HIGH EFFICIENCY TRANS-SPLICING FOR REPLACEMENT OF TARGETED RNA SEQUENCES IN HUMAN CELLS

FIELD OF THE DISCLOSURE

[1] The present disclosure is directed, generally, to gene therapy, molecular biology, and compositions and methods for altering the sequence composition of RNA molecules.

BACKGROUND

[2] There is a long-felt but unmet need in the art for replacing disease-causing RNA sequences. The disclosure provides compositions and methods for replacement of specific RNA sequences in RNA molecules in human cells with high efficiency. Specifically, the present disclosure provides compositions and methods for replacement of chosen RNA sequences within target RNAs using RNA trans-splicing molecules that carry sequences that increase trans-splicing efficiency to treat a disease in the context of a human gene therapy

[3] Effective treatment of human genetic disease necessitates efficient replacement of defective genetic sequences in human cells. RNA trans-splicing has been proposed as a human gene therapeutic but has not experienced success in clinical trials due to low efficiency. The present disclosure describes improvements to RNA trans-splicing molecules that could address this long-felt but unmet need.

SUMMARY

[4] The present disclosure provides, in some examples, a composition comprising a trans-splicing RNA molecule comprising (a) at least one domain that promotes trans-splicing ("Intronic

Domain") that comprises one or more trans-splicing enhancer sequences, (b) at least one binding domain ("Antisense Domain") that contains or consists of a sequence complementary to a pre-mRNA present in a human cells ("Target RNA"), and (c) a coding domain that is inserted into the Target RNA via trans-splicing ("Replacement Domain"). The trans-splicing enhancer sequences increase splicing efficiency so that the trans-splicing RNA molecule can exchange sequences within the Target RNA with the Replacement Domain with high efficiency. In other examples, the present disclosure provides a composition comprising a nucleic acid sequence encoding the trans-splicing RNA molecule.

[5] In some examples, the trans-splicing enhancer sequences comprise 5'-$X_1X_2X_3X_4X_5X_6$-3' wherein X_1 is uracil (U) or guanine (G); X_2 is adenine (A), uracil (U) or guanine (G); X_3 is adenine (A), uracil (U) and guanine (G); X_4 is adenine (A), uracil (U), cytosine (C) or guanine (G); X_5 is adenine (A), cytosine (C), uracil (U) or guanine (G); and X_6 is adenine (A), uracil (U) or guanine (G).

[6] In some examples, the trans-splicing enhancer sequences comprise $X_1X_2X_3X_4X_5X_6$ wherein; X_1 is selected from the group including adenine (A), uracil (U) and guanine (G); X_2 is selected from the group including adenine (A), uracil (U) and guanine (G); X_3 is selected from the group including adenine (A), uracil (U) and guanine (G); X_4 is selected from the group including adenine (A), uracil (U) and guanine (G); X_5 is selected from the group including adenine (A), uracil (U) and guanine (G); and X_6 is selected from the group including uracil (U) and guanine (G).

[7] In some examples, the trans-splicing enhancer sequences comprise $X_1X_2X_3X_4X_5X_6$ wherein; X_1 is selected from the group including adenine (A), uracil (U) and guanine (G); X_2 is selected from the group including uracil (U) and guanine (G); X_3 is selected from the group

including adenine (A), uracil (U) and guanine (G); X_4 is selected from the group including uracil (U) and guanine (G); X_5 is selected from the group including uracil (U) and guanine (G); and X_6 is selected from the group including uracil (U) and guanine (G).

[8] In some examples, the trans-splicing enhancer sequences are directly adjacent to the Replacement domain.

