoint is “bp”.

rtm_pos	ABCA4 CDS before splice site mitigation	ABCA4 CDS after splice site mitigation	correction	ABCA4 CDS before splice site mitigation _avg_cpm	ABCA4 CDS after splice site mitigation _avg_cpm	Annotation
3086	ACATGCTGTTCT ATGCCAG (SEQ ID NO: 133)	ACATGCTGTT CTATGCACAA (SEQ ID NO: 92)	splice-site	1480.81474 5	1.742685513	ABCA4 ex21
2931	CACCTTGTCAT CCTCACAG (SEQ ID NO: 134)	CACCTTGTC ATCCTCACTG (SEQ ID NO: 94)	splice-site	256.316569 6	0	ABCA4 ex20
2576	TCGCTTGGTACC TTGATCAG (SEQ ID NO: 135)	TCGCTTGTA CCTTGATCAA (SEQ ID NO: 96)	splice-site	142.190876 3	0.961889921 3	ABCA4 ex16
3029	GCCTTGTCATGT GTCCACAG (SEQ ID NO: 136)	GCCTTGTCAT GTGTCCACAG (SEQ ID NO: 97)	not	87.6323617 1	183.2327206	ABCA4 ex20
127	CTGGTCTTGATC TGGTTAAG (SEQ ID NO: 137)	CTGGTCTTGA TCTGGTTACG (SEQ ID NO: 100)	splice-site	48.2768858 1	0	ABCA4 ex1
2348	TCCTGTGCTTCG CCTGGCAG (SEQ ID NO: 138)	TCCTGTGCTT CGCCTGGCAA (SEQ ID NO: 102)	splice-site	46.3682731	0	ABCA4 ex15

1835	GCGGGTTTGCCT ATCTGCAG (SEQ ID NO: 139)	GCGGGTTTGC ATATCTGCAA (SEQ ID NO: 104)	splice- site,branch point	40.5448602 7	0	ABCA4 ex13
1218	CACTCCTGATTC ACCTGCAG (SEQ ID NO: 140)	CACTCCTGAT TCACCTGCTG (SEQ ID NO: 106)	splice-site	34.906711	0.961889921 3	ABCA4 ex9
3083	AGCACATGCTGT TCTATGCC (SEQ ID NO: 141)	AGCACATGCT GTTCTATGCA (SEQ ID NO: 108)	not	29.0454237 8	0	ABCA4 ex21
2110	TGTACCTGGTTC CTGGACAG (SEQ ID NO: 142)	TGTACCTGGT TCCTGGACTC (SEQ ID NO: 110)	splice-site	27.6329014 5	0.871342756 6	ABCA4 ex14
2628	TTGGTACTTTCTT CTACAAG (SEQ ID NO: 143)	TTGGTACTTTCT TTCTACAAG (SEQ ID NO: 111)	not	23.5686579 5	25.48416316	ABCA4 ex17
3103	CAGCTGAAAGGA AAGTCCCA (SEQ ID NO: 144)	CAACTGAAAG GAAAGTCCCA (SEQ ID NO: 114)	not	20.9169367 8	26.29908764	ABCA4 ex21
3093	GTTCTATGCCCA GCTGAAAG (SEQ ID NO: 145)	GTTCTATGCA CAACTGAAAG (SEQ ID NO: 116)	not	20.0243849 5	64.52533239	ABCA4 ex21
701	CCCTGTGCTCCC TCTCCCAG (SEQ ID NO: 146)	CCCTGTGCTC CCTCTCCCAG (SEQ ID NO: 117)	not	18.7851861 1	15.44665702	ABCA4 ex6
528	CATCGGCCTGTC TGACTCAG (SEQ ID NO: 147)	CATCGGCCTG TCTGACTCTG (SEQ ID NO: 120)	splice-site	15.5636799 9	0	ABCA4 ex5
1451	TCCTAAACTTCC TCTACAAG (SEQ ID NO: 148)	TCCTAAACTTCC CTCTACAAG (SEQ ID NO: 121)	not	14.4978046 6	0	
2015	CTGTCTCCATGA CTGTGAAG (SEQ ID NO: 149)	CTGTCTCCAT GACTGTGAAG (SEQ ID NO: 123)	not	12.2923594 1	12.23292748	
3179	AGCGGAATGAAG AGGCTCAG (SEQ ID NO: 150)	AGCGGAATGA AGAGGCTCAG (SEQ ID NO: 125)	not	11.3380530 6	9.16616339	

2254	ATGCTGTGCTTT CTGCTCAG (SEQ ID NO: 151)	ATGCTGTGCT TTCTGCTCAG (SEQ ID NO: 126)	not	9.19436232 8	44.59728553	
1124	TTTGTAAATGCATT GATCCAG (SEQ ID NO: 152)	TTTGTAAATGCA TTGATCCAG (SEQ ID NO: 128)	not	8.57476291 2	14.84695575	
1907	TTGGAATCTACC TCCAGCAG (SEQ ID NO: 153)	TTGGAATCTA CCTCCAGCAG (SEQ ID NO: 130)	not	8.17830145 4	24.70336757	
1090	CCTATCTATTCTT ATGACAG (SEQ ID NO: 154)	CCTATCTATTC TTATGACAG (SEQ ID NO: 671)	not	6.93910262 1	22.96068206	
555	CCTTCTGATCAA CTCTCAAG (SEQ ID NO: 155)	CCTTCTGATC AACTCTCAAG (SEQ ID NO: 672)	not	6.54264116 3	27.31739584	
2801	GGGTATGCGTGA AGAATCTG (SEQ ID NO: 156)	GGGTATGCGT GAAGAATCTG (SEQ ID NO: 673)	not	6.14617970 6	0	
3104	AGCTGAAAGGAA AGTCCCAG (SEQ ID NO: 157)	AACTGAAAGG AAAGTCCCAG (SEQ ID NO: 674)	not	6.03461072 6	0.871342756 6	
3147	AGCCATGTTGGA GGACACAG (SEQ ID NO: 158)	AGCCATGTTG GAGGACACAG (SEQ ID NO: 675)	not	5.86128722 7	7.332930712	
1333	TGGTACTTCTTT GACAACAG (SEQ ID NO: 159)	TGGTACTTCTT TGACAACAG (SEQ ID NO: 676)	not	5.74971824 8	11.96128599	
3116	AGTCCCAGGAG GAGGCCAG (SEQ ID NO: 160)	AGTCCCAGGA GGAGGCCCA G (SEQ ID NO: 677)	not	5.13011883 1	0.961889921 3	
2082	TCAGGGTGTCTC CAATGCAG (SEQ ID NO: 161)	TCAGGGTGTCTC TCCAATGCAG (SEQ ID NO: 678)	not	5.13011883 1	9.075616225	
3084	GCACATGCTGTT CTATGCCC (SEQ ID NO: 162)	GCACATGCTG TTCTATGCAC (SEQ ID NO: 679)	not	5.13011883 1	0	

1904	CAGTTGGAATCT ACCTCCAG (SEQ ID NO: 163)	CAGTTGGAAT CTACCTCCAG (SEQ ID NO: 680)	not	4.90698087 2	7.932631974	
2874	CTTCTACGAGAA CCAGATCA (SEQ ID NO: 164)	CTTCTACGAG AACCAGATCA (SEQ ID NO: 681)	not	4.51051941 5	5.409150869	
3088	ATGCTGTTCTAT GCCCAGCT (SEQ ID NO: 165)	ATGCTGTTCT ATGCACAACCT (SEQ ID NO: 682)	not	4.28738145 6	1.833232678	
3040	TGTCCACAGCAC AACATCCT (SEQ ID NO: 166)	TGTCCACAGC ACAACATCCT (SEQ ID NO: 683)	not	3.89091999 8	0.871342756 6	
250	AACAATCCCTGT TTTCAAAG (SEQ ID NO: 167)	AACAATCCCT GTTTTCAAAG (SEQ ID NO: 684)	not	3.89091999 8	2.61402827	
2820	GGTAAAGATTTT TGAGCCCT (SEQ ID NO: 168)	GGTGAAGATT TTTGAGCCCT (SEQ ID NO: 685)	branchpoint	3.77935101 9	2.795122599	
3160	GACACAGGCCTC CACCACAA (SEQ ID NO: 169)	GACACAGGCC TCCACCACAA (SEQ ID NO: 686)	not	3.60602752	2.704575435	
1936	CCCTGCTTCGTG GACGATTC (SEQ ID NO: 170)	CCCTGCTTCG TGGACGATTC (SEQ ID NO: 687)	not	3.38288956 1	2.795122599	
1396	GTAAAAGACTTT TTGAATAG (SEQ ID NO: 171)	GTAAAAGACT TTTTGAATAG (SEQ ID NO: 688)	not	3.38288956 1	4.628355277	
2703	AAAGACCGAGCC CCTAACAG (SEQ ID NO: 172)	AAAGACCGAG CCCCTAACAG (SEQ ID NO: 689)	not	3.38288956 1	2.885669764	
2959	CCTCCAACCTCT GGGACTGT (SEQ ID NO: 173)	CCTCCAACCT CTGGGACTGT (SEQ ID NO: 690)	not	3.38288956 1	0.871342756 6	
783	TGTGCTTCCCAC ACTCCTAG (SEQ ID NO: 174)	TGTGCTTCCC ACACTCCTAG (SEQ ID NO: 691)	not	3.27132058 2	2.795122599	

2522	GCTTCCTGCTGT CCATGCAG (SEQ ID NO: 175)	GCTTCCTGCT GTCCATGCAG (SEQ ID NO: 692)	not	3.04818262 3	21.03690221	
3097	TATGCCCAGCTG AAAGGAAA (SEQ ID NO: 176)	TATGCACAAC TGAAAGGAAA (SEQ ID NO: 693)	not	3.04818262 3	2.795122599	
176	AATGCCATTTC CCAACAAG (SEQ ID NO: 177)	AATGCCATTT CCCCAACAAG (SEQ ID NO: 694)	not	3.04818262 3	0.871342756 6	
3085	CACATGCTGTT TATGCCCA (SEQ ID NO: 178)	CACATGCTGT TCTATGCACA (SEQ ID NO: 695)	not	2.98642810 3	0	
649	GAGCGCTTCATC ATCTTCAG (SEQ ID NO: 179)	GAGCGCTTCA TCATCTTCAG (SEQ ID NO: 696)	not	2.98642810 3	0.871342756 6	
3118	TCCCAGGAGGA GGCCCAGCT (SEQ ID NO: 180)	TCCCAGGAGG AGGCCAGCT (SEQ ID NO: 697)	not	2.76329014 5	0.871342756 6	
3098	ATGCCAGCTGA AAGGAAAG (SEQ ID NO: 181)	ATGCACAAC GAAAGGAAAG (SEQ ID NO: 698)	not	2.76329014 5	1.923779843	
3128	AGGCCAGCTG GAGATGGAA (SEQ ID NO: 182)	AGGCCAGCT GGAGATGGAA (SEQ ID NO: 699)	not	2.76329014 5	1.833232678	
1982	TCCCTATCTTCA TGGTGCTG (SEQ ID NO: 183)	TCCCTATCTTC ATGGTGCTG (SEQ ID NO: 700)	not	2.65172116 5	0.961889921 3	
2586	CCTTGATCAGGT GTTTCCAG (SEQ ID NO: 184)	CCTTGATCAA GTGTTTCCAG (SEQ ID NO: 701)	not	2.54015218 6	0.871342756 6	
2845	CGGCCAGCTGT GGACCGTCT (SEQ ID NO: 185)	CGGCCAGCTG TGGACCGTCT (SEQ ID NO: 702)	not	2.54015218 6	0	
1782	TTGGGATTCTGG TCCCAGAG (SEQ ID NO: 186)	TTGGGATTCT GGTCCCAGAG (SEQ ID NO: 703)	not	2.36682868 7	0	

2901	CCTGGGCCACAA TGGAGCTG (SEQ ID NO: 187)	CCTGGGCCAC AATGGAGCTG (SEQ ID NO: 704)	not	2.25525970 7	3.757012521	
310	AACAACTCCATC TTGGCAAG (SEQ ID NO: 188)	AACAACTCCA TCTTGGCAAG (SEQ ID NO: 705)	not	2.25525970 7	8.204273468	
2137	ATCATGTCTGATG AGCATCTT (SEQ ID NO: 189)	ATCATGTCTGA TGAGCATCTT (SEQ ID NO: 706)	not	2.25525970 7	0	
3067	CACCTCACGGTG GCTGAGCA (SEQ ID NO: 190)	CACCTCACGG TGGCTGAGCA (SEQ ID NO: 707)	not	2.25525970 7	0	
3090	GCTGTTCTATGC CCAGCTGA (SEQ ID NO: 191)	GCTGTTCTAT GCACAACTGA (SEQ ID NO: 708)	not	2.14369072 8	0.871342756 6	
3133	CAGCTGGAGATG GAAGCCAT (SEQ ID NO: 192)	CAGCTGGAGA TGGAAGCCAT (SEQ ID NO: 709)	not	2.14369072 8	0.961889921 3	
3066	CCACCTCACGGT GGCTGAGC (SEQ ID NO: 193)	CCACCTCACG GTGGCTGAGC (SEQ ID NO: 710)	not	2.14369072 8	0	
3152	TGTTGGAGGACA CAGGCCTC (SEQ ID NO: 194)	TGTTGGAGGA CACAGGCCTC (SEQ ID NO: 711)	not	2.03212174 8	0	
554	ACCTTCTGATCA ACTCTCAA (SEQ ID NO: 195)	ACCTTCTGAT CAACTCTCAA (SEQ ID NO: 712)	not	2.03212174 8	0	
2870	TCACCTTCTACG AGAACCAG (SEQ ID NO: 196)	TCACCTTCTA CGAGAACCAG (SEQ ID NO: 713)	not	1.85879825	1.833232678	
352	CTCCTCATGAAT GCACCAGA (SEQ ID NO: 197)	CTCCTCATGA ATGCACCAGA (SEQ ID NO: 714)	not	1.85879825	0	
1203	GGGAAAAATCCT GTACACTC (SEQ ID NO: 198)	GGGAAAAATC CTGTACACTC (SEQ ID NO: 715)	not	1.85879825	0	

1966	ATCCTGAACCGC TGTTTCCC (SEQ ID NO: 199)	ATCCTGAACC GCTGTTTCCC (SEQ ID NO: 716)	not	1.85879825	0	
3057	CCTGTTCCACCA CCTCACGG (SEQ ID NO: 200)	CCTGTTCCAC CACCTCACGG (SEQ ID NO: 717)	not	1.85879825	0.871342756 6	
2996	ACATTGAAACCA GCCTGGAT (SEQ ID NO: 201)	ACATTGAAAC CAGCCTGGAT (SEQ ID NO: 718)	not	1.85879825	0	
2372	GCATGACCGCTG AGCTGAAG (SEQ ID NO: 202)	GCATGACCGC TGAGCTGAAG (SEQ ID NO: 719)	not	1.85879825	0	
2966	CCTCTGGGACTG TGCTCGTT (SEQ ID NO: 203)	CCTCTGGGAC TGTGCTCGTT (SEQ ID NO: 720)	not	1.85879825	0	
2818	CTGGTAAAGATT TTTGAGCC (SEQ ID NO: 204)	CTGGTGAAGA TTTTTGAGCC (SEQ ID NO: 721)	branchpoint	1.74722927	1.923779843	
2811	GAAGAATCTGGT AAAGATTT (SEQ ID NO: 205)	GAAGAATCTG GTGAAGATTT (SEQ ID NO: 722)	branchpoint	1.74722927	0	
3080	CTGAGCACATGC TGTTCTAT (SEQ ID NO: 206)	CTGAGCACAT GCTGTTCTAT (SEQ ID NO: 723)	not	1.74722927	0	
3081	TGAGCACATGCT GTTCTATG (SEQ ID NO: 207)	TGAGCACATG CTGTTCTATG (SEQ ID NO: 724)	not	1.74722927	0.961889921 3	
2832	TGAGCCCTGTGG CCGGCCAG (SEQ ID NO: 208)	TGAGCCCTGT GGCCGGCCA G (SEQ ID NO: 725)	not	1.74722927	0	
3125	AGGAGGCCCCAG CTGGAGATG (SEQ ID NO: 209)	AGGAGGCCCA GCTGGAGATG (SEQ ID NO: 726)	not	1.74722927	0	
1860	GGTTGAACAGG GGATCACAA (SEQ ID NO: 210)	GGTTGAACAG GGGATCACAA (SEQ ID NO: 727)	not	1.74722927	0	

3123	GGAGGAGGCC AGCTGGAGA (SEQ ID NO: 211)	GGAGGAGGC CCAGCTGGAG A (SEQ ID NO: 728)	not	1.74722927	0	
2774	AGAGAGAACACC CCGGGTGG (SEQ ID NO: 212)	AGAGAGAACACC CCCGGGTGG G (SEQ ID NO: 729)	not	1.63566029 1	7.932631974	
3071	TCACGGTGGCTG AGCACATG (SEQ ID NO: 213)	TCACGGTGGC TGAGCACATG (SEQ ID NO: 730)	not	1.63566029 1	0.961889921 3	
2600	TTCCAGGAGACT ATGGAACC (SEQ ID NO: 214)	TTCCAGGAGA CTATGGAACC (SEQ ID NO: 731)	not	1.63566029 1	0	
3119	CCCAGGAGGAG GCCCAGCTG (SEQ ID NO: 215)	CCCAGGAGGA GGCCCAGCTG (SEQ ID NO: 732)	not	1.63566029 1	0	
2713	CCCCTAACAGAG GAAACGGA (SEQ ID NO: 216)	CCCCTAACAG AGGAAACGGA (SEQ ID NO: 733)	not	1.63566029 1	0	
2721	AGAGGAAACGG AGGATCCAG (SEQ ID NO: 217)	AGAGGAAACG GAGGATCCAG (SEQ ID NO: 734)	not	1.63566029 1	0	
2857	GACCGTCTGAAC ATCACCTT (SEQ ID NO: 218)	GACCGTCTGA ACATCACCTT (SEQ ID NO: 735)	not	1.63566029 1	0	
3046	CAGCACAACATC CTGTTCCA (SEQ ID NO: 219)	CAGCACAACA TCCTGTTCCA (SEQ ID NO: 736)	not	1.63566029 1	0	
3087	CATGCTGTTCTA TGCCCAGC (SEQ ID NO: 220)	CATGCTGTTC TATGCACAAC (SEQ ID NO: 737)	not	1.63566029 1	0	
2676	TTCAACCAGAGA AGAAAGAG (SEQ ID NO: 221)	TTCAACCAGA GAAGAAAGAG (SEQ ID NO: 738)	not	1.63566029 1	0	
2831	TTGAGCCCTGTG GCCGCCA (SEQ ID NO: 222)	TTGAGCCCTG TGGCCGCCA (SEQ ID NO: 739)	not	1.52409131 1	3.847559685	

148	AATGCCAACCCG CTCTACAG (SEQ ID NO: 223)	AATGCCAACC CGCTCTACAG (SEQ ID NO: 740)	not	1.52409131 1	2.61402827	
580	CCAGAGCAGTTC GCTCATGG (SEQ ID NO: 224)	CCAGAGCAGT TCGCTCATGG (SEQ ID NO: 741)	not	1.52409131 1	0	
643	CTCCTGGAGCG CTTCATCAT (SEQ ID NO: 225)	CTCCTGGAGC GCTTCATCAT (SEQ ID NO: 742)	not	1.52409131 1	0	
703	CTGTGCTCCCTC TCCCAGGG (SEQ ID NO: 226)	CTGTGCTCCC TCTCCCAGGG (SEQ ID NO: 743)	not	1.52409131 1	0	
3070	CTCACGGTGGCT GAGCACAT (SEQ ID NO: 227)	CTCACGGTGG CTGAGCACAT (SEQ ID NO: 744)	not	1.52409131 1	0.961889921 3	
2783	ACCCCGGGTGG GTTCTGTTGGG (SEQ ID NO: 228)	ACCCCGGGTG GGTTCCTGGG (SEQ ID NO: 745)	not	1.52409131 1	0	
3075	GGTGGCTGAGC ACATGCTGT (SEQ ID NO: 229)	GGTGGCTGAG CACATGCTGT (SEQ ID NO: 746)	not	1.52409131 1	0	
2852	CTGTGGACCGTC TGAACATC (SEQ ID NO: 230)	CTGTGGACCG TCTGAACATC (SEQ ID NO: 747)	not	1.52409131 1	0	
3107	TGAAAGGAAAGT CCCAGGAG (SEQ ID NO: 231)	TGAAAGGAAA GTCCCAGGAG (SEQ ID NO: 748)	not	1.52409131 1	2.795122599	
2270	TCAGCACCTTCT TCTCCAAG (SEQ ID NO: 232)	TCAGCACCTT CTTCTCCAAG (SEQ ID NO: 749)	not	1.52409131 1	1.833232678	
2988	GGGAAGGGACA TTGAAACCA (SEQ ID NO: 233)	GGGAAGGGA CATTGAAACC A (SEQ ID NO: 750)	not	1.23919883 3	0	
3149	CCATGTTGGAGG ACACAGGC (SEQ ID NO: 234)	CCATGTTGGA GGACACAGGC (SEQ ID NO: 751)	not	1.23919883 3	0	

698	ATGCCCTGTGCT CCCTCTCC (SEQ ID NO: 235)	ATGCCCTGTG CTCCCTCTCC (SEQ ID NO: 752)	not	1.23919883 3	0	
3106	CTGAAAGGAAAG TCCCAGGA (SEQ ID NO: 236)	CTGAAAGGAA AGTCCCAGGA (SEQ ID NO: 753)	not	1.23919883 3	0	
2691	AAGAGCCCTGGA AAAGACCG (SEQ ID NO: 237)	AAGAGCCCTG GAAAAGACCG (SEQ ID NO: 754)	not	1.23919883 3	0	
3135	GCTGGAGATGG AAGCCATGT (SEQ ID NO: 238)	GCTGGAGATG GAAGCCATGT (SEQ ID NO: 755)	not	1.23919883 3	0.871342756 6	
3115	AAGTCCCAGGAG GAGGCCCA (SEQ ID NO: 239)	AAGTCCCAGG AGGAGGCCCA (SEQ ID NO: 756)	not	1.23919883 3	1.923779843	
2974	ACTGTGCTCGTT GGGGGAAG (SEQ ID NO: 240)	ACTGTGCTCG TTGGGGGAAG (SEQ ID NO: 757)	not	1.23919883 3	0	
2789	GGTGGGTTCTCG GGGTATGC (SEQ ID NO: 241)	GGTGGGTTCC TGGGGTGTGC (SEQ ID NO: 758)	not	1.23919883 3	0	
2134	TCCATCATGTCG ATGAGCAT (SEQ ID NO: 242)	TCCATCATGT CGATGAGCAT (SEQ ID NO: 759)	not	1.23919883 3	0	
2700	GGAAAAGACCGA GCCCCTAA (SEQ ID NO: 243)	GGAAAAGACC GAGCCCCTAA (SEQ ID NO: 760)	not	1.23919883 3	0	
566	ACTCTCAAGTCC GTCCAGAG (SEQ ID NO: 244)	ACTCTCAAGT CCGTCCAGAG (SEQ ID NO: 761)	not	1.23919883 3	0	
3137	TGGAGATGGAAG CCATGTTG (SEQ ID NO: 245)	TGGAGATGGA AGCCATGTTG (SEQ ID NO: 762)	not	1.23919883 3	0.961889921 3	
3096	CTATGCCAGCT GAAAGGAA (SEQ ID NO: 246)	CTATGCACAA CTGAAAGGAA (SEQ ID NO: 763)	not	1.23919883 3	0.871342756 6	

