

HTT TRANS-SPLICING MOLECULES

CROSS REFERENCE

This application claims the benefit of U.S. Provisional Patent Application No. 63/485,142, filed on February 15, 2023 and U.S. Provisional Patent Application No. 63/485,146, filed on February 15, 2023,
5 each of which is incorporated herein by reference in its entirety.

SEQUENCE LISTING

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 February 14, 2024, is
10 named 61313-709_601_SL.xml and is 521,958 bytes in size.

BACKGROUND

Huntington's Disease (HD) is a movement-associated progressive neurodegenerative disorder. The disease is associated with loss of pyramidal neurons in the cortex, loss of medium spiny
15 neurons in the striatum, and loss of hypothalamic neurons. The genetic cause of HD is the autosomal dominant inheritance of an expanded CAG trinucleotide repeat in exon 1 of the *HTT* gene, wherein the presence of over 40 repeats of CAG in this region is disease-causing.

The *HTT* locus is large, spanning 180 kb and consisting of 67 exons and expression of the *HTT*
20 gene is required for normal development. Although HTT protein is widely expressed, the brain is most severely impacted by pathological expansion of CAG trinucleotide repeats, with early effects noted in the striatum and motor cortex. It is generally believed that the underlying mechanism of HD pathogenesis is the somatic CAG repeat expansion in *HTT* that occurs in affected brain regions (e.g., striatum) of HD patients.

25 Patients with HD have progressive neurodegeneration leading to death, typically 10 to 20 years after disease onset. There are currently no disease-modifying treatments for HD. Current treatments are limited to providing symptomatic relief.

30 SUMMARY

Nucleic acid trans-splicing molecules which include a coding domain comprising one or more *HTT* exons, a splice site, and a binding domain that binds a target intron of a *HTT* pre-mRNA are described herein. In some embodiments, constructs comprising nucleic acid trans-splicing molecules may further include additional sequences that encode anti-sense RNA that binds to
35 target pre-mRNA and in so doing blocks cis-splicing, thereby promoting trans-splicing of the nucleic acid trans-splicing molecule to the pre-mRNA target (e.g., *HTT* pre-mRNA). In some embodiments, nucleic acid trans-splicing molecules may further comprise a second binding

domain that binds to a target intron in a second target pre-mRNA (e.g., *MSH3* pre-mRNA), whereby trans-splicing into the target intron of the second target pre-mRNA generates a hybrid pre-mRNA comprising the one or more *HTT* exons and exons of the second target pre-mRNA (e.g., *MSH3* pre-mRNA), wherein the hybrid pre-mRNA is processed to hybrid mRNA that is targeted for degradation by, e.g., nonsense-mediated decay. In some embodiments, constructs comprising nucleic acid trans-splicing molecules may further include additional sequences that encode small nuclear RNA (snRNA) that blocks normal processing of a second target pre-mRNA (e.g., *MSH3* pre-mRNA) via, e.g., introducing a premature stop codon into the second target pre-mRNA (e.g., *MSH3* pre-mRNA) that targets the second target's mRNA (e.g., *MSH3* mRNA) for nonsense-mediated decay. In some embodiments, constructs comprising nucleic acid trans-splicing molecules may further comprise elements encoding anti-sense RNA that blocks normal processing of a second target pre-mRNA (e.g., *MSH3* pre-mRNA) via, e.g., blocking *MSH3* splice junctions or annealing to the 5' UTR or initial coding sequence of *MSH3* to block *MSH3* translation. In some embodiments, constructs comprising nucleic acid trans-splicing molecules may further include additional sequences that encode microRNA (miRNA) specific for the endogenous *HTT* mRNA or specific for *MSH3* mRNA that promote cleavage of the endogenous *HTT* mRNA or the *MSH3* mRNA, respectively, thereby reducing the levels of endogenous *HTT* mRNA or *MSH3* mRNA. Compositions comprising such nucleic acid trans-splicing molecules are also encompassed, as are compositions comprising combinations of nucleic acid trans-splicing molecules with additional therapeutic agents (e.g., anti-sense oligonucleotides or anti-sense RNA encoding constructs). Nucleic acid trans-splicing molecules and compositions comprising same may be used alone or in combination with additional therapeutic agents in methods for treating Huntington's Disease (HD). Nucleic acid trans-splicing molecules and compositions comprising same are also used, alone or in combination with additional therapeutic agents, in treating HD or in the preparation of medicaments for treating HD.

Disclosed herein is an *HTT* nucleic acid trans-splicing molecule comprising: (a) a coding domain comprising *HTT* exon 1 and *HTT* exon 2; (b) a splicing domain; and (c) a binding domain that binds a target intron of an *HTT* pre-mRNA, wherein the target intron comprises intron 2. In some embodiments, the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 60-81, or a sequence having at least 90% identity to any one of SEQ ID NOs: 60-81.

Also disclosed herein is an *HTT* nucleic acid trans-splicing molecule comprising: (a) a coding domain comprising *HTT* exons 1-3; (b) a splicing domain; and (c) a binding domain that binds a target intron of an *HTT* pre-mRNA, wherein the target intron comprises intron 3. In some

embodiments, the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 158-174, or a sequence having at least 90% identity to any one of SEQ ID NOs: 158-174.

5 Also disclosed herein is an HTT nucleic acid trans-splicing molecule comprising: (a) a coding domain comprising HTT exon 1; (b) a splicing domain; and (c) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 1, and wherein the binding domain comprises any one of SEQ ID NOs: 8-21.

10 In some embodiments, in any of the HTT nucleic acid trans-splicing molecules described above, the coding domain comprises, consists essentially of, or consists of HTT exon 1; HTT exon 1 and HTT exon 2; or HTT exons 1-3. In some embodiments, the coding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 3, 59, 157, or 349-353 or a sequence having at least 90% identity to SEQ ID NOs: 3, 59, 157, or 349-353. In some
15 embodiments, the coding domain, the splicing domain, and the binding domain are operatively linked in a 5'-to-3' direction.

In some embodiments, the HTT nucleic acid trans-splicing molecules described above further comprise a linker, wherein the coding domain, splicing domain, linker, and binding domain are
20 operatively linked in a 5'-to-3' direction. In some embodiments, the linker comprises, consists essentially of, or consists of: a sequence ranging from 20 to 50 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine; a sequence ranging from 20 to 45 nucleotides long, wherein the linker comprises 60-80%
25 guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine; or a sequence ranging from 22 to 42 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine. In some embodiments, the linker
30 comprises, consists essentially of, or consists of any one of: SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38; SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39; SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; or SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41. In some embodiments, the linker comprises, consists essentially of, or consists of any one of:
35 SEQ ID NO: 37 or a sequence having at least 90% identity to SEQ ID NO: 37; SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42; SEQ ID NO: 43 or a sequence having

at least 90% identity to SEQ ID NO: 43; SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44; SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45; SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46; SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106; SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107; SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108; SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109; SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110; SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111; SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112; SEQ ID NO: 197 or a sequence having at least 90% identity to SEQ ID NO: 197; or SEQ ID NO: 198 or a sequence having at least 90% identity to SEQ ID NO: 198.

In some embodiments, the HTT nucleic acid trans-splicing molecules described above further comprise a triple helix terminator, wherein the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator are operatively linked in a 5'-to-3' direction. In some embodiments, the triple helix terminator comprises, consists essentially of, or consists of SEQ ID NO: 5 or a sequence having at least 90% identity to SEQ ID NO: 5. In some embodiments, the triple helix terminator comprises, consists essentially of, or consists of SEQ ID NO: 6.

In some embodiments, the HTT nucleic acid trans-splicing molecules described above further comprise a 5' untranslated region (5' UTR), wherein the 5' UTR, the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction. In some embodiments, the 5' UTR is an *HTT* 5' UTR. In some embodiments, the *HTT* 5' UTR comprises, consists essentially of, or consists of any one of SEQ ID NO: 136 or 192 or a sequence having at least 90% identity to any one of SEQ ID NO: 136 or 192.

In some embodiments, the HTT nucleic acid trans-splicing molecules described above further comprise a sequence encoding an epitope tag, wherein the 5' UTR, when present, the epitope tag, the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction. In some embodiments, the sequence encoding the epitope tag comprises, consists essentially of, or consists of a SEQ ID NO: 4.

Also disclosed herein is an HTT nucleic acid trans-splicing molecule comprising: (a) a coding domain comprising HTT exon 1 and HTT exon 2; (b) a splicing domain; (c) a linker; and (d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 2. In some embodiments, the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 60-81, or a sequence having at least 90% identity to any one of SEQ ID NOs: 60-81.

Also disclosed herein is an HTT nucleic acid trans-splicing molecule comprising: (a) a coding domain comprising HTT exons 1-3; (b) a splicing domain; (c) a linker; and (d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 3. In some embodiments, the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 158-174, or a sequence having at least 90% identity to any one of SEQ ID NOs: 158-174.

Also disclosed herein is an HTT nucleic acid trans-splicing molecule comprising: (a) a coding domain comprising HTT exon 1; (b) a splicing domain; (c) a linker; and (d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 1. In some embodiments, the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 8-21, or a sequence having at least 90% identity to any one of SEQ ID NOs: 8-21.

In some embodiments the linker comprises, consists essentially of, or consists of any one of SEQ ID NOs: SEQ ID NO: 37 or a sequence having at least 90% identity to SEQ ID NO: 37; SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38; SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39; SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41; SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42; SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43; SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44; SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45; SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46; SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106; SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107; SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108; SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109; SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110; SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111; SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112; SEQ ID NO: 197 or a sequence having at least 90% identity to SEQ ID NO: 197; or SEQ ID NO: 198 or a sequence having at least 90% identity to SEQ ID NO:

198. In some embodiments, the HTT nucleic acid trans-splicing molecule further comprises a triple helix terminator, wherein the coding domain, the splicing domain, the linker, the binding domain, and the triple helix terminator are operatively linked in a 5'-to-3' direction; and optionally, further comprising a 5' UTR, wherein the 5' UTR, when present, the coding domain, the splicing domain, the linker, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.

Also disclosed herein is a nucleic acid trans-splicing molecule comprising a linker, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 20 to 50 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine. In some embodiments, the linker comprises, consists essentially of, or consists of a sequence ranging from 20 to 45 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine. In some embodiments, the linker comprises, consists essentially of, or consists of a sequence ranging from 22 to 42 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine. In some embodiments, the linker comprises, consists essentially of, or consists of: SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38; SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39; SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; or SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41.

Also disclosed herein is a nucleic acid trans-splicing molecule comprising a linker, wherein the linker comprises, consists essentially of, or consists of: SEQ ID NO: 37 or a sequence having at least 90% identity to SEQ ID NO: 37; SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42; SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43; SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44; SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45; SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46; SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106; SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107; SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108; SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109; SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110; SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111; SEQ ID NO: 112 or a sequence having at least

90% identity to SEQ ID NO: 112; SEQ ID NO: 197 or a sequence having at least 90% identity to SEQ ID NO: 197; or SEQ ID NO: 198 or a sequence having at least 90% identity to SEQ ID NO: 198.

5 In some embodiments of the HTT nucleic acid trans-splicing molecules described above, the HTT nucleic acid trans-splicing molecule further comprises a binding domain that binds a target intron of an *MSH3* pre-mRNA. In some embodiments, the *MSH3* target intron comprises any one of intron 5 or intron 15 of *MSH3*. In some embodiments, the binding domain that binds a target intron of an *MSH3* pre-mRNA comprises, consists essentially of, or consists of any one of
10 SEQ ID NOs: 140, 142, 144, 146, 209, or 210 or a sequence having at least 90% identity to any one of SEQ ID NOs: 140, 142, 144, 146, 209, or 210. In some embodiments, the nucleic acid trans-splicing molecule comprises any one of SEQ ID NOs: 149 – 154 or SEQ ID NOs: 212-223, or a sequence having at least 90% identity to any one of SEQ ID NOs: 149 – 154 or SEQ ID NOs: 212-223.

15 In some embodiments, any of the HTT nucleic acid trans-splicing molecules describe above further comprise a nucleic acid sequence encoding a pri-miRNA that comprises a microRNA (miRNA) sequence specific for exon 1 of endogenous HTT mRNA, wherein exon 1 of the nucleic acid trans-splicing molecule comprises a change in nucleotide sequence that impairs binding of
20 the miRNA to mRNA encoded at least in part by the nucleic acid trans-splicing molecule. In some embodiments, the miRNA sequence comprises any one of SEQ ID NOs: 339 or 342 or a nucleic acid sequence having at least 90% identity to any one of SEQ ID NOs: 339 or 342. In some embodiments, the nucleic acid sequence encoding the pri-miRNA comprises any one of SEQ ID NOs: 341 or 344. In some embodiments, the pri-miRNA comprises a mir-33 scaffold
25 sequence. In some embodiments, the pri-miRNA comprises a mir-30a scaffold sequence, a mir-30a loop sequence, a mir-155 scaffold sequence, a mir-155 loop sequence, a mir-33 scaffold sequence, or a mir-33 loop sequence. In some embodiments, the mir-30a scaffold sequence comprises a 5' scaffold sequence set forth in SEQ ID NO: 227 or a 3' scaffold sequence set forth in SEQ ID NO: 228; wherein the mir-30a loop sequence comprises SEQ ID NO: 229; wherein the
30 mir-155 scaffold sequence comprises a 5' scaffold sequence set forth in SEQ ID NO: 230 or a 3' scaffold sequence set forth in SEQ ID NO: 231; wherein the mir-155 loop sequence comprises SEQ ID NO: 232; wherein the mir-33 scaffold sequence comprises a 5' scaffold sequence set forth in SEQ ID NO: 259 or a 3' scaffold sequence set forth in SEQ ID NO: 260; or wherein the mir-33 loop sequence comprises SEQ ID NO: 261.