[9] In some examples, the trans-splicing enhancer sequences are 1 nucleotide, 2 nucleotides, 3 nucleotides, 4 nucleotides, 5 nucleotides, 6 nucleotides, 7 nucleotides, 8 nucleotides, 9 nucleotides, 10 nucleotides, 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, 30 nucleotides, 31 nucleotides, 31 nucleotides, 32 nucleotides, 33 nucleotides, 34 nucleotides, 35 nucleotides, 36 nucleotides, 37 nucleotides, 38 nucleotides, 39 nucleotides, 40 nucleotides, 41 nucleotides, 42 nucleotides, 43 nucleotides, 44 nucleotides, 45 nucleotides, 46 nucleotides, 47 nucleotides, 48 nucleotides, 49 nucleotides, 50 nucleotides, 55 nucleotides, 60 nucleotides, 65 nucleotides, 70 nucleotides, 75 nucleotides, 80 nucleotides, 85 nucleotides, 90 nucleotides, 95 nucleotides, 100 nucleotides, 110 nucleotides, 120 nucleotides, 130 nucleotides, 140 nucleotides, 150 nucleotides, 160 nucleotides, 170 nucleotides, 180 nucleotides, 190 nucleotides, 200 nucleotides, 250 nucleotides, 300 nucleotides, 400 nucleotides, 500 nucleotides, more than 500 nucleotides, or any number of nucleotides in between distant from the first nucleotide of the Replacement Domain in the 5’ direction.

[10] In some examples, the trans-splicing enhancer sequences are 1 nucleotide, 2 nucleotides, 3 nucleotides, 4 nucleotides, 5 nucleotides, 6 nucleotides, 7 nucleotides, 8

nucleotides, 9 nucleotides, 10 nucleotides, 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, 30 nucleotides, 31 nucleotides, 31 nucleotides, 32 nucleotides, 33 nucleotides, 34 nucleotides, 35 nucleotides, 36 nucleotides, 37 nucleotides, 38 nucleotides, 39 nucleotides, 40 nucleotides, 41 nucleotides, 42 nucleotides, 43 nucleotides, 44 nucleotides, 45 nucleotides, 46 nucleotides, 47 nucleotides, 48 nucleotides, 49 nucleotides, 50 nucleotides, 55 nucleotides, 60 nucleotides, 65 nucleotides, 70 nucleotides, 75 nucleotides, 80 nucleotides, 85 nucleotides, 90 nucleotides, 95 nucleotides, 100 nucleotides, 110 nucleotides, 120 nucleotides, 130 nucleotides, 140 nucleotides, 150 nucleotides, 160 nucleotides, 170 nucleotides, 180 nucleotides, 190 nucleotides, 200 nucleotides, 250 nucleotides, 300 nucleotides, 400 nucleotides, 500 nucleotides, more than 500 nucleotides, or any number of nucleotides in between distant from the last nucleotide of the Replacement Domain in the 3' direction.

[11] In some examples, the Intronic Domain comprises 1 trans-splicing enhancer sequence. In some examples, the Intronic Domain comprises 2 or more trans-splicing enhancer sequences. In some examples, the Intronic Domain comprises 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 30, 40, 50, 75, 100, 200, 300 or more trans-splicing enhancer sequences.

[12] In some examples, the Antisense Domain is complementary to a gene (corresponding accession numbers in brackets, corresponding disease in parentheses) that encodes a Target RNA that sometimes carry disease-causing mutations and is selected from the group consisting of: TNFRSF13B [ENSG00000240505] (common variable immune deficiency); ADA, CECR1 [ENSG00000196839, ENSG00000093072] (Adenosine deaminase deficiency); IL2RG

ENSG00000131504, ENSG00000109501] (autosomal dominant hearing impairment); POU3F3 [ENSG00000198914] (nonsyndromic hearing loss).

[13] In some examples, the Replacement domain is derived or isolated from the Target RNA.

[14] In some examples, the Intronic Domains carry binding sites that are preferentially-targeted by RNA-binding proteins with disease-causing mutations. In some examples, the dissociation constant of these mutated RNA-binding proteins and the Intronic Domain is lower than the dissociation constant of the non-mutated RNA-binding protein and the Intronic Domain.

[15] In some examples, the trans-splicing RNA further comprises a 5’ untranslated region. In some examples, the 5’ untranslated region increases the stability of the trans-splicing nucleic acid. In some examples, the 5’ untranslated region reduces the stability of the trans-splicing nucleic acid. In some examples, the 5’ untranslated region alters the localization of the trans-splicing nucleic acid. In some examples, the 5’ untranslated region alters the processing of the trans-splicing nucleic acid.

[16] In some examples, the trans-splicing RNA further comprises a 3' untranslated region. In some examples, the 3' untranslated region increases the stability of the trans-splicing nucleic acid. In some examples, the 3' untranslated region reduces the stability of the trans-splicing nucleic acid. In some examples, the 3' untranslated region alters the localization of the trans-splicing nucleic acid. In some examples, the 3' untranslated region alters the processing of the trans-splicing nucleic acid.