19	AGCGGCTTCGTG AGACAGAT (SEQ ID NO: 247)	TCCGGCTTCG TGAGACAGAT (SEQ ID NO: 764)	not	1.23919883 3	0	
3092	TGTTCTATGCCC AGCTGAAA (SEQ ID NO: 248)	TGTTCTATGC ACAACTGAAA (SEQ ID NO: 765)	not	1.23919883 3	0	
1877	CAAGGAGCCAG GTGCAGGCG (SEQ ID NO: 249)	CAAGGAGCCA GGTGCAGGC G (SEQ ID NO: 766)	not	1.23919883 3	0	
857	TGTCACCAAGAA TTCAAGAG (SEQ ID NO: 250)	TGTCACCAAG AATTCAAGAG (SEQ ID NO: 767)	not	1.12762985 4	0.961889921 3	
2846	GGCCAGCTGTG GACCGTCTG (SEQ ID NO: 251)	GGCCAGCTGT GGACCGTCTG (SEQ ID NO: 768)	not	1.12762985 4	0	
2977	GTGCTCGTTGGG GGAAGGGA (SEQ ID NO: 252)	GTGCTCGTTG GGGAAGGG A (SEQ ID NO: 769)	not	1.12762985 4	0	
616	CTGAAGGACATC GCCTGCAG (SEQ ID NO: 253)	CTGAAGGACA TCGCCTGCAG (SEQ ID NO: 770)	not	1.12762985 4	0	
2729	CGGAGGATCCA GAGCACCCA (SEQ ID NO: 254)	CGGAGGATCC AGAGCACCCA (SEQ ID NO: 771)	not	1.12762985 4	0	
2834	AGCCCTGTGGC CGGCCAGCT (SEQ ID NO: 255)	AGCCCTGTGG CCGGCCAGCT (SEQ ID NO: 772)	not	1.12762985 4	0	
3140	AGATGGAAGCCA TGTTGGAG (SEQ ID NO: 256)	AGATGGAAGC CATGTTGGAG (SEQ ID NO: 773)	not	1.12762985 4	0	
2989	GGAAGGGACATT GAAACCAG (SEQ ID NO: 257)	GGAAGGGACA TTGAAACCAG (SEQ ID NO: 774)	not	1.12762985 4	0	
2938	TCCATCCTCACA GGCCTGCT (SEQ ID NO: 258)	TCCATCCTCA CTGGCCTGCT (SEQ ID NO: 775)	not	1.12762985 4	0	

2972	GGA CTGTGCTCG TTGGGGGA (SEQ ID NO: 259)	GGACTGTGCT CGTTGGGGGA (SEQ ID NO: 776)	not	1.12762985 4	0.961889921 3	
3127	GAGGCCAGCT GGAGATGGA (SEQ ID NO: 260)	GAGGCCAGC TGGAGATGGA (SEQ ID NO: 777)	not	1.12762985 4	0	
192	CAAGGCGATGC CCTCAGCAG (SEQ ID NO: 261)	CAAGGCGATG CCCTCAGCAG (SEQ ID NO: 778)	not	1.12762985 4	0.871342756 6	
2652	GTATTGGCTTGG CGGTGAAG (SEQ ID NO: 262)	GTATTGGCTT GGCGGTGAAG (SEQ ID NO: 779)	not	1.12762985 4	3.757012521	
1469	AGGGCCCTCGG GAAAGCCAG (SEQ ID NO: 263)	AGGGCCCTCG GGAAAGCCAG (SEQ ID NO: 780)	not	1.12762985 4	0	
3162	CACAGGCCTCCA CCACAAGC (SEQ ID NO: 264)	CACAGGCCTC CACCACAAGC (SEQ ID NO: 781)	not	1.12762985 4	2.885669764	
2705	AGACCGAGCCC CTAACAGAG (SEQ ID NO: 265)	AGACCGAGCC CCTAACAGAG (SEQ ID NO: 782)	not	1.12762985 4	0	
2665	GGTGAAGGGTG TTCAACCAG (SEQ ID NO: 266)	GGTGAAGGGT GTTCAACCAG (SEQ ID NO: 783)	not	1.12762985 4	0.961889921 3	
2927	CCACCACCTTGT CCATCCTC (SEQ ID NO: 267)	CCACCACCTT GTCCATCCTC (SEQ ID NO: 784)	not	1.12762985 4	0	
2822	TAAAGATTTTGA GCCCTGT (SEQ ID NO: 268)	TGAAGATTTTT GAGCCCTGT (SEQ ID NO: 785)	branchpoint	1.12762985 4	0	
3124	GAGGAGGCCCA GCTGGAGAT (SEQ ID NO: 269)	GAGGAGGCC CAGCTGGAGA T (SEQ ID NO: 786)	not	1.12762985 4	0	
3129	GGCCCAGCTGG AGATGGAAG (SEQ ID NO: 270)	GGCCCAGCTG GAGATGGAAG (SEQ ID NO: 787)	not	1.12762985 4	0	

2993	GGGACATTGAAA CCAGCCTG (SEQ ID NO: 271)	GGGACATTGA AACCAGCCTG (SEQ ID NO: 788)	not	1.01606087 4	0	
3101	CCCAGCTGAAAG GAAAGTCC (SEQ ID NO: 272)	CACAACTGAA AGGAAAGTCC (SEQ ID NO: 789)	not	1.01606087 4	0	
2650	TCGTATTGGCTT GGCGGTGA (SEQ ID NO: 273)	TCGTATTGGC TTGGCGGTGA (SEQ ID NO: 790)	not	1.01606087 4	0	
2816	ATCTGGTAAAGA TTTTTGAG (SEQ ID NO: 274)	ATCTGGTGAA GATTTTTGAG (SEQ ID NO: 791)	branchpoint	1.01606087 4	3.757012521	
2361	CTGGCAGGACC GCATGACCG (SEQ ID NO: 275)	CTGGCAAGAC CGCATGACCG (SEQ ID NO: 792)	not	1.01606087 4	0	
592	GCTCATGGAGTC CCGGACCT (SEQ ID NO: 276)	GCTCATGGAG TCCCGGACCT (SEQ ID NO: 793)	not	1.01606087 4	0	
490	AAAGATGAAGAA AACTGAC (SEQ ID NO: 277)	AAAGATGAAG AAACTGAC (SEQ ID NO: 794)	not	1.01606087 4	0	
2260	TGCTTTCTGCTC AGCACCTT (SEQ ID NO: 278)	TGCTTTCTGC TCAGCACCTT (SEQ ID NO: 795)	not	1.01606087 4	0	
2104	ATTTGGTGTACC TGGTTCCT (SEQ ID NO: 279)	ATTTGGTGTA CCTGGTTCCT (SEQ ID NO: 796)	not	1.01606087 4	0	
2285	CCAAGGCCAGTC TGGCAGCA (SEQ ID NO: 280)	CCAAGGCCAG TCTGGCAGCA (SEQ ID NO: 797)	not	1.01606087 4	0	
2806	TGCGTGAAGAAT CTGGTAAA (SEQ ID NO: 281)	TGCGTGAAGA ATCTGGTGAA (SEQ ID NO: 798)	not	1.01606087 4	1.833232678	
2808	CGTGAAGAATCT GGTAAAGA (SEQ ID NO: 282)	CGTGAAGAAT CTGGTGAAGA (SEQ ID NO: 799)	not	1.01606087 4	0.961889921 3	

2257	CTGTGCTTTCTG CTCAGCAC (SEQ ID NO: 283)	CTGTGCTTTC TGCTCAGCAC (SEQ ID NO: 800)	not	1.01606087 4	0.961889921 3	
2778	AGAACACCCCG GGTGGGTTTC (SEQ ID NO: 284)	AGAACACCCC GGGTGGGTTTC (SEQ ID NO: 801)	not	1.01606087 4	0.871342756 6	
3049	CACAACATCCTG TTCCACCA (SEQ ID NO: 285)	CACAACATCC TGTTCCACCA (SEQ ID NO: 802)	not	1.01606087 4	1.923779843	
2052	GGAGTTGCGACT GAAGGAGA (SEQ ID NO: 286)	GGAGTTGCGA CTGAAGGAGA (SEQ ID NO: 803)	not	1.01606087 4	0	
3121	CAGGAGGAGGC CCAGCTGGA (SEQ ID NO: 287)	CAGGAGGAG GCCCAGCTGG A (SEQ ID NO: 804)	not	1.01606087 4	0	
767	TCTTCAAGCTCT TCCGTGTG (SEQ ID NO: 288)	TCTTCAAGCT CTTCCGTGTG (SEQ ID NO: 805)	not	1.01606087 4	0.871342756 6	
1861	GTTGAACAGGG GATCACAAG (SEQ ID NO: 289)	GTTGAACAGG GGATCACAAG (SEQ ID NO: 806)	not	1.01606087 4	0	
702	CCTGTGCTCCCT CTCCCAGG (SEQ ID NO: 290)	CCTGTGCTCC CTCTCCCAGG (SEQ ID NO: 807)	not	1.01606087 4	0	
3000	TGAAACCAGCCT GGATGCAG (SEQ ID NO: 291)	TGAAACCAGC CTGGATGCAG (SEQ ID NO: 808)	not	1.01606087 4	0.871342756 6	
2098	GCAGTGATTTGG TGTAACCTG (SEQ ID NO: 292)	GCAGTGATTT GGTGTAACCTG (SEQ ID NO: 809)	not	1.01606087 4	1.742685513	
2782	CACCCCGGGTG GGTTCCTGG (SEQ ID NO: 293)	CACCCCGGGT GGGTTCTGG (SEQ ID NO: 810)	not	1.01606087 4	0.871342756 6	
2862	TCTGAACATCAC CTTCTACG (SEQ ID NO: 294)	TCTGAACATC ACCTTCTACG (SEQ ID NO: 811)	not	1.01606087 4	0	

2273	GCACCTTCTTCT CCAAGGCC (SEQ ID NO: 295)	GCACCTTCTT CTCCAAGGCC (SEQ ID NO: 812)	not	1.01606087 4	0	
569	CTCAAGTCCGTC CAGAGCAG (SEQ ID NO: 296)	CTCAAGTCCG TCCAGAGCAG (SEQ ID NO: 813)	not	1.01606087 4	0.961889921 3	
676	CGCGGGGCAAA GACGGTGCG (SEQ ID NO: 297)	CGCGGGGCA AAGACGGTGC G (SEQ ID NO: 814)	not	1.01606087 4	0	
3141	GATGGAAGCCAT GTTGGAGG (SEQ ID NO: 298)	GATGGAAGCC ATGTTGGAGG (SEQ ID NO: 815)	not	1.01606087 4	0	
2627	CTTGGTACTTTC TTCTACAA (SEQ ID NO: 299)	CTTGGTACTTT CTTCTACAA (SEQ ID NO: 816)	not	1.01606087 4	1.923779843	
2009	TCTACTCTGTCT CCATGACT (SEQ ID NO: 300)	TCTACTCTGT CTCCATGACT (SEQ ID NO: 817)	not	1.01606087 4	0	
1850	TGCAGGACATGG TTGAACAG (SEQ ID NO: 301)	TGCAAGACAT GGTTGAACAG (SEQ ID NO: 818)	not	1.01606087 4	0	
1887	GGTGCAGGCGG AGGCTCCAG (SEQ ID NO: 302)	GGTGCAGGC GGAGGCTCCA G (SEQ ID NO: 819)	not	1.01606087 4	0	
2315	GTGTCATCTATT TCACCCTC (SEQ ID NO: 303)	GTGTCATCTA TTTCAACCCTG (SEQ ID NO: 820)	not	1.01606087 4	0	
1359	GATGAACATGAT CAGAGATA (SEQ ID NO: 304)	GATGAACATG ATCAGAGATA (SEQ ID NO: 821)	not	1.01606087 4	0	
2787	CGGGTGGGTTC CTGGGGTAT (SEQ ID NO: 305)	CGGGTGGGTT CCTGGGGTGT (SEQ ID NO: 822)	not	1.01606087 4	0.961889921 3	
2953	CTGCTGCCTCCA ACCTCTGG (SEQ ID NO: 306)	CTGCTGCCTC CAACCTCTGG (SEQ ID NO: 823)	not	1.01606087 4	0	

2924	AAACCACCACCT TGTCATC (SEQ ID NO: 307)	AAACCACCAC CTTGTCATC (SEQ ID NO: 824)	not	0.61959941 66	0	
3100	GCCCAGCTGAAA GGAAAGTC (SEQ ID NO: 308)	GCACAACTGA AAGGAAAGTC (SEQ ID NO: 825)	not	0.61959941 66	0	
2687	AAGAAAGAGCCC TGAAAAAG (SEQ ID NO: 309)	AAGAAAGAGC CCTGGAAAAG (SEQ ID NO: 826)	not	0.61959941 66	0	
2446	GTTCGCTTTGAA GAGCAAGG (SEQ ID NO: 310)	GTTCGCTTTG AAGAGCAAGG (SEQ ID NO: 827)	not	0.61959941 66	0	
3113	GAAAGTCCCAGG AGGAGGCC (SEQ ID NO: 311)	GAAAGTCCCA GGAGGAGGC C (SEQ ID NO: 828)	not	0.61959941 66	0	
2661	TGGCGGTGAAG GGTGTTCAG (SEQ ID NO: 312)	TGGCGGTGAAG GGGTGTTCAG (SEQ ID NO: 829)	not	0.61959941 66	0	
2681	CCAGAGAAGAAA GAGCCCTG (SEQ ID NO: 313)	CCAGAGAAGA AAGAGCCCTG (SEQ ID NO: 830)	not	0.61959941 66	0	
2786	CCGGGTGGGT CCTGGGGTA (SEQ ID NO: 314)	CCGGGTGGGT TCCTGGGGTG (SEQ ID NO: 831)	not	0.61959941 66	0	
1880	GGAGCCAGGTG CAGGCGGAG (SEQ ID NO: 315)	GGAGCCAGGT GCAGGCGGA G (SEQ ID NO: 832)	not	0.61959941 66	0	
2632	TACTTTCTTCTAC AAGAGTC (SEQ ID NO: 316)	TACTTTCTTCT ACAAGAGTC (SEQ ID NO: 833)	not	0.61959941 66	0	
1999	CTGGCATGGATC TACTCTGT (SEQ ID NO: 317)	CTGGCATGGA TCTACTCTGT (SEQ ID NO: 834)	not	0.61959941 66	0	
3284	CAAGCGGCGTG GACCCTTAC (SEQ ID NO: 318)	CAAGCGGCGT GGACCCTTAC (SEQ ID NO: 835)	not	0.61959941 66	0	

2543	TGATGCTCCTTG ATGCTGCT (SEQ ID NO: 319)	TGATGCTCCT TGATGCTGCT (SEQ ID NO: 836)	not	0.61959941 66	0	
2792	GGGTTCTGGG GTATGCGTG (SEQ ID NO: 320)	GGGTTCTGG GGTGTGCGTG (SEQ ID NO: 837)	not	0.61959941 66	0	
2853	TGTGGACCGTCT GAACATCA (SEQ ID NO: 321)	TGTGGACCGT CTGAACATCA (SEQ ID NO: 838)	not	0.61959941 66	0.871342756 6	
714	CTCCCAGGGCA CCCTACAGT (SEQ ID NO: 322)	CTCCCAGGGC ACCCTACAGT (SEQ ID NO: 839)	not	0.61959941 66	0	
2158	CTCCTGACGATA TTCATCAT (SEQ ID NO: 323)	CTCCTGACGA TATTCATCAT (SEQ ID NO: 840)	not	0.61959941 66	0	
1158	TCCTTTAACCAA AATCGCTT (SEQ ID NO: 324)	TCCTTTAACCA AAATCGCTT (SEQ ID NO: 841)	not	0.61959941 66	0	
2869	ATCACCTTCTAC GAGAACCA (SEQ ID NO: 325)	ATCACCTTCTA CGAGAACCA (SEQ ID NO: 842)	not	0.61959941 66	0.871342756 6	
2772	CGAGAGAGAACA CCCCGGGT (SEQ ID NO: 326)	CGAGAGAGAA CACCCCGGGT (SEQ ID NO: 843)	not	0.61959941 66	2.61402827	
3158	AGGACACAGGC CTCCACCAC (SEQ ID NO: 327)	AGGACACAGG CCTCCACCAC (SEQ ID NO: 844)	not	0.61959941 66	0	
3230	GCGTGGCCATTG CCTTCGTG (SEQ ID NO: 328)	GCGTGGCCAT TGCCTTCGTG (SEQ ID NO: 845)	not	0.61959941 66	0	
2825	AGATTTTGTGAGC CCTGTGGC (SEQ ID NO: 329)	AGATTTTGA GCCCTGTGGC (SEQ ID NO: 846)	not	0.61959941 66	0	
3156	GGAGGACACAG GCCTCCACC (SEQ ID NO: 330)	GGAGGACACA GGCCTCCACC (SEQ ID NO: 847)	not	0.61959941 66	0	

978	TGACCTCCTGTG TGGCTACC (SEQ ID NO: 331)	TGACCTCCTG TGTGGCTACC (SEQ ID NO: 848)	not	0.61959941 66	0	
2663	GCGGTGAAGGG TGTTCAACC (SEQ ID NO: 332)	GCGGTGAAGG GTGTTCAACC (SEQ ID NO: 849)	not	0.61959941 66	0	
189	CAACAAGGCGAT GCCCTCAG (SEQ ID NO: 333)	CAACAAGGCG ATGCCCTCAG (SEQ ID NO: 850)	not	0.61959941 66	0	
2201	GCGACCCATTCA TCCTCTTC (SEQ ID NO: 334)	GCGACCCATT CATCCTCTTC (SEQ ID NO: 851)	not	0.61959941 66	0	
332	TATATCGAGATT TTCAAGAA (SEQ ID NO: 335)	TATATCGAGA TTTCAAGAA (SEQ ID NO: 852)	not	0.61959941 66	0	
1873	ATCACAAGGAGC CAGGTGCA (SEQ ID NO: 336)	ATCACAAGGA GCCAGGTGCA (SEQ ID NO: 853)	not	0.61959941 66	0	
1291	AGGAAGTTGGTC AAAGCCTG (SEQ ID NO: 337)	CGGAAGTTGG TCAAAGCCTG (SEQ ID NO: 854)	not	0.61959941 66	0	
1413	TAGGCAGCTTGG TGAAGAAG (SEQ ID NO: 338)	TAGGCAGCTT GGTGAAGAAG (SEQ ID NO: 855)	not	0.61959941 66	0	
2826	GATTTTTGAGCC CTGTGGCC (SEQ ID NO: 339)	GATTTTTGAG CCCTGTGGCC (SEQ ID NO: 856)	not	0.61959941 66	0.961889921 3	
661	ATCTTCAGCCAG AGGCGCGG (SEQ ID NO: 340)	ATCTTCAGCC AGAGGCGCG G (SEQ ID NO: 857)	not	0.61959941 66	0	
670	CAGAGGCGCGG GGCAAAGAC (SEQ ID NO: 341)	CAGAGGCGC GGGGCAAAGA C (SEQ ID NO: 858)	not	0.61959941 66	0	
2550	CCTTGATGCTGC TGCTATG (SEQ ID NO: 342)	CCTTGATGCT GCTGTGTATG (SEQ ID NO: 859)	not	0.61959941 66	0	

2304	AGCCTGTAGTGG TGTCATCT (SEQ ID NO: 343)	AGCCTGTAGT GGTGTCTATCT (SEQ ID NO: 860)	not	0.61959941 66	0	
1633	CTCTCTCTACTG GAGGAAAA (SEQ ID NO: 344)	CTCTCTCTACT GGAGGAAAA (SEQ ID NO: 861)	not	0.61959941 66	0	
1546	CTCCGCCTGGTC AATCAATA (SEQ ID NO: 345)	CTCCGCCTGG TCAATCAATA (SEQ ID NO: 862)	not	0.61959941 66	0.961889921 3	
212	GAATGCTGCCGT GGCTCCAG (SEQ ID NO: 346)	GAATGCTGCC GTGGCTCCAG (SEQ ID NO: 863)	not	0.61959941 66	0	
2575	CTCGCTTGGTAC CTTGATCA (SEQ ID NO: 347)	CTCGCTTGGT ACCTTGATCA (SEQ ID NO: 864)	not	0.61959941 66	0	
1498	ATGGCCAACTTC GACTGGAG (SEQ ID NO: 348)	ATGGCCAACT TCGACTGGAG (SEQ ID NO: 865)	not	0.61959941 66	0	
1759	ACCAATAAGATT AAAGACAG (SEQ ID NO: 349)	ACCAATAAGA TTAAAGACAG (SEQ ID NO: 866)	not	0.61959941 66	0	
2975	CTGTGCTCGTTG GGGGAAGG (SEQ ID NO: 350)	CTGTGCTCGT TGGGGGAAG G (SEQ ID NO: 867)	not	0.61959941 66	0	
3012	GGATGCAGTCC GGCAGAGCC (SEQ ID NO: 351)	GGATGCAGTC CGGCAGAGCC (SEQ ID NO: 868)	not	0.61959941 66	0	
3142	ATGGAAGCCATG TTGGAGGA (SEQ ID NO: 352)	ATGGAAGCCA TGTTGGAGGA (SEQ ID NO: 869)	not	0.61959941 66	0	
2690	AAAGAGCCCTGG AAAAGACC (SEQ ID NO: 353)	AAAGAGCCCT GGAAAAGACC (SEQ ID NO: 870)	not	0.61959941 66	0.961889921 3	
642	CCTCCTGGAGC GCTTCATCA (SEQ ID NO: 354)	CCTCCTGGAG CGCTTCATCA (SEQ ID NO: 871)	not	0.61959941 66	0	

2028	TGTGAAGAGCAT CGTCTTGG (SEQ ID NO: 355)	TGTGAAGAGC ATCGTCTTGG (SEQ ID NO: 872)	not	0.61959941 66	0	
2364	GCAGGACCGCA TGACCGCTG (SEQ ID NO: 356)	GCAAGACCGC ATGACCGCTG (SEQ ID NO: 873)	not	0.61959941 66	0	
2716	CTAACAGAGGAA ACGGAGGA (SEQ ID NO: 357)	CTAACAGAGG AAACGGAGGA (SEQ ID NO: 874)	not	0.61959941 66	0	
3024	GCAGAGCCTTG GCATGTGTC (SEQ ID NO: 358)	GCAGAGCCTT GGCATGTGTC (SEQ ID NO: 875)	not	0.61959941 66	0	
1820	TCCGGTACATCT GGGGCGGG (SEQ ID NO: 359)	TCCGGTACAT CTGGGGCGG G (SEQ ID NO: 876)	not	0.61959941 66	0	
3039	GTGTCCACAGCA CAACATCC (SEQ ID NO: 360)	GTGTCCACAG CACAACATCC (SEQ ID NO: 877)	not	0.61959941 66	0	
2645	AAGAGTCGTATT GGCTTGGC (SEQ ID NO: 361)	AAGAGTCGTA TTGGCTTGGC (SEQ ID NO: 878)	not	0.61959941 66	0	
2696	CCCTGGAAAAGA CCGAGCCC (SEQ ID NO: 362)	CCCTGGAAAA GACCGAGCCC (SEQ ID NO: 879)	not	0.61959941 66	0	
2759	ACGACTCCTTCT TCGAGAGA (SEQ ID NO: 363)	ACGACTCCTT CTTCGAGAGA (SEQ ID NO: 880)	not	0.61959941 66	0.871342756 6	
2362	TGGCAGGACCG CATGACCGC (SEQ ID NO: 364)	TGGCAAGACC GCATGACCGC (SEQ ID NO: 881)	not	0.61959941 66	0	
983	TCCTGTGTGGCT ACCCCGAG (SEQ ID NO: 365)	TCCTGTGTGG CTACCCCGAG (SEQ ID NO: 882)	not	0.61959941 66	0	
740	AAGACACTCTGT ATGCCAAC (SEQ ID NO: 366)	AAGACACTCT GTATGCCAAC (SEQ ID NO: 883)	not	0.61959941 66	0	