35

Also disclosed herein is an MSH3 exon skipping nucleic acid construct comprising, operatively linked: (a) a sequence encoding an antisense RNA that promotes exon skipping of a target exon of MSH3 pre-mRNA, wherein the target exon is any one of MSH3 exons 2-4, 6-8, or 15, wherein the target exon comprises a 5' exon-intron junction and a 3' exon-intron junction sequence; and

5 (b) a sequence encoding a small nuclear RNA (snRNA) sequence. In some embodiments, the MSH3 exon skipping nucleic acid construct further comprises a U1 promoter and a U1 terminator operatively linked to (a) and (b). In some embodiments, the snRNA is a modified snRNA. In some embodiments, the modified snRNA comprises a U7 Sm OPT sequence or a U2 snRNA sequence. In some embodiments, the antisense RNA targets either the 5' exon-intron

10 junction or the 3' exon-intron junction of the target exon. In some embodiments, the antisense RNA comprises, consists essentially of, or consists of any one of SEQ ID NOs: 274, 275, 276, 277, 278, 279, 280, 300, 302, 301, 303, 281, 282, 306, 308, 305, 307, 311, 313, 310, 312, 316, 318, 315, 317, 321, 323, 320, or 322 or a sequence having at least 90% identity to any one of SEQ ID NOs: 274, 275, 276, 277, 278, 279, 280, 300, 302, 301, 303, 281, 282, 306, 308, 305,

15 307, 311, 313, 310, 312, 316, 318, 315, 317, 321, 323, 320, or 322. In some embodiments, an MSH3 exon skipping nucleic acid construct described above comprises any one of SEQ ID NOs: 284, 285, 286, 287, 288, 289, 290, 325, 326, 291, 292, 328, 329, 331, 332, 334, 335, 337, and 338.

20 In some embodiments, the antisense RNA targets both the 5' exon-intron junction and the 3' exon-intron junction. In some embodiments, the antisense RNA comprises a sequence that is at least 80% complementary to the entire sequence of the target exon.

In some embodiments, the antisense RNA further comprises: (a) a sequence that is at least 80% complementary to a 5-nucleotide sequence upstream of the 5' exon-intron junction; and (b) a

25 sequence that is at least 80% complementary to a 5-nucleotide sequence downstream of the 3' exon-intron junction. In some embodiments, the antisense RNA comprises any one of SEQ ID NOs: 299, 304, 309, 314, or 319, or a sequence having at least 90% identity to any one of SEQ ID NOs: 299, 304, 309, 314, or 319. In some embodiments, an MSH3 exon skipping nucleic acid

30 construct described above comprises any one of SEQ ID NOs: 324, 327, 330, 333, or 336.

In some embodiments, the antisense RNA comprises, operatively linked in a 5' to 3' direction: (a) a sequence that targets the 3' exon-intron junction; (b) a linker sequence of at least 15 nucleotides that does not anneal to the target exon; and (c) a sequence that targets the 5' exon-

35 intron junction. In some embodiments, the linker sequence is less than 50% complementary to all sequences of the target exon of the same length as the linker. In some embodiments, the

antisense RNA comprises any one of SEQ ID NOs: 300, 301, 302, 303, 305, 306, 307, 308, 310, 311, 312, 313, 315, 316, 317, 318, 320, 321, 322, or 323, or any combination thereof. In some embodiments, the MSH3 exon skipping nucleic acid construct comprises any one of SEQ ID NOs: 325, 326, 328, 329, 331, 332, 334, 335, 337, or 338.

5

In some embodiments, the antisense RNA targets MSH3 exon 7. In some embodiments, the MSH3 exon skipping nucleic acid construct comprises SEQ ID NO: 309. In some embodiments, the MSH3 exon skipping nucleic acid construct comprises from 5' to 3': (a) SEQ ID NO: 310 (In7/Ex7 asRNA), SEQ ID NO: 298 (linker), and SEQ ID NO: 311 (In7/Ex7 asRNA); or (b) SEQ ID NO: 312 (In7/Ex7 asRNA), SEQ ID NO: 298 (linker), and SEQ ID NO: 313 (In7/Ex7 asRNA). In some embodiments, the antisense RNA comprises any one of SEQ ID NOs: 309, 310, 311, 312, or 313, or any combination thereof or a sequence having at least 90% identity to any one of SEQ ID NOs: 309, 310, 311, 312, or 313. In some embodiments, the MSH3 exon skipping nucleic acid construct comprises at least one of SEQ ID NOs: 330-332, or any combination thereof.

15

Also disclosed herein is an MSH3 miRNA nucleic acid construct comprising a sequence encoding a pri-miRNA that comprises a scaffold sequence, a loop sequence, and a miRNA sequence that targets endogenous MSH3 mRNA, wherein: (a) the scaffold sequence is derived from mir-30a, mir-33, or mir-155; (b) the loop sequence is derived from mir-22, mir-30a, mir-33, or mir-155; and (c) the miRNA sequence comprises any one of SEQ ID NOs: 224, 244, 246, 248, 250, 252, 254, 256, or 257 or a sequence having at least 90% identity to any one of SEQ ID NOs: 224, 244, 246, 248, 250, 252, 254, 256, or 257. In some embodiments, the scaffold sequence comprises any one of SEQ ID NOs: 227, 228, 230, 231, 259 or 260. In some embodiments, the loop sequence comprises any one of SEQ ID NOs: 229, 232, or 261. In some embodiments, the pri-miRNA sequence comprises any one of SEQ ID NOs: 234, 235, 238-241, or 262-269. In some embodiments, the sequence encoding the pri-miRNA is operatively linked to a U6 promoter or a CMV promoter.

25

30

Also disclosed herein is an MSH3 nucleic acid trans-splicing molecule comprising: (a) coding domain sequence; (b) a splicing domain; and (c) a binding domain that binds a target intron of an MSH3 pre-mRNA; wherein the coding domain sequence is not an MSH3 coding domain sequence. In some embodiments, the coding domain sequence comprises a sequence that results in a frameshift in a mature MSH3 mRNA when trans-spliced into the MSH3 pre-mRNA. In some embodiments, the coding domain sequence comprises one or more of exons 1, 2, and 3 of HTT. In some embodiments, the target intron of the MSH3 pre-mRNA is intron 5 or intron 15. In some embodiments, the binding domain comprises any one of SEQ ID NOs: 140, 142, 144,

35

146, 209, or 210, or a sequence having at least 90% identity to any one of SEQ ID NOs: 140, 142, 144, 146, 209, or 210.

Also disclosed herein is an HTT trans-splicing and MSH3 exon skipping nucleic acid construct
5 comprising: (a) any of the HTT nucleic acid trans-splicing molecules described above; and (b)
any of the MSH3 exon skipping nucleic acid constructs described above. In some embodiments,
(a) and (b) are comprised on a single vector. In some embodiments, the single vector is an AAV
vector. In some embodiments, the HTT trans-splicing and MSH3 exon skipping nucleic acid
construct comprises any one of SEQ ID NOs: 356, 357, 363, or 364. In some embodiments, the
10 AAV vector is a scAAV or ssAAV vector. In some embodiments, the HTT trans-splicing and
MSH3 exon skipping nucleic acid construct comprises any one of SEQ ID NOs: 369, 370, and
371.

Also disclosed herein is an HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic
15 acid construct comprising: (a) any of the HTT nucleic acid trans-splicing molecules described
above; and (b) any of the MSH3 exon skipping nucleic acid constructs described above. In some
embodiments, (a) and (b) are comprised on a single vector. In some embodiments, the HTT
trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct comprises any one
of SEQ ID NOs: 358 or 359. In some embodiments, the single vector is an AAV vector.

20

Also disclosed herein is an HTT trans-splicing and MSH3 miRNA nucleic acid construct
comprising: (a) any of the HTT nucleic acid trans-splicing molecules described above; and (b)
any of the MSH3 miRNA nucleic acid constructs described above. In some embodiments, (a)
and (b) are comprised on a single vector. In some embodiments, the HTT trans-splicing and
25 MSH3 miRNA nucleic acid construct comprises any one of SEQ ID NOs: 354 or 355. In some
embodiments, the vector is an AAV vector.

Also disclosed is an AAV vector comprising any of the HTT nucleic acid trans-splicing molecules
described above. In some embodiments, the AAV vector comprises any one of SEQ ID NOs:
30 356, 357, 363, or 364.

Also disclosed is an AAV vector comprising any of the HTT nucleic acid trans-splicing molecules
described above or any of the nucleic acid trans-splicing molecules described above.

35 Also disclosed is a ribonucleic acid trans-splicing molecule comprising any one of SEQ ID NOs:
23-36, 47-56, 83-105, 113-125, 175-191, or 199-206.

Also disclosed is a ribonucleic acid trans-splicing molecule transcribed from any of the HTT nucleic acid trans-splicing molecules described above or from any of the nucleic acid trans-splicing molecules described above.

5

In some embodiments of the HTT nucleic acid trans-splicing molecules described above, the HTT pre-mRNA comprises at least one mutation associated with Huntington's Disease (HD). In some embodiments, at least one mutation associated with HD comprises an expansion of CAG repeats in an *HTT* gene allele. In some embodiments, the expansion of CAG repeats in an *HTT* gene allele comprises greater than 35 CAG repeats. In some embodiments, the at least one mutation associated with HD is autosomal dominant. In some embodiments, the at least one mutation associated with HD is expressed in at least one of cortical pyramidal neurons, striatal medium spiny neurons, or hypothalamic neurons.

10

Also disclosed is a vector comprising any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above; or any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above.

20

Also disclosed is a vector comprising any of the HTT nucleic acid trans-splicing molecules described above. In some embodiments, the vector comprises a 5' regulatory domain operatively linked 5' to the coding domain. In some embodiments, the 5' regulatory domain is operatively linked to a 5' untranslated region. In some embodiments, the 5' regulatory domain comprises a constitutive promoter or a tissue specific promoter. In some embodiments, the constitutive promoter is a CMV promoter or a CAGGS promoter.

30

Also disclosed is a proviral plasmid comprising any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon

35

skipping nucleic acid constructs described above; or any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above.

Also disclosed is an adeno-associated virus (AAV) comprising any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above; or any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above.

Also disclosed is an adeno-associated virus (AAV) comprising any of the HTT nucleic acid trans-splicing molecules described above, wherein the AAV optionally comprises a 5' regulatory domain operatively linked 5' to the nucleic acid trans-splicing molecule. In some embodiments, the AAV comprises a 5' regulatory domain operatively linked 5' to the coding domain. In some embodiments, the 5' regulatory domain is operatively linked to a 5' untranslated region. In some embodiments, the 5' regulatory domain comprises a constitutive promoter. In some embodiments, the constitutive promoter is a CMV promoter or a CAGGS promoter. In some embodiments, the AAV exhibits neuronal tropism. In some embodiments, the AAV is AAV9, AAV8, AAV5, AAV2, AAV7, or AAV2.7m8, AAV-retro, AAV1, AAV4, or AAV-PHP.eB.

Also disclosed is a composition comprising any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of the vectors described above; any of the proviral plasmids described above; or any of the AAVs described above. In some embodiments, the composition further comprises a pharmaceutically acceptable excipient. In some embodiments, the composition further comprises at least one anti-sense oligonucleotide or a construct that encodes at least one anti-sense RNA that inhibits cis-splicing of the HTT pre-mRNA. In some embodiments, the at least one anti-sense oligonucleotide comprises any one of SEQ ID NOs: 126-135 or the construct that encodes the at least one anti-sense RNA binds to a

target sequence bound by any one of SEQ ID NOs: 126-135. In some embodiments, the at least one anti-sense oligonucleotide comprises SEQ ID NO: 131 or the construct that encodes the at least one anti-sense RNA binds to a target sequence bound by SEQ ID NO: 131.

5 Also disclosed is a method of expressing biologically active HTT in a target cell to restore functional levels of HTT protein in the target cell, the method comprising transducing the target cell with any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described
10 above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of the vectors described above; any of the proviral plasmids described above; any of the AAVs
15 described above; or any of the compositions of described above. In some embodiments, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, or at least 45% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced. In some embodiments, at least 50%, at least 55%, at least 60%, at
20 least 65%, at least 70%, at least 75%, at least 80%, or at least 85% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced. In some embodiments, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced. In some embodiments, functional levels of HTT are restored in the target cell by expressing biologically functional HTT protein and/or mutant *HTT* RNA and related transcripts
25 (e.g., *HTT1a*) are reduced.