[17] In some examples, the Replacement Domain is comprised of sequence derived or isolated from a human gene. In some examples of the compositions of the disclosure, the sequence comprising the Replacement Domain has at least 50%, 55%, 60%, 65%, 70%, 75%,

80%, 87%, 90%, 95%, 97%, 99% or any percentage in between of identity with a human gene. In some examples, the Replacement Domain has 100% identity with a sequence derived or isolated from a human gene. In some examples, the Replacement Domain comprises or consists of 2 nucleotides, 5 nucleotides, 10 nucleotides, 20 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 60 nucleotides, 70 nucleotides, 80 nucleotides, 90 nucleotides, 100 nucleotides, 110 nucleotides, 120 nucleotides, 130 nucleotides, 140 nucleotides, 150 nucleotides, 160 nucleotides, 170 nucleotides, 180 nucleotides, 190 nucleotides, 200 nucleotides, 210 nucleotides, 220 nucleotides, 230 nucleotides, 240 nucleotides, 250 nucleotides, 260 nucleotides, 270 nucleotides, more than 270 nucleotides, or any number of nucleotides in between.

[18] In some examples of the compositions of the disclosure, the sequence comprising the Antisense Domain has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or any percentage in between of complementarity to the Target RNA sequence. In some examples, the Antisense Domain has 100% complementarity to the Target RNA sequence. In some examples, the Antisense Domain comprises or consists of 20 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 60 nucleotides, 70 nucleotides, 80 nucleotides, 90 nucleotides, 100 nucleotides, 110 nucleotides, 120 nucleotides, 130 nucleotides, 140 nucleotides, 150 nucleotides, 160 nucleotides, 170 nucleotides, 180 nucleotides, 190 nucleotides, 200 nucleotides, 210 nucleotides, 220 nucleotides, 230 nucleotides, 240 nucleotides, 250 nucleotides, 260 nucleotides, 270 nucleotides, more than 270 nucleotides, or any number of nucleotides in between the complementary to the Target RNA sequence.

[19] In some examples of the compositions of the disclosure, the sequence encoding the trans-splicing RNA further comprises a sequence encoding a promoter capable of expressing the trans-splicing RNA in a eukaryotic cell.

[20] In some examples of the compositions of the disclosure, the eukaryotic cell is an animal cell. In some examples, the animal cell is a mammalian cell. In some examples, the animal cell is a human cell.

[21] In some examples of the compositions and methods of the disclosure, a vector of the disclosure is a viral vector. In some examples, the viral vector comprises a sequence isolated or derived from a retrovirus. In some examples, the viral vector comprises a sequence isolated or derived from a lentivirus. In some examples, the viral vector comprises a sequence isolated or derived from an adenovirus. In some examples, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some examples, the viral vector is replication incompetent. In some examples, the viral vector is isolated or recombinant. In some examples, the viral vector is self-complementary.

[22] In some examples of the compositions and methods of the disclosure, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some examples, the viral vector comprises an inverted terminal repeat sequence or a capsid sequence that is isolated or derived from an AAV of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 or AAV12. In some examples, the viral vector is replication incompetent. In some examples, the viral vector is isolated or recombinant (rAAV). In some examples, the viral vector is self-complementary (scAAV).

[23] In some examples of the compositions and methods of the disclosure, a vector of the disclosure is a non-viral vector. In some examples, the vector comprises or consists of a nanoparticle, a micelle, a liposome or lipoplex, a polymersome, a polyplex, an exosome or a dendrimer. In some examples, the vector is an expression vector or recombinant expression

system. As used herein, the term “recombinant expression system” refers to a genetic construct for the expression of certain genetic material formed by recombination.

[24] In some examples of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, an expression control element. An “expression control element” as used herein refers to any sequence that regulates the expression of a coding sequence, such as a gene. Exemplary expression control elements include but are not limited to promoters, enhancers, microRNAs, post-transcriptional regulatory elements, polyadenylation signal sequences, 5’ or 3’ untranslated regions, and introns.