907	CTGTGGGTGACC AGGCCCT (SEQ ID NO: 367)	CTGTGGGTGA CCAGGCCCT (SEQ ID NO: 884)	not	0.61959941 66	0	
750	GTATGCCAACGT GGACTTCT (SEQ ID NO: 368)	GTATGCCAAC GTGGACTTCT (SEQ ID NO: 885)	not	0.61959941 66	0	
2788	GGGTGGGTTCCT GGGGTATG (SEQ ID NO: 369)	GGGTGGGTTC CTGGGGTGTG (SEQ ID NO: 886)	not	0.61959941 66	0.961889921 3	
3248	TGGGCGACGCC AAGGTTGTG (SEQ ID NO: 370)	TGGGCGACGC CAAGTTGTG (SEQ ID NO: 887)	not	0.61959941 66	0	
2798	CTGGGGTATGC GTGAAGAAT (SEQ ID NO: 371)	CTGGGGTGTG CGTGAAGAAT (SEQ ID NO: 888)	not	0.61959941 66	0	
3143	TGGAAGCCATGT TGGAGGAC (SEQ ID NO: 372)	TGGAAGCCAT GTTGGAGGAC (SEQ ID NO: 889)	not	0.61959941 66	0	
2138	TCATGTCGATGA GCATCTTC (SEQ ID NO: 373)	TCATGTCGAT GAGCATCTTC (SEQ ID NO: 890)	not	0.61959941 66	0	
2940	CATCCTCACAGG CCTGCTGC (SEQ ID NO: 374)	CATCCTCACT GGCCTGCTGC (SEQ ID NO: 891)	not	0.61959941 66	0	
3300	TTACAGCAGAAG ATCCATCT (SEQ ID NO: 375)	TTACAGCAGA AGATCCATCT (SEQ ID NO: 892)	not	0.61959941 66	0	
2091	CTCCAATGCAGT GATTTGGT (SEQ ID NO: 376)	CTCCAATGCA GTGATTTGGT (SEQ ID NO: 893)	not	0.61959941 66	0	
591	CGCTCATGGAGT CCCGGACC (SEQ ID NO: 377)	CGCTCATGGA GTCCCGGACC (SEQ ID NO: 894)	not	0.61959941 66	0	
2532	GTCCATGCAGAT GATGCTCC (SEQ ID NO: 378)	GTCCATGCAG ATGATGCTCC (SEQ ID NO: 895)	not	0.61959941 66	0	

3110	AAGGAAAGTCCC AGGAGGAG (SEQ ID NO: 379)	AAGGAAAGTC CCAGGAGGAG (SEQ ID NO: 896)	not	0.61959941 66	3.575918191	
288	TCCTGGAATTGT GTCAAACCT (SEQ ID NO: 380)	TCCTGGAATT GTGTCAAACCT (SEQ ID NO: 897)	not	0.61959941 66	0	
1758	AACCAATAAGAT TAAAGACA (SEQ ID NO: 381)	AACCAATAAG ATTAAAGACA (SEQ ID NO: 898)	not	0.61959941 66	0	
2314	GGTGTCATCTAT TTCACCCT (SEQ ID NO: 382)	GGTGTCATCT ATTTCACCCT (SEQ ID NO: 899)	not	0.61959941 66	0.871342756 6	
2710	GAGCCCCTAACA GAGGAAAC (SEQ ID NO: 383)	GAGCCCCTAA CAGAGGAAAC (SEQ ID NO: 900)	not	0.61959941 66	0	
2192	TACATTACAGCG ACCCATTC (SEQ ID NO: 384)	TACATTACAG CGACCCATTC (SEQ ID NO: 901)	not	0.61959941 66	0	
1868	AGGGGATCACAA GGAGCCAG (SEQ ID NO: 385)	AGGGGATCAC AAGGAGCCAG (SEQ ID NO: 902)	not	0.61959941 66	0	
2809	GTGAAGAATCTG GTAAAGAT (SEQ ID NO: 386)	GTGAAGAATC TGGTGAAGAT (SEQ ID NO: 903)	not	0.61959941 66	0	
3091	CTGTTCTATGCC CAGCTGAA (SEQ ID NO: 387)	CTGTTCTATG CACAACCTGAA (SEQ ID NO: 904)	not	0.61959941 66	0.961889921 3	
753	TGCCAACGTGGA CTTCTTCA (SEQ ID NO: 388)	TGCCAACGTG GACTTCTTCA (SEQ ID NO: 905)	not	0.61959941 66	0	
980	ACCTCCTGTGTG GCTACCCC (SEQ ID NO: 389)	ACCTCCTGTG TGGCTACCCC (SEQ ID NO: 906)	not	0.61959941 66	0	
1461	CCTCTACAAGGG CCCTCGGG (SEQ ID NO: 390)	CCTCTACAAG GGCCCTCGG G (SEQ ID NO: 907)	not	0.61959941 66	0	

2413	CCGGTGGCATT GGATTGG (SEQ ID NO: 391)	CCGGTGGCAT TTGGATTGG (SEQ ID NO: 908)	not	0.61959941 66	0	
3213	CATGCAGAGAAA ACTGAGCG (SEQ ID NO: 392)	CATGCAGAGA AACTGAGCG (SEQ ID NO: 909)	not	0.61959941 66	0	
3089	TGCTGTTCTATG CCCAGCTG (SEQ ID NO: 393)	TGCTGTTCTAT GCACAACTG (SEQ ID NO: 910)	not	0.61959941 66	0	
2860	CGTCTGAACATC ACCTTCTA (SEQ ID NO: 394)	CGTCTGAACA TCACCTTCTA (SEQ ID NO: 911)	not	0.61959941 66	0	
3068	ACCTCACGGTGG CTGAGCAC (SEQ ID NO: 395)	ACCTCACGGT GGCTGAGCAC (SEQ ID NO: 912)	not	0.61959941 66	0	
2841	TGGCCGGCCAG CTGTGGACC (SEQ ID NO: 396)	TGGCCGGCCA GCTGTGGACC (SEQ ID NO: 913)	not	0.61959941 66	0.871342756 6	
2965	ACCTCTGGGACT GTGCTCGT (SEQ ID NO: 397)	ACCTCTGGGA CTGTGCTCGT (SEQ ID NO: 914)	not	0.50803043 71	0.961889921 3	
2859	CCGTCTGAACAT CACCTTCT (SEQ ID NO: 398)	CCGTCTGAAC ATCACCTTCT (SEQ ID NO: 915)	not	0.50803043 71	0	
2283	CTCCAAGGCCAG TCTGGCAG (SEQ ID NO: 399)	CTCCAAGGCC AGTCTGGCAG (SEQ ID NO: 916)	not	0.50803043 71	0.871342756 6	
3131	CCCAGCTGGAG ATGGAAGCC (SEQ ID NO: 400)	CCCAGCTGGA GATGGAAGCC (SEQ ID NO: 917)	not	0.50803043 71	0	
2946	CACAGGCCTGCT GCCTCAA (SEQ ID NO: 401)	CACTGGCCTG CTGCCTCAA (SEQ ID NO: 918)	not	0.50803043 71	0	
2730	GGAGGATCCAG AGCACCCAG (SEQ ID NO: 402)	GGAGGATCCA GAGCACCCAG (SEQ ID NO: 919)	not	0.50803043 71	0	

2971	GGGACTGTGCTC GTTGGGGG (SEQ ID NO: 403)	GGGACTGTGC TCGTTGGGGG (SEQ ID NO: 920)	not	0.50803043 71	0	
1797	CAGAGCTGATCC CGTGGAAG (SEQ ID NO: 404)	CAGAGCTGAT CCCGTGGAAG (SEQ ID NO: 921)	not	0.50803043 71	0	
2624	TTCCTTGGTACT TTCTTCTA (SEQ ID NO: 405)	TTCCTTGGTA CTTTCTTCTA (SEQ ID NO: 922)	not	0.50803043 71	0	
1472	GCCCTCGGGAA AGCCAGGCT (SEQ ID NO: 406)	GCCCTCGGGA AAGCCAGGCT (SEQ ID NO: 923)	not	0.50803043 71	1.923779843	
1765	AAGATTAAAGAC AGGTATTG (SEQ ID NO: 407)	AAGATTAAAG ACAGATATTG (SEQ ID NO: 924)	not	0.50803043 71	0	
1385	GGAACCCAACAG TAAAAGAC (SEQ ID NO: 408)	GGAACCCAAC AGTAAAAGAC (SEQ ID NO: 925)	not	0.50803043 71	0	
1553	TGGTCAATCAAT ACCTGGAG (SEQ ID NO: 409)	TGGTCAATCA ATACCTGGAG (SEQ ID NO: 926)	not	0.50803043 71	0	
2824	AAGATTTTTGAG CCCTGTGG (SEQ ID NO: 410)	AAGATTTTTGA GCCCTGTGG (SEQ ID NO: 927)	not	0.50803043 71	0.961889921 3	
273	CACCCCAGGAG AATCTCCTG (SEQ ID NO: 411)	CACCCCAGGA GAATCTCCTG (SEQ ID NO: 928)	not	0.50803043 71	0	
3060	GTTCCACCACCT CACGGTGG (SEQ ID NO: 412)	GTTCCACCAC CTCACGGTGG (SEQ ID NO: 929)	not	0.50803043 71	0	
2149	AGCATCTTCCTC CTGACGAT (SEQ ID NO: 413)	AGCATCTTCC TCCTGACGAT (SEQ ID NO: 930)	not	0.50803043 71	0	
2881	GAGAACCAGATC ACCGCATT (SEQ ID NO: 414)	GAGAACCAGA TCACCGCATT (SEQ ID NO: 931)	not	0.50803043 71	0	

2797	CCTGGGGTATGC GTGAAGAA (SEQ ID NO: 415)	CCTGGGGTGT GCGTGAAGAA (SEQ ID NO: 932)	not	0.50803043 71	0	
2733	GGATCCAGAGCA CCCAGAAG (SEQ ID NO: 416)	GGATCCAGAG CACCCAGAAG (SEQ ID NO: 933)	not	0.50803043 71	0	
798	CCTAGACAGCCG TTCTCAAG (SEQ ID NO: 417)	CCTAGACAGC CGTTCTCAAG (SEQ ID NO: 934)	not	0.50803043 71	0	
2849	CAGCTGTGGACC GTCTGAAC (SEQ ID NO: 418)	CAGCTGTGGA CCGTCTGAAC (SEQ ID NO: 935)	not	0.50803043 71	0	
3117	GTCCCAGGAGG AGGCCAGC (SEQ ID NO: 419)	GTCCCAGGAG GAGGCCAGC (SEQ ID NO: 936)	not	0.50803043 71	2.795122599	
1780	TATTGGGATTCT GGTCCCAG (SEQ ID NO: 420)	TATTGGGATT CTGGTCCCAG (SEQ ID NO: 937)	not	0.50803043 71	0	
2817	TCTGGTAAAGAT TTTTGAGC (SEQ ID NO: 421)	TCTGGTGAAG ATTTTGAAGC (SEQ ID NO: 938)	branchpoint	0.50803043 71	0.871342756 6	
2894	CCGCATTCTGG GCCACAAT (SEQ ID NO: 422)	CCGCATTCTT GGGCCACAAT (SEQ ID NO: 939)	not	0.50803043 71	1.833232678	
2807	GCGTGAAGAATC TGGTAAAG (SEQ ID NO: 423)	GCGTGAAGAA TCTGGTGAAG (SEQ ID NO: 940)	not	0.50803043 71	0	
3008	GCCTGGATGCA GTCCGGCAG (SEQ ID NO: 424)	GCCTGGATGC AGTCCGGCAG (SEQ ID NO: 941)	not	0.50803043 71	0	
2785	CCCGGGTGGGT TCCTGGGGT (SEQ ID NO: 425)	CCCGGGTGG GTTCTGGGG T (SEQ ID NO: 942)	not	0.50803043 71	0	
608	ACCTGGCGCTGA AGGACATC (SEQ ID NO: 426)	ACCTGGCGCT GAAGGACATC	not	0.50803043 71	0	

		(SEQ ID NO: 943)				
1713	CCACGTGAAGTA TAAGATCC (SEQ ID NO: 427)	CCACGTGAAG TATAAGATCC (SEQ ID NO: 944)	not	0.50803043 71	0	
2944	CTCACAGGCCTG CTGCCTCC (SEQ ID NO: 428)	CTCACTGGCC TGCTGCCTCC (SEQ ID NO: 945)	not	0.50803043 71	0	
2066	AGGAGACTTTGA AAAAATCAG (SEQ ID NO: 429)	AGGAGACATT GAAAAATCAG (SEQ ID NO: 946)	not	0.50803043 71	0.871342756 6	
2639	TTCTACAAGAGT CGTATTGG (SEQ ID NO: 430)	TTCTACAAGA GTCGTATTGG (SEQ ID NO: 947)	not	0.50803043 71	0	
237	CTTCTGCAATGT GAACAATC (SEQ ID NO: 431)	CTTCTGCAAT GTGAACAATC (SEQ ID NO: 948)	not	0.50803043 71	0	
1870	GGGATCACAAG GAGCCAGGT (SEQ ID NO: 432)	GGGATCACAA GGAGCCAGGT (SEQ ID NO: 949)	not	0.50803043 71	0	
3099	TGCCCAGCTGAA AGGAAAGT (SEQ ID NO: 433)	TGCACAACCTG AAAGGAAAGT (SEQ ID NO: 950)	not	0.50803043 71	0	
1841	TTGCCTATCTGC AGGACATG (SEQ ID NO: 434)	TTGCATATCT GCAAGACATG (SEQ ID NO: 951)	branchpoint	0.50803043 71	0	
17	GCAGCGGCTTC GTGAGACAG (SEQ ID NO: 435)	GATCCGGCTT CGTGAGACAG (SEQ ID NO: 952)	not	0.50803043 71	82.58601767	
3151	ATGTTGGAGGAC ACAGGCCT (SEQ ID NO: 436)	ATGTTGGAGG ACACAGGCCT (SEQ ID NO: 953)	not	0.50803043 71	0	
1400	AAGACTTTTTGA AATAGGCAG (SEQ ID NO: 437)	AAGACTTTTTG AATAGGCAG (SEQ ID NO: 954)	not	0.50803043 71	3.575918191	

3062	TCCACCACCTCA CGGTGGCT (SEQ ID NO: 438)	TCCACCACCT CACGGTGGCT (SEQ ID NO: 955)	not	0.50803043 71	0	
2769	CTTCGAGAGAGA ACACCCCG (SEQ ID NO: 439)	CTTCGAGAGA GAACACCCCG (SEQ ID NO: 956)	not	0.50803043 71	0.871342756 6	
2032	AAGAGCATCGTC TTGGAGAA (SEQ ID NO: 440)	AAGAGCATCG TCTTGGAGAA (SEQ ID NO: 957)	not	0.50803043 71	0	
181	CATTTCCTCAAC AAGGCGAT (SEQ ID NO: 441)	CATTTCCTCA ACAAGGCGAT (SEQ ID NO: 958)	not	0.50803043 71	0	
854	ATATGTCACCAA GAATTCAA (SEQ ID NO: 442)	ATATGTCACC AAGAATTCAA (SEQ ID NO: 959)	not	0.50803043 71	0	
1899	GGCTCCAGTTGG AATCTACC (SEQ ID NO: 443)	GGCTCCAGTT GGAATCTACC (SEQ ID NO: 960)	not	0.50803043 71	0	
2426	GATTGCGCACTG AGTACCTG (SEQ ID NO: 444)	GATTGCGCAC TGAGTACCTG (SEQ ID NO: 961)	not	0.50803043 71	0	
3077	TGGCTGAGCACA TGCTGTTC (SEQ ID NO: 445)	TGGCTGAGCA CATGCTGTTC (SEQ ID NO: 962)	not	0.50803043 71	0	
2286	CAAGGCCAGTCT GGCAGCAG (SEQ ID NO: 446)	CAAGGCCAGT CTGGCAGCAG (SEQ ID NO: 963)	not	0.50803043 71	0	
3308	GAAGATCCATCT GGGATCTG (SEQ ID NO: 447)	GAAGATCCAT CTGGGATCTG (SEQ ID NO: 964)	not	0.50803043 71	0	
927	CATGCAGAATGG TGGTCCAG (SEQ ID NO: 448)	CATGCAGAAT GGTGGTCCAG (SEQ ID NO: 965)	not	0.50803043 71	0.871342756 6	
3114	AAAGTCCCAGGA GGAGGCC (SEQ ID NO: 449)	AAAGTCCCAG GAGGAGGCC C (SEQ ID NO: 966)	not	0.50803043 71	0	

908	TGTGGGTGACCA GGCCCCCTC (SEQ ID NO: 450)	TGTGGGTGAC CAGGCCCCCTC (SEQ ID NO: 967)	not	0.50803043 71	0	
2127	CAGCTTCTCCAT CATGTCTCGA (SEQ ID NO: 451)	CTCCTTCTCC ATCATGTCTCGA (SEQ ID NO: 968)	not	0.50803043 71	0	
1886	AGGTGCAGGCG GAGGCTCCA (SEQ ID NO: 452)	AGGTGCAGGC GGAGGCTCCA (SEQ ID NO: 969)	not	0.50803043 71	0	
2603	CAGGAGACTATG GAACCCCA (SEQ ID NO: 453)	CAGGAGACTA TGGAAACCCCA (SEQ ID NO: 970)	not	0.50803043 71	0	
3109	AAAGGAAAGTCC CAGGAGGA (SEQ ID NO: 454)	AAAGGAAAGT CCCAGGAGGA (SEQ ID NO: 971)	not	0.50803043 71	0.871342756 6	
2851	GCTGTGGACCGT CTGAACAT (SEQ ID NO: 455)	GCTGTGGACC GTCTGAACAT (SEQ ID NO: 972)	not	0.50803043 71	0	
3108	GAAAGGAAAGTC CCAGGAGG (SEQ ID NO: 456)	GAAAGGAAAG TCCCAGGAGG (SEQ ID NO: 973)	not	0.50803043 71	0	
738	AGAAGACACTCT GTATGCCA (SEQ ID NO: 457)	AGAAGACACT CTGTATGCCA (SEQ ID NO: 974)	not	0.50803043 71	0	
2614	GGAACCCCACTT CCTTGGTA (SEQ ID NO: 458)	GGAACCCCACT TTCCTTGGTA (SEQ ID NO: 975)	not	0.50803043 71	0	
3095	TCTATGCCCAGC TGAAAGGA (SEQ ID NO: 459)	TCTATGCACA ACTGAAAGGA (SEQ ID NO: 976)	not	0.50803043 71	0	
3016	GCAGTCCGGCA GAGCCTTGG (SEQ ID NO: 460)	GCAGTCCGGC AGAGCCTTGG (SEQ ID NO: 977)	not	0.50803043 71	0	
448	GAGAGAATTGCA GGAAGAGG (SEQ ID NO: 461)	GAGAGAATTG CAGGAAGAGG (SEQ ID NO: 978)	not	0.50803043 71	0	

1510	GACTGGAGGGA CATATTTAA (SEQ ID NO: 462)	GACTGGAGGG ACATATTTAA (SEQ ID NO: 979)	not	0.50803043 71	0	
2503	ACGGAAGGGGA CGAATTCAG (SEQ ID NO: 463)	ACGGAAGGG GACGAATTCA G (SEQ ID NO: 980)	not	0.50803043 71	0	
2781	ACACCCCGGGT GGGTTCTG (SEQ ID NO: 464)	ACACCCCGGG TGGTTCTG (SEQ ID NO: 981)	not	0.50803043 71	0	
607	GACCTGGCGCT GAAGGACAT (SEQ ID NO: 465)	GACCTGGCGC TGAAGGACAT (SEQ ID NO: 982)	not	0.50803043 71	0	
3201	CCTATCAGGCGG CATGCAGA (SEQ ID NO: 466)	CCTATCAGGC GGCATGCAGA (SEQ ID NO: 983)	not	0.50803043 71	0	
1298	TGGTCAAAGCCT GGGAAGAA (SEQ ID NO: 467)	TGGTCAAAGC CTGGGAAGAA (SEQ ID NO: 984)	not	0.50803043 71	0	
1243	AGGATACTGAAG AATGCCAA (SEQ ID NO: 468)	AGGATACTGA AGAATGCCAA (SEQ ID NO: 985)	not	0.50803043 71	0	
1570	GAGTGCTTGGTC CTGGATAA (SEQ ID NO: 469)	GAGTGCTTGG TCCTGGATAA (SEQ ID NO: 986)	not	0.50803043 71	0	
2920	GGGAAAACCACC ACCTTGTC (SEQ ID NO: 470)	GGGAAAACCA CCACCTTGTC (SEQ ID NO: 987)	not	0.50803043 71	0	
2914	GGAGCTGGGAA AACCACCAC (SEQ ID NO: 471)	GGAGCTGGGA AAACCACCAC (SEQ ID NO: 988)	not	0.50803043 71	0.871342756 6	
1819	TTCCGGTACATC TGGGGCGG (SEQ ID NO: 472)	TTCCGGTACA TCTGGGGCGG (SEQ ID NO: 989)	not	0.50803043 71	0	
2601	TCCAGGAGACTA TGGAACCC (SEQ ID NO: 473)	TCCAGGAGAC TATGGAACCC (SEQ ID NO: 990)	not	0.50803043 71	0	

3047	AGCACAAACATCC TGTTCCAC (SEQ ID NO: 474)	AGCACAAACAT CCTGTTCCAC (SEQ ID NO: 991)	not	0.50803043 71	0	
2800	GGGGTATGCGT GAAGAATCT (SEQ ID NO: 475)	GGGGTGTGC GTGAAGAATC T (SEQ ID NO: 992)	not	0.50803043 71	0	
1834	GGCGGGTTTGC CTATCTGCA (SEQ ID NO: 476)	GGCGGGTTTG CATATCTGCA (SEQ ID NO: 993)	branchpoint	0.50803043 71	0	
309	TAACAACTCCAT CTTGCAA (SEQ ID NO: 477)	TAACAACTCC ATCTTGCAA (SEQ ID NO: 994)	not	0.50803043 71	0	
2823	AAAGATTTTTGA GCCCTGTG (SEQ ID NO: 478)	GAAGATTTTT GAGCCCTGTG (SEQ ID NO: 995)	not	0.50803043 71	0	
3229	AGCGTGGCCATT GCCTTCGT (SEQ ID NO: 479)	AGCGTGGCCA TTGCCTTCGT (SEQ ID NO: 996)	not	0.50803043 71	0	
2564	TCTATGGCTTAC TCGCTTGG (SEQ ID NO: 480)	TGTATGGCTT GCTCGCTTGG (SEQ ID NO: 997)	branchpoint	0.50803043 71	0	
713	TCTCCCAGGGCA CCCTACAG (SEQ ID NO: 481)	TCTCCCAGGG CACCTACAG (SEQ ID NO: 998)	not	0.50803043 71	0.961889921 3	
1433	GTATTACTGCTG AAGCCATC (SEQ ID NO: 482)	GCATTACTGC GGAGGCCATC (SEQ ID NO: 999)	branchpoint	0.50803043 71	0	
2714	CCCTAACAGAGG AAACGGAG (SEQ ID NO: 483)	CCCTAACAGA GGAAACGGAG (SEQ ID NO: 1000)	not	0.50803043 71	0	
23	GCTTCGTGAGAC AGATACAG (SEQ ID NO: 484)	GCTTCGTGAG ACAGATACAG (SEQ ID NO: 1001)	not	0.50803043 71	0	
1597	AGCTACAATGAT GAAACTCA (SEQ ID NO: 485)	AGCTACAATG ATGAAACTCA (SEQ ID NO: 1002)	not	0.50803043 71	0	