Also disclosed is a method of reducing expression of HTT comprising a polyglutamine repeat exceeding 35 consecutive glutamine residues in a subject, the method comprising transfecting or transducing a target cell, more particularly a neuron, in the subject with any of the HTT
30 nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT
35 miRNA, and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of the vectors

described above; any of the proviral plasmids described above; any of the AAVs described above; or any of the compositions of described above.

Also disclosed is a method of correcting at least one mutation in an HTT exon sequence in an
5 HTT pre-mRNA in a target cell of a subject, the method comprising administering to the subject
any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid
trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid
constructs described above; any of the MSH3 miRNA nucleic acid constructs described above;
any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-
10 splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT
trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above;
any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of
the vectors described above; any of the proviral plasmids described above; any of the AAVs
described above; or any of the compositions of described above.

15
Also disclosed is a method of treating Huntington's disease (HD) in a subject in need thereof,
the method comprising administering to the subject any of the HTT nucleic acid trans-splicing
molecules described above; any of the nucleic acid trans-splicing molecules described above;
any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3
20 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing
molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid
constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon
skipping nucleic acid constructs described above; any of the HTT trans-splicing and MSH3
miRNA nucleic acid constructs described above; any of the vectors described above; any of the
25 proviral plasmids described above; any of the AAVs described above; or any of the
compositions of described above in a therapeutically effective amount.

In some embodiments of the methods described above, the method comprises administration of
any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid
30 trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid
constructs described above; any of the MSH3 miRNA nucleic acid constructs described above;
any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-
splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT
trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above;
35 any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of
the vectors described above; any of the proviral plasmids described above; any of the AAVs

described above; or any of the compositions of described above to the subject's brain. In some embodiments, the subject is a mammal, preferentially a rodent, non-human primate, or a human. In some embodiments, the subject is genetically predisposed to have HD or has been diagnosed with HD.

5

Also disclosed is any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of the vectors described above; any of the proviral plasmids described above; any of the AAVs described above; or any of the compositions of described above for use in preventing or treating HD in a subject in need thereof.

Also disclosed herein is any of the HTT nucleic acid trans-splicing molecules described above; any of the nucleic acid trans-splicing molecules described above; any of the MSH3 exon skipping nucleic acid constructs described above; any of the MSH3 miRNA nucleic acid constructs described above; any of the MSH3 nucleic acid trans-splicing molecules described above; any of the HTT trans-splicing and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid constructs described above; any of the HTT trans-splicing and MSH3 miRNA nucleic acid constructs described above; any of the vectors described above; any of the proviral plasmids described above; any of the AAVs described above; or any of the compositions of described above for use in the preparation of a medicament for the treatment or prevention of HD in a subject in need thereof.

Also disclosed is a method comprising introducing into a cell a nucleic acid trans-splicing molecule configured to splice to both a first target pre-mRNA and a second target pre-mRNA, wherein the splicing to the first target pre-mRNA corrects a defect in the first target pre-mRNA, and wherein the splicing to the second target pre-mRNA introduces a defect in the second target pre-mRNA. In some embodiments, the nucleic acid trans-splicing molecule comprises a first binding domain configured to target an intron of the first target pre-mRNA and a second binding domain configured to target an intron of the second target pre-mRNA. In some embodiments, the nucleic acid trans-splicing molecule further comprises a coding domain

sequence comprising a functional sequence of one or more exons of the first target pre-mRNA that corrects the defect in the first target pre-mRNA. In some embodiments, the defect in the second target pre-mRNA comprises a frameshift in the coding sequence of the second target pre-mRNA. In some embodiments, the frameshift creates a premature termination codon in the second target pre-mRNA. In some embodiments, the defect comprises the endogenous start codon of the second target pre-mRNA being eliminated. In some embodiments, the defect comprises an inserted 5' UTR that prevents translation of the protein encoded by the second target pre-mRNA. In some embodiments, the defect comprises an inserted 3' UTR that destabilizes the pre-mRNA or prevents export of the second target pre-mRNA from the nucleus. In some embodiments, the defect comprises elimination of a 5' cap or a 3' polyA tail from the second target pre-mRNA. In some embodiments, the defect causes nonsense-mediated decay of the second target pre-mRNA. In some embodiments, the introducing causes the abundance of gene product of the second target pre-mRNA in the cell to be reduced compared to the abundance of the gene product before the introducing.

Also disclosed is a method of reducing the abundance of a protein in a cell, the method comprising introducing into the cell a nucleic acid trans-splicing molecule that introduces a defect into a pre-mRNA that encodes the protein. In some embodiments, the defect comprises one or more of the following: (a) a frameshift introduced into the coding sequence of the pre-mRNA; (b) elimination of the endogenous start codon of the pre-mRNA; (c) introduction of a premature stop codon into the coding sequence of the pre-mRNA; (d) replacement of the endogenous coding sequence of the pre-mRNA with an alternative coding sequence; (e) insertion of a 5' UTR that prevents translation of the endogenous coding sequence of the pre-mRNA; (f) insertion of a 3' UTR that destabilizes the pre-mRNA; (g) insertion of 3' UTR that prevents the pre-mRNA from being exported from the nucleus; (h) elimination of a 5' cap from the pre-mRNA; or (i) elimination of a 3' polyA tail from the pre-mRNA. In some embodiments, the protein is MSH3. In some embodiments, the nucleic acid trans-splicing molecule comprises a binding domain that binds to an intron of the pre-mRNA. In some embodiments, the nucleic acid trans-splicing molecule comprises a heterologous coding domain sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts the range of CAG trinucleotide repeats found in exon 1 of the *HTT* gene. As shown therein, 8-35 CAG trinucleotide repeats reflect the phenotypically normal (wildtype) range found in humans exhibiting no signs of disease associated with the *HTT* gene. 35-39 CAG

trinucleotide repeats are associated with incomplete penetrance of HD. The presence of over 40 CAG trinucleotide repeats in this region of exon 1 of the *HTT* gene is disease-causing for HD.

FIG. 2 depicts somatic CAG repeat expansion of the expanded CAG trinucleotide repeat in exon 1 of the *HTT* gene, which is the underlying mechanism of HD pathogenesis.

FIG. 3 presents exon-editing by *HTT* pre-mRNA trans-splicing as a therapeutic approach for HD. Intron 1-targeting, intron 2-targeting, or intron 3-targeting Exon Editors were designed and tested for efficiency of replacement of the mutant *HTT* exon 1.

FIG. 4 depicts an exemplary general structure of constructs encoding RNA Exon Editors targeting *HTT* intron 1. The exemplary exon editor produced by the depicted construct comprises a 5' UTR, exon 1 coding sequence, splice donor site, a linker, a binding domain, and a terminator sequence. The binding domain was varied to target different positions along intron 1 of *HTT*. In some embodiments, the promoter is a CMV promoter; in some embodiments, the 5' UTR is the *HTT* 5' UTR; in some embodiments, the linker is the 40mer linker; in some embodiments, the *HTT* 5' UTR is combined with the 40mer linker; in some embodiments, an epitope tag is included, an example of which is an N-terminal 3X FLAG tag for on-target protein detection. The aforementioned embodiments may be combined, wherein at least one of these embodiments is included in an RNA Exon Editor and any and all combinations thereof, including a combination of all these embodiments in one RNA Exon Editor.

FIG. 5 presents a graph depicting *HTT* intron 1-targeting RNA Exon Editors, which exhibit varying levels of trans-splicing efficiencies (% replacement) depending on where the binding domain targets in the intron. HEK293 cells were transfected with *HTT* intron 1-targeting RNA Exon Editors that target various regions of intron 1. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The nomenclature for the binding domain is: (nucleotide base position in the intron of the 5' end start position of the binding domain, numbering according to SEQ ID NO: 1)_(length (nt) of the binding domain). For example, 701_150 indicates that the binding domain is the reverse complement sequence of bases 701-850 in intron 1, numbering according to SEQ ID NO: 1. Results shown relate to exemplary *HTT* intron 1-targeting RNA Exon Editors comprising the indicated binding domain targets, wherein the 5' UTR comprises the *HTT* 5' UTR (SEQ ID NO: 136) and the linker comprises the 40mer linker (SEQ ID NO: 37). NBD_150 is a control editor in which the binding domain targeting *HTT* is replaced with a binding domain that does not target *HTT*.

FIG. 6 pictorially depicts the positions and trans-splicing efficiencies (% replacement) of HTT intron 1-targeting RNA Exon Editors. Varying levels of trans-splicing efficiencies (% replacement) were determined that reflected where the binding domain targets in the intron. HEK293 cells were transfected with *HTT* intron 1-targeting RNA Exon Editors that target various regions of intron 1. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The nomenclature for the binding domain is: (nucleotide base position in the intron of the 5' end start position of the binding domain, numbering according to SEQ ID NO: 1)_(length (nt) of the binding domain). For example, 701_150 indicates that the binding domain is the reverse complement sequence of bases 701-850 in intron 1, numbering according to SEQ ID NO: 1. Results shown relate to exemplary *HTT* intron 1-targeting RNA Exon Editors comprising the indicated binding domain targets, wherein the 5' UTR comprises the *HTT* 5' UTR and the linker comprises the 40mer linker.

FIG. 7 depicts a general structure of exemplary constructs encoding HTT intron 1-targeting RNA Exon Editors. An exemplary exon editor is shown wherein expression is driven by a CMV promoter. The exemplary exon editor depicted comprises the HTT 5' UTR, N-terminal 3X FLAG tag, exon 1 coding sequence, a splice donor site, a linker, a binding domain (HTT_intron1_11704_100), and a terminator sequence.

FIG. 8 shows the activity of exemplary HTT intron 1-targeting RNA Exon Editors that included various linkers. As shown therein, some linkers increased trans-splicing efficiency relative to the 40mer linker in HTT intron 1 (HTT_intron1_11704_100) Exon Editors. HEK293 cells were transfected with HTT intron 1-targeting RNA Exon Editors that comprised the indicated linkers. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR.

FIG. 9 depicts a general structure of exemplary constructs encoding RNA Exon Editors targeting *HTT* intron 2. The exemplary exon editor depicted comprises a 5' UTR, exons 1-2 coding sequences, a splice donor site, a linker, a binding domain, and a terminator sequence. The binding domain was varied to target different positions along intron 2 of *HTT*. In some embodiments, the 5' UTR is the *HTT* 5' UTR; in some embodiments, the linker is the 40mer linker; in some embodiments, the *HTT* 5' UTR is combined with the 40mer linker. In some embodiments, the promoter is a CMV promoter; in some embodiments, the 5' UTR is the *HTT* 5' UTR; in some embodiments, the linker is the 40mer linker; in some embodiments, the CMV promoter is combined with the *HTT* 5' UTR; in some embodiments, the CMV promoter is combined with the *HTT* 5' UTR and the 40mer linker; in some embodiments, an epitope tag is

included, an example of which is an N-terminal 3X FLAG tag for on-target protein detection. The aforementioned embodiments may be combined, wherein at least one of these embodiments is included in an RNA Exon Editor and any and all combinations thereof, including a combination of all these embodiments in one RNA Exon Editor.

5

FIG. 10 shows the activity of various exemplary *HTT* intron 2-targeting RNA Exon Editors, which exhibit varying levels of trans-splicing efficiencies (% replacement) depending on where the binding domain targets in the intron. HEK293 cells were transfected with exemplary *HTT* intron 2-targeting RNA Exon Editors that target various regions of intron 2. Cells were harvested 48
10 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The nomenclature for the binding domain is: (nucleotide base position in the intron of the 5' end start position of the binding domain, numbering according to SEQ ID NO: 57)_(length (nt) of the binding domain). Results shown relate to exemplary *HTT* intron 2-targeting RNA Exon Editors comprising the indicated binding domain targets, wherein the 5' UTR comprises the *HTT* 5' UTR
15 and the linker comprises the 40mer linker. NBD, control editor in which the binding domain targeting *HTT* is replaced with a binding domain that does not target *HTT*. Splice mutant, control editor lacking a functional splice donor site.

FIG. 11 pictorially depicts the positions and trans-splicing efficiencies (% replacement) of *HTT*
20 intron 2-targeting RNA Exon Editors. Varying levels of trans-splicing efficiencies (% replacement) were determined that reflected where the binding domain targets the intron. HEK293 cells were transfected with *HTT* intron 2-targeting RNA Exon Editors that target various regions of intron 2. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The nomenclature for the binding domain is: (nucleotide base
25 position in the intron of the 5' end start position of the binding domain, numbering according to SEQ ID NO: 57)_(length (nt) of the binding domain). Results shown relate to exemplary *HTT* intron 2-targeting RNA Exon Editors comprising the indicated binding domain targets, wherein the 5' UTR comprises the *HTT* 5' UTR and the linker comprises the 40mer linker.