[25] Expression control elements may be constitutive, inducible, repressible, or tissue-specific, for example. A “promoter” is a control sequence that is a region of a polynucleotide sequence at which initiation and rate of transcription are controlled. It may contain genetic elements at which regulatory proteins and molecules may bind such as RNA polymerase and other transcription factors. In some examples, expression control by a promoter is tissue-specific. Non-limiting exemplary promoters include CMV, CBA, CAG, Cbh, EF-1a, PGK, UBC, GUSB, UCOE, hAAT, TBG, Desmin, MCK, C5-12, NSE, Synapsin, PDGF, MecP2, CaMKII, mGluR2, NFL, NFH, nβ2, PPE, ENK, EAAT2, GFAP, MBP, H1 and U6 promoters. In some examples, the promoter is a sequence isolated or derived from a promoter capable of driving expression of a transfer RNA (tRNA). In some examples, the promoter is isolated or derived from an alanine tRNA promoter, an arginine tRNA promoter, an asparagine tRNA promoter, an aspartic acid tRNA promoter, a cysteine tRNA promoter, a glutamine tRNA promoter, a glutamic acid tRNA promoter, a glycine tRNA promoter, a histidine tRNA promoter, an isoleucine tRNA promoter, a leucine tRNA promoter, a lysine tRNA promoter, a methionine tRNA promoter, a phenylalanine tRNA promoter, a proline tRNA promoter, a serine tRNA promoter, a threonine tRNA promoter,

a tryptophan tRNA promoter, a tyrosine tRNA promoter, or a valine tRNA promoter. In some examples, the promoter is isolated or derived from a valine tRNA promoter.

[26] An "enhancer" is a region of DNA that can be bound by activating proteins to increase the likelihood or frequency of transcription. Non-limiting exemplary enhancers and post-transcriptional regulatory elements include the CMV enhancer and WPRE.

[27] In some examples of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, vector elements such as an IRES or 2A peptide sites for configuration of "multicistronic" or "polycistronic" or "bicistronic" or tricistronic" constructs, i.e., having double or triple or multiple coding areas or exons, and as such will have the capability to express from mRNA two or more proteins from a single construct. Multicistronic vectors simultaneously express two or more separate proteins from the same mRNA. The two strategies most widely used for constructing multicistronic configurations are through the use of an IRES or a 2A self-cleaving site. An "IRES" refers to an internal ribosome entry site or portion thereof of viral, prokaryotic, or eukaryotic origin which are used within polycistronic vector constructs. In some examples, an IRES is an RNA element that allows for translation initiation in a cap-independent manner. The term "self-cleaving peptides" or "sequences encoding self-cleaving peptides" or "2A self-cleaving site" refer to linking sequences which are used within vector constructs to incorporate sites to promote ribosomal skipping and thus to generate two polypeptides from a single promoter, such self-cleaving peptides include without limitation, T2A, and P2A peptides or sequences encoding the self-cleaving peptides.

[28] In some examples, the vector is a viral vector. In some examples, the vector is an adenoviral vector, an adeno-associated viral (AAV) vector, or a lentiviral vector. In some

examples, the vector is a retroviral vector, an adenoviral/retroviral chimera vector, a herpes simplex viral I or II vector, a parvoviral vector, a reticuloendotheliosis viral vector, a polioviral vector, a papillomaviral vector, a vaccinia viral vector, or any hybrid or chimeric vector incorporating favorable aspects of two or more viral vectors. In some examples, the vector further comprises one or more expression control elements operably linked to the polynucleotide. In some examples, the vector further comprises one or more selectable markers. In some examples, the AAV vector has low toxicity. In some examples, the AAV vector does not incorporate into the host genome, thereby having a low probability of causing insertional mutagenesis. In some examples, the AAV vector can encode a range of total polynucleotides from .3 kb to 4.75 kb. In some examples, exemplary AAV vectors that may be used in any of the herein described compositions, systems, methods, and kits can include an AAV1 vector, a modified AAV1 vector, an AAV2 vector, a modified AAV2 vector, an AAV3 vector, a modified AAV3 vector, an AAV4 vector, a modified AAV4 vector, an AAV5 vector, a modified AAV5 vector, an AAV6 vector, a modified AAV6 vector, an AAV7 vector, a modified AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV.rh10 vector, a modified AAV.rh10 vector, an AAV.rh32/33 vector, a modified AAV.rh32/33 vector, an AAV.rh43 vector, a modified AAV.rh43 vector, an AAV.rh74 vector, a modified AAV.rh74 vector, an AAV.rh64R1 vector, and a modified AAV.rh64R1 vector and any combinations or equivalents thereof. In some examples, the lentiviral vector is an integrase-competent lentiviral vector (ICLV). In some examples, the lentiviral vector can refer to the transgene plasmid vector as well as the transgene plasmid vector in conjunction with related plasmids (e.g., a packaging plasmid, a rev expressing plasmid, an envelope plasmid) as well as a lentiviral-based particle capable of introducing exogenous nucleic acid into a cell through a viral or viral-like entry mechanism.