2343	ACACATCCTGTG CTTCGCCT (SEQ ID NO: 486)	CCACATCCTG TGCTTCGCCT (SEQ ID NO: 1003)	not	0.50803043 71	0	
2351	TGTGCTTCGCCT GGCAGGAC (SEQ ID NO: 487)	TGTGCTTCGC CTGGCAAGAC (SEQ ID NO: 1004)	not	0.50803043 71	0	
2642	TACAAGAGTCGT ATTGGCTT (SEQ ID NO: 488)	TACAAGAGTC GTATTGGCTT (SEQ ID NO: 1005)	not	0.50803043 71	0	
2239	TCCACTGCCACC ATCATGCT (SEQ ID NO: 489)	TCCACTGCCA CCATCATGCT (SEQ ID NO: 1006)	not	0.50803043 71	0	
2223	GTTCTTGTTGGC TTTCTCCA (SEQ ID NO: 490)	GTTCTTGTTG GCTTTCTCCA (SEQ ID NO: 1007)	not	0.50803043 71	0	
441	TCACCCGGAGA GAATTGCAG (SEQ ID NO: 491)	TCACCCGGAG AGAATTGCAG (SEQ ID NO: 1008)	not	0.50803043 71	0.871342756 6	
1189	AAGCCTTTGCTG ATGGGAAA (SEQ ID NO: 492)	AAGCCTTTGC TGATGGGAAA (SEQ ID NO: 1009)	not	0.50803043 71	0	
2555	ATGCTGCTGTCT ATGGCTTA (SEQ ID NO: 493)	ATGCTGCTGT GTATGGCTTG (SEQ ID NO: 1010)	branchpoint	0.50803043 71	0	
1636	TCTCTACTGGAG GAAAACAT (SEQ ID NO: 494)	TCTCTACTGG AGGAAAACAT (SEQ ID NO: 1011)	not	0.50803043 71	0	
229	CAGGGGATCTTC TGCAATGT (SEQ ID NO: 495)	CAGGGGATCT TCTGCAATGT (SEQ ID NO: 1012)	not	0.50803043 71	0.871342756 6	
705	GTGCTCCCTCTC CCAGGGCA (SEQ ID NO: 496)	GTGCTCCCTC TCCCAGGGCA (SEQ ID NO: 1013)	not	0.50803043 71	0	
1747	GTGGTGGAGAAA ACCAATAA (SEQ ID NO: 497)	GTGGTGGAGA AAACCAATAA (SEQ ID NO: 1014)	not	0.50803043 71	0	

1318	GTAGGGCCCCA GATCTGGTA (SEQ ID NO: 498)	GTGGGGCCC CAGATCTGGT A (SEQ ID NO: 1015)	not	0.50803043 71	0.961889921 3	
3015	TGCAGTCCGGCA GAGCCTTG (SEQ ID NO: 499)	TGCAGTCCGG CAGAGCCTTG (SEQ ID NO: 1016)	not	0.50803043 71	0.961889921 3	
2103	GATTTGGTGTAC CTGGTTCC (SEQ ID NO: 500)	GATTTGGTGT ACCTGGTTCC (SEQ ID NO: 1017)	not	0.50803043 71	0	
3122	AGGAGGAGGCC CAGCTGGAG (SEQ ID NO: 501)	AGGAGGAGG CCCAGCTGGA G (SEQ ID NO: 1018)	not	0.50803043 71	1.833232678	
1555	GTCAATCAATAC CTGGAGTG (SEQ ID NO: 502)	GTCAATCAAT ACCTGGAGTG (SEQ ID NO: 1019)	not	0.50803043 71	0	
1725	TAAGATCCGAAT GGACATAG (SEQ ID NO: 503)	TAAGATCCGA ATGGACATAG (SEQ ID NO: 1020)	not	0.50803043 71	0	
3078	GGCTGAGCACAT GCTGTTCT (SEQ ID NO: 504)	GGCTGAGCAC ATGCTGTTCT (SEQ ID NO: 1021)	not	0.50803043 71	0	
1681	GACATGTATCCC TGGACCAG (SEQ ID NO: 505)	GACATGTATC CCTGGACCAG (SEQ ID NO: 1022)	not	0.50803043 71	1.833232678	
2023	ATGACTGTGAAG AGCATCGT (SEQ ID NO: 506)	ATGACTGTGA AGAGCATCGT (SEQ ID NO: 1023)	not	0.50803043 71	0	
3082	GAGCACATGCTG TTCTATGC (SEQ ID NO: 507)	GAGCACATGC TGTTCTATGC (SEQ ID NO: 1024)	not	0.50803043 71	0	
103	CTCGTGTGGCCT TTATCTTT (SEQ ID NO: 508)	CTCGTGTGGC CTTTATCTTT (SEQ ID NO: 1025)	not	0.50803043 71	0	
2935	TTGTCCATCCTC ACAGGCCT (SEQ ID NO: 509)	TTGTCCATCC TCACTGGCCT (SEQ ID NO: 1026)	not	0.50803043 71	0	

51	CTGGAAGAACTG GACCCTGC (SEQ ID NO: 510)	CTGGAAGAAC TGGACCCTGC (SEQ ID NO: 1027)	not	0.50803043 71	0	
175	GAATGCCATTTC CCCAACAA (SEQ ID NO: 511)	GAATGCCATT TCCCCAACAA (SEQ ID NO: 1028)	not	0.50803043 71	0	
2630	GGTACTTTCTTC TACAAGAG (SEQ ID NO: 512)	GGTACTTTCTT CTACAAGAG (SEQ ID NO: 1029)	not	0.50803043 71	0	
2308	TGTAGTGGTGTC ATCTATTT (SEQ ID NO: 513)	TGTAGTGGTG TCATCTATTT (SEQ ID NO: 1030)	not	0	1.742685513	
1272	TGAAGAACTGGA ACACGTTA (SEQ ID NO: 514)	TGAAGAACTG GAACACGTTC (SEQ ID NO: 1031)	not	0	0.871342756 6	
644	TCCTGGAGCGCT TCATCATC (SEQ ID NO: 515)	TCCTGGAGCG CTTCATCATC (SEQ ID NO: 1032)	not	0	0.871342756 6	
1162	TTAACCAAAATC GCTTGGAG (SEQ ID NO: 516)	TTAACCAAAAT CGCTTGGAG (SEQ ID NO: 1033)	not	0	0.871342756 6	
2917	GCTGGGAAAACC ACCACCTT (SEQ ID NO: 517)	GCTGGGAAAA CCACCACCTT (SEQ ID NO: 1034)	not	0	0.871342756 6	
2889	GATCACCGCATT CCTGGGCC (SEQ ID NO: 518)	GATCACCGCA TTCCTGGGCC (SEQ ID NO: 1035)	not	0	0.871342756 6	
3004	ACCAGCCTGGAT GCAGTCCG (SEQ ID NO: 519)	ACCAGCCTGG ATGCAGTCCG (SEQ ID NO: 1036)	not	0	1.742685513	
2349	CCTGTGCTTCGC CTGGCAGG (SEQ ID NO: 520)	CCTGTGCTTC GCCTGGCAAG (SEQ ID NO: 1037)	not	0	7.242383547	
3002	AAACCAGCCTGG ATGCAGTC (SEQ ID NO: 521)	AAACCAGCCT GGATGCAGTC (SEQ ID NO: 1038)	not	0	0.871342756 6	

3076	GTGGCTGAGCA CATGCTGTT (SEQ ID NO: 522)	GTGGCTGAGC ACATGCTGTT (SEQ ID NO: 1039)	not	0	2.704575435	
2813	AGAATCTGGTAA AGATTTTT (SEQ ID NO: 523)	AGAATCTGGT GAAGATTTTT (SEQ ID NO: 1040)	branchpoint	0	0.871342756 6	
2773	GAGAGAGAACAC CCCGGGTG (SEQ ID NO: 524)	GAGAGAGAAC ACCCCGGGTG (SEQ ID NO: 1041)	not	0	3.575918191	
619	AAGGACATCGCC TGCAGCGA (SEQ ID NO: 525)	AAGGACATCG CCTGCAGCGA (SEQ ID NO: 1042)	not	0	2.61402827	
2957	TGCCTCCAACCT CTGGGACT (SEQ ID NO: 526)	TGCCTCCAAC CTCTGGGACT (SEQ ID NO: 1043)	not	0	1.742685513	
2949	AGGCCTGCTGC CTCCAACCT (SEQ ID NO: 527)	TGGCCTGCTG CCTCCAACCT (SEQ ID NO: 1044)	not	0	1.833232678	
2929	ACCACCTTGTC ATCCTCAC (SEQ ID NO: 528)	ACCACCTTG CCATCCTCAC (SEQ ID NO: 1045)	not	0	4.537808112	
2320	ATCTATTTACAC CTCTACCT (SEQ ID NO: 529)	ATCTATTTAC CCTGTACCT (SEQ ID NO: 1046)	not	0	0.871342756 6	
2814	GAATCTGGTAAA GATTTTTG (SEQ ID NO: 530)	GAATCTGGTG AAGATTTTTG (SEQ ID NO: 1047)	branchpoint	0	0.871342756 6	
2297	TGGCAGCAGCCT GTAGTGGT (SEQ ID NO: 531)	TGGCAGCAGC CTGTAGTGGT (SEQ ID NO: 1048)	not	0	0.871342756 6	
3073	ACGGTGGCTGA GCACATGCT (SEQ ID NO: 532)	ACGGTGGCTG AGCACATGCT (SEQ ID NO: 1049)	not	0	2.795122599	
1785	GGATTCTGGTCC CAGAGCTG (SEQ ID NO: 533)	GGATTCTGGT CCCAGAGCTG (SEQ ID NO: 1050)	not	0	0.871342756 6	

3094	TTCTATGCCCAG CTGAAAGG (SEQ ID NO: 534)	TTCTATGCAC AACTGAAAGG (SEQ ID NO: 1051)	not	0	0.871342756 6	
173	ATGAATGCCATT TCCCCAAC (SEQ ID NO: 535)	ATGAATGCCA TTTCCCCAAC (SEQ ID NO: 1052)	not	0	0.871342756 6	
2194	CATTACAGCGAC CCATTCAT (SEQ ID NO: 536)	CATTACAGCG ACCCATTCAT (SEQ ID NO: 1053)	not	0	1.742685513	
2795	TTCCTGGGGTAT GCGTGAAG (SEQ ID NO: 537)	TTCCTGGGGT GTGCGTGAAG (SEQ ID NO: 1054)	not	0	2.61402827	
13	GGTGGCAGCGG CTTCGTGAG (SEQ ID NO: 538)	GGTGGATCCG GCTTCGTGAG (SEQ ID NO: 1055)	not	0	0.871342756 6	
2242	ACTGCCACCATC ATGCTGTG (SEQ ID NO: 539)	ACTGCCACCA TCATGCTGTG (SEQ ID NO: 1056)	not	0	0.871342756 6	
2863	CTGAACATCACC TTCTACGA (SEQ ID NO: 540)	CTGAACATCA CCTTCTACGA (SEQ ID NO: 1057)	not	0	1.833232678	
2435	CTGAGTACCTGG TTCGCTTT (SEQ ID NO: 541)	CTGAGTACCT GGTTCGCTTT (SEQ ID NO: 1058)	not	0	0.871342756 6	
1225	GATTCACCTGCA GCACGAAG (SEQ ID NO: 542)	GATTCACCTG CTGCACGAAG (SEQ ID NO: 1059)	not	0	2.61402827	
901	GACTTGCTGTGG GTGACCAG (SEQ ID NO: 543)	GACTTGCTGT GGGTGACCAG (SEQ ID NO: 1060)	not	0	0.871342756 6	
1123	TTTTGTAATGCAT TGATCCA (SEQ ID NO: 544)	TTTTGTAATGC ATTGATCCA (SEQ ID NO: 1061)	not	0	0.871342756 6	
2130	CTTCTCCATCAT GTCGATGA (SEQ ID NO: 545)	CTTCTCCATC ATGTCGATGA (SEQ ID NO: 1062)	not	0	0.871342756 6	

3011	TGGATGCAGTCC GGCAGAGC (SEQ ID NO: 546)	TGGATGCAGT CCGGCAGAGC (SEQ ID NO: 1063)	not	0	0.871342756 6	
2081	ATCAGGGTGTCT CCAATGCA (SEQ ID NO: 547)	ATCAGGGTGT CTCCAATGCA (SEQ ID NO: 1064)	not	0	0.871342756 6	
1092	TATCTATTCTTAT GACAGAA (SEQ ID NO: 548)	TATCTATTCTT ATGACAGAA (SEQ ID NO: 1065)	not	0	1.742685513	
3120	CCAGGAGGAGG CCCAGCTGG (SEQ ID NO: 549)	CCAGGAGGAG GCCCAGCTGG (SEQ ID NO: 1066)	not	0	1.833232678	
3111	AGGAAAGTCCCA GGAGGAGG (SEQ ID NO: 550)	AGGAAAGTCC CAGGAGGAG G (SEQ ID NO: 1067)	not	0	1.742685513	
2839	TGTGGCCGGCC AGCTGTGGA (SEQ ID NO: 551)	TGTGGCCGGC CAGCTGTGGA (SEQ ID NO: 1068)	not	0	0.871342756 6	
2677	TCAACCAGAGAA GAAAGAGC (SEQ ID NO: 552)	TCAACCAGAG AAGAAAGAGC (SEQ ID NO: 1069)	not	0	0.871342756 6	
2821	GTAAAGATTTT GAGCCCTG (SEQ ID NO: 553)	GTGAAGATTT TTGAGCCCTG (SEQ ID NO: 1070)	branchpoint	0	0.871342756 6	
2784	CCCCGGGTGGG TTCCTGGGG (SEQ ID NO: 554)	CCCCGGGTG GGTTCCTGGG G (SEQ ID NO: 1071)	not	0	0.871342756 6	
1129	AATGCATTGATC CAGAGCCT (SEQ ID NO: 555)	AATGCATTGA TCCAGAGCCT (SEQ ID NO: 1072)	not	0	0.871342756 6	
918	CAGGCCCCTCAT GCAGAATG (SEQ ID NO: 556)	CAGGCCCCTC ATGCAGAATG (SEQ ID NO: 1073)	not	0	0.871342756 6	
351	ACTCCTCATGAA TGCACCAG (SEQ ID NO: 557)	ACTCCTCATG AATGCACCAG (SEQ ID NO: 1074)	not	0	0.871342756 6	

2540	AGATGATGCTCC TTGATGCT (SEQ ID NO: 558)	AGATGATGCT CCTTGATGCT (SEQ ID NO: 1075)	not	0	0.871342756 6	
3056	TCCTGTTCCACC ACCTCACG (SEQ ID NO: 559)	TCCTGTTCCA CCACCTCACG (SEQ ID NO: 1076)	not	0	1.742685513	
1465	TACAAGGGCCCT CGGGAAAG (SEQ ID NO: 560)	TACAAGGGCC CTCGGGAAAG (SEQ ID NO: 1077)	not	0	0.871342756 6	
2838	CTGTGGCCGGC CAGCTGTGG (SEQ ID NO: 561)	CTGTGGCCGG CCAGCTGTGG (SEQ ID NO: 1078)	not	0	0.871342756 6	
3065	ACCACCTCACGG TGGCTGAG (SEQ ID NO: 562)	ACCACCTCAC GGTGGCTGAG (SEQ ID NO: 1079)	not	0	0.871342756 6	
1150	GAGTCAAATCCT TTAACCAA (SEQ ID NO: 563)	GAGTCAAATC CTTTAACCAA (SEQ ID NO: 1080)	not	0	0.871342756 6	
1959	CATGATCATCCT GAACCGCT (SEQ ID NO: 564)	CATGATCATC CTGAACCGCT (SEQ ID NO: 1081)	not	0	0.871342756 6	
2656	TGGCTTGGCGGT GAAGGGTG (SEQ ID NO: 565)	TGGCTTGGCG GTGAAGGGTG (SEQ ID NO: 1082)	not	0	0.871342756 6	
101	AACTCGTGTGGC CTTTATCT (SEQ ID NO: 566)	AACTCGTGTG GCCTTTATCT (SEQ ID NO: 1083)	not	0	0.871342756 6	
1807	CCCGTGGAAGAT TTCCGGTA (SEQ ID NO: 567)	CCCGTGGAAG ATTTCCGGTA (SEQ ID NO: 1084)	not	0	0.871342756 6	
1270	TTTGAAGAACTG GAACACGT (SEQ ID NO: 568)	TTTGAAGAAC TGGAACACGT (SEQ ID NO: 1085)	not	0	0.871342756 6	
2738	CAGAGCACCCA GAAGGAATA (SEQ ID NO: 569)	CAGAGCACCC AGAAGGAATA (SEQ ID NO: 1086)	not	0	0.871342756 6	

2827	ATTTTGTGAGCCC TGTGGCCG (SEQ ID NO: 570)	ATTTTGTGAGC CCTGTGGCCG (SEQ ID NO: 1087)	not	0	0.871342756 6	
2074	TTGAAAAATCAG GGTGTCTC (SEQ ID NO: 571)	TTGAAAAATCA GGGTGTCTC (SEQ ID NO: 1088)	not	0	0.871342756 6	
1414	AGGCAGCTTGGT GAAGAAGG (SEQ ID NO: 572)	AGGCAGCTTG GTGAAGAAGG (SEQ ID NO: 1089)	not	0	0.871342756 6	
2383	GAGCTGAAGAAG GCTGTGAG (SEQ ID NO: 573)	GAGCTGAAGA AGGCTGTGAG (SEQ ID NO: 1090)	not	0	0.871342756 6	
2961	TCCAACCTCTGG GACTGTGC (SEQ ID NO: 574)	TCCAACCTCT GGGACTGTGC (SEQ ID NO: 1091)	not	0	0.871342756 6	
1355	CACAGATGAACA TGATCAGA (SEQ ID NO: 575)	CACAGATGAA CATGATCAGA (SEQ ID NO: 1092)	not	0	0.871342756 6	
987	GTGTGGCTACCC CGAGGGAG (SEQ ID NO: 576)	GTGTGGCTAC CCCGAGGGA G (SEQ ID NO: 1093)	not	0	0.871342756 6	
2195	ATTACAGCGACC CATTATC (SEQ ID NO: 577)	ATTACAGCGA CCCATTATC (SEQ ID NO: 1094)	not	0	0.871342756 6	
2072	CTTTGAAAAATC AGGGTGTC (SEQ ID NO: 578)	CATTGAAAAAT CAGGGTGTC (SEQ ID NO: 1095)	not	0	0.871342756 6	
2899	TTCCTGGGCCAC AATGGAGC (SEQ ID NO: 579)	TTCCTGGGCC ACAATGGAGC (SEQ ID NO: 1096)	not	0	0.871342756 6	
2674	TGTTCAACCAGA GAAGAAAG (SEQ ID NO: 580)	TGTTCAACCA GAGAAGAAAG (SEQ ID NO: 1097)	not	0	0.871342756 6	
1987	ATCTTCATGGTG CTGGCATG (SEQ ID NO: 581)	ATCTTCATGG TGCTGGCATG (SEQ ID NO: 1098)	not	0	0.871342756 6	

498	AGAAACACTGAC ACTATTTTC (SEQ ID NO: 582)	AGAAACACTG ACACTATTTTC (SEQ ID NO: 1099)	not	0	0.871342756 6	
1955	CTTTCATGATCA TCCTGAAC (SEQ ID NO: 583)	CTTTCATGATC ATCCTGAAC (SEQ ID NO: 1100)	not	0	0.871342756 6	
2777	GAGAACACCCC GGGTGGGTT (SEQ ID NO: 584)	GAGAACACCC CGGGTGGGTT (SEQ ID NO: 1101)	not	0	0.871342756 6	
96	GGTGGAACCTCGT GTGGCCTT (SEQ ID NO: 585)	GGTGGAACCTC GTGTGGCCTT (SEQ ID NO: 1102)	not	0	0.871342756 6	
1146	CCTGGAGTCAAA TCCTTTAA (SEQ ID NO: 586)	CCTGGAGTCA AATCCTTTAA (SEQ ID NO: 1103)	not	0	0.871342756 6	
121	TTATTTCTGGTCT TGATCTG (SEQ ID NO: 587)	TTATTTCTGGT CTTGATCTG (SEQ ID NO: 1104)	not	0	0.871342756 6	
2495	ACAGTCCCACGG AAGGGGAC (SEQ ID NO: 588)	ACAGTCCCAC GGAAGGGGA C (SEQ ID NO: 1105)	not	0	0.871342756 6	
2203	GACCCATTTCATC CTCTTCCT (SEQ ID NO: 589)	GACCCATTCA TCCTCTTCCT (SEQ ID NO: 1106)	not	0	0.871342756 6	
3170	TCCACCACAAGC GGAATGAA (SEQ ID NO: 590)	TCCACCACAA GCGGAATGAA (SEQ ID NO: 1107)	not	0	0.961889921 3	
3169	CTCCACCACAAG CGGAATGA (SEQ ID NO: 591)	CTCCACCACA AGCGGAATGA (SEQ ID NO: 1108)	not	0	2.885669764	
2086	GGTGTCTCCAAT GCAGTGAT (SEQ ID NO: 592)	GGTGTCTCCA ATGCAGTGAT (SEQ ID NO: 1109)	not	0	0.961889921 3	
240	CTGCAATGTGAA CAATCCCT (SEQ ID NO: 593)	CTGCAATGTG AACAATCCCT (SEQ ID NO: 1110)	not	0	1.923779843	