FIGs. 12A and 12B depict exemplary *HTT* intron 2-targeting RNA Exon Editors targeting the region upstream of the branchpoint. A) RNA exon editors were designed to target upstream of the intron 2 branchpoint and vary in binding domain length. Exon Editor expression is driven by a CMV promoter. The Exon Editors comprise the *HTT* 5' UTR, N-terminal 3X FLAG tag for on-target protein detection, exons 1-2 coding sequence, a splice donor site, the 41mer_2 linker, the
30 indicated binding domain, and a terminator sequence. The nomenclature for the binding domain is: (nucleotide base position in the intron of the 5' end start position of the binding domain,
35

numbering according to SEQ ID NO: 57)_(length (nt) of the binding domain). The binding domain was varied to target different lengths upstream of the branchpoint in intron 2. B) The length of the binding domain influences targeting efficiencies, based on % replacement calculations. HEK293 cells were transfected with HTT intron 2-targeting RNA Exon Editors that
5 varied in their binding domain lengths. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR.

FIG. 13 depicts a general structure of exemplary constructs encoding HTT intron 2-targeting RNA Exon Editors. The Exon Editor expression is driven by a CMV promoter. Exemplary Exon
10 Editors comprise the HTT 5' UTR, N-terminal 3X FLAG tag, exon 1-2 coding sequences, a splice donor site, a linker, a binding domain (HTT_intron2_12061_150), and a terminator sequence.

FIG. 14 shows that Exon Editors comprising the indicated linkers do not exhibit significantly different trans-splicing efficiencies relative to the 40mer linker in HTT intron 2
15 (HTT_intron2_12061_150). HEK293 cells were transfected with HTT intron 2-targeting RNA Exon Editors that varied with respect to the linkers included therein. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR.

FIG. 15 presents a cartoon depicting a trans-splicing reaction and competition thereof with respect to cis-splicing. Abbreviations used therein: MALAT1 terminator (term); binding domain
20 (BD); linker (L); splice site (SS). The cis-spliced molecule resulting from cis-splicing comprises the CAG repeat expansion. The Chimeric trans-spliced molecule resulting from 5' trans-splicing mediated by the HTT 5' RTM comprises a HTT exon 1 comprising a normal number of CAG repeats (8-35). Abbreviations used therein: MALAT1 terminator (term); binding domain (BD);
25 linker (L); splice site (SS); codon optimized (C/O).

FIG. 16 presents a cartoon showing anti-sense oligonucleotides (ASOs) designed to block competing cis-splicing sites (ASO8-10), as well as cis-splicing sites for the upstream exon
30 (ASO2-7). Each of these ASOs was co-transfected with the indicated HTT intron 2-targeting Exon Editor (HTT_intron2_12061_150) and assayed for trans-splicing efficiency *in vitro* in HEK293 cells. Abbreviations used therein: MALAT1 terminator (term); binding domain (BD); linker (L); splice site (SS).

FIG. 17 shows percent replacement activity of an exemplary HTT intron 2-targeting Exon Editor
35 (HTT_intron2_12061_150) in combination with the indicated ASOs. ASO6, which was designed to block the cis-splicing of the upstream intron, led to an improvement in trans-splicing

efficiencies *in vitro*. HEK293 cells were co-transfected with an exemplary HTT intron 2-targeting RNA Exon Editor Construct (REEC) and ASOs designed to block the competing cis-splicing site or ASOs designed to block the splicing of the upstream intron. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR.

5

FIG. 18 shows representative RT-qPCR and Western blot images (probed for the N-terminal FLAG epitope) of whole cell lysates from HEK293 cells transfected with *HTT* RNA Exon Editors comprising the indicated elements. It is noteworthy that on-target protein detection levels correlate with trans-splicing efficiency as represented by % Replacement of *HTT* RNA. HEK293 cells were transfected with N-terminally FLAG-tagged *HTT* Exon Editors that have a range of activity based on RT-qPCR assays (Upper panel). Anti-FLAG antibody was used to detect protein generated following successful trans-splicing in whole cell lysates (Lower panel). Anti-Huntingtin protein antibody was used to detect native and ONT protein generated following successful trans-splicing in whole cell lysates (Lower panel). The intensity of the FLAG ONT band scales with the relative performance of the Exon Editor based on qPCR.

10

15

FIG. 19 presents a schematic depicting a potential mechanism of a hybrid therapeutic approach designed to treat HD. The hybrid therapeutic approach combines agents that inhibit somatic CAG expansion (e.g., by MSH3 reduction) with RNA Exon Editors targeting *HTT* pre-mRNA. RNA Exon Editors targeting *HTT* pre-mRNA serve to replace any mutant *HTT* RNA that might be produced from DNA that “escapes” inhibition of the somatic expansion process.

20

FIG. 20 presents a schematic depicting a potential mechanism of action of a tandem binding domain RNA Exon Editor targeting *HTT* and *MSH3* pre-mRNA. In an exemplary embodiment, expression of an Exon Editor is driven by a CMV promoter. Such exemplary Exon Editors may comprise the *HTT* 5' UTR, N-terminal 3X FLAG tag, *HTT* exon 1 coding sequence, splice donor site, a linker, an *MSH3* binding domain (targeting, e.g., intron 5 or intron 15 of *MSH3* pre-mRNA), an *HTT* binding domain (e.g., HTT_intron1_11704_100), and a terminator sequence. The *HTT* binding domain will target the Exon Editor to produce the corrected *HTT* RNA after successful trans-splicing, while the *MSH3* binding domain will target the Exon Editor to produce a chimeric *HTT* exon 1-*MSH3* RNA molecule with a premature stop codon which will be subject to nonsense-mediated decay (NMD) and lead to the subsequent reduction of *MSH3* expression.

25

30

FIGs. 21A-21C present results showing that tandem binding domain RNA Exon Editors targeting *HTT* and *MSH3* exhibit successful trans-splicing to both pre-mRNAs. RT-qPCR of A) *HTT* on-target (ONT) trans-splicing efficiency (via *HTT* binding domain), B) *HTT*-*MSH3* chimeric trans-

35

splicing efficiency (via *MSH3* binding domain), and C) *MSH3* RNA transcript expression level, on HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 1 and *MSH3* intron 5.

5 FIGs. 22A-22C present results showing that tandem binding domain RNA Exon Editors targeting *HTT* and *MSH3* exhibit successful trans-splicing to both pre-mRNAs. RT-qPCR of A) *HTT* on-target (ONT) trans-splicing efficiency (via *HTT* binding domain), B) *HTT-MSH3* chimeric trans-splicing efficiency (via *MSH3* binding domain), and C) *MSH3* RNA transcript expression level, on HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 1
10 and *MSH3* intron 15.

FIG. 23 depicts a general structure of exemplary constructs encoding *HTT* intron 3-targeting RNA Exon Editors. The exemplary exon editor depicted comprises a 5' UTR, exons 1-3 coding sequences, a splice donor site, a linker, a binding domain, and a terminator sequence. The
15 binding domain was varied to target different positions along intron 3 of *HTT*. In some embodiments, the 5' UTR is the *HTT* 5' UTR; in some embodiments, the linker is the 40mer linker. In some embodiments, the promoter is a CMV promoter; in some embodiments, the 5' UTR is the *HTT* 5' UTR; in some embodiments, the linker is the 40mer linker; in some
20 embodiments, the CMV promoter is combined with the *HTT* 5' UTR; in some embodiments, the CMV promoter is combined with the *HTT* 5' UTR and the 40mer linker; in some embodiments, an epitope tag is included, an example of which is an N-terminal 3X FLAG tag for on-target protein detection. The aforementioned embodiments may be combined, wherein at least one of these embodiments is included in an RNA Exon Editor and any and all combinations thereof, including a combination of all these embodiments in one RNA Exon Editor.

25
FIG. 24 shows the activity of various exemplary *HTT* intron 3-targeting RNA Exon Editors, which show varying levels of trans-splicing efficiencies (% replacement) depending on the binding site within the intron targeted by the binding domain. HEK293 cells were transfected with *HTT* intron 3-targeting RNA Exon Editors that target various regions of intron 3. Cells were harvested 48
30 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The nomenclature for the binding domain is: (nucleotide base position in intron 3 of the 5' end start position of the binding domain, numbering according to SEQ ID NO: 155)_(length (nt) of the binding domain).

35 FIG. 25 shows a direct comparison of *HTT* intron 2-targeting Exon Editors with *HTT* intron 3-targeting Exon Editors. HEK293 cells were transfected with *HTT* intron 2-targeting and intron 3-

targeting RNA Exon Editors in parallel. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. The nomenclature for the binding domain is: (nucleotide base position in intron 3 of the 5' end start position of the binding domain)_(length (nt) of the binding domain).

5

FIG. 26 shows that self-splicing mitigation does not affect trans-splicing efficiency of HTT Exon Editors. Table 1 presents cryptic splice sites identified in the original exon editor and the sequence changes made in self-splicing mitigated exon editors. HEK293 cells were transfected with HTT intron 2-targeting (HTT_intron2_12061_150) RNA Exon Editors with or without self-splicing mitigation. Cells were harvested 48 hours post-transfection, assayed for trans-splicing efficiencies by RT-qPCR.

10

FIG. 27 shows a representative RT-qPCR and Western blot image of lysates from HEK293 cells transfected with HTT RNA Exon Editors testing 2 promoters and different 5' UTR combinations.

15

HEK293 cells were transfected with N-terminally FLAG-tagged HTT Exon Editors driven by either the CMV or CAGGS promoter, with or without the HTT 5'UTR, and testing wild-type (GTAAGT) splice site targeting intron 2 (HTT_intron2_12061_150), splice mutant targeting intron 2 (HTT_intron2_12061_150), or wild-type splice site with a non-targeting binding domain (NBD). RNA from these cells were subject to RT-qPCR (Upper panel). α -FLAG antibody was used to detect protein generated following successful trans-splicing (which comprises an N-terminal FLAG epitope) in whole cell lysates (Lower panel). ONT: successfully trans-spliced on-target HTT protein. NSP: non-spliced protein.

20

FIG. 28 shows that non-spliced protein (NSP) is reduced in a combinatorial manner by the inclusion of three tandem repeats of U1 snRNA binding site (3X UBS; SEQ ID NO: 345) and an AU-rich element (ARE; SEQ ID NO: 346) in an exemplary 5' HTT intron 1-targeting Exon Editor. HEK293 cells were transfected with HTT intron 1-targeting (HTT_intron1_11704_100) RNA Exon Editors that were varied at their linker region to include the indicated NSP reduction elements. Cells were harvested 48 hours post-transfection and subjected to Western Blot analysis. ONT: successfully trans-spliced on-target HTT protein. NSP: non-spliced protein.

30

FIG. 29 shows that non-spliced protein (NSP) is reduced in a combinatorial manner by the inclusion of three tandem repeats of U1 snRNA binding site (3X UBS; SEQ ID NO: 345) and an AU-rich element (ARE; SEQ ID NO: 346) in the 5' HTT intron 2-targeting Exon Editor. HEK293 cells were transfected with HTT intron 2-targeting (HTT_intron2_12061_150) RNA Exon Editors that were varied at their linker region to include the indicated NSP reduction elements. Cells

35

were harvested 48 hours post-transfection, assayed for trans-splicing efficiencies by RT-qPCR (upper panel) or subjected to Western Blot analysis (lower panel).

FIG. 30 shows a mechanism of action of a tandem binding domain RNA Exon Editor targeting *HTT* intron 2 and *MSH3* pre-mRNA. The Exon Editor expression is driven by a CMV promoter and contains the HTT 5' UTR, N-terminal 3X FLAG tag, HTT exon 1 and exon 2 coding sequence, splice donor site, the linker, MSH3 binding domain (targeting intron 5 or intron 15 of *MSH3* pre-mRNA), HTT binding domain (HTT_intron2_12061_150), and a terminator sequence. The HTT binding domain will target the Exon Editor to produce the corrected HTT RNA after successful trans-splicing, while the MSH3 binding domain will target the Exon Editor to produce a chimeric HTT exon 1+2 -MSH3 RNA molecule with a premature stop codon which will be subject to nonsense-mediated decay (NMD) and lead to the subsequent reduction of MSH3 expression.

FIGs. 31A and 31B present RT-qPCR profiles of HTT trans-splicing and HTT-MSH3 chimera production (via MSH3 trans-splicing) in tandem binding domain Exon Editors. RT-qPCR of A) HTT on-target (ONT) trans-splicing efficiency (via *HTT* intron 2-targeting binding domain) and B) HTT-MSH3 chimeric trans-splicing efficiency (via *MSH3* intron 5-targeting binding domain), in HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 2 and *MSH3* intron 5. The binding domain targeting *HTT* intron 2 was HTT_intron2_12061_150 for all Exon Editors tested here, while MSH3_intron5_213_100 and MSH3_intron5_188_150 were tested for the *MSH3* intron 5-targeting binding domains. Binding domains were positioned in tandem and the order of the binding domains was as indicated. An Exon Editor was also tested for each MSH3 binding domain with the MALAT1 triple helix placed between the two tandem binding domains.