Lentiviral vectors are well-known in the art (see, e.g., Trono D. (2002) Lentiviral vectors, New York: Spring-Verlag Berlin Heidelberg and Durand et al. (2011) Viruses 3(2):132-159 doi: 10.3390/v3020132). In some examples, exemplary lentiviral vectors that may be used in any of the herein described compositions, systems, methods, and kits can include a human immunodeficiency virus (HIV) 1 vector, a modified human immunodeficiency virus (HIV) 1 vector, a human immunodeficiency virus (HIV) 2 vector, a modified human immunodeficiency virus (HIV) 2 vector, a sooty mangabey simian immunodeficiency virus (SIVSM) vector, a modified sooty mangabey simian immunodeficiency virus (SIVSM) vector, a African green monkey simian immunodeficiency virus (SIVAGM) vector, a modified African green monkey simian immunodeficiency virus (SIVAGM) vector, an equine infectious anemia virus (EIAV) vector, a modified equine infectious anemia virus (EIAV) vector, a feline immunodeficiency virus (FIV) vector, a modified feline immunodeficiency virus (FIV) vector, a Visna/maedi virus (VNV/VMV) vector, a modified Visna/maedi virus (VNV/VMV) vector, a caprine arthritis-encephalitis virus (CAEV) vector, a modified caprine arthritis-encephalitis virus (CAEV) vector, a bovine immunodeficiency virus (BIV), or a modified bovine immunodeficiency virus (BIV).

[29] In some examples, the trans-splicing nucleic acid is RNA, DNA, a DNA/RNA hybrid, and/or comprises at least one of a nucleic acid analog, a chemically-modified nucleic acid, or a chimera composed of two or more nucleic acids or nucleic acid analogs. As used herein, the term "nucleic acid analog" refers to a compound having structural similarity to a canonical purine or pyrimidine base occurring in DNA or RNA. The nucleic acid analog may contain a modified sugar and/or a modified nucleobase, as compared to a purine or pyrimidine base occurring naturally in DNA or RNA. In some examples, the nucleic acid analog is a 2'-deoxyribonucleoside, 2'-ribonucleoside, 2'-deoxyribonucleotide or a 2'-ribonucleotide, wherein

the nucleobase includes a modified base (such as, for example, xanthine, uridine, oxanine (oxanosine), 7-methlguanosine, dihydrouridine, 5-methylcytidine, C3 spacer, 5-methyl dC, 5-hydroxybutynl-2'-deoxyuridine, 5-nitroindole, 5-methyl iso-deoxycytosine, iso deoxyguanosine, deoxyuradine, iso deoxycytidine, other 0-1 purine analogs, N-6-hydroxylaminopurine, nebularine, 7-deaza hypoxanthine, other 7-deazapurines, and 2-methyl purines). In some examples, the nucleic acid analog may be selected from the group consisting of inosine, 7-deaza-2'-deoxyinosine, 2'-aza-2'-deoxyinosine, PNA-inosine, morpholino-inosine, LNA-inosine, phosphoramidate-inosine, 2'-O-methoxyethyl-inosine, and 2'-OMe-inosine. In other examples the nucleic acid analog is a nucleic acid mimic (such as, for example, artificial nucleic acids and xeno nucleic acids (XNA).