2653	TATTGGCTTGGC GGTGAAGG (SEQ ID NO: 594)	TATTGGCTTG GCGGTGAAGG (SEQ ID NO: 1111)	not	0	0.961889921 3	
2268	GCTCAGCACCTT CTTCTCCA (SEQ ID NO: 595)	GCTCAGCACCC TTCTTCTCCA (SEQ ID NO: 1112)	not	0	0.961889921 3	
353	TCCTCATGAATG CACCAGAG (SEQ ID NO: 596)	TCCTCATGAA TGCACCAGAG (SEQ ID NO: 1113)	not	0	1.923779843	
2073	TTTGAAAAATCA GGGTGTCT (SEQ ID NO: 597)	ATTGAAAAATC AGGGTGTCT (SEQ ID NO: 1114)	not	0	0.961889921 3	
2271	CAGCACCTTCTT CTCCAAGG (SEQ ID NO: 598)	CAGCACCTTC TTCTCCAAGG (SEQ ID NO: 1115)	not	0	1.923779843	
2970	TGGGACTGTGCT CGTTGGGG (SEQ ID NO: 599)	TGGGACTGTG CTCGTTGGGG (SEQ ID NO: 1116)	not	0	0.961889921 3	
2815	AATCTGGTAAAG ATTTTTGA (SEQ ID NO: 600)	AATCTGGTGA AGATTTTTGA (SEQ ID NO: 1117)	branchpoint	0	0.961889921 3	
3132	CCAGCTGGAGAT GGAAGCCA (SEQ ID NO: 601)	CCAGCTGGAG ATGGAAGCCA (SEQ ID NO: 1118)	not	0	0.961889921 3	
2097	TGCAGTGATTTG GTGTACCT (SEQ ID NO: 602)	TGCAGTGATT TGGTGTACCT (SEQ ID NO: 1119)	not	0	0.961889921 3	
1894	GCGGAGGCTCC AGTTGGAAT (SEQ ID NO: 603)	GCGGAGGCTC CAGTTGGAAT (SEQ ID NO: 1120)	not	0	0.961889921 3	
617	TGAAGGACATCG CCTGCAGC (SEQ ID NO: 604)	TGAAGGACAT CGCCTGCAGC (SEQ ID NO: 1121)	not	0	2.885669764	
3027	GAGCCTTGGCAT GTGTCCAC (SEQ ID NO: 605)	GAGCCTTGGC ATGTGTCCAC (SEQ ID NO: 1122)	not	0	1.923779843	

1977	CTGTTTCCCTAT CTTCATGG (SEQ ID NO: 606)	CTGTTTCCCT ATCTTCATGG (SEQ ID NO: 1123)	not	0	1.923779843	
2835	GCCCTGTGGCC GGCCAGCTG (SEQ ID NO: 607)	GCCCTGTGGC CGGCCAGCTG (SEQ ID NO: 1124)	not	0	0.961889921 3	
205	TCAGCAGGAATG CTGCCGTG (SEQ ID NO: 608)	TCAGCAGGAA TGCTGCCGTG (SEQ ID NO: 1125)	not	0	0.961889921 3	
3161	ACACAGGCCTCC ACCACAAG (SEQ ID NO: 609)	ACACAGGCCT CCACCACAAG (SEQ ID NO: 1126)	not	0	1.923779843	
2864	TGAACATCACCT TCTACGAG (SEQ ID NO: 610)	TGAACATCAC CTTCTACGAG (SEQ ID NO: 1127)	not	0	1.923779843	
3055	ATCCTGTTCCAC CACCTCAC (SEQ ID NO: 611)	ATCCTGTTCC ACCACCTCAC (SEQ ID NO: 1128)	not	0	0.961889921 3	
1931	CCTACCCCTGCT TCGTGGAC (SEQ ID NO: 612)	CCTACCCCTG CTTCGTGGAC (SEQ ID NO: 1129)	not	0	0.961889921 3	
2872	ACCTTCTACGAG AACCAGAT (SEQ ID NO: 613)	ACCTTCTACG AGAACCAGAT (SEQ ID NO: 1130)	not	0	1.923779843	
2950	GGCCTGCTGCCT CCAACCTC (SEQ ID NO: 614)	GGCCTGCTGC CTCCAACCTC (SEQ ID NO: 1131)	not	0	0.961889921 3	
92	TTGTGGTGGAAC TCGTGTGG (SEQ ID NO: 615)	TTGTGGTGGA ACTCGTGTGG (SEQ ID NO: 1132)	not	0	1.923779843	
3050	ACAACATCCTGT TCCACCAC (SEQ ID NO: 616)	ACAACATCCT GTTCCACCAC (SEQ ID NO: 1133)	not	0	0.961889921 3	
1874	TCACAAGGAGCC AGGTGCAG (SEQ ID NO: 617)	TCACAAGGAG CCAGGTGCAG (SEQ ID NO: 1134)	not	0	0.961889921 3	

2984	TTGGGGGAAGG GACATTGAA (SEQ ID NO: 618)	TTGGGGGAAG GGACATTGAA (SEQ ID NO: 1135)	not	0	1.923779843	
2107	TGGTGTACCTGG TTCCTGGA (SEQ ID NO: 619)	TGGTGTACCT GGTTCCTGGA (SEQ ID NO: 1136)	not	0	0.961889921 3	
33	ACAGATACAGCT TTTGCTCT (SEQ ID NO: 620)	ACAGATACAG CTTTTGCTCT (SEQ ID NO: 1137)	not	0	1.923779843	
2999	TTGAAACCAGCC TGGATGCA (SEQ ID NO: 621)	TTGAAACCAG CCTGGATGCA (SEQ ID NO: 1138)	not	0	0.961889921 3	
2836	CCCTGTGGCCG GCCAGCTGT (SEQ ID NO: 622)	CCCTGTGGCC GGCCAGCTGT (SEQ ID NO: 1139)	not	0	0.961889921 3	
2262	CTTTCTGCTCAG CACCTTCT (SEQ ID NO: 623)	CTTTCTGCTC AGCACCTTCT (SEQ ID NO: 1140)	not	0	0.961889921 3	
2640	TCTACAAGAGTC GTATTGGC (SEQ ID NO: 624)	TCTACAAGAG TCGTATTGGC (SEQ ID NO: 1141)	not	0	1.923779843	
2717	TAACAGAGGAAA CGGAGGAT (SEQ ID NO: 625)	TAACAGAGGA AACGGAGGAT (SEQ ID NO: 1142)	not	0	0.961889921 3	
3038	TGTGTCCACAGC ACAACATC (SEQ ID NO: 626)	TGTGTCCACA GCACAACATC (SEQ ID NO: 1143)	not	0	0.961889921 3	
1579	GTCCTGGATAAG TTTGAAAG (SEQ ID NO: 627)	GTCCTGGATA AGTTTGAAAG (SEQ ID NO: 1144)	not	0	0.961889921 3	
2307	CTGTAGTGGTGT CATCTATT (SEQ ID NO: 628)	CTGTAGTGGT GTCATCTATT (SEQ ID NO: 1145)	not	0	0.961889921 3	
3138	GGAGATGGAAG CCATGTTGG (SEQ ID NO: 629)	GGAGATGGAAG GCCATGTTGG (SEQ ID NO: 1146)	not	0	0.961889921 3	

2892	CACCGCATTCT GGGCCACA (SEQ ID NO: 630)	CACCGCATTC CTGGGCCACA (SEQ ID NO: 1147)	not	0	0.961889921 3	
1558	AATCAATACCTG GAGTGCTT (SEQ ID NO: 631)	AATCAATACCT GGAGTGCTT (SEQ ID NO: 1148)	not	0	1.923779843	
2955	GCTGCCTCCAAC CTCTGGGA (SEQ ID NO: 632)	GCTGCCTCCA ACCTCTGGGA (SEQ ID NO: 1149)	not	0	0.961889921 3	
3146	AAGCCATGTTGG AGGACACA (SEQ ID NO: 633)	AAGCCATGTT GGAGGACACA (SEQ ID NO: 1150)	not	0	1.923779843	
2799	TGGGGTATGCGT GAAGAATC (SEQ ID NO: 634)	TGGGGTGTGC GTGAAGAATC (SEQ ID NO: 1151)	not	0	0.961889921 3	
91	TTTGTGGTGGAA CTCGTGTG (SEQ ID NO: 635)	TTTGTGGTGG AACTCGTGTG (SEQ ID NO: 1152)	not	0	0.961889921 3	
1495	GACATGGCCAAC TTCGACTG (SEQ ID NO: 636)	GACATGGCCA ACTTCGACTG (SEQ ID NO: 1153)	not	0	0.961889921 3	
1969	CTGAACCGCTGT TTCCCTAT (SEQ ID NO: 637)	CTGAACCGCT GTTTCCCTAT (SEQ ID NO: 1154)	not	0	0.961889921 3	
2340	GCCACACATCCT GTGCTTCG (SEQ ID NO: 638)	GCCCCACATC CTGTGCTTCG (SEQ ID NO: 1155)	not	0	1.923779843	
578	GTCCAGAGCAGT TCGCTCAT (SEQ ID NO: 639)	GTCCAGAGCA GTTTCGTCAT (SEQ ID NO: 1156)	not	0	0.961889921 3	
2491	GGGAACAGTCC CACGGAAGG (SEQ ID NO: 640)	GGGAACAGTC CCACGGAAGG (SEQ ID NO: 1157)	not	0	0.961889921 3	
1826	ACATCTGGGGC GGGTTTGCC (SEQ ID NO: 641)	ACATCTGGGG CGGGTTTGCA (SEQ ID NO: 1158)	not	0	0.961889921 3	

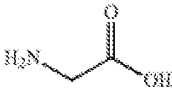
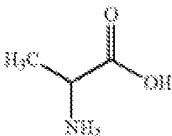
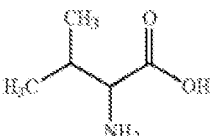
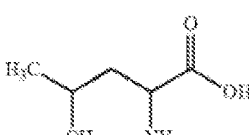
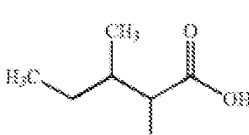
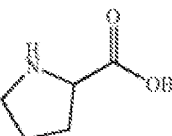
563	TCAACTCTCAAG TCCGTCCA (SEQ ID NO: 642)	TCAACTCTCA AGTCCGTCCA (SEQ ID NO: 1159)	not	0	0.961889921 3	
1535	CTGATCGCACCC TCCGCCTG (SEQ ID NO: 643)	CTGATCGCAC CCTCCGCCTG (SEQ ID NO: 1160)	not	0	0.961889921 3	
1444	GAAGCCATCCTA AACTTCCT (SEQ ID NO: 644)	GAGGCCATCC TAAACTTCCT (SEQ ID NO: 1161)	not	0	0.961889921 3	
3223	AAACTGAGCGTG GCCATTGC (SEQ ID NO: 645)	AAACTGAGCG TGGCCATTGC (SEQ ID NO: 1162)	not	0	0.961889921 3	
1500	GGCCAACTTCGA CTGGAGGG (SEQ ID NO: 646)	GGCCAACTTC GACTGGAGGG (SEQ ID NO: 1163)	not	0	1.923779843	
3200	ACCTATCAGGCG GCATGCAG (SEQ ID NO: 647)	ACCTATCAGG CGGCATGCAG (SEQ ID NO: 1164)	not	0	0.961889921 3	
787	CTTCCCACACTC CTAGACAG (SEQ ID NO: 648)	CTTCCCACAC TCCTAGACAG (SEQ ID NO: 1165)	not	0	0.961889921 3	
2213	TCCTCTTCCTGT TCTTGTTG (SEQ ID NO: 649)	TCCTCTTCCT GTTCTTGTTG (SEQ ID NO: 1166)	not	0	0.961889921 3	
2411	CTCCGGTGGCAT TTGGATTT (SEQ ID NO: 650)	CTCCGGTGGC ATTTGGATTT (SEQ ID NO: 1167)	not	0	0.961889921 3	
3035	GCATGTGTCCAC AGCACAAC (SEQ ID NO: 651)	GCATGTGTCC ACAGCACAAC (SEQ ID NO: 1168)	not	0	0.961889921 3	
2812	AAGAATCTGGTA AAGATTTT (SEQ ID NO: 652)	AAGAATCTGG TGAAGATTTT (SEQ ID NO: 1169)	branchpoint	0	0.961889921 3	
626	TCGCCTGCAGC GAGGCCCTC (SEQ ID NO: 653)	TCGCCTGCAG CGAGGCCCTC (SEQ ID NO: 1170)	not	0	0.961889921 3	

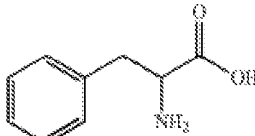
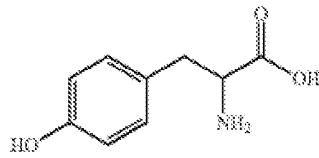
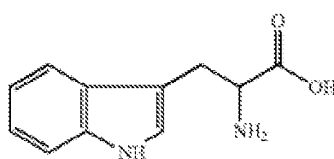
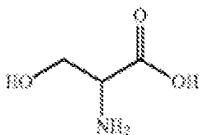
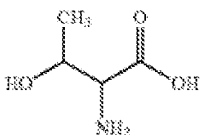
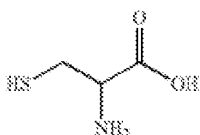
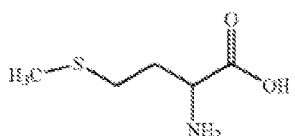
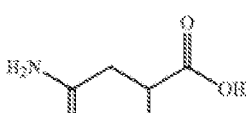
2005	TGGATCTACTCT GTCTCCAT (SEQ ID NO: 654)	TGGATCTACT CTGTCTCCAT (SEQ ID NO: 1171)	not	0	0.961889921 3	
65	CCCTGCGGAAAA GGCAAAAG (SEQ ID NO: 655)	CCCTGCGGAA AAGGCAAAAG (SEQ ID NO: 1172)	not	0	0.961889921 3	
2928	CACCACCTTGTC CATCCTCA (SEQ ID NO: 656)	CACCACCTTG TCCATCCTCA (SEQ ID NO: 1173)	not	0	0.961889921 3	
2758	CACGACTCCTTC TTCGAGAG (SEQ ID NO: 657)	CACGACTCCT TCTTCGAGAG (SEQ ID NO: 1174)	not	0	0.961889921 3	
2295	TCTGGCAGCAGC CTGTAGTG (SEQ ID NO: 658)	TCTGGCAGCA GCCTGTAGTG (SEQ ID NO: 1175)	not	0	0.961889921 3	
3139	GAGATGGAAGC CATGTTGGA (SEQ ID NO: 659)	GAGATGGAAG CCATGTTGGA (SEQ ID NO: 1176)	not	0	1.923779843	
169	CATCATGAATGC CATTTCCC (SEQ ID NO: 660)	CATCATGAAT GCCATTTCCC (SEQ ID NO: 1177)	not	0	0.961889921 3	
2964	AACCTCTGGGAC TGTGCTCG (SEQ ID NO: 661)	AACCTCTGGG ACTGTGCTCG (SEQ ID NO: 1178)	not	0	0.961889921 3	
1354	ACACAGATGAAC ATGATCAG (SEQ ID NO: 662)	ACACAGATGA ACATGATCAG (SEQ ID NO: 1179)	not	0	0.961889921 3	
1416	GCAGCTTGGTGA AGAAGGTA (SEQ ID NO: 663)	GCAGCTTGGT GAAGAAGGCA (SEQ ID NO: 1180)	not	0	0.961889921 3	
2902	CTGGGCCACAAT GGAGCTGG (SEQ ID NO: 664)	CTGGGCCACA ATGGAGCTGG (SEQ ID NO: 1181)	not	0	0.961889921 3	
954	TACAAAGCTGAT GGGCATCC (SEQ ID NO: 665)	TACAAAGCTG ATGGGCATCC (SEQ ID NO: 1182)	not	0	0.961889921 3	

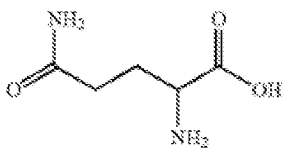
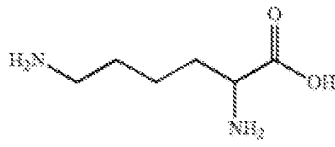
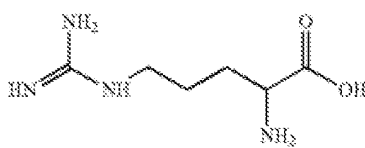
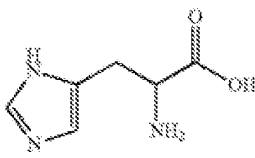
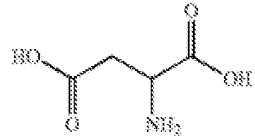
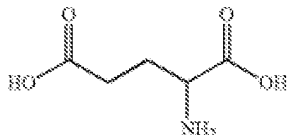
2942	TCCTCACAGGCC TGCTGCCT (SEQ ID NO: 666)	TCCTCACTGG CCTGCTGCCT (SEQ ID NO: 1183)	not	0	0.961889921 3	
94	GTGGTGGAAC GTGTGGCC (SEQ ID NO: 667)	GTGGTGGAAC TCGTGTGGCC (SEQ ID NO: 1184)	not	0	0.961889921 3	
3064	CACCACCTCACG GTGGCTGA (SEQ ID NO: 668)	CACCACCTCA CGGTGGCTGA (SEQ ID NO: 1185)	not	0	0.961889921 3	
2232	GGCTTTCTCCAC TGCCACCA (SEQ ID NO: 669)	GGCTTTCTCC ACTGCCACCA (SEQ ID NO: 1186)	not	0	0.961889921 3	
2660	TTGGCGGTGAAG GGTGTTC (SEQ ID NO: 670)	TTGGCGGTGA AGGGTGTTC (SEQ ID NO: 1187)	not	0	0.961889921 3	

[0336] “Canonical amino acids” refer to those 20 amino acids found naturally in the human body shown in Table 6 below.

Table 6: Depicts the canonical amino acids with each of their three letter abbreviations, one letter abbreviations, structures, and corresponding codons:

non-polar, aliphatic residues				
Glycine	Gly	G		GGU GGC GGA GGG
Alanine	Ala	A		GCU GCC GCA GCG
Valine	Val	V		GUU GUC GUA GUG
Leucine	Leu	L		UUA UUG CUU CUC CUA CUG
Isoleucine	Ile	I		AUU AUC AUA
Proline	Pro	P		CCU CCC CCA CCG

aromatic residues				
Phenylalanine	Phe	F		UUU UUC
Tyrosine	Tyr	Y		UAU UAC
Tryptophan	Trp	W		UGG
polar, non-charged residues				
Serine	Ser	S		UCU UCC UCA UCG AGU AGC
Threonine	Thr	T		ACU ACC ACA ACG
Cysteine	Cys	C		UGU UGC
Methionine	Met	M		AUG
Asparagine	Asn	N		AAU AAC

Glutamine	Gln	Q		CAA CAG
positively charged residues				
Lysine	Lys	K		AAA AAG
Arginine	Arg	R		CGU CGC CGA CGG AGA AGG
Histidine	His	H		CAU CAC
negatively charged residues				
Aspartate	Asp	D		GAU GAC
Glutamate	Glu	E		GAA GAG

Example 8. Evaluation of Exemplary ABCA4 trans-splicing molecules in vivo

[0337] The primary objective of this study is to confirm persistent Exon Editor expression and efficacy at the 1- and 3-month timepoints following subretinal administration of an exemplary AAV8-ABCA4 RNA Exon Editor construct at a dose of 4.3×10^{10} (4.3E10) vg/eye oculus uterque (OU; both eyes) in African Green monkey (AGM). The exemplary AAV8-ABCA4 RNA Exon Editor construct comprises SEQ ID NO: 77. The Exon Editor encoded by the exemplary AAV8-ABCA4 RNA Exon Editor construct comprises the same regulatory elements (enhancer, promoter and 5' UTR) as described in FIG. 13, including the double STOP codon that prevents NSP translation beyond the SD, but it does not comprise the 3xU1 and ARE features, and therefore produces significantly higher levels of the NSP compared to a construct that includes a potentiator (3xU1) and a translational

potentiator (ARE). The AAV8-ABCA4 RNA Exon Editor construct titer was measured using a ddPCR assay with a forward primer and probe targeting the terminal exon 22 within the Exon Editor's ABCA4 CDS and a reverse primer spanning the 40mer linker and double STOP codon sequence in the Exon Editor construct. This same ddPCR assay has been used to calculate the titer of other exemplary ABCA4 RNA Exon Editor constructs.

[0338] RT-qPCR was used to measure expression of edited *ABCA4* RNA, native NHP *ABCA4* RNA, and the exemplary AAV8-ABCA4 RNA Exon Editor construct, all normalized to the housekeeping gene *RNF20*. The percentage of *ABCA4* RNA replacement at 4 weeks (1 month) ranged from 14.4% to 31.5%, and at 12 weeks (3 months) ranged from 26.3% to 37.8%. The range of Exon Editor copy number per 10 ng RNA at 4 weeks was 1.3E6 to 0.6E6 and at 12 weeks 2.1E6 to 3.0E6. These results demonstrate that transduction of AAV packaged ABCA4 RNA Exon Editor constructs described herein result in durable expression of encoded RNA exon editors and durable activity as reflected by % replacement observed. Indeed, the % replacement is increasing over time. See FIGs. 22A and 22B.

[0339] In-life ocular assessment by slit-lamp microscopy showed mild-to-moderate inflammation in a few animals associated with the subretinal injections that peaked at Day 7 and was mostly resolved by Day 28. Fundus images showed ocular changes consistent with subretinal injections. Intraocular pressure (IOP), retinal thickness, and retinal volume remained stable for all animals over the course of the study. Confocal scanning laser ophthalmoscope (cSLO) and color fundus photography of eyes treated with the exemplary AAV8-ABCA4 RNA Exon Editor construct or vehicle exhibited comparable degrees of change in the areas of injection.

[0340] Twice daily general observations did not reveal clinical signs indicative of adverse systemic effects of the exemplary AAV8-ABCA4 RNA Exon Editor construct. Body weights, respiratory rates, heart rates and body temperature were stable during the study and no AAV8-ABCA4 RNA Exon Editor construct-related pattern was evident among the variations that occurred. No clinical pathology [clinical chemistry, complete blood count (CBC), coagulation profile] abnormalities were attributed to the administration of AAV8-ABCA4 RNA Exon Editor construct.

[0341] Histopathological analysis confirmed that the lateral geniculate nuclei, optic tract, optic chiasm, spleen, and liver samples from all animals were within normal limits. No signs of immune infiltrates or abnormalities in retinal anatomy were observed upon ocular histopathological analysis in any vehicle or AAV8-ABCA4 RNA Exon Editor construct-treated animal. Minimal to mild hypertrophy and migration of retinal pigment epithelium (RPE) at the level ora serrata (site of vitrectomy) was noted in two animals treated with the AAV8-ABCA4 RNA Exon Editor construct and one treated with vehicle. These findings are consistent with the injection procedure and considered not to be test article (construct) related.

[0342] This study demonstrated that the selected regulatory elements and 40mer linker and BD shared in common with ABCA4-01 (See, e.g., Examples 8-10) confer persistent expression and exon editing activity out to a 3-month time-point in AGM. See FIGs. 22A and 22B.

Example 9. Investigative Toxicology Study: Evaluation of Exemplary ABCA4 trans-splicing molecule ABCA4-01 in vivo

[0343] Investigative Toxicology Study: Evaluation of an exemplary AAV8-ABCA4 RNA Exon Editor construct ABCA4-01 (SEQ ID NO: 90) (which comprises an RNA Exon Editor comprising SEQ ID NO: 78) with respect to exon editing efficiency and investigative safety assessment in Cynomolgus Macaque at 1- and 6- months following a single subretinal administration of 3.5E11 or 1E12 vg/eye has been completed.

[0344] The primary objective of this study was to confirm in vivo safety and efficacy of the exemplary AAV8-ABCA4 RNA Exon Editor (ABCA4-01; SEQ ID NO: 90), which comprises SEQ ID NO: 78) at 1- and 6-month timepoints. Two dose levels were being evaluated in this study; 3.5E11 and 1E12 vg/eye, given as a single subretinal injection per eye (OU) in male cynomolgus macaques. The 1-month groups had n=2 for each dose level and n=2 for the vehicle group. The 6-month group had n=3 for each dose level and n=2 for vehicle. All dosing has been completed for this study.