FIGs. 32A and 32B present RT-qPCR profiles of HTT trans-splicing and HTT-MSH3 chimera production (via MSH3 trans-splicing) in tandem binding domain Exon Editors. RT-qPCR of A) HTT on-target (ONT) trans-splicing efficiency (via *HTT* intron 2-targeting binding domain), and B) HTT-MSH3 chimeric trans-splicing efficiency (via *MSH3* intron 15-targeting binding domain), in HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 2 and *MSH3* intron 15. The binding domain targeting *HTT* intron 2 was HTT_intron2_12061_150 for all Exon Editors tested here, while MSH3_intron15_6523_120 and MSH3_intron15_6498_150 were tested for the *MSH3* intron 15-targeting binding domains. Binding domains were positioned in tandem and the order of the binding domains was as indicated. An Exon Editor was

also tested for each MSH3 binding domain with the MALAT1 triple helix placed between the two tandem binding domains.

FIG. 33 presents a cartoon depicting MSH3 knockdown by a miRNA targeting the *MSH3* mRNA.

5 MSH3 can be knocked down by miRNAs that target the *MSH3* mRNA and degrade the transcript.

FIG. 34 shows Western blot analysis of MSH3 exon 23-targeting RNAi constructs. Constructs (SEQ ID NOs: 234, 235, 238-241) comprising MSH3 exon 23-targeting miRNA active sequence
10 TTAATCCATAACTCCTTGC (SEQ ID NO: 224) were analyzed, as well as control constructs (SEQ ID NOs: 236, 237, 242, and 243). ImageJ analysis was performed on the Western blots to analyze MSH3 protein knockdown (upper panel). U6 promoter-driven shRNAs and CMV promoter-driven pri-miRNA mimics were designed and tested. Variations include the strand placement (5' arm or 3' arm) of the guide strand, including a bulge in the stem structure, and
15 varying the miRNA scaffold. Negative controls include constructs that contain a non-targeting sequence or a no hairpin loop control.

FIG. 35 shows RT-qPCR and Western blot analysis of MSH3-targeting RNAi constructs.

Constructs encoding miRNAs targeting different regions of the *MSH3* transcript were analyzed.

20 ImageJ analysis was performed on the Western blots to analyze MSH3 protein knockdown. CMV promoter-driven pri-miRNA mimics targeting different exonic sequences of MSH3 were designed and tested.

FIG. 36 presents a cartoon depicting MSH3 knockdown by small nuclear RNA (snRNA)-based
25 antisense RNA (asRNA). MSH3 can be inactivated by antisense RNAs encoded in a snRNA scaffold that anneal to *MSH3* splice junctions, preventing exon inclusion. This leads to exon skipping and the generation of a premature stop codon, ultimately causing NMD of the MSH3 transcript. The illustration here depicts an example of a MSH3 splice modulator targeting the skipping of exon 2.

30

FIG. 37 depicts relative MSH3 RNA expression levels of exon 1 – exon 2 and exon 2 – exon 3 junctions in MSH3 splice modulators targeting exon 2 skipping (bottom) and cartoon illustrating asRNA target region in MSH3 pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 284-287. SEQ ID NO: 284: U7SmOPT containing MSH3 In1/Ex2 asRNA (SEQ ID NO: 274);
35 SEQ ID NO: 285 (U7SmOPT containing MSH3 Ex2/In2 asRNA (SEQ ID NO: 275); SEQ ID NO:

286: U7SmOPT containing MSH3 In1/Ex2 + Ex2/In2 asRNA (SEQ ID NO: 276); SEQ ID NO: 287: MSH3 In1/Ex2 + Ex2/In2 long (160 nt) asRNA (SEQ ID NO: 277).

FIG. 38 depicts relative MSH3 RNA expression levels of exon 2 – exon 3 and exon 3 – exon 4 junctions in MSH3 splice modulators targeting exon 3 skipping (bottom) and cartoon illustrating asRNA target region in MSH3 pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 288-290. SEQ ID NO: 288: U7SmOPT containing MSH3 In2/Ex3 asRNA (SEQ ID NO: 278); SEQ ID NO: 289: U7SmOPT containing MSH3 Ex3/In3 asRNA (SEQ ID NO: 279); SEQ ID NO: 290: U7SmOPT containing MSH3 In2/Ex3 + Ex3/In3 asRNA (SEQ ID NO: 280).

FIG. 39 depicts relative MSH3 RNA expression levels of exon 3 – exon 4 and exon 4 – exon 5 junctions in MSH3 splice modulators targeting exon 4 skipping (bottom) and cartoon illustrating asRNA target region in MSH3 pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 291-293. SEQ ID NO: 291: U7SmOPT containing MSH3 In3/Ex4 asRNA (SEQ ID NO: 281); SEQ ID NO: 292: U7SmOPT containing MSH3 Ex4/In4 asRNA (SEQ ID NO: 282); SEQ ID NO: 293: U7SmOPT containing MSH3 In3/Ex4 + Ex4/In4 asRNA (SEQ ID NO: 283).

FIG. 40 depicts relative MSH3 RNA expression levels of exon 2 – exon 3 and exon 3 – exon 4 junctions in MSH3 splice modulators targeting exon 3 skipping (bottom) and cartoon illustrating asRNA target region in MSH3 pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 290, 324-326. SEQ ID NO: 290: U7SmOPT containing MSH3 In2/Ex3 + Ex3/In3 asRNA (SEQ ID NO: 280); SEQ ID NO: 324: MSH3 U7 SmOPT Splice modulator containing In3/Ex3/In2 asRNA (SEQ ID NO: 299); SEQ ID NO: 325: MSH3 U7 SmOPT Splice modulator containing In3/Ex3-1 + linker + Ex3/In2-1 asRNA (SEQ ID NO: 300 + 298 + 301); SEQ ID NO: 326: MSH3 U2 Splice modulator containing In3/Ex3-2 + linker + Ex3/In2-2 asRNA (SEQ ID NO: 302 + 298 + 303).

FIG. 41 depicts relative MSH3 RNA expression levels of exon 5 – exon 6 and exon 6 – exon 7 junctions in MSH3 splice modulators targeting exon 6 skipping (bottom) and cartoon illustrating asRNA target region in MSH3 pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 327-329. SEQ ID NO: 327: MSH3 U7 SmOPT Splice modulator containing In6/Ex6/In5 asRNA (SEQ ID NO: 304); SEQ ID NO: 328: MSH3 U7 SmOPT Splice modulator containing In6/Ex6-1 + linker + Ex6/In5-1 asRNA (SEQ ID NO: 305 + 298 + 306); SEQ ID NO: 329: MSH3 U2 Splice modulator containing In6/Ex6-2 + linker + Ex6/In5-2 asRNA (SEQ ID NO: 307 + 298 + 308).

FIG. 42 depicts relative *MSH3* RNA expression levels of exon 6 – exon 7 and exon 7 – exon 8 junctions in *MSH3* splice modulators targeting exon 7 skipping (bottom) and cartoon illustrating asRNA target region in *MSH3* pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 330-332. The anti-sense RNA (asRNA) comprised in SEQ ID NO: 330 is *MSH3* In7/Ex7/In6 asRNA (asRNA region SEQ ID NO: 309); the asRNA comprised in SEQ ID NO: 331 are In7/Ex7-1 (asRNA region SEQ ID NO: 310) + linker (SEQ ID NO: 298) + Ex7/In6-1 (asRNA region SEQ ID NO: 311); the asRNA comprised in SEQ ID NO: 332 are In7/Ex7-2 (asRNA region SEQ ID NO: 312) + linker (SEQ ID NO: 298) + Ex7/In6-2 (asRNA region SEQ ID NO: 313).

FIG. 43 depicts relative *MSH3* RNA expression levels of exon 7 – exon 8 and exon 8 – exon 9 junctions in *MSH3* splice modulators targeting exon 8 skipping (bottom) and cartoon illustrating asRNA target region in *MSH3* pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 333-335. SEQ ID NO: 333: *MSH3* U7 SmOPT Splice modulator containing In8/Ex8/In7 asRNA (SEQ ID NO: 314); SEQ ID NO: 334: *MSH3* U7 SmOPT Splice modulator containing In8/Ex8-1 + linker + Ex8/In7-1 asRNA (SEQ ID NO: 315 + 298 + 316); SEQ ID NO: 335: *MSH3* U2 Splice modulator containing In8/Ex8-2 + linker + Ex8/In7-2 asRNA (SEQ ID NO: 317 + 298 + 318).

FIG. 44 depicts relative *MSH3* RNA expression levels of exon 14 – exon 15 and exon 15 – exon 16 junctions in *MSH3* splice modulators targeting exon 15 skipping (bottom) and cartoon illustrating asRNA target region in *MSH3* pre-mRNA (top). The splice modulator transcripts are SEQ ID NOs: 336-338. SEQ ID NO: 336: *MSH3* U7 SmOPT Splice modulator containing In15/Ex15/In14 asRNA (SEQ ID NO: 319); SEQ ID NO: 337: *MSH3* U7 SmOPT Splice modulator containing In15/Ex15-1 + linker + Ex15/In14-1 asRNA (SEQ ID NO: 320 + 298 + 321); SEQ ID NO: 338: *MSH3* U2 Splice modulator containing In15/Ex15-2 + linker + Ex15/In14-2 asRNA (SEQ ID NO: 322 + 298 + 323).

FIGs. 45A, 45B, and 45C shows that *MSH3* exon 7 splice modulators show reduction of *MSH3* RNA and protein levels. HEK293 cells were transfected with snRNA-based splice modulators designed to skip *MSH3* exon 7. Cells were harvested 48 hours post-transfection, assayed for *MSH3* knockdown by RT-qPCR (FIG. 45B) or subjected to Western Blot analysis (FIG. 45C). The splice modulator transcripts are SEQ ID NOs: 330-332. SEQ ID NO: 330: *MSH3* U7 SmOPT Splice modulator containing In7/Ex7/In6 asRNA (SEQ ID NO: 309); SEQ ID NO: 331: *MSH3* U7 SmOPT Splice modulator containing In7/Ex7-1 + linker + Ex7/In6-1 asRNA (SEQ ID NO: 310 + 298 + 311); SEQ ID NO: 332: *MSH3* U2 Splice modulator containing In7/Ex7-2 + linker + Ex7/In6-2 asRNA (SEQ ID NO: 312 + 298 + 313).

FIG. 46 presents a combination strategy to correct mutant HTT by trans-splicing and knockdown unedited *HTT* species (including HTT1a) with a microRNA (miRNA).

5 FIG. 47 shows a vectorized hybrid molecule that combines a HTT Exon Editor with a miRNA targeting unedited *HTT* mRNA. A short-hairpin RNA (shRNA) or microRNA (miRNA) designed to reduce HTT gene expression can be added to an Exon Editor within the same cistron (e.g., in an intron of the Exon Editor) or as a separate cistron with its own regulatory sequences. The RNAi can reduce expression of the unedited target (e.g., mutant *HTT*). For the purposes of selectively
10 reducing unedited *HTT* with RNAi, the Exon Editor comprises a *HTT* CDS comprising a sequence-altered portion that renders edited *HTT* resistant to the shRNA or miRNA.

FIGs. 48A and 48B present trans-splicing and HTT knockdown profiles of HTT Exon Editor, HTT miRNA-1, and a dual hybrid molecule of HTT Exon Editor and HTT miRNA-1. A) % trans-spliced
15 (edited) *HTT* transcripts in all *HTT* transcripts, and B) copy numbers of unedited and trans-spliced (edited) *HTT* transcripts in each of the treatments. HEK293 cells were transfected with HTT intron 2-targeting RNA Exon Editor, HTT miRNA-1, and a dual hybrid molecule of HTT intron 2-targeting Exon Editor (SEQ ID NO: 204) + HTT miRNA-1 (HTT miRNA-1 encoding sequence SEQ ID NO: 341 containing HTT miRNA-1 active sequence SEQ ID NO: 339). Cells
20 were harvested 48 hours post-transfection and RNA was subjected to RT-qPCR. Dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204.

FIGs. 49A-49C present trans-splicing and HTT knockdown profiles of molecules containing HTT
25 miRNA-1 and HTT miRNA-2 in vitro. HEK293 cells were transfected with HTT intron 2-targeting RNA Exon Editor, and dual hybrid molecules of HTT intron 2-targeting Exon Editor + HTT miRNA-1 or HTT miRNA-2. Cells were harvested 48 hours post-transfection and RNA was subjected to RT-qPCR analysis for A) trans-splicing profile, B) HTT knockdown profile, and C) *HTT* copy number analysis. Dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204. Dual hybrid
30 molecule of HTT Exon Editor and HTT miRNA-2 (SEQ ID NO: 355), which dual hybrid comprises SEQ ID NO: 344 and SEQ ID NO: 204.