[29a] In one aspect, the present disclosure provides a composition comprising a trans-splicing nucleic acid, comprising: (a) one or more replacement domains that encode a therapeutic sequence operably linked to; (b) one or more intronic domains that promote RNA splicing of the one or more replacement domains, wherein the one or more intronic domains each comprises a trans-splicing enhancer sequence; and (c) one or more antisense domains that bind to a target RNA molecule, wherein the trans-splicing enhancer sequence consists of a chain of RNA nucleobases comprising at least one RNA motif having the formula $X_1X_2X_3X_4X_5X_6$, wherein: X_1 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_2 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_3 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_4 is selected from the group consisting of adenine (A), uracil (U), and cytosine (C); X_5 is selected from the group consisting of adenine (A), cytosine (C), uracil (U) and guanine (G); and X_6 is selected from the group

consisting of adenine (A), uracil (U) and guanine (G), wherein each of the one or more intronic domains comprises 2 or more trans-splicing enhancer sequences.

[29b] In another aspect, the present disclosure provides a composition comprising a trans-splicing nucleic acid, comprising: (a) one or more replacement domains that encode a therapeutic sequence operably linked to; (b) one or more intronic domains that promote RNA splicing of the one or more replacement domains, wherein the one or more intronic domains each comprise a trans-splicing enhancer sequence, and (c) one or more antisense domains that bind to a target RNA molecule, , wherein the trans-splicing enhancer sequence consists of a chain of RNA nucleobases comprising at least one RNA motif having the formula $X_1X_2X_3X_4X_5X_6$, wherein: X_1 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_2 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_3 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_4 is selected from the group consisting of adenine (A), uracil (U), cytosine (C) and guanine (G); X_5 is selected from the group consisting of adenine (A), cytosine (C), uracil (U) and guanine (G); and X_6 is selected from the group consisting of adenine (A), uracil (U) and guanine (G), wherein each of the one or more intronic domains comprises 6 or more trans-splicing enhancer sequences.

[29c] In yet another aspect, the present disclosure provides a vector comprising or encoding the composition as defined herein.

[29d] In yet another aspect, the present disclosure provides a cell comprising the vector as defined herein.

[29e] In yet another aspect, the present disclosure provides use of the composition as defined herein or the vector as defined herein in the manufacture of a medicament for treating a genetic disease in a subject in need thereof.

[29f] In yet another aspect, the present disclosure provides use of the composition as defined herein or the vector as defined herein in the manufacture of a medicament for correcting a genetic defect in a subject in need thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[30] FIGURE 1 illustrates the unmet need addressed by the present invention and provides a schematic of the invention. FIGURE 1A illustrates the concept of human genetic disease where mutated ("defective") DNA sequences are transcribed into RNA which directly contribute to disease ("RNA pathogenicity") or are translated into disease-causing protein ("translation of pathogenic protein"). FIGURE 1B illustrates the state-of-the-art RNA trans-splicing technology where a mutation-carrying RNA molecule is targeted by a trans-splicing RNA that corrects the mutation but with low efficiency. This low efficiency is typically insufficient to halt or reverse progression of disease. FIGURE 1C illustrates the present invention where the trans-splicing molecule carries trans-splicing enhancer sequences that increase the efficiency of the trans-splicing reaction. This supports halting or reversal of disease progression and/or elimination of key disease phenotypes thereby providing an effective therapeutic for human genetic disease.

CLAIMS

1. A composition comprising a trans-splicing nucleic acid, comprising: (a) one or more replacement domains that encode a therapeutic sequence operably linked to; (b) one or more intronic domains that promote RNA splicing of the one or more replacement domains, wherein the one or more intronic domains each comprises a trans-splicing enhancer sequence; and (c) one or more antisense domains that bind to a target RNA molecule, wherein the trans-splicing enhancer sequence consists of a chain of RNA nucleobases comprising at least one RNA motif having the formula $X_1X_2X_3X_4X_5X_6$, wherein: X_1 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_2 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_3 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_4 is selected from the group consisting of adenine (A), uracil (U), and cytosine (C); X_5 is selected from the group consisting of adenine (A), cytosine (C), uracil (U) and guanine (G); and X_6 is selected from the group consisting of adenine (A), uracil (U) and guanine (G), wherein each of the one or more intronic domains comprises 2 or more trans-splicing enhancer sequences.