[0345] As described herein, ABCA4-01 vector genome DNA copy (VGC) biodistribution, RNA expression, exon editing activity (% RNA replacement), and the edited human-NHP chimeric ABCA4 protein was measured in neural retinal samples.

[0346] In addition, VGC was evaluated in optic chiasm, optic tract, optic nerve. The following non-ocular tissues were collected: brain, dorsal root ganglion (lumbar), heart, kidney, liver, gallbladder, lateral geniculate nucleus, lung, lymph nodes (mandibular and mesenteric), spleen, and testis. The initial plan was to analyze all non-ocular tissues from only the 1E12 vg/eye dose group at the 1-month timepoint to assess the need for downstream and future analysis requirements.

[0347] As described herein, RT-qPCR (or ddPCR) was used to measure edited *ABCA4* RNA, native NHP *ABCA4* RNA, and ABCA4-01 Exon Editor RNA expression, all normalized to the housekeeper *RNF20*. Beyond neural retinal samples, only tissues positive for ABCA4-01 VGC were evaluated for ABCA4-01 Exon Editor RNA expression.

[0348] One eye from each animal was evaluated by a certified pathologist and a signed histopathological report was generated. All standard systemic tissues (including adrenal gland, any gross lesions, aorta (ascending), bone marrow, brain, esophagus, gallbladder, heart (left ventricular wall), kidney, large intestine, lateral geniculate nucleus, liver, lung, mesenteric lymph node, optic chiasm, optic tract, pancreas, sciatic nerve, skeletal muscle, skin, small intestine (jejunum), spinal cord, spleen, stomach, mandibular lymph node, testis, thymus, and urinary bladder) in support of non-clinical toxicology assessment were collected, fixed and embedded in paraffin blocks. From all animals any gross lesions, bone marrow smear (rib), dorsal root ganglion, heart, kidney, lateral geniculate nucleus, liver, lung, optic chiasm, optic nerve, optic tract, and spleen were intended to be brought to slide, stained with H&E and evaluated by a certified pathologist. All other histology samples were banked for potential future investigation if needed.

Preliminary Observations from Investigative Toxicology Study:

[0349] Results at 1-month post-injection at the 3.5E11 dose of ABCA4-01 revealed 54% replacement in vivo. These results underscore the potential for therapeutic efficacy for ABCA4-01 at least because the observed level of % replacement suggests that therapeutically effective amounts of trans-spliced ABCA4 protein were generated in the context of photoreceptor and RPE cells to restore wildtype ABCA4 functionality in cells in which endogenous ABCA4 activity is reduced or absent.

[0350] To date, ABCA4-01 at both doses has been well tolerated. Preliminary findings of ocular exams up to day 35 of the study are summarized here. Each eye was monitored at 2-3 days, 1 week, 2 weeks and 4 weeks or 5 weeks after dosing for signs of inflammation. Aqueous cells, vitreous cells, conjunctival status, iridial hyperemia, chemosis, discharge, eyelids, PLR, corneal opacity, keratic precipitates, corneal vascularization, aqueous flare, aqueous cells color, vitreous cells color, vitreous haze, lens, and fundus were examined. Also, intraocular pressure was measured.

[0351] There was mild-to-moderate ocular inflammation seen in some eyes and at some time points attributed to the subretinal procedure. A treatment plan was in place to treat with anti-inflammatory agents depending on ocular exam results. The lower threshold (aqueous cell score >3 or 3 with aqueous flare, no signs of anterior vitreous cells or vitreal haze) triggered subconjunctival injection with dexamethasone (0.1 ml of 10 mg/ml) and the higher threshold (aqueous cell scores greater than 3 or 3 with aqueous flare and signs of anterior vitreous cells or vitreal haze) intravitreal injection with triamcinolone. No treatments have been required to date in any animals that received the 3.5E11 vg/eye dose of ABCA4-01. In one vehicle-treated animal, triamcinolone was administered (OD only) at 72h. In the 1E12 dose/eye groups, one animal required dexamethasone at Day 2 and triamcinolone on Day 7 (OS only) and dexamethasone treatment on Day 29 (OD only), a second animal required OU triamcinolone treatment on Day 8, and a third animal received triamcinolone on Day 29 (OD only). In each case follow-up ocular exams showed reduction of inflammation. Of note, by design animals in this study did not receive any prophylactic corticosteroid or other immunosuppressive regimen to prevent masking of any unfavorable immune responses, especially against the ABCA4-01 encoded Exon Editor. The inflammation described above was consistent with subretinal injection and response to higher AAV8 doses. Thus, this investigative study was well suited to help define the high dose level in subsequent studies.

[0352] Intraocular pressure remained stable throughout the study in all groups. Imaging of the eyes with cSLO/OCT showed fundus changes consistent with subretinal injection and did not show changes attributed to ABCA4-01 administration (study ongoing, day 8 data evaluated).

[0353] Anti-AAV8 NABs were evaluated prior to ABCA4-01 administration in all animals. These titers ranged from below 1:10 to $\geq 1:100$. The pre-existing NAb titer data was used to distribute animals across the three groups. At 4- or 5-weeks, vehicle dosed animals had NAb titers within the range $\geq 1:10$ to $\geq 1:100$ while those dosed with ABCA4-01 at either dose ranged from $\geq 1:100$ to $\geq 1:1000$ consistent with an expected adaptive immune response against a minor amount of AAV8 capsid that can leak from the subretinal bleb into systemic circulation.

[0354] Cage side monitoring of animal health and measuring of body weights showed no signs of adverse events.

Example 10. Definitive GLP Toxicology Study: Evaluation of Exemplary ABCA4 trans-splicing molecule in vivo

[0355] Evaluation of an exemplary AAV8-ABCA4 RNA Exon Editor construct ABCA4-01 (SEQ ID NO: 90) (which comprises SEQ ID NO: 78) with respect to exon editing efficiency and investigative safety assessment in Cynomolgus Macaque at 1- and 6- months following a single subretinal administration of 3.5E11 or 1E12 vg/eye has been completed.

[0356] Definitive GLP Toxicology Study: The goal of this study was to evaluate the subretinal tolerability of an exemplary AAV8-ABCA4 RNA Exon Editor ABCA4-01 (SEQ ID NO: 90), which comprises SEQ ID NO: 78, in Cynomolgus Macaque at 3 and 6 months following a single subretinal administration of different doses of ABCA4-01 that ranged from 4E10 to 1E12 vg/eye. The doses selected for this study were 3.5E11 or 1E12 vg/eye. Both male and female NHPs were enrolled and dosed with 100µl per subretinal bleb, OU. The study was designed to evaluate n=3 animals/timepoint for each AAV8-ABCA4 RNA Exon Editor ABCA4-01 (SEQ ID NO: 90) and dose level group, and n=2 vehicle animals/timepoint.

[0357] Endpoints included ocular and systemic safety assessments both in-life and samples collected at necropsy. In-life assessments of general and ocular health were performed as follows: Detailed observations (weekly), body weights (pretest, day -1 and weekly thereafter), food consumption (weekly), ophthalmology (pretest, weeks 1 and 5 and prior to terminal necropsy; indirect ophthalmoscopy and slit lamp biomicroscopy), intraocular pressure measurements (pretest, weeks 1 and 5 and prior to terminal necropsy), confocal laser ophthalmoscopy (pretest, week 5 and prior to terminal necropsy), optical coherence tomography (pretest, week 5 and prior to terminal necropsy), clinical pathology (hematology, coagulation, clinical chemistry: twice pre-test, day 1 pre-dose, days 2, 8, 15, 22, 29, 35, 43, 50, 57 and monthly thereafter until necropsy).

[0358] For each animal, one eye was examined for histopathology. The contralateral eye was dissected, and neural retinal samples and RPE/Choroid samples were flash frozen. The neural retina sample mapped to the treated area of the retina were analyzed for ABCA4-01 VGC biodistribution, RNA expression, and exon editing activity (% RNA replacement). Pending neural retinal sample availability (after sufficient DNA and RNA extraction is confirmed for required analyses), the edited human-NHP chimeric ABCA4 protein was evaluated by a qualified IA-LC-MS/MS assay.

[0359] In addition, VGC was evaluated only in tissues that are positive for ABCA4-01 VGC in the 1E12 vg/eye dose group at 1-month from the Investigative Toxicology Study described above. Likewise, ABCA4-01 RNA expression was measured in tissues positive for ABCA4-01 VGC.

[0360] All standard systemic tissues (including adrenal gland, any gross lesions, aorta, bone marrow, brain, esophagus, gallbladder, heart (left ventricular wall), kidney, large intestine, lateral geniculate nucleus, liver, lung, mesenteric lymph node, optic chiasm, optic tract, pancreas, sciatic

nerve, skeletal muscle, skin, small intestine (jejunum), spinal cord, spleen, stomach, mandibular lymph node, testis or ovaries, thymus, and urinary bladder) in support of non-clinical toxicology assessment were collected, fixed and embedded in paraffin blocks.

[0361] Eyes from all animals were evaluated by a certified pathologist. The same non-ocular tissue types found to be positive for ABCA4-01 VGC in the Investigative Toxicology Study were mounted on slides, stained with H&E, and evaluated. All other samples were stored for potential future evaluation if needed.

[0362] In the investigative and definitive GLP studies described above the following assessments were conducted: ABCA4-01 DNA biodistribution and RNA expression, exon edited ABCA4 RNA and protein, clinical pathology (chemistry, hematology, coagulation), in-life ocular measurements (slit lamp, optical coherence tomography (OCT)/cSLO, intraocular pressure), and histopathology. Specific to any immunological response risk, in both the investigative and GLP toxicology studies, anti-AAV8 neutralizing antibodies were measured from serum at baseline and post-dose. In addition, in the GLP studies, baseline, in-life, and terminal circulating cytokines and complement were evaluated and histopathology assessment in any tissues with confirmed ABCA4-01 DNA biodistribution was conducted. Results from this study indicate that treatment with ABCA4-01 is safe and not associated with any adverse events.

[0363] Further to the above, FIGs. 30A-30C present results demonstrating robust ABCA4 RNA and protein replacement in vivo in NHP six (6) months after treatment with an exemplary AAV8 RNA exon editor construct (ABCA4-01; SEQ ID NO: 90). Viral genome copy (VGC) (x-axis) and exon editor RNA copies numbers (y-axis) are shown in FIG. 30A; % ABCA4 RNA replacement is shown in FIG. 30B; and ABCA4 protein replacement is shown in FIG. 30C for cynomolgus animals (n = 3) at 6-months following treatment with ABCA4-01 (results for an exemplary dose of 3.5E11 vg/eye are presented). DNA and RNA levels were evaluated using Droplet Digital Polymerase Chain Reaction (ddPCR) based assays. Percent total protein (edited ABCA4 relative to the sum of edited and NHP native protein) was quantified using a qualified immuno-affinity mass spectrometry (IA-MS) assay. The exemplary AAV8 RNA exon editor construct used in this in vivo study comprises an RNA exon editor comprising SEQ ID NO: 90. Results for four (4) NHPs are shown and designated in (B) and (C) as 1-4: 1 (vehicle treated); 2-4 (treated with 3.5E11 vg/eye).

[0364] In brief, the present inventors treated wild-type cynomolgus macaques with a single subretinal dose of ABCA4-01 (3.5E11 vg/eye). AAV vector genome DNA copy (VGC) biodistribution, Exon Editor RNA expression, and the percent of edited ABCA4 RNA and protein at a 6-month timepoint were then evaluated. Exon Editor RNA expression correlated well with VGC (FIG. 30A) and resulted in robust % RNA replacement (conversion of native ABCA4 RNA copies to edited copies), calculated as the percent of edited ABCA4 RNA out of the total (edited + native) ABCA4 RNA population present in the neural retina samples (FIG. 30B). As shown in FIG. 30B, the percent RNA replacement ranged from about 40%-60% for this dose of ABCA4-01. This resulted in therapeutic levels of the resultant human-NHP chimeric ABCA4 protein as measured by a qualified IA-MS assay

(FIG. 30C). More specifically, human NHP chimeric ABCA4 protein, which results from ABCA4 RNA editing in these animals, was found to be present at 20-40% of the total (human-NHP + NHP) ABCA4 protein in the tested samples. The present inventors have, moreover, demonstrated up to 66% RNA replacement and up to 45% human NHP chimeric ABCA4 protein in wild-type cynomolgus macaques treated with a higher dose of ABCA4-01 delivered via single subretinal injection. These results demonstrate that ABCA4-01 achieves therapeutically relevant levels of edited ABCA4 protein expression in NHP, which exceed the levels of rescued ABCA4 protein previously shown to confer therapeutic benefit in an *Abca4* KO mouse model.

Example 11: AAV vector plasmid comprising RNA exon editor SEQ ID NO: 78

[0365] The plasmid (SEQ ID NO: 82) encoding ABCA4-01 is a 13,232 base pair (bp) AAV vector plasmid comprising the RNA exon editor (SEQ ID NO: 78) flanked by AAV2 ITRs (SEQ ID NO: 90). The plasmid backbone contains a Kanamycin resistance (KanR) gene, a bacterial origin of replication and two non-coding regions derived from the Opsin 1 intron 1. These latter sequences are described as 'stuffer' sequences and are designed to increase the size of the plasmid beyond the packaging limit for AAV and therefore reduce the level of possible encapsidated plasmid DNA backbone components. A schematic representation of AAV vector plasmid is presented in FIG. 23 and additional details regarding same are presented in Table 5.

Table 5: Genetic Elements Within the AAV Vector Plasmid comprising SEQ ID NO: 82

Feature	Position in plasmid	Length (bp)
KanR	810-1	810
Stuffer: Human Opsin 1 Intron 1 (2 nd half)	1185-4127	2943
5' AAV2 ITR	4284-4413	130
CMV immediate early enhancer	4486-4789	304
CMV immediate early promoter	4790-4993	204
ABCA4 5' UTR	4994-5099	106
ABCA4 coding sequence (CDS) Exons 1-22	5100-8427	3328
Splice donor (SD)	8428-8433	6
Double STOP codon	8433-8438	6
40mer Linker	8439-8478	40
3xU1 sites	8479-8549	71
GM-CSF ARE	8550-8595	46
Binding Domain (BD)	8596-8745	150
MALAT1 triple helix terminator	8746-8847	102
Anti-Mut1_mascRNA	8848-8905	58
3' AAV2 ITR	8944-9073	130
Stuffer: Human Opsin 1 Intron 1 (1 st half)	9091-12040	2950
Ori	12467-13055	589

Example 12. Evaluation of Exemplary ABCA4 trans-splicing molecules ex vivo

[0366] To evaluate ABCA4 exon editing in human photoreceptors, the present inventors treated human donor retinal explants with ABCA4-01, an exemplary AAV8 RNA exon editor construct comprising SEQ ID NO: 90. The present inventors have found that incubating explants from multiple donors with ABCA4-01 (3.9E11 vg) for 21 days results in approximately 20-30% RNA replacement of the endogenous exons 1-22 with the same corresponding exons encoded by ABCA4-01 (FIG. 31). The donor explants in this experiment came from individuals without any previously diagnosed retinal disease. Of note, the same lot of ABCA4-01 was also used in the investigative NHP study mentioned above (FIG. 30), further strengthening the present inventors' confidence in the translatability of ABCA4-01 results from NHP to human subjects.

Example 13. Sequence Confirmation of the Exon-edited ABCA4 mRNA

[0367] To confirm that the ABCA4 mature mRNA generated from exon editing has the sequence predicted to result from a precise trans-splicing reaction, including the correct Exon 22-Exon23 junction, the present inventors isolated and sequenced the edited ABCA4 RT PCR product from ABCA4 KO cells that were transiently transfected with a plasmid comprising SEQ ID NO: 78 that encodes for the ABCA4-01 Exon Editor. Briefly, confirmation of the full-length exon edited mRNA was achieved via RT-PCR on RNA isolated from transfected ABCA4 KO, cells followed by TOPO cloning and sequencing. Sanger sequencing confirmed the correct mature mRNA sequence in multiple, randomly selected clones. In addition to confirming the correct Exon 22-Exon 23 junction, these sequencing results confirmed that the full-length exon-edited ABCA4 mature mRNA molecule has the expected sequence over its entire length, including the Exon Editor derived CDS spanning Exons 1 to 22 and the endogenous sequence spanning Exons 23 to 50. These results provide additional confidence that ABCA4 exon editing mediated by treatment with ACDN-01 results in the intended, correct ABCA4 mature mRNA sequence.

ENUMERATED EMBODIMENTS I

1. A nucleic acid trans-splicing molecule comprising:
 - (a) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27; and
 - (b) a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA; and
 - (c) optionally, a coding domain sequence (CDS) comprising ABCA4 exons;
 - (d) optionally, a splicing domain;

(e) optionally, a potentiator comprising at least one of GGTAAGT or SEQ ID NO: 61 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to GGTAAGT or SEQ ID NO: 61;

and

(f) optionally, a translational potentiator comprising a sequence consisting of SEQ ID NO: 63 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 63,

wherein the CDS; the splicing domain, when present; the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain are operatively linked in a 5'-to-3' direction.

2. The nucleic acid trans-splicing molecule of embodiment 1, wherein the binding domain anneals to an intron in the endogenous ABCA4 pre-mRNA.
3. The nucleic acid trans-splicing molecule of embodiment 2, wherein the intron in the endogenous ABCA4 pre-mRNA is intron 22.
4. The nucleic acid trans-splicing molecule of any one of embodiments 1-3, wherein the binding domain comprises, consists essentially of, or consists of SEQ ID NO: 18 or SEQ ID NO: 20.
5. The nucleic acid trans-splicing molecule of any of embodiments 1-4, wherein the binding domain consists essentially of or consists of SEQ ID NO: 18 or SEQ ID NO: 20.
6. The nucleic acid trans-splicing molecule of any one of embodiments 1-5, wherein the binding domain consists essentially of or consists of SEQ ID NO: 20.
7. The nucleic acid trans-splicing molecule of any one of embodiments 1-6, wherein the CDS, when present, comprises ABCA4 exons 1-22 or variants thereof.
8. The nucleic acid trans-splicing molecule of any one of embodiments 1-7, wherein the CDS comprises, consists essentially of, or consists of any one of SEQ ID Nos: 53-59.
9. A nucleic acid trans-splicing molecule of any one of embodiments 1-8, comprising:
 - (a) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
 - (b) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93%

identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27; and

(c) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and

(d) optionally, a splicing domain comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 4, 6, 8, or 10;

(e) optionally, a 5' untranslated region;

(f) optionally, a 3' transcription terminator domain; and

(g) optionally, at least one stop codon operatively linked 3' to the splicing domain, wherein, wherein the 5' untranslated region, when present; the CDS; the splicing domain, when present; the at least one stop codon, when present; the linker domain; the potentiator, when present; the translational potentiator, when present; the binding domain; and the 3' transcription terminator domain, when present, are operatively linked in a 5'-to-3' direction.

10. The nucleic acid trans-splicing molecule of embodiment 9, wherein the potentiator comprises three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to GGTAAGT or SEQ ID NO: 61, or any combination thereof.

11. The nucleic acid trans-splicing molecule of any one of embodiments 9 or 10, wherein, the potentiator comprises SEQ ID NO: 62 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 62.

12. The nucleic acid trans-splicing molecule of any one of embodiments 9-11, wherein the splicing domain comprises, consists essentially of, or consists of SEQ ID NO: 4 (GTAAGT).

13. The nucleic acid trans-splicing molecule of any one of embodiments 9-12, wherein the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64.

14. The nucleic acid trans-splicing molecule of any one of embodiments 9-13, wherein the 3' transcription terminator domain encodes a ribonucleic acid that forms a triple helical structure that caps the 3' end of the trans-splicing molecule.

15. The nucleic acid trans-splicing molecule of any one of embodiments 9-14, wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 33 or SEQ ID NO: 66.

16. The nucleic acid trans-splicing molecule of embodiment 15, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 80% identity with SEQ ID NO: 66; at least 85% identity with SEQ ID NO: 66; at least 90% identity with SEQ ID NO: 66; at least 91% identity with SEQ ID NO: 66; at least 92% identity with SEQ ID NO: 66; at least 93% identity with SEQ ID NO: 66; at least 94% identity with SEQ ID NO: 66; at least 95% identity with SEQ ID NO: 66; at least 96% identity with SEQ ID NO: 66; at least 97% identity with SEQ ID NO: 66; at least 98% identity with SEQ ID NO: 66; at least 99% identity with SEQ ID NO: 66; or 100% identical to SEQ ID NO: 66.

17. The nucleic acid trans-splicing molecule of any one of embodiments 1-16, comprising, operatively linked in a 5'-to-3' direction:

- (a) a 5' untranslated region;
- (b) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
- (c) a splicing domain;
- (d) at least one stop codon;
- (e) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27;
- (f) at least one potentiator;
- (g) a translational potentiator;
- (h) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 20;
- and
- (i) a 3' transcriptional terminator domain.

18. The nucleic acid trans-splicing molecule of embodiment 17, wherein

- (a) the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64;
- (b) the splicing domain comprises, consists essentially of, or consists of SEQ ID NO: 4 (GTAAGT);
- (c) the at least one stop codon comprises two stop codons, wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence and wherein the sequence of the splicing domain in combination with the two stop codons comprises, consists essentially of, or consists of SEQ ID NO: 68;

- (d) the at least one potentiator comprises, consists essentially of, or consists of three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61, or any combination thereof;
- (e) the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63; and
- (f) the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

19. A nucleic acid trans-splicing molecule of any one of embodiments 1-18, wherein the CDS comprises, consists essentially of, or consists of SEQ ID NO: 56.

20. The nucleic acid trans-splicing molecule of embodiment 18 or embodiment 19, wherein the nucleic acid trans-splicing molecule comprises, consists essentially of, or consists of SEQ ID NO: 69, SEQ ID NO: 78, or SEQ ID NO: 90.

21. The nucleic acid trans-splicing molecule of any one of embodiments 1-16, wherein the nucleic acid trans-splicing molecule comprises, consists essentially of, or consists of SEQ ID NO: 70 or SEQ ID NO: 71.

22. The nucleic acid trans-splicing molecule of any one of embodiments 1-16, wherein the nucleic acid trans-splicing molecule comprises a nucleic acid sequence having at least 80% identity with any one of SEQ ID NOs: 40-44, 46, or 48-50.

23. A ribonucleic acid trans-splicing molecule transcribed from the nucleic acid trans-splicing molecule of any one of embodiments 1-22.

24. A nucleic acid trans-splicing molecule, comprising:

- (a) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27; and
- (b) optionally, a potentiator comprising, consisting essentially of, or consisting of at least one of GGTAAGT or 61, or any combination thereof; and
- (c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63, wherein the linker domain; the potentiator, when present, and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

25. A nucleic acid trans-splicing molecule of embodiment 24, wherein the potentiator, when present, comprises, consists essentially of, or consists of SEQ ID NO: 62.

26. The nucleic acid trans-splicing molecule of any one of embodiments 24 or 25, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain, are operatively linked in a 5'-to-3' direction.
27. The nucleic acid trans-splicing molecule of embodiment 26, further comprising a CDS comprising ABCA4 exons, wherein the CDS; the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.
28. A 5' nucleic acid trans-splicing molecule, comprising:
- (a) a potentiator comprising, consisting essentially of, or consisting of at least one of GGTAAGT or SEQ ID NO: 61 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to GGTAAGT or SEQ ID NO: 61; and
 - (b) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27;
 - (c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63,
- wherein the linker domain, when present; the potentiator; and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.
29. A 5' nucleic acid trans-splicing molecule, comprising:
- (a) a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62 or a sequence having at least 90% identity to SEQ ID NO: 62; and
 - (b) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27;
 - (c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63,
- wherein the linker domain, when present; the potentiator; and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.
30. The nucleic acid trans-splicing molecule of embodiment 28 or embodiment 29, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the linker domain, when present; the potentiator; the translational potentiator, when present; and the binding domain, are operatively linked in a 5'-to-3' direction.