FIG. 50 shows that HTT miRNA-1 successfully knockdowns unedited *HTT* transcripts with
35 minimal interaction with the Exon Editor and edited *HTT* transcript. HEK293 cells were transfected with a HTT Exon Editor +/- HTT miRNA-1 dual hybrid molecules. Cells were

harvested 48 hours post-transfection, assayed for HTT knockdown and trans-splicing efficiencies by RT-qPCR (upper panels) or subjected to Western Blot analysis (lower panels). Dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204.

5

FIG. 51 presents a cartoon depicting reduction of MSH3 by splice modulation in combination with HTT trans-splicing.

FIGs. 52A and 52B present results showing the performance of HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecules. HEK293 cells were transfected with a HTT Exon Editor +/- MSH3 Splice Modulator. Cells were harvested 48 hours post-transfection and assayed for A) trans-splicing profiles and B) MSH3 knockdown profiles by RT-qPCR. HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecules: MSH3 Splice Modulator + HTT Exon Editor (SEQ ID NO: 356), which dual hybrid comprises SEQ ID NO: 331 and SEQ ID NO: 204.

15

FIGs. 53A-D present a comparison of self-complementary AAV (scAAV) to single-stranded AAV (ssAAV) using a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule. HEK293 cells were transduced with scAAV or ssAAV expressing a HTT Exon Editor + MSH3 exon 7-skipping Splice Modulator dual hybrid molecule. AAV2 serotype was used. Cells were harvested 48 hours after transduction and subjected to RT-qPCR and Western Blot analysis. Sequences indicated were inserted between the ITRs of AAV2, whether ssAAV or scAAV. HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule (SEQ ID NO: 357), which dual hybrid comprises SEQ ID NO: 331 and SEQ ID NO: 204, in a head-to-head orientation.

FIG. 54 presents results showing the performance of HTT Exon Editor + HTT miRNA + MSH3 Splice Modulator triple hybrid molecules compared to its controls. HEK293 cells were transfected with HTT Exon Editor +/- HTT miRNA-1 or HTT miRNA-2 +/- MSH3 Splice Modulator. Cells were harvested 48 hours post-transfection and assayed for Top panel) % trans-spliced HTT transcripts, Middle panel) HTT knockdown profiles, and Bottom panel) MSH3 knockdown profiles by RT-qPCR. Control hybrid molecules comprising an Exon Editor with a splice donor mutation, a Splice Modulator that contains a scrambled asRNA sequence, or a miRNA that contains a scrambled asRNA, were also tested. "1" for HTT miRNA indicates HTT miRNA-1, "2" for HTT miRNA indicates HTT miRNA-2 was used. SM, splice mutant. Scr, scrambled control. miR-33 was used for the miRNA scaffold. HTT Exon Editor (SEQ ID NO: 204); dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204; HTT Exon Editor + MSH3 Splice Modulator dual hybrid

35

- molecules: MSH3 Splice Modulator + HTT Exon Editor (SEQ ID NO: 356), which dual hybrid comprises SEQ ID NO: 331 and SEQ ID NO: 204; HTT Exon Editor + HTT miRNA-1 + MSH3 Splice Modulator triple hybrid molecule (SEQ ID NO: 358), which triple hybrid comprises SEQ ID NO: 331, SEQ ID NO: 341, and SEQ ID NO: 204; HTT Exon Editor + HTT miRNA-2 + MSH3
- 5 Splice Modulator triple hybrid molecule (SEQ ID NO: 359), which triple hybrid comprises SEQ ID NO: 331, SEQ ID NO: 344, and SEQ ID NO: 204; dual hybrid molecule of HTT Exon Editor and HTT miRNA-2 (SEQ ID NO: 355), which dual hybrid comprises SEQ ID NO: 344 and SEQ ID NO: 204.
- 10 FIG. 55 shows trans-splicing profiles of HD molecules in iCell GlutaNeurons as measured by RT-ddPCR and Western blotting. iCell GlutaNeurons were transduced with a HTT Exon Editor + HTT miRNA dual hybrid molecule or a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule, both packaged in AAV2.7m8. Cells were harvested for RT-ddPCR and Western Blots after 18-21 days. Sequences indicated were inserted between the ITRs of AAV2.7m8. Dual
- 15 hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204; HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule (SEQ ID NO: 357), which dual hybrid comprises SEQ ID NO: 331 and SEQ ID NO: 204, in a head-to-head orientation; HTT Exon Editor + HTT miRNA-1 + MSH3 Splice Modulator triple hybrid molecule (SEQ ID NO: 358), which triple hybrid comprises SEQ ID NO:
- 20 331, SEQ ID NO: 341, and SEQ ID NO: 204.

- FIG. 56 shows HTT miRNA knockdown profiles in iCell GlutaNeurons as measured by RT-ddPCR and Western blotting. iCell GlutaNeurons were transduced with a HTT Exon Editor + HTT miRNA dual hybrid molecule or a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule, both
- 25 packaged in AAV2.7m8. Cells were harvested for RT-ddPCR and Western Blots after 18 days. Sequences indicated were inserted between the ITRs of AAV2.7m8. Dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204; HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule (SEQ ID NO: 357), which dual hybrid comprises SEQ ID NO: 331 and SEQ ID NO: 204, in a
- 30 head-to-head orientation.

- FIG. 57 shows a MSH3 knockdown profile of the MSH3 Splice Modulator in iCell GlutaNeurons as measured by RT-ddPCR and Western blotting. iCell GlutaNeurons were transduced with a
- 35 HTT Exon Editor + HTT miRNA-1 dual hybrid molecule or a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule, both packaged in AAV2.7m8. Cells were harvested for RT-ddPCR and Western Blots after 18 days. Sequences indicated were inserted between the ITRs

of AAV2.7m8. Dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), which dual hybrid comprises SEQ ID NO: 341 and SEQ ID NO: 204; HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule (SEQ ID NO: 357), which dual hybrid comprises SEQ ID NO: 331 and SEQ ID NO: 204, in a head-to-head orientation.

5

FIG. 58 shows HTT trans-splicing profiles in the BAC-CAG mouse brain. Neonatal ICV injections (at 1E+11 or 3E+11 vg/animal) in BAC-CAG mice were performed with the indicated HD molecules packaged in AAV9. Mouse cortex and striatum were harvested 4 weeks post-injection and the efficiencies of HTT Exon Replacement by trans-splicing profiled by RT-ddPCR and

10 Western Blotting. SEQ ID NO: 369: scAAV, Mouse Msh3 Splice Modulator + CMVp::Exon Editor (Dual Hybrid of SEQ ID NO: 362 and SEQ ID NO: 204, head-to-head orientation); SEQ ID NO: 370: ssAAV, Mouse Msh3 Splice Modulator + CMVp::Exon Editor (Dual Hybrid of SEQ ID NO: 362 and SEQ ID NO: 204, head-to-head orientation); SEQ ID NO: 371: ssAAV, Mouse Msh3 Splice Modulator + CAGGSp::Exon Editor (Dual Hybrid of SEQ ID NO: 362 and SEQ ID NO: 204, head-to-head orientation).

15

FIG. 59 depicts an observed relationship between Exon Editor RNA copy number and trans-splicing efficiency (% HTT replacement) in the BAC-CAG mouse brain. SEQ ID NO: 369: scAAV, Mouse Msh3 Splice Modulator + CMVp::Exon Editor (Dual Hybrid of SEQ ID NO: 362 and SEQ ID NO: 204, head-to-head orientation); SEQ ID NO: 370: ssAAV, Mouse Msh3 Splice Modulator + CMVp::Exon Editor (Dual Hybrid of SEQ ID NO: 362 and SEQ ID NO: 204, head-to-head orientation); SEQ ID NO: 371: ssAAV, Mouse Msh3 Splice Modulator + CAGGSp::Exon Editor (Dual Hybrid of SEQ ID NO: 362 and SEQ ID NO: 204, head-to-head orientation).

20

25

DETAILED DESCRIPTION

The following examples are provided to illustrate some embodiments of the present disclosure, but are not intended to limit the scope of the disclosure; it will be understood by their exemplary nature that other procedures, methodologies, or techniques known to those skilled in the art may alternatively be used.

30

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 *HTT* gene. Such mutations comprise an expanded CAG trinucleotide repeat in exon 1 of the *HTT* gene. The compositions and methods described herein employ pre-mRNA trans-splicing molecules for gene therapy (e.g., *in vivo* gene therapy, as, e.g., delivered by adeno-associated virus) to treat diseases caused by an expanded CAG trinucleotide repeat in *HTT*, such as HD.

35

The compositions and methods described herein also employ pre-mRNA trans-splicing molecules for gene therapy (e.g., *in vivo* gene therapy) in combination with other therapeutic agents described herein to treat diseases caused by at least one mutation in *HTT*, such as HD.

- 5 Compositions and methods described herein also involve therapeutic molecules that knock down expression of MSH3. The MSH3 knockdown constructs can be used alone in therapeutic methods or can be used in combination with HTT-correcting therapeutic agents and/or HTT knockdown therapeutic agents. MSH3 knockdown constructs can also be used in combination with other therapeutic agents other than HTT-correcting therapeutic agents, such as other
10 therapeutic agents that are designed to correct trinucleotide repeat expansion disorders.

As described herein, HD is an inherited progressive neurodegenerative disorder, for which only palliative therapy is available. Such therapy includes drugs, physical therapy, and counseling, which interventions provide some symptomatic relief. Although presentation of HD varies, the
15 disease is characterized by a progressive loss in the ability to control movement, regulate emotions, and maintain cognitive attributes. HD typically presents in people in their 30s and 40s. The disease is associated with loss of pyramidal neurons in the cortex, loss of medium spiny neurons in the striatum, and loss of hypothalamic neurons. The genetic cause of HD is the autosomal dominant inheritance of an expanded CAG trinucleotide repeat in exon 1 of the *HTT*
20 gene, wherein the presence of over 40 repeats of CAG in this region is disease-causing. See FIG. 1.

The *HTT* locus is large, spanning 180 kb and consisting of 67 exons. Expression of the *HTT* gene is required for normal development. Although HTT protein is widely expressed, the brain is
25 most severely impacted by pathological expansion of CAG trinucleotide repeats, with early pathological effects observed in the striatum and motor cortex. The underlying mechanism of HD pathogenesis is the somatic CAG repeat expansion in *HTT* that occurs in affected brain regions (e.g., striatum) of HD patients. See FIG. 2. Human genetic evidence implicates genes in the DNA mismatch repair pathway (ex. *MSH2*, *MSH3*, *FAN1*, *MLH1*) in controlling this process
30 and modifying the clinical outcome of HD.

Despite considerable effort directed to developing HD therapeutics by many biotech and pharmaceutical companies, there are currently no disease-modifying treatments for HD. The present inventors sought to address this need using a variety of different modalities, each of
35 which may be implemented alone or in combination to provide therapeutic intervention for HD.

Exon replacement by HTT pre-mRNA trans-splicing

Exon replacement by pre-mRNA trans-splicing is well suited as a therapeutic approach for HD because of its ability to replace mutant *HTT* exon 1 while sparing the wild-type copy of *HTT*. The approach can, moreover, theoretically treat 100% of the HD population since it can address the full spectrum of genetic variability in HD patients. In other words, it is not limited to addressing specific HD patient-specific variations (e.g., SNPs), but rather serves as a pan-specific therapeutic agent capable of correcting genetic defects across the entirety of an exon/s.

As described herein, the present inventors designed and tested RNA Exon Editors that target intron 1, intron 2, or intron 3 of the *HTT* pre-mRNA. Accordingly, these RNA Exon Editors can replace all of exon 1, exons 1 and 2, or exons 1-3 of *HTT* mRNA, thereby correcting any and all pathogenic mutations present in exon 1, exons 1 and 2, or exons 1-3 (e.g., correcting the expanded CAG trinucleotide repeat in exon 1 to a wild-type number), respectively. As shown in FIG. 3, RNA Exon Editing via trans-splicing enables the replacement of mutant *HTT* exon 1 with wild-type *HTT* exon 1 using any one of an Intron 1-targeting Exon Editor, an Intron 2-targeting Exon Editor, or an Intron 3-targeting Exon Editor.