2. A composition comprising a trans-splicing nucleic acid, comprising: (a) one or more replacement domains that encode a therapeutic sequence operably linked to; (b) one or more intronic domains that promote RNA splicing of the one or more replacement domains, wherein the one or more intronic domains each comprise a trans-splicing enhancer sequence, and (c) one or more antisense domains that bind to a target RNA molecule, , wherein the trans-splicing enhancer sequence consists of a chain of RNA nucleobases comprising at least one RNA motif having the formula $X_1X_2X_3X_4X_5X_6$, wherein: X_1 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_2 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_3 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_4 is selected from the group consisting of adenine (A), uracil (U), cytosine (C) and guanine (G); X_5 is selected from the group consisting of adenine (A), cytosine (C), uracil (U) and guanine (G); and X_6 is selected from the group consisting of adenine (A), uracil (U) and guanine (G), wherein each of the one or more intronic domains comprises 6 or more trans-splicing enhancer sequences.

3. The composition of claim 1, wherein the trans-splicing enhancer sequence consists of a chain of RNA nucleobases comprising at least one RNA motif having the formula

$X_1X_2X_3X_4X_5X_6$, wherein: X_1 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_2 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_3 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_4 is selected from the group consisting of adenine (A) and uracil (U); X_5 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); and X_6 is selected from the group consisting of uracil (U) and guanine (G).

4. The composition of claim 1, wherein the trans-splicing enhancer sequence consists of a chain of RNA nucleobases comprising at least one RNA motif having the formula $X_1X_2X_3X_4X_5X_6$, wherein: X_1 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_2 is selected from the group consisting of uracil (U) and guanine (G); X_3 is selected from the group consisting of adenine (A), uracil (U) and guanine (G); X_4 is uracil (U); X_5 is selected from the group consisting of uracil (U) and guanine (G); and X_6 is selected from the group consisting of uracil (U) and guanine (G).

5. The composition of any one of claims 1, 3, and 4, wherein each of the one or more intronic domains comprises 6 or more trans-splicing enhancer sequences.

6. The composition of any one of claims 1-5, wherein the trans-splicing enhancer sequence is less than 300 bases from a 3’ splice site of the trans-splicing nucleic acid.

7. The composition of any one of claims 1-6, wherein the trans-splicing enhancer sequence is less than 300 bases from a 5’ splice site of the trans-splicing nucleic acid.

8. The composition of any one of claims 1-7, wherein each of the one or more intronic domains comprises 10 or more trans-splicing enhancer sequences.

9. The composition of any one of claims 1-8, wherein the trans-splicing nucleic acid further comprises a 3’ untranslated region that increases trans-splicing efficiency.

10. The composition of any one of claims 1-9, wherein the trans-splicing nucleic acid further comprises a 5’ untranslated region that increases trans-splicing efficiency.

11. The composition of any one of claims 1-10, wherein the one or more replacement domains each comprises a gene expression-enhancing element.

12. The composition of claim 11, wherein the gene expression-enhancing element comprises a sequence derived or isolated from the group consisting of: Woodchuck Hepatitis Virus (WHV) Posttranscriptional Regulatory Element (WPRE), triplex from MALAT1, the PRE of Hepatitis B virus (HPRE), and an iron response.

13. The composition of any one of claims 1-12, further comprising an RNA-binding protein that strengthens the interaction between the trans-splicing nucleic acid and the target RNA molecule and increases trans-splicing efficiency.

14. The composition of any one of claims 1-13, wherein the trans-splicing nucleic acid is RNA, DNA, a DNA/RNA hybrid, a nucleic acid analog, a chemically-modified nucleic acid, or a chimera composed of two or more nucleic acids or nucleic acid analogs.

15. The composition of any one of claims 1-14, wherein the trans-splicing nucleic acid further comprises a heterologous promoter.

16. A vector comprising or encoding the composition of any one of claims 1-15.

17. The vector of claim 16, wherein the vector is selected from the group consisting of: adeno-associated virus, retrovirus, lentivirus, adenovirus, nanoparticle, micelle, liposome, lipoplex, polymersome, polyplex, and dendrimer.

18. A cell comprising the vector of claim 16 or claim 17.

19. Use of the composition of any one of claims 1-15 or the vector of claim 16 or claim 17 in the manufacture of a medicament for treating a genetic disease in a subject in need thereof.

20. Use of the composition of any one of claims 1-15 or the vector of claim 16 or claim 17 in the manufacture of a medicament for correcting a genetic defect in a subject in need thereof.