31. The nucleic acid trans-splicing molecule of any one of embodiments 28-30, further comprising a CDS comprising ABCA4 exons,
wherein the CDS; the linker domain, when present; the potentiator; the translational potentiator, when present; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.
32. A nucleic acid trans-splicing molecule, comprising:
(a) a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63 or a sequence having at least 90% identity to SEQ ID NO: 63;
(b) optionally, a potentiator comprising, consisting essentially of, or consisting of at least one of GGTAAGT or 61, or any combination thereof; and
(c) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27, wherein the linker domain, when present; the potentiator, when present; and the translational potentiator, are operatively linked in a 5'-to-3' direction.
33. A nucleic acid trans-splicing molecule of embodiment 32, wherein the potentiator, when present, comprises, consists essentially of, or consists of SEQ ID NO: 62.
34. The nucleic acid trans-splicing molecule of embodiment 32 or embodiment 33, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the linker domain, when present; the potentiator, when present; the translational potentiator; and the binding domain, are operatively linked in a 5'-to-3' direction.
35. The nucleic acid trans-splicing molecule of any one of embodiments 32-34, further comprising a CDS comprising ABCA4 exons,
wherein the CDS; the linker domain, when present; the potentiator, when present; the translational potentiator; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.
36. A nucleic acid trans-splicing molecule, comprising:
a cassette comprising a linker domain and a potentiator, wherein the cassette comprises, consists essentially of, or consists of SEQ ID NO: 72.
37. A nucleic acid trans-splicing molecule, comprising:
a cassette comprising a linker domain, a potentiator, and a translational potentiator, wherein the cassette comprises, consists essentially of, or consists of SEQ ID NO: 73.

38. A nucleic acid trans-splicing molecule, comprising:
- (a) a 3' transcription terminator domain comprising, consisting essentially of, or consisting of SEQ ID NO: 66 or a sequence having at least 95% identity to SEQ ID NO: 66.
39. The nucleic acid trans-splicing molecule of embodiment 38, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 96% identity with SEQ ID NO: 66; at least 97% identity with SEQ ID NO: 66; at least 98% identity with SEQ ID NO: 66; at least 99% identity with SEQ ID NO: 66; or 100% identical to SEQ ID NO: 66.
40. The nucleic acid trans-splicing molecule of embodiment 38 or embodiment 39 wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

ENUMERATED EMBODIMENTS II

1. A nucleic acid trans-splicing molecule comprising:
- (a) a coding domain sequence (CDS) comprising ABCA4 exons;
 - (b) a linker domain comprising SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27; and
 - (c) a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the CDS, the linker domain, and the binding domain are operatively linked in a 5'-to-3' direction.
2. A nucleic acid trans-splicing molecule comprising:
- (a) a CDS comprising, consisting essentially of, or consisting of a variant of any one of SEQ ID NOs: 53-55, wherein the variant of any one of SEQ ID NOs: 53-55 comprises at least one nucleotide variation in at least one cryptic splice site listed in Table 3, wherein the at least one nucleotide variation reduces cryptic splice site use at each of the at least one cryptic splice sites comprising at least one nucleotide variation; and
 - (b) optionally, a linker domain; and
 - (c) optionally, a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the CDS; the linker domain, when present; and the binding domain, when present; are operatively linked in a 5'-to-3' direction.
3. The nucleic acid trans-splicing molecule of embodiment 2, wherein the variant of SEQ ID NO: 55 comprises any one of SEQ ID NOs: 56-59.

4. The nucleic acid trans-splicing molecule of embodiment 3, wherein the variant of SEQ ID NO: 55 comprises SEQ ID NO: 56.
5. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the linker domain, when present, comprises a sequence having at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27.
6. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the linker domain, when present, comprises, consists essentially of, or consists of SEQ ID NO: 27.
7. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the binding domain, when present, anneals to an intron in the endogenous ABCA4 pre-mRNA.
8. The nucleic acid trans-splicing molecule of embodiment 7, wherein the intron in the endogenous ABCA4 pre-mRNA is intron 22.
9. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the binding domain, when present, comprises SEQ ID NO: 18 or SEQ ID NO: 20.
10. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the binding domain consists essentially of or consists of SEQ ID NO: 18 or SEQ ID NO: 20.
11. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the binding domain, when present, consists essentially of or consists of SEQ ID NO: 20.
12. The nucleic acid trans-splicing molecule of any one of embodiments 1 or 5-11, wherein the CDS comprises ABCA4 exons 1-22 or variants thereof.
13. The nucleic acid trans-splicing molecule of embodiment 12, wherein the CDS comprises, consists essentially of, or consists of any one of SEQ ID NOs: 53-59.
14. A nucleic acid trans-splicing molecule comprising:
 - (a) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27;

- (b) a binding domain comprising, consisting essentially of, or consisting of SEQ NO: 18 or ID SEQ ID NO: 20; and
- (c) optionally, a CDS comprising ABCA4 exons;
- wherein the CDS, when present; the linker domain; and the binding domain are operatively linked in a 5'-to-3' direction.
15. The nucleic acid trans-splicing molecule of embodiment 14, wherein the CDS, when present, comprises, consists essentially of, or consists of any one of SEQ ID NOs: 53-59.
16. A nucleic acid trans-splicing molecule comprising:
- (a) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
- (b) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27; and
- (c) optionally, a binding domain comprising, consisting essentially of, or consisting of SEQ NO: 18 or SEQ ID NO: 20,
- wherein the CDS; the linker domain; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.
17. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a potentiator comprising at least one sequence comprising GGTAAGT or SEQ ID NO: 61.
18. The nucleic acid trans-splicing molecule of embodiment 17, wherein the potentiator is operably linked 3' to the linker domain.
19. The nucleic acid trans-splicing molecule of any one of embodiments 17 or 18, wherein the potentiator comprises three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61.
20. The nucleic acid trans-splicing molecule of any one of embodiments 17-19, wherein the potentiator comprises three sequences arranged in tandem, wherein the three sequences comprise SEQ ID NO: 62.
21. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a translational potentiator, wherein the translational potentiator comprises SEQ ID NO:

63 or a sequence having at least 90% identity to SEQ ID NO: 63, wherein SEQ ID NO: 63 or the sequence having at least 90% identity to SEQ ID NO: 63 is operably linked 3' to the potentiator.

22. The nucleic acid trans-splicing molecule of embodiment 21, wherein the translational potentiator has at least 91% identity to SEQ ID NO: 63, has at least 92% identity to SEQ ID NO: 63, has at least 93% identity to SEQ ID NO: 63, has at least 94% identity to SEQ ID NO: 63, has at least 95% identity to SEQ ID NO: 63, has at least 96% identity to SEQ ID NO: 63, has at least 97% identity to SEQ ID NO: 63, has at least 98% identity to SEQ ID NO: 63, or has at least 99% identity to SEQ ID NO: 63.

23. The nucleic acid trans-splicing molecule of any one of embodiments 21 or 22, wherein the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63.

24. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a splicing domain comprising any one of GTAAGT, GTAAGG, GTAAGC, or GTAAC.

25. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a splicing domain comprising, consisting essentially of, or consisting of GTAAGT.

26. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a 5' untranslated region.

27. The nucleic acid trans-splicing molecule of embodiment 26, wherein the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64.

28. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a 3' transcription terminator domain.

29. The nucleic acid trans-splicing molecule of embodiment 28, wherein the 3' transcription terminator domain encodes a ribonucleic acid that forms a triple helical structure that protects the 3' end of the trans-splicing molecule.

30. The nucleic acid trans-splicing molecule of embodiment 28 or embodiment 29, wherein the 3' transcriptional terminator domain comprises a wildtype MALAT1 sequence (SEQ ID NO: 33) or a nucleic acid sequence having at least 80% identity with SEQ ID NO: 33.

31. The nucleic acid trans-splicing molecule of embodiment 30, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 80% identity with SEQ ID

NO: 33; at least 85% identity with SEQ ID NO: 33; at least 90% identity with SEQ ID NO: 33; at least 91% identity with SEQ ID NO: 33; at least 92% identity with SEQ ID NO: 33; at least 93% identity with SEQ ID NO: 33; at least 94% identity with SEQ ID NO: 33; at least 95% identity with SEQ ID NO: 33; at least 96% identity with SEQ ID NO: 33; at least 97% identity with SEQ ID NO: 33; at least 98% identity with SEQ ID NO: 33; at least 99% identity with SEQ ID NO: 33; or 100% identical to SEQ ID NO: 33.

32. The nucleic acid trans-splicing molecule of embodiment 28 or embodiment 29 wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 33 or SEQ ID NO: 66.

33. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising at least one stop codon operatively linked 3' to the splicing domain.

34. The nucleic acid trans-splicing molecule of embodiment 33, comprising, operatively linked in a 5'-to-3' direction:

- (a) a 5' untranslated region;
- (b) a CDS consisting essentially of or consisting of any one of SEQ ID NOs: 53-59;
- (c) a splicing domain;
- (d) at least one stop codon;
- (e) a linker domain consisting essentially of or consisting of SEQ ID NO: 27;
- (f) at least one potentiator;
- (g) a translational potentiator;
- (h) a binding domain consisting essentially of or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and
- (i) a 3' transcriptional terminator domain.

35. The nucleic acid trans-splicing molecule of embodiment 34, wherein

- (a) the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64;
- (b) the splicing domain comprises, consists essentially of, or consists of GTAAGT;
- (c) the at least one stop codon is present and comprises, consists essentially of, or consists of two stop codons, wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence;
- (d) the at least one potentiator comprises, consists essentially of, or consists of three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61;

- (e) the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63; and
- (f) the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

36. A nucleic acid trans-splicing molecule comprising:

- (a) a 5' untranslated region;
 - (b) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
 - (c) a splicing domain;
 - (d) optionally, at least one stop codon;
 - (e) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27;
 - (f) at least one potentiator;
 - (g) a translational potentiator;
 - (h) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and
 - (i) a 3' transcriptional terminator domain,
- wherein the 5' untranslated region; the CDS; the splicing domain; the at least one stop codon, when present; the linker domain; the at least one potentiator; the translational potentiator; the binding domain; and the 3' transcriptional terminator domain are operatively linked in a 5'-to-3' direction.

37. The nucleic acid trans-splicing molecule of embodiment 36 wherein

- (a) the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64;
- (b) the splicing domain comprises, consists essentially of, or consists of GTAAGT;
- (c) the at least one stop codon is present and comprises, consists essentially of, or consists of two stop codons wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence;
- (d) the at least one potentiator comprises, consists essentially of, or consists of three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61;
- (e) the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63; and
- (f) the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

38. A nucleic acid trans-splicing molecule comprising, operatively linked in a 5'-to-3' direction:

- (a) a 5' untranslated region comprising, consisting essentially of, or consisting of SEQ ID NO: 13 or SEQ ID NO: 64;
- (b) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
- (c) a splicing domain comprising, consisting essentially of, or consisting of GTAAGT;
- (d) two stop codons, wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence and wherein the sequence of the splicing domain in combination with the two stop codons comprises, consists essentially of, or consists of SEQ ID NO: 68;
- (e) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27;
- (f) a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62;
- (g) a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63;
- (h) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and
- (i) a 3' transcriptional terminator domain comprising, consisting essentially of, or consisting of SEQ ID NO: 66.

39. The nucleic acid trans-splicing molecule of embodiment 37 or embodiment 38, wherein the nucleic acid trans-splicing molecule comprises, consists essentially of, or consists of SEQ ID NO: 69, SEQ ID NO: 78, or SEQ ID NO: 90.

40. The nucleic acid trans-splicing molecule of any one of embodiments 1-20, wherein the nucleic acid trans-splicing molecule comprises, consists essentially of, or consists of SEQ ID NO: 70 or SEQ ID NO: 71.

41. The nucleic acid trans-splicing molecule of any one of embodiments 1-12, wherein the nucleic acid trans-splicing molecule comprises a nucleic acid sequence having at least 80% identity with any one of SEQ ID NOs: 40-44, 46, or 48-50.

42. A ribonucleic acid trans-splicing molecule transcribed from the nucleic acid trans-splicing molecule of any one of embodiments 1-41.

43. The nucleic acid trans-splicing molecule of any one of embodiments 1-41, wherein the endogenous ABCA4 pre-mRNA comprises at least one mutation associated with an ABCA4-associated retinal dystrophy.

44. The nucleic acid trans-splicing molecule of embodiment 43, wherein the ABCA4-associated retinal dystrophy is recessively inherited.
45. The nucleic acid trans-splicing molecule of embodiment 43 or 44, wherein the ABCA4-associated retinal dystrophy is associated with allelic pathogenic mutations spanning exons 1-22 of the ABCA4 gene and a phenotype consistent with cone-rod dystrophy (e.g., Cone-Rod Dystrophy-3), Stargardt macular dystrophy (e.g., Stargardt Disease-1), fundus flavimaculatus, Retinitis Pigmentosa-19, or Age-Related Macular Degeneration-2.
46. The nucleic acid trans-splicing molecule of any one of embodiments 1-45, wherein the endogenous ABCA4 pre-mRNA comprising the at least one mutation is expressed in a photoreceptor cell or a retinal pigment epithelial cell.
47. A vector comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-46.
48. The vector of embodiment 47, wherein the vector comprises a 5' regulatory domain operatively linked 5' to the CDS.
49. The vector of embodiment 48, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.
50. The vector of any one of embodiments 48-49, wherein the 5' regulatory domain comprises a constitutive promoter or a tissue specific promoter.
51. The vector of embodiment 50, wherein the constitutive promoter is a CMV promoter.
52. A proviral plasmid comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-46.
53. An adeno-associated virus (AAV) comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-46, wherein the AAV optionally comprises a 5' regulatory domain operatively linked 5' to the nucleic acid trans-splicing molecule.
54. The AAV of embodiment 53, wherein the AAV comprises a 5' regulatory domain operatively linked 5' to the CDS.
55. The AAV of any one of embodiments 53 or 54, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.

56. The AAV of any one of embodiments 53-55, wherein the 5' regulatory domain comprises a constitutive promoter.
57. The AAV of embodiment 56, wherein the constitutive promoter is a CMV promoter.
58. The AAV of any one of embodiments 53-57, wherein the AAV preferentially targets a photoreceptor cell and/or a retinal pigment epithelial cell.
59. The AAV of any one of embodiments 53-58, wherein the AAV is AAV8, AAV5, or AAV2.
60. A composition comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, or the AAV of any one of embodiments 53-59.
61. The composition of embodiment 60, comprising a pharmaceutically acceptable excipient.
62. A method of expressing biologically active ABCA4 in a target cell to restore functional levels of ABCA4 protein in the target cell, the method comprising transducing the target cell with the nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61.
63. The method of embodiment 62, wherein at least 5%, at least 10%, or at least 15% of the target ABCA4 mRNA in the target cell is replaced.
64. The method of embodiment 63, wherein at least 20% of the target ABCA4 mRNA in the target cell is replaced.
65. The method of embodiment 64, wherein at least 35% of the target ABCA4 mRNA in the target cell is replaced.
66. The method of any one of embodiments 62-65, wherein functional levels of ABCA4 are restored in the target cell by expressing full-length ABCA4 protein having biological function.
67. A method of reducing ABCA4-associated lipofuscin and/or A2E accumulation in a subject, the method comprising transfecting or transducing a target retinal cell in the subject with the nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of

embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61.

68. A method of correcting at least one mutation in an ABCA4 exon sequence in a target cell of a subject, the method comprising administering to the subject the nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61.

69. A method of treating an ABCA4-associated retinal dystrophy in a subject in need thereof, the method comprising administering to the subject the nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61 in a therapeutically effective amount.

70. The method of any one of embodiments 67-69, the method comprising subretinal administration of the nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61 to the subject.

71. The method of any one of embodiments 67-70, wherein the subject is a non-human primate or a human.

72. The method of any one of embodiments 67-71, wherein the subject has been diagnosed with an ABCA4-associated retinal dystrophy.

73. The nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61 for use in treating an ABCA4-associated retinal dystrophy in a subject in need thereof.

74. The nucleic acid trans-splicing molecule of any one of embodiments 1-46, the vector of any one of embodiments 47-51, the proviral plasmid of embodiment 52, the AAV of any one of embodiments 53-59, or the composition of embodiment 60 or 61 for use in the preparation of a medicament for the treatment of an ABCA4-associated retinal dystrophy in a subject in need thereof.

75. The method of any one of embodiments 67-72 or the use of any one of embodiments 73 or 74, wherein the ABCA4-associated retinal dystrophy is associated with allelic pathogenic

mutations spanning Exons 1-22 of the ABCA4 gene and a phenotype consistent with cone-rod dystrophy, Stargardt macular dystrophy, or fundus flavimaculatus.

76. A nucleic acid trans-splicing molecule, comprising:

- (a) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27; and
- (b) optionally, a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62; and
- (c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63, wherein the linker domain; the potentiator, when present, and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

77. The nucleic acid trans-splicing molecule of embodiment 76, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA,

wherein the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain, are operatively linked in a 5'-to-3' direction.

78. The nucleic acid trans-splicing molecule of embodiment 76 or embodiment 77, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

79. A 5' nucleic acid trans-splicing molecule, comprising:

- (a) a potentiator comprising, consisting essentially of, or consisting of at least one of GGTAAGT or SEQ ID NO: 61 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to GGTAAGT or SEQ ID NO: 61; and
 - (b) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27;
 - (c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63,
- wherein the linker domain, when present; the potentiator; and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

80. A 5' nucleic acid trans-splicing molecule, comprising:

- (a) a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62 or a sequence having at least 90% identity to SEQ ID NO: 62; and
 - (b) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27;
 - (c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63,
- wherein the linker domain, when present; the potentiator; and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

81. The nucleic acid trans-splicing molecule of embodiment 79 or embodiment 80, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the linker domain, when present; the potentiator; the translational potentiator, when present; and the binding domain, are operatively linked in a 5'-to-3' direction.

82. The nucleic acid trans-splicing molecule of any one of embodiments 79-81, further comprising a CDS comprising ABCA4 exons, wherein the CDS; the linker domain, when present; the potentiator; the translational potentiator, when present; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

83. A nucleic acid trans-splicing molecule, comprising:

- (a) a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63 or a sequence having at least 90% identity to SEQ ID NO: 63;
- (b) optionally, a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62; and
- (c) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27, wherein the linker domain, when present; the potentiator, when present; and the translational potentiator, are operatively linked in a 5'-to-3' direction.

84. The nucleic acid trans-splicing molecule of embodiment 83, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the linker domain, when present; the potentiator, when present; the translational potentiator; and the binding domain, are operatively linked in a 5'-to-3' direction.

85. The nucleic acid trans-splicing molecule of embodiment 83 or embodiment 84, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the linker domain, when present; the potentiator, when present; the translational potentiator; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

86. A nucleic acid trans-splicing molecule, comprising:

a cassette comprising a linker domain and a potentiator, wherein the cassette comprises, consists essentially of, or consists of SEQ ID NO: 72.

87. A nucleic acid trans-splicing molecule, comprising:

a cassette comprising a linker domain, a potentiator, and a translational potentiator, wherein the cassette comprises, consists essentially of, or consists of SEQ ID NO: 73.

88. A nucleic acid trans-splicing molecule, comprising:

(a) a 3' transcription terminator domain comprising, consisting essentially of, or consisting of SEQ ID NO: 66 or a sequence having at least 95% identity to SEQ ID NO: 66.

89. The nucleic acid trans-splicing molecule of embodiment 88, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 96% identity with SEQ ID NO: 66; at least 97% identity with SEQ ID NO: 66; at least 98% identity with SEQ ID NO: 66; at least 99% identity with SEQ ID NO: 66; or 100% identical to SEQ ID NO: 66.

90. The nucleic acid trans-splicing molecule of embodiment 88 or embodiment 89 wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

91. The nucleic acid trans-splicing molecule of claim 89 or claim 90, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the cassette; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

92. A nucleic acid trans-splicing molecule, comprising:

(a) a 3' transcription terminator domain comprising, consisting essentially of, or consisting of SEQ ID NO: 66 or a sequence having at least 95% identity to SEQ ID NO: 66.

93. The nucleic acid trans-splicing molecule of claim 92, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 96% identity with SEQ ID NO: 66; at least 97% identity with SEQ ID NO: 66; at least 98% identity with SEQ ID NO: 66; at least 99% identity with SEQ ID NO: 66; or 100% identical to SEQ ID NO: 66.

94. The nucleic acid trans-splicing molecule of claim 92 or claim 93, wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

ENUMERATED EMBODIMENTS III

1. A nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain sequence (CDS) comprising ABCA4 exons having at least one nucleotide variation relative to a wild-type ABCA4 sequence at a cryptic splice site; and
 - (b) a binding domain that is complementary to a binding site within an endogenous ABCA4 pre-mRNA.
2. The nucleic acid trans-splicing molecule of embodiment 1, wherein the cryptic splice site is a cryptic splice site listed in Table 3.
3. The nucleic acid trans-splicing molecule of embodiment 1 or 2, wherein the at least one nucleotide variation is a synonymous mutation that reduces or eliminates splicing at the cryptic splice site.
4. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 3, wherein the CDS comprises one or more of SEQ ID NOs: 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 127, 129, 131, 673, 674, 679, 682, 685, 693, 695, 698, 701, 708, 721, 722, 737, 758, 763, 764, 765, 775, 785, 789, 791, 798, 799, 792, 818, 820, 822, 825, 831, 837, 854, 859, 873, 881, 886, 888, 891, 903, 904, 910, 918, 924, 932, 938, 940, 945, 946, 950, 951, 952, 968, 976, 992, 993, 995, 997, 998, 999, 1003, 1004, 1010, 1015, 1026, 1031, 1040, 1047, 1070, 1117, or 1169.
5. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 4, wherein the CDS comprises exons 1-22 of ABCA4 having the at least one nucleotide variation.
6. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 5, wherein the CDS comprises any one of SEQ ID NOs: 56-59.
7. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 6, wherein the CDS comprises SEQ ID NO: 56.

8. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 7, further comprising a linker domain comprising a sequence having at least 85% identity to SEQ ID NO: 27, wherein the CDS, the linker domain, and the binding domain are operatively linked in a 5'-to-3' direction.
9. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 8, wherein the binding domain anneals to a sequence within intron 22 of the endogenous ABCA4 pre-mRNA.
10. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 9, wherein the binding domain comprises SEQ ID NO: 18 or 20.
11. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 10, further comprising a potentiator comprising at least one sequence comprising GGTAAGT or SEQ ID NO: 61, wherein the potentiator is operably linked 3' to the linker domain.
12. The nucleic acid trans-splicing molecule of embodiment 11, wherein the potentiator comprises three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61.
13. The nucleic acid trans-splicing molecule of embodiment 12, wherein the potentiator comprises three sequences arranged in tandem and operably linked, wherein the three sequences together comprise SEQ ID NO: 62.
14. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 13, further comprising a translational potentiator, wherein the translational potentiator comprises a sequence having at least 85% identity to SEQ ID NO: 63 operably linked 3' to the potentiator.
15. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 14, further comprising a splicing domain comprising the sequence GTAAGT.
16. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 15, further comprising a 5' untranslated region.
17. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 16, further comprising a 3' transcription terminator domain comprising a MALAT1 sequence having at least 85% identity with SEQ ID NO: 33 or SEQ ID NO: 66.
18. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 17 comprising, operatively linked in a 5'-to-3' direction:

- (a) a 5' untranslated region comprising SEQ ID NO: 13 or SEQ ID NO: 64;
 - (b) a CDS comprising any one of SEQ ID NOs: 53-59;
 - (c) a splicing domain comprising the sequence GTAAGT;
 - (d) two stop codons, wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence and wherein the sequence of the splicing domain in combination with the two stop codons comprises SEQ ID NO: 68 (GTAAGTAGTGA);
 - (e) a linker domain comprising SEQ ID NO: 27;
 - (f) a potentiator comprising SEQ ID NO: 62;
 - (g) a translational potentiator comprising SEQ ID NO: 63;
 - (h) a binding domain comprising SEQ ID NO: 18 or SEQ ID NO: 20; and
 - (i) a 3' transcriptional terminator domain comprising SEQ ID NO: 66.
19. The nucleic acid trans-splicing molecule of any one of embodiments 1 to 18 comprising SEQ ID NO: 69, SEQ ID NO: 78, or SEQ ID NO: 90.
20. A ribonucleic acid trans-splicing molecule transcribed from the nucleic acid trans-splicing molecule of embodiment 19.
21. A vector or adeno-associated virus (AAV) comprising the nucleic acid trans-splicing molecule of any one of embodiments 1 to 19.
22. The vector of embodiment 21, wherein the vector comprises SEQ ID NO: 82.
23. The AAV of embodiment 21 or 22, wherein the AAV preferentially targets photoreceptor cells and/or retinal pigment epithelial cells.
24. The AAV of embodiment 23, wherein the AAV is AAV8, AAV5, or AAV2.
25. A pharmaceutical composition comprising the vector or the AAV of any one of embodiments 21 to 24 and a pharmaceutically acceptable excipient.
26. A method of treating an ABCA4-associated retinal dystrophy in a subject in need thereof, the method comprising administering to the subject the composition of embodiment 25 in a therapeutically effective amount.
27. A method of expressing biologically active ABCA4 in a target cell, the method comprising transfecting or transducing the target cell with a nucleic acid trans-splicing molecule comprising:

- (a) a coding domain sequence (CDS) comprising ABCA4 exons having at least one nucleotide variation relative to a wild-type ABCA4 sequence at a cryptic splice site; and
- (b) a binding domain that is complementary to a binding site within an endogenous ABCA4 pre-mRNA.

28. The method of embodiment 27, wherein the CDS comprises one or more of SEQ ID NOs: 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 127, 129, 131, 673, 674, 679, 682, 685, 693, 695, 698, 701, 708, 721, 722, 737, 758, 763, 764, 765, 775, 785, 789, 791, 798, 799, 792, 818, 820, 822, 825, 831, 837, 854, 859, 873, 881, 886, 888, 891, 903, 904, 910, 918, 924, 932, 938, 940, 945, 946, 950, 951, 952, 968, 976, 992, 993, 995, 997, 998, 999, 1003, 1004, 1010, 1015, 1026, 1031, 1040, 1047, 1070, 1117, or 1169.

29. A method of treating an ABCA4-associated retinal dystrophy in a subject in need thereof, the method comprising administering to the subject a nucleic acid trans-splicing molecule comprising:
- (a) a coding domain sequence (CDS) comprising ABCA4 exons having at least one nucleotide variation relative to a wild-type ABCA4 sequence at a cryptic splice site; and
 - (b) a binding domain that is complementary to a binding site within an endogenous ABCA4 pre-mRNA.

30. The method of embodiment 29, wherein the ABCA4-associated retinal dystrophy is associated with one or more allelic pathogenic mutations within exons 1-22 of the ABCA4 gene and a phenotype consistent with cone-rod dystrophy (e.g., Cone-Rod Dystrophy-3), Stargardt macular dystrophy (e.g., Stargardt Disease-1), fundus flavimaculatus, Retinitis Pigmentosa-19, or Age-Related Macular Degeneration-2.

OTHER EMBODIMENTS

[0368] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

[0369] While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure that come within known or customary practice within the art to which the invention pertains and may be applied to the essential features hereinbefore set forth, and follows in the scope of the claims.

[0370] Other embodiments are within the claims.

CLAIMS

1. A nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain sequence (CDS) comprising ABCA4 exons;
 - (b) a linker domain comprising SEQ ID NO: 27 or a sequence having at least 85% identity to SEQ ID NO: 27; and
 - (c) a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the CDS, the linker domain, and the binding domain are operatively linked in a 5'-to-3' direction.
2. A nucleic acid trans-splicing molecule comprising:
 - (a) a CDS comprising, consisting essentially of, or consisting of a variant of any one of SEQ ID NOs: 53-55, wherein the variant of any one of SEQ ID NOs: 53-55 comprises at least one nucleotide variation in at least one cryptic splice site listed in Table 3, wherein the at least one nucleotide variation reduces cryptic splice site use at each of the at least one cryptic splice sites comprising at least one nucleotide variation; and
 - (b) a linker domain; and
 - (c) a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA, wherein the CDS; the linker domain, and the binding domain, are operatively linked in a 5'-to-3' direction.
3. The nucleic acid trans-splicing molecule of claim 2, wherein the variant of SEQ ID NO: 55 comprises any one of SEQ ID NOs: 56-59.
4. The nucleic acid trans-splicing molecule of claim 3, wherein the variant of SEQ ID NO: 55 comprises SEQ ID NO: 56.
5. The nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the linker domain comprises a sequence having at least 85% identity, at least 86% identity, at least 87% identity, at least 88% identity, or at least 89% identity, 90% identity, 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27.
6. The nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the linker domain comprises, consists essentially of, or consists of SEQ ID NO: 27.
7. The nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the binding domain anneals to an intron in the endogenous ABCA4 pre-mRNA.

8. The nucleic acid trans-splicing molecule of claim 7, wherein the intron in the endogenous ABCA4 pre-mRNA is intron 22.
9. The nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the binding domain comprises SEQ ID NO: 18 or SEQ ID NO: 20.
10. The nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the binding domain consists essentially of or consists of SEQ ID NO: 18 or SEQ ID NO: 20.
11. The nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the binding domain consists essentially of or consists of SEQ ID NO: 20.
12. The nucleic acid trans-splicing molecule of any one of claims 1 or 5-11, wherein the CDS comprises ABCA4 exons 1-22 or variants thereof.
13. The nucleic acid trans-splicing molecule of claim 12, wherein the CDS comprises, consists essentially of, or consists of any one of SEQ ID NOs: 53-59.
14. A nucleic acid trans-splicing molecule comprising:
- (a) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27;
 - (b) a binding domain comprising, consisting essentially of, or consisting of SEQ NO: 18 or ID SEQ ID NO: 20; and
 - (c) a CDS comprising ABCA4 exons;
- wherein the CDS, the linker domain, and the binding domain are operatively linked in a 5'-to-3' direction.
15. The nucleic acid trans-splicing molecule of claim 14, wherein the CDS comprises, consists essentially of, or consists of any one of SEQ ID NOs: 53-59.
16. A nucleic acid trans-splicing molecule comprising:
- (a) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
 - (b) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity,

at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to SEQ ID NO: 27; and

(c) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 20, wherein the CDS, the linker domain, and the binding domain are operatively linked in a 5'-to-3' direction.

17. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a potentiator comprising at least one sequence comprising GGTAAGT or SEQ ID NO: 61.

18. The nucleic acid trans-splicing molecule of claim 17 wherein the potentiator is operably linked 3' to the linker domain.

19. The nucleic acid trans-splicing molecule of any one of claims 17 or 18, wherein the potentiator comprises three sequences arranged in tandem and are operably linked, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61.

20. The nucleic acid trans-splicing molecule of any one of claims 17-19 wherein the potentiator comprises three sequences arranged in tandem, wherein the three sequences comprise SEQ ID NO: 62.

21. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a translational potentiator, wherein the translational potentiator comprises SEQ ID NO: 63 or a sequence having at least 90% identity to SEQ ID NO: 63, wherein SEQ ID NO: 63 or the sequence having at least 90% identity to SEQ ID NO: 63 is operably linked 3' to the potentiator.

22. The nucleic acid trans-splicing molecule of claim 21, wherein the translational potentiator has at least 91% identity to SEQ ID NO: 63, has at least 92% identity to SEQ ID NO: 63, has at least 93% identity to SEQ ID NO: 63, has at least 94% identity to SEQ ID NO: 63, has at least 95% identity to SEQ ID NO: 63, has at least 96% identity to SEQ ID NO: 63, has at least 97% identity to SEQ ID NO: 63, has at least 98% identity to SEQ ID NO: 63, or has at least 99% identity to SEQ ID NO: 63.

23. The nucleic acid trans-splicing molecule of any one of claims 21 or 22, wherein the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63.

24. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a splicing domain comprising any one of GTAAGT, GTAAGG, GTAAGC, or GTAAC.

25. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a splicing domain comprising, consisting essentially of, or consisting of GTAAGT.
26. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a 5' untranslated region.
27. The nucleic acid trans-splicing molecule of claim 26, wherein the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64.
28. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a 3' transcription terminator domain.
29. The nucleic acid trans-splicing molecule of claim 28, wherein the 3' transcription terminator domain encodes a ribonucleic acid that forms a triple helical structure that protects the 3' end of the trans-splicing molecule.
30. The nucleic acid trans-splicing molecule of claim 28 or claim 29, wherein the 3' transcriptional terminator domain comprises a wildtype MALAT1 sequence (SEQ ID NO: 33) or a nucleic acid sequence having at least 80% identity with SEQ ID NO: 33.
31. The nucleic acid trans-splicing molecule of claim 30, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 80% identity with SEQ ID NO: 33; at least 85% identity with SEQ ID NO: 33; at least 90% identity with SEQ ID NO: 33; at least 91% identity with SEQ ID NO: 33; at least 92% identity with SEQ ID NO: 33; at least 93% identity with SEQ ID NO: 33; at least 94% identity with SEQ ID NO: 33; at least 95% identity with SEQ ID NO: 33; at least 96% identity with SEQ ID NO: 33; at least 97% identity with SEQ ID NO: 33; at least 98% identity with SEQ ID NO: 33; at least 99% identity with SEQ ID NO: 33; or 100% identical to SEQ ID NO: 33.
32. The nucleic acid trans-splicing molecule of claim 28 or claim 29 wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 33 or SEQ ID NO: 66.
33. The nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising at least one stop codon operatively linked 3' to the splicing domain.
34. The nucleic acid trans-splicing molecule of claim 33, comprising, operatively linked in a 5'-to-3' direction:
- (a) a 5' untranslated region;
 - (b) a CDS consisting essentially of or consisting of any one of SEQ ID NOs: 53-59;

- (c) a splicing domain;
- (d) at least one stop codon;
- (e) a linker domain consisting essentially of or consisting of SEQ ID NO: 27;
- (f) at least one potentiator;
- (g) a translational potentiator;
- (h) a binding domain consisting essentially of or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and
- (i) a 3' transcriptional terminator domain.

35. The nucleic acid trans-splicing molecule of claim 34, wherein

- (a) the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64;
- (b) the splicing domain comprises, consists essentially of, or consists of GTAAGT;
- (c) the at least one stop codon comprises, consists essentially of, or consists of two stop codons, wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence;
- (d) the at least one potentiator comprises, consists essentially of, or consists of three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61;
- (e) the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63; and
- (f) the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

36. A nucleic acid trans-splicing molecule comprising:

- (a) a 5' untranslated region;
- (b) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;
- (c) a splicing domain;
- (d) optionally, at least one stop codon;
- (e) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27;
- (f) at least one potentiator;
- (g) a translational potentiator;
- (h) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and
- (i) a 3' transcriptional terminator domain,

wherein the 5' untranslated region; the CDS; the splicing domain; the at least one stop codon, when present; the linker domain; the at least one potentiator; the translational potentiator; the binding domain; and the 3' transcriptional terminator domain are operatively linked in a 5'-to-3' direction.

37. The nucleic acid trans-splicing molecule of claim 36 wherein

(a) the 5' untranslated region comprises, consists essentially of, or consists of SEQ ID NO: 13 or SEQ ID NO: 64;

(b) the splicing domain comprises, consists essentially of, or consists of GTAAGT;

(c) the at least one stop codon is present and comprises, consists essentially of, or consists of two stop codons wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence;

(d) the at least one potentiator comprises, consists essentially of, or consists of three sequences arranged in tandem, wherein each of the three sequences comprises GGTAAGT or SEQ ID NO: 61;

(e) the translational potentiator comprises, consists essentially of, or consists of SEQ ID NO: 63; and

(f) the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.

38. A nucleic acid trans-splicing molecule comprising, operatively linked in a 5'-to-3' direction:

(a) a 5' untranslated region comprising, consisting essentially of, or consisting of SEQ ID NO: 13 or SEQ ID NO: 64;

(b) a CDS comprising, consisting essentially of, or consisting of any one of SEQ ID NOs: 53-59;

(c) a splicing domain comprising, consisting essentially of, or consisting of GTAAGT;

(d) two stop codons, wherein the sequence of the first of the two stop codons overlaps in part with the splicing domain sequence and wherein the sequence of the splicing domain in combination with the two stop codons comprises, consists essentially of, or consists of SEQ ID NO: 68;

(e) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27;

(f) a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62;

(g) a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63;

(h) a binding domain comprising, consisting essentially of, or consisting of SEQ ID NO: 18 or SEQ ID NO: 20; and

(i) a 3' transcriptional terminator domain comprising, consisting essentially of, or consisting of SEQ ID NO: 66.

39. The nucleic acid trans-splicing molecule of claim 37 or claim 38, wherein the nucleic acid trans-splicing molecule comprises, consists essentially of, or consists of SEQ ID NO: 69, SEQ ID NO: 78, or SEQ ID NO: 90.
40. The nucleic acid trans-splicing molecule of any one of claims 1-20, wherein the nucleic acid trans-splicing molecule comprises, consists essentially of, or consists of SEQ ID NO: 70 or SEQ ID NO: 71.
41. The nucleic acid trans-splicing molecule of any one of claims 1-12, wherein the nucleic acid trans-splicing molecule comprises a nucleic acid sequence having at least 80% identity with any one of SEQ ID NOs: 40-44, 46, or 48-50.
42. A ribonucleic acid trans-splicing molecule transcribed from the nucleic acid trans-splicing molecule of any one of claims 1-41.
43. The nucleic acid trans-splicing molecule of any one of claims 1-41, wherein the endogenous ABCA4 pre-mRNA comprises at least one mutation associated with an ABCA4-associated retinal dystrophy.
44. The nucleic acid trans-splicing molecule of claim 43, wherein the ABCA4-associated retinal dystrophy is recessively inherited.
45. The nucleic acid trans-splicing molecule of claim 43 or 44, wherein the ABCA4-associated retinal dystrophy is associated with allelic pathogenic mutations spanning exons 1-22 of the ABCA4 gene and a phenotype consistent with cone-rod dystrophy (e.g., Cone-Rod Dystrophy-3), Stargardt macular dystrophy (e.g., Stargardt Disease-1), fundus flavimaculatus, Retinitis Pigmentosa-19, or Age-Related Macular Degeneration-2.
46. The nucleic acid trans-splicing molecule of any one of claims 1-45, wherein the endogenous ABCA4 pre-mRNA comprising the at least one mutation is expressed in a photoreceptor cell or a retinal pigment epithelial cell.
47. A vector comprising the nucleic acid trans-splicing molecule of any one of claims 1-46.
48. The vector of claim 47, wherein the vector comprises a 5' regulatory domain operatively linked 5' to the CDS.
49. The vector of claim 48, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.

50. The vector of any one of claims 48-49, wherein the 5' regulatory domain comprises a constitutive promoter or a tissue specific promoter.
51. The vector of claim 50, wherein the constitutive promoter is a CMV promoter.
52. A proviral plasmid comprising the nucleic acid trans-splicing molecule of any one of claims 1-46.
53. An adeno-associated virus (AAV) comprising the nucleic acid trans-splicing molecule of any one of claims 1-46, wherein the AAV optionally comprises a 5' regulatory domain operatively linked 5' to the nucleic acid trans-splicing molecule.
54. The AAV of claim 53, wherein the AAV comprises a 5' regulatory domain operatively linked 5' to the CDS.
55. The AAV of any one of claims 53 or 54, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.
56. The AAV of any one of claims 53-55, wherein the 5' regulatory domain comprises a constitutive promoter.
57. The AAV of claim 56, wherein the constitutive promoter is a CMV promoter.
58. The AAV of any one of claims 53-57, wherein the AAV preferentially targets a photoreceptor cell and/or a retinal pigment epithelial cell.
59. The AAV of any one of claims 53-58, wherein the AAV is AAV8, AAV5, or AAV2.
60. A composition comprising the nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, or the AAV of any one of claims 53-59.
61. The composition of claim 60, comprising a pharmaceutically acceptable excipient.
62. A method of expressing biologically active ABCA4 in a target cell to restore functional levels of ABCA4 protein in the target cell, the method comprising transducing the target cell with the nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61.

63. The method of claim 62, wherein at least 5%, at least 10%, or at least 15% of the target ABCA4 mRNA in the target cell is replaced.
64. The method of claim 63, wherein at least 20% of the target ABCA4 mRNA in the target cell is replaced.
65. The method of claim 64, wherein at least 35% of the target ABCA4 mRNA in the target cell is replaced.
66. The method of any one of claims 62-65, wherein functional levels of ABCA4 are restored in the target cell by expressing full-length ABCA4 protein having biological function.
67. A method of reducing ABCA4-associated lipofuscin and/or A2E accumulation in a subject, the method comprising transfecting or transducing a target retinal cell in the subject with the nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61.
68. A method of correcting at least one mutation in an ABCA4 exon sequence in a target cell of a subject, the method comprising administering to the subject the nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61.
69. A method of treating an ABCA4-associated retinal dystrophy in a subject in need thereof, the method comprising administering to the subject the nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61 in a therapeutically effective amount.
70. The method of any one of claims 67-69, the method comprising subretinal administration of the nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61 to the subject.
71. The method of any one of claims 67-70, wherein the subject is a non-human primate or a human.
72. The method of any one of claims 67-71, wherein the subject has been diagnosed with an ABCA4-associated retinal dystrophy.

73. The nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61 for use in treating an ABCA4-associated retinal dystrophy in a subject in need thereof.

74. The nucleic acid trans-splicing molecule of any one of claims 1-46, the vector of any one of claims 47-51, the proviral plasmid of claim 52, the AAV of any one of claims 53-59, or the composition of claim 60 or 61 for use in the preparation of a medicament for the treatment of an ABCA4-associated retinal dystrophy in a subject in need thereof.

75. The method of any one of claims 67-72 or the use of any one of claims 73 or 74, wherein the ABCA4-associated retinal dystrophy is associated with allelic pathogenic mutations spanning Exons 1-22 of the ABCA4 gene and a phenotype consistent with cone-rod dystrophy, Stargardt macular dystrophy, or fundus flavimaculatus.

76. A nucleic acid trans-splicing molecule, comprising:

(a) a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27; and

(b) optionally, a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62; and

(c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63, wherein the linker domain; the potentiator, when present, and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

77. The nucleic acid trans-splicing molecule of claim 76, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA,

wherein the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain, are operatively linked in a 5'-to-3' direction.

78. The nucleic acid trans-splicing molecule of claim 76 or claim 77, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the linker domain; the potentiator, when present; the translational potentiator, when present; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

79. A 5' nucleic acid trans-splicing molecule, comprising:

(a) a potentiator comprising, consisting essentially of, or consisting of at least one of GGTAAGT or SEQ ID NO: 61 or a sequence having at least 90% identity, at least 91% identity, at least 92% identity, at least 93% identity, at least 94% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to GGTAAGT or SEQ ID NO: 61; and

(b) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27;

(c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63,

wherein the linker domain, when present; the potentiator; and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

80. A 5' nucleic acid trans-splicing molecule, comprising:

(a) a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62 or a sequence having at least 90% identity to SEQ ID NO: 62; and

(b) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27;

(c) optionally, a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63,

wherein the linker domain, when present; the potentiator; and the translational potentiator, when present, are operatively linked in a 5'-to-3' direction.

81. The nucleic acid trans-splicing molecule of claim 79 or claim 80, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA,

wherein the linker domain, when present; the potentiator; the translational potentiator, when present; and the binding domain, are operatively linked in a 5'-to-3' direction.

82. The nucleic acid trans-splicing molecule of any one of claims 79-81, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the linker domain, when present; the potentiator; the translational potentiator, when present; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

83. A nucleic acid trans-splicing molecule, comprising:

(a) a translational potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 63 or a sequence having at least 90% identity to SEQ ID NO: 63;

(b) optionally, a potentiator comprising, consisting essentially of, or consisting of SEQ ID NO: 62; and

(c) optionally, a linker domain comprising, consisting essentially of, or consisting of SEQ ID NO: 27 or a sequence having at least 90% identity to SEQ ID NO: 27, wherein the linker domain, when present; the potentiator, when present; and the translational potentiator, are operatively linked in a 5'-to-3' direction.

84. The nucleic acid trans-splicing molecule of claim 83, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA,

wherein the linker domain, when present; the potentiator, when present; the translational potentiator; and the binding domain, are operatively linked in a 5'-to-3' direction.

85. The nucleic acid trans-splicing molecule of claim 83 or claim 84, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the linker domain, when present; the potentiator, when present; the translational potentiator; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

86. A nucleic acid trans-splicing molecule, comprising:

a cassette comprising a linker domain and a potentiator, wherein the cassette comprises, consists essentially of, or consists of SEQ ID NO: 72.

87. The nucleic acid trans-splicing molecule of claim 86, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA,

wherein the cassette and the binding domain, are operatively linked in a 5'-to-3' direction.

88. The nucleic acid trans-splicing molecule of claim 86 or claim 87, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the cassette; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

89. A nucleic acid trans-splicing molecule, comprising:

a cassette comprising a linker domain, a potentiator, and a translational potentiator, wherein the cassette comprises, consists essentially of, or consists of SEQ ID NO: 73.

90. The nucleic acid trans-splicing molecule of claim 89, further comprising a binding domain that anneals to a binding site within endogenous ABCA4 pre-mRNA,

wherein the cassette and the binding domain, are operatively linked in a 5'-to-3' direction.

91. The nucleic acid trans-splicing molecule of claim 89 or claim 90, further comprising a CDS comprising ABCA4 exons,

wherein the CDS; the cassette; and the binding domain, when present, are operatively linked in a 5'-to-3' direction.

92. A nucleic acid trans-splicing molecule, comprising:

(a) a 3' transcription terminator domain comprising, consisting essentially of, or consisting of SEQ ID NO: 66 or a sequence having at least 95% identity to SEQ ID NO: 66.

93. The nucleic acid trans-splicing molecule of claim 92, wherein the 3' transcriptional terminator domain comprises a nucleic acid sequence having at least 96% identity with SEQ ID NO: 66; at least 97% identity with SEQ ID NO: 66; at least 98% identity with SEQ ID NO: 66; at least 99% identity with SEQ ID NO: 66; or 100% identical to SEQ ID NO: 66.

94. The nucleic acid trans-splicing molecule of claim 92 or claim 93, wherein the 3' transcriptional terminator domain comprises, consists essentially of, or consists of SEQ ID NO: 66.