In some embodiments, an RNA Exon Editor targeting intron 1 of the *HTT* pre-mRNA comprises a CMV promoter (SEQ ID NO: 137) (used here for *in vitro* experimentation in HEK293 cells), a 5' UTR (e.g., 5' UTR of *HTT* (SEQ ID NO: 136 or 192)), an epitope tag (e.g., a 3X FLAG tag (SEQ ID NO: 4)) for the detection of on-target (ONT) HTT protein generated following successful trans-splicing, an *HTT* exon 1 (SEQ ID NO: 3) (which may be codon-modified followed by the native sequence), a splice donor sequence (GTAAGT), a linker (e.g., the 40mer linker (SEQ ID NO: 37)), a binding domain targeting any one of various regions in *HTT* intron 1 pre-mRNA, and a triple helix terminator (e.g. MALAT1 terminator (SEQ ID NO: 5 or a modified version thereof, e.g., SEQ ID NO: 6)). See, e.g., FIG. 4. In some embodiments, the RNA Exon Editor targeting intron 1 of the *HTT* pre-mRNA does not comprise one or more of a CMV promoter (SEQ ID NO: 137), a 5' UTR (e.g., 5' UTR of *HTT* (SEQ ID NO: 136 or 192)), an epitope tag (e.g., a 3X FLAG tag (SEQ ID NO: 4)), the splice donor sequence GTAAGT, the 40mer linker (SEQ ID NO: 37), a triple helix terminator (e.g. MALAT1 terminator (SEQ ID NO: 5 or a modified version thereof, e.g., SEQ ID NO: 6)). In some embodiments, the CMV promoter may be replaced with a different promoter. Promoters may be chosen to have properties suited to *in vivo* studies and for a therapeutic agent comprising an RNA Exon Editor. In some embodiments, an exemplary RNA Exon Editor does not comprise a FLAG tag or any epitope tag. In such embodiments, a therapeutic agent comprising an RNA Exon Editor may not comprise any epitope tag, which may reduce the potential for immunogenicity. FIG. 5 shows that the level of *HTT* exon replacement

(trans-splicing efficiency) varies depending on where the binding domain targets in intron 1. Results presented herein demonstrate that targeting the 3' end of intron 1 near the branchpoint correlates with increased trans-splicing efficiency into the *HTT* pre-mRNA. See FIG. 6.

5 The present inventors next sought to improve trans-splicing efficiency of *HTT* intron 1-targeting Exon Editors even further by selecting an exemplary binding domain, *HTT_intron1_11704_100*, and determining the effects of different linkers operably linked to this binding domain in the context of RNA Exon Editors. See, e.g., FIG. 7 which presents a schematic of the arrangement of an exemplary RNA Exon Editor comprising a binding domain, *HTT_intron1_11704_100*. Results
10 presented herein demonstrate that several linkers identified by extensive trial and error procedures (23mer GU linker, 33mer GU linker, 34mer GU linker, 41mer GU linker, 40mer_2, 41mer_2, 68mer, 84mer, and 60mer) confer higher levels of trans-splicing efficiency relative to that conferred by the 40mer (SEQ ID NO: 37). See, e.g., FIG. 8. The sequences of the 23mer GU linker (~69-70% guanine content), 33mer GU linker (~66-67% guanine content), 34mer GU
15 linker (~67-68% % guanine content), 41mer GU linker (~73-74% % guanine content), 40mer_2, 41mer_2, 68mer, 84mer, and 60mer correspond to SEQ ID NOs: 38-46, respectively.

The present inventors next tested RNA Exon Editors that target *HTT* intron 2. See FIG. 9 for a general schematic of *HTT* intron 2-targeting RNA Exon Editors. Similar to intron 1-targeting Exon
20 Editors, a binding domain scan in intron 2 was initially performed to identify binding domains that bind to target regions in intron 2 that are associated with and promote high levels of trans-splicing efficiency. See, e.g., FIG. 10. The present inventors determined that the 3' end of the intron, 10-20 nucleotides (nt) upstream of the branchpoint A, is a particularly effective region to target for trans-splicing into *HTT* intron 2. See, e.g., FIG. 11. Additional analyses examining the
25 association of binding domain length in the region of the intron upstream of the branchpoint with functionality identified binding domain lengths ranging from 125-200 nt as having the highest relative trans-splicing efficiencies. See, e.g., FIG. 12.

The present inventors further investigated combinatorial functionality of binding domains
30 operably linked to different linkers as shown pictorially in FIG. 13. The present inventors examined the effects of a variety of exemplary linkers on the trans-splicing activity of a high level performing binding domain for *HTT* intron 2 (*HTT_intron2_12061_150*). As shown in FIG. 14, the trans-splicing activity of RNA exon editors comprising *HTT_intron2_12061_150* operably linked to the indicated linkers did not vary significantly. Indeed, only a minimal effect was observed
35 when *HTT_intron2_12061_150* was operably linked to different linkers, suggesting that trans-splicing using *HTT_intron2_12061_150* is highly optimized.

Results presented herein showing that none of the linkers tested significantly improved trans-splicing efficiency compared to that conferred by the 40mer linker suggests that *HTT*_intron2_12061_150 Exon Editors have achieved high efficiency via binding affinity and recruitment of the relevant spliceosome machinery. The results also indicate that different linker sequences may be included in an *HTT* RNA Exon Editor that is effective to induce trans-splicing. The present inventors tested experimentally whether inhibition of cis-splicing might contribute to enhanced trans-splicing by designing anti-sense oligonucleotides (ASOs) that specifically block the competing cis-splicing site (ASO8 (SEQ ID NO: 133), ASO9 (SEQ ID NO: 134), ASO10 (SEQ ID NO: 135)), as well as cis-splicing sites involved in the splicing of the upstream intron (ASO2-7 (SEQ ID NOs: 127-132, respectively)), and co-expressing each of these ASOs with an RNA Exon Editor targeting intron 2. See FIG. 16. FIG. 17 shows that *HTT* intron 2-targeting Exon Editor (*HTT*_intron2_12061_150) in combination with ASO6 (SEQ ID NO: 131), which blocks cis-splicing of the upstream intron, led to an improvement in trans-splicing efficiencies *in vitro*. Some embodiments described herein involve strategies to block these cis-splicing events using anti-sense RNA expressed in the same plasmid as the *HTT* RNA Exon Editor. See also the Combinatorial Embodiments section below.

In order to confirm the presence of a *bona fide* chimeric *HTT* protein generated following trans-splicing, Western blot analysis was performed on cellular lysates made from HEK293 cells transfected with select *HTT* Exon Editors wherein the blots were probed using antibodies that specifically bind the N-terminal FLAG epitope. As shown in FIG. 18, Western blotting confirmed on-target (ONT) trans-splicing and resultant generation of on-target protein generated following trans-splicing by detecting FLAG-tagged *HTT* protein having the predicted molecular weight. Notably, the intensity of the anti-FLAG ONT protein band correlated with trans-splicing efficiency (% RNA replacement). See FIG. 18.

The present inventors also designed and tested RNA Exon Editors that target *HTT* intron 3. See FIG. 23 for a general schematic of *HTT* intron 3-targeting RNA Exon Editors. Similar to intron 1 and intron 2-targeting Exon Editors, a binding domain scan was initially performed to identify binding domains that bind to target regions in intron 3 that are associated with and promote high levels of trans-splicing efficiency. See, e.g., FIG. 24. The present inventors determined that the 3' end of the intron, 5-15 nt upstream of the branchpoint A, is a particularly effective region to target for trans-splicing into *HTT* intron 3.

35

To compare how intron 2-targeting Exon Editors perform relative to intron 3-targeting Exon Editors, the present inventors performed transfections of Exon Editors containing the highest performing binding domains for each target intron in parallel and analyzed the RNA by RT-ddPCR. Results presented herein demonstrate that HTT intron 3-targeting Exon Editors exhibit
5 slightly higher trans-splicing efficiency when compared to intron 2-targeting Exon Editors. See, e.g., FIG. 25.

To investigate cryptic self-splicing (either by cis-splicing of AAV concatemers or inter-molecular trans-splicing) in HTT Exon Editors and constructs comprising same, cryptic splice sites were
10 identified using in silico prediction of self-splicing sites (see Table 1) to mitigate such unwanted events. An exemplary cryptic splice site mitigated HTT intron 2-targeting Exon Editor (SEQ ID NO: 203) was generated (in which 4 of the identified self-splicing sites were altered) and tested for trans-splicing efficiency and compared to that of the original (unaltered) Exon Editor (SEQ ID NO: 95). Results presented herein confirm that the minor sequence changes applied do not
15 influence trans-splicing activity of the exemplary cryptic splice site mitigated Exon Editor. See, e.g., FIG. 26.

Initial rounds of HTT Exon Editor molecules were expressed from a CMV promoter (SEQ ID NO: 137) and contained the native HTT 5' UTR (SEQ ID NO: 136). The present inventors next
20 explored how the Exon Editors perform when expressed from a CAGGS promoter (SEQ ID NO: 196). Since the CAGGS promoter contains a 5' UTR (CBA exon + rabbit beta-globin exon), testing of Exon Editors with or without the native HTT 5' UTR in combination with the CAGGS promoter were included in this study. Of note, although results are presented in the context of HTT intron 2-targeting Exon Editors, the effects of these modifications are applicable to Exon
25 Editors targeting other introns.

In HEK293 cell transfection experiments, the present inventors found that the CAGGS promoter-driven Exon Editor that contains both the CAGGS 5' UTR and the HTT 5' UTR possessed similar % replacement of HTT mRNA to that driven by a CMV promoter. See, e.g., FIG. 27. When the
30 HTT 5' UTR was removed from the CAGGS promoter-driven Exon Editor, a decrease in % replacement of HTT RNA was observed and a significant increase in ONT protein expression was observed by Western Blot, suggesting the presence of a protein translation activator element in the CAGGS 5' UTR. Concurrently, the present inventors observed increased non-spliced protein (NSP) and background (possible OFT) protein signals with the removal of the
35 HTT 5' UTR.

To explore NSP reduction strategies in the context of 5' HTT Exon Editors, a variety of elements were incorporated into HTT intron 1-targeting Exon Editors and HTT intron 2-targeting Exon Editors and tested to evaluate their impact, singly and in combination, on Exon Editor performance. In the context of a minimal NSP context (i.e., stop codon in or immediately following the splice donor sequence in the Exon Editor to minimize neoantigens), these NSP reduction strategies include: 1) three tandem repeats of U1 snRNA binding site (3X UBS) (SEQ ID NO: 345) and 2) an AU-rich element (ARE) (SEQ ID NO: 346).

In HTT intron 1-targeting Exon Editors, the present inventors demonstrated that the inclusion of 3X UBS in the linker domain reduced NSP levels by ~75% relative to that of the baseline 40mer-only linker control. Inclusion of ARE in the linker reduced the levels of NSP by ~40% relative to that of the baseline 40mer-only linker control. The combination of 3X UBS and ARE in the Exon Editor reduced the NSP level by ~88% relative to that of the baseline 40mer-only linker control. See, e.g., FIG. 28.

Similarly, in HTT intron 2-targeting Exon Editors, the present inventors demonstrated that the inclusion of 3X UBS in the linker domain reduced NSP levels by ~38% relative to that of the baseline 40mer-only linker control. Inclusion of ARE in the linker reduced the levels of NSP by ~33% relative to that of the baseline 40mer-only linker control. The combination of 3X UBS and ARE in the Exon Editor worked combinatorially in reducing the NSP level by approximately 66% as compared to the baseline 40mer-only linker control. See, e.g., FIG. 29.

RNA exon editors targeting an *HTT* intron (e.g., intron 1, intron 2, or intron 3) may include target-specific elements such as, for example, a binding domain specific for an *HTT* target intron and a coding domain sequence that encodes an *HTT* coding sequence (e.g., a sequence encoding all or part of exon 1, exon 2, or exon 3, or any combination thereof (e.g., exons 1 and 2 or exons 1-3)). RNA exon editors targeting an *HTT* intron may include one or more target-independent elements that may improve the functioning of an exon editor. Target-independent elements may include, for example, an epitope tag, a linker, a splice donor sequence, one or more repeats of a U1 snRNA binding site, an AU-rich element, or a terminator. Embodiments described herein may include one or more of such target-independent elements. Embodiments described herein may exclude one or more of such target-independent elements. RNA exon editors that lack one or more target-independent elements described herein may be capable of effecting exon editing. For example, in some embodiments, RNA exon editors described herein do not include a triple helix terminator, a MALAT-1 terminator, or any terminator sequence. In some embodiments, RNA exon editors described herein do not include a CMV promoter or CAGGS promoter. In

some embodiments, RNA exon editors described herein do not include any 5' UTR sequence, or do not include an *HTT* 5' UTR sequence. In some embodiments, RNA exon editors described herein do not include a GTAAGT splice donor sequence. In some embodiments, RNA exon editors described herein do not include any epitope tag, or do not include a 3X FLAG tag. In some embodiments, RNA exon editors described herein do not include any of the linkers described herein, including any of SEQ ID NOs: 37-46. In some embodiments, RNA exon editors described herein do not include three tandem repeats of U1 snRNA binding site (3X UBS) (SEQ ID NO: 345). In some embodiments, RNA exon editors described herein do not include an AU-rich element (ARE) (SEQ ID NO: 346).

Inhibition of somatic CAG repeat expansion

There is mounting evidence that the underlying mechanism of pathogenesis for HD is somatic instability of the CAG repeat tract. This supports pursuing factors that modify somatic expansion as viable therapeutic targets. The present inventors designed and tested strategies to inhibit somatic CAG expansion. See, e.g., FIG. 19. Recent Genome Wide Association Studies (GWAS) performed in HD patients identified many components of the DNA mismatch repair pathway as key genetic modifiers associated with the rate of disease progression. For example, in HD patients, single nucleotide variants (SNVs) in the MSH3 gene that reduce its expression are linked to delayed age of onset of the disease. MSH3 is speculated to be a good target for a HD therapeutic and evidence suggests that MSH3 knockdown leads to inhibition of somatic CAG repeat expansion both in vitro and in vivo. Here, the present inventors pursued several approaches to reduce MSH3 levels.

The present inventors explored using trans-splicing in a novel manner to reduce expression levels of a target gene, by trans-splicing into the target pre-mRNA to produce an RNA that is quickly degraded or that cannot produce a functional protein. The present inventors envisioned that either a 5' Exon Editor or a 3' Exon Editor can achieve this outcome. A 5' Exon Editor can replace one or more of the 5' exons of the target mRNA to remove the start codon and some or all of the coding sequence, replacing it with an alternative coding sequence or a non-coding sequence. A 3' Exon Editor can replace one or more of the 3' exons of the target, replacing them with an alternative coding sequence or a non-coding sequence. In the resulting chimeric mRNA, there may be no translation of any amino acids from the target mRNA. 5' Exon Editors can also insert 5' untranslated regions (UTRs) that prevent translation. 3' Exon Editors can insert 3' UTRs that de-stabilize the transcript or prevent it from being exported from the nucleus or translated. The portion of the target mRNA remaining after the trans-splicing reaction is presumed to be quickly degraded, since it will lack either a 5' cap or a 3' polyA tail.

The present inventors reasoned that Exon Editors that edit a target, such as *HTT*, can at the same time knockdown a second gene, such as *MSH3*, through the inclusion of two different binding domains, one targeting *HTT* and the other *MSH3*. An embodiment of how this dual effect can be achieved is that a 5' *HTT* exon editor containing a 5' portion of *HTT* can both trans-splice to *HTT* to reconstitute functional wild-type *HTT* while also trans-splicing to *MSH3* to produce a non-functional *MSH3* mRNA. In an embodiment of which, the chimeric non-functional *HTT-MSH3* mRNA encodes an *HTT* polypeptide that terminates at the first in-frame stop codon in the *MSH3* portion of the mRNA. Thus, a normal, wildtype *MSH3* pre-mRNA is transformed into a chimeric non-functional *HTT-MSH3* mRNA that encodes few to no amino acids from the *MSH3* coding sequence and is likely to undergo nonsense-mediated decay. As shown herein, this approach works with binding domains targeting several different *MSH3* introns and indeed, would work with binding domains targeting any *MSH3* intron. Moreover, although the examples presented here relate to a 5' Exon Editor, a similar approach could be implemented with a 3' Exon Editor with a reasonable expectation for success.

Accordingly, the present inventors have designed and tested strategies to treat HD at an earlier stage of pathogenesis by inhibiting somatic CAG expansion. Embodiments described herein, such as those described for inhibiting somatic CAG expansion, may be used alone or in conjunction with RNA exon editor-mediated trans-splicing to correct pathogenic mutations. Although inhibition of somatic expansion alone is a viable therapeutic approach for HD, the present inventors have tested hybrid therapeutic approaches and dual action/hybrid molecules designed to inhibit somatic CAG expansion and correct *HTT* exon 1 RNA. See, e.g., FIGs. 19 and 20. Accordingly, embodiments described herein, such as those designed to inhibit somatic CAG expansion, may be used alone or in conjunction with RNA exon editor-mediated trans-splicing to correct pathogenic mutations.

Further to the above, to inhibit an underlying mechanism of HD pathogenesis (i.e., somatic CAG expansion) while also replacing mutant *HTT* exon 1, the present inventors designed a set of dual action/hybrid RNA Exon Editors that contain targeting sequences for both *HTT* and *MSH3* pre-mRNAs. In some embodiments, Exon Editor expression is driven by a CMV promoter. In some embodiments, an exemplary Exon Editor comprises the *HTT* 5' UTR, N-terminal 3X FLAG tag, an *HTT* exon 1 coding sequence, a splice domain (splice donor site), a linker, an *MSH3* binding domain, an *HTT* binding domain, and a terminator sequence. The *HTT* binding domain targets the Exon Editor to the *HTT* pre-mRNA and leads to the production of corrected *HTT* RNA after successful trans-splicing. The *MSH3* binding domain targets the Exon Editor to the *MSH3* pre-

mRNA, leading to the production of a chimeric *HTT* exon 1 - *MSH3* RNA molecule with a premature stop codon (due to a frameshift resulting from the replacement of upstream *MSH3* exons with *HTT* Exon 1), which is subject to nonsense-mediated decay (NMD) and thus, leads to the subsequent reduction of *MSH3* expression. See, e.g., FIG. 20.

5

Using Exon Editors that contain tandem binding domains, the present inventors' experiments showed successful trans-splicing into both *HTT* and *MSH3* pre-mRNAs, as measured by RT-qPCR designed to detect chimeric "corrected" *HTT* RNA and chimeric *HTT* exon 1-*MSH3* RNA molecules, respectively. See, e.g., FIGs. 21 and 22. The sequences of the *MSH3* binding domains used in the tested constructs are set forth in SEQ ID NOs: 140 and 142 (FIG. 21) and in SEQ ID NOs: 144 and 146 (FIG. 22), the sequences of the *HTT* binding domains used in the tested constructs are set forth in SEQ ID NOs: 20, 141, and 145 (FIGs. 21, 22), and the sequence of the *HTT* CDS in the tested constructs is set forth in SEQ ID NO: 3. It is noteworthy that the chimeric *HTT* exon 1 - *MSH3* RNA levels might be underestimated due to NMD. To address this, the present inventors also measured *MSH3* RNA transcript levels.

10

15

In another embodiment, the present inventors also designed and tested tandem binding domain Exon Editors that target *HTT* intron 2 (using *HTT_intron2_12061_150* binding domain (SEQ ID NO: 72)) in conjunction with *MSH3* pre-mRNA. See, e.g., FIGs. 30-32. The sequences of the *MSH3* binding domains used in the tested constructs are set forth in SEQ ID NOs: 140, 209, 144, and 210 and the sequence of the *HTT* CDS in the tested constructs is set forth in SEQ ID NO: 59.

20

RT-qPCR analysis of tandem binding domain Exon Editors targeting *HTT* intron 2 and *MSH3* intron 5 showed successful trans-splicing into both *HTT* and *MSH3*, as measured by RT-qPCR designed to detect chimeric corrected *HTT* RNA and chimeric *HTT* exon 2-*MSH3* exon 6 RNA molecules. See, e.g., FIGs. 31A and 31B. *MSH3_intron5_188_150* (SEQ ID NO: 209) had higher efficiency of *MSH3* trans-splicing compared to *MSH3_intron5_213_100* (SEQ ID NO: 140). Similarly, tandem binding domain Exon Editors targeting *HTT* intron 2 and *MSH3* intron 15 showed successful trans-splicing into both *HTT* and *MSH3*. See, e.g., FIGs. 32A and 32B. It is noteworthy that the chimeric *HTT* exon 2 - *MSH3* RNA levels might be underestimated due to NMD.

25

30

Results presented herein demonstrate that tandem binding domain Exon Editors are effective and are likely to have therapeutic benefits. Furthermore, these experiments identified efficient binding domains for trans-splicing targeting of *MSH3* pre-mRNA. Beyond their use in a tandem

35

binding domain context as described here, these MSH3-targeting binding domains can also be used in the context of an Exon Editor with an NMD-inducing coding sequencing; such Exon Editors can be used alone to target MSH3 knockdown or be packaged in the same AAV vector with an HTT-targeting Exon Editor and co-delivered to the target tissue for a combinatorial approach.

In some embodiments, an MSH3 exon editor can be used to knock down MSH3 expression, wherein the MSH3 exon editor has a binding domain targeting an intron of *MSH3* pre-mRNA, as described above, but does not include a binding domain targeting *HTT* pre-mRNA or any other pre-mRNA. Dual binding domain RNA exon editors described above may be modified by removing a binding domain that targets an *HTT* intron. In some embodiments, *MSH3* exon editors may be administered without also administering a treatment that targets or corrects another gene sequence. In some embodiments, such MSH3 exon editors may be used in a method of treating or preventing a trinucleotide repeat expansion disorder.

MSH3 knockdown by RNAi

RNA interference (RNAi) is a natural mechanism by which double-stranded RNA (dsRNA) induces gene silencing in a sequence-specific manner by targeting mRNA for degradation. In an initial design to identify vectorized RNAi constructs that could efficiently knockdown MSH3, various constructs based on one miRNA sequence (*MSH3* Ex23-targeting, TTAATCCATAACTCCTTGC; SEQ ID NO: 224) were designed. Pol III promoter-driven shRNA were constructed as potential positive controls. These U6 promoter-driven shRNAs were tested alongside constructs that express Pol II promoter (CMV promoter)-driven primary miRNA (pri-miRNA)-like transcripts targeting *MSH3*. See, e.g., FIG. 33.

As shown in, e.g., FIG. 34, Western blot analysis revealed that for this *MSH3* Ex23-targeting miRNA (TTAATCCATAACTCCTTGC; SEQ ID NO: 224), around 50% MSH3 protein knockdown was observed when expressed in a miR-30a scaffold context. Furthermore, efficiencies of knockdown were not improved when placing the guide RNA on the 5' arm or adding a bulge to the stem structure (37% and 23% knockdown, respectively).

Additional designs focused on varying the miRNA active sequence showed that in addition to the exon 23-targeting sequence tested in FIG. 34, exon 22-targeting (SEQ ID NO: 246), exon 9-targeting (SEQ ID NO: 248), exon 12-targeting (SEQ ID NO: 256), and exon 15-targeting (SEQ ID NO: 257) miRNA sequences performed equally well with around 50% MSH3 protein knockdown observed. See, e.g., FIG. 35. This MSH3 RNAi modality alone can be therapeutically

relevant for many repeat expansion disorders and/or could be combined with an Exon Editor for enhanced therapeutic potential.

In some embodiments, MSH3-targeting RNAi constructs disclosed herein may be used in conjunction with RNA exon editors that target HTT, as described above, or that target other genes associated with trinucleotide repeat expansion diseases. In other embodiments, MSH3-targeting RNAi constructs disclosed herein may be used independently of HTT-targeting therapeutic approaches. In some embodiments, MSH3-targeting RNAi constructs disclosed herein may be administered to treat or prevent a trinucleotide repeat expansion disease.

Splice modulation of MSH3

MSH3 can be inactivated by incorporating sequences that are complementary to MSH3 splice junctions into snRNA sequences, such as the U7 snRNA. The present inventors designed and tested strategies to block cis-splicing events by antisense RNA expressed from the same plasmid as the HTT-targeting RNA Exon Editor. Modified snRNAs such as the U7 SmOPT were designed by 1) changing the targeting sequence (e.g., the histone binding sequence at the 5' region of U7 snRNA) to the complementary sequence of the gene to be modified, and 2) changing the binding site (AAUUUGUCUAG; SEQ ID NO: 367; U7 Sm WT) for U7 snRNP specific proteins (Lsm10 and Lsm11) to the consensus sequence derived from major spliceosomal uridine-rich small nuclear ribonucleoproteins (U snRNPs) (AAUUUUUGGAG; SEQ ID NO: 368; U7 Sm OPT), leading to the formation of a spliceosomal-type heptameric protein core wrapped around U7 Sm OPT. In some embodiments, asRNA molecules described herein consist of the U1 promoter, snRNA (with asRNA sequence targeting *MSH3* intron-exon boundaries and consensus Sm binding site), and a U1 terminator. See, e.g., FIG. 36.

The present inventors designed and tested MSH3 splice modulators targeting the skipping of *MSH3* exons 2, 3 and 4. For each set, asRNA sequences against the upstream intron – exon boundary (e.g. intron 1/exon 2 boundary for the exon 2 splice modulator (SEQ ID NO: 274)), downstream exon – intron boundary (e.g. exon 2/intron 2 boundary for the exon 2 splice modulator (SEQ ID NO: 275)), and tandem sequences (SEQ ID NOs: 276 and 277) consisting of the two asRNAs were tested. Sequences of exemplary MSH3 splice modulators that mediate exon 2 skipping are shown in SEQ ID NOs: 284-286, which comprise MSH3 splice modulator elements SEQ ID NOs: 274-276, respectively. For MSH3 exon 2 skipping molecules, an asRNA molecule targeting both intron 1/exon 2 and exon 2/intron 2 boundaries without the U7SmOPT snRNA scaffold was also tested. RT-qPCR assays measuring the levels of spliced *MSH3* mRNA products showed that U7SmOPT splice modulators (but not the asRNA sequence that does not

contain the U7SmOPT scaffold) exhibited intended skipping of the target exons. Sequences of exemplary MSH3 splice modulators that mediate exon 3 skipping are shown in SEQ ID NOs: 288-290, which comprise MSH3 splice modulator elements SEQ ID NOs: 278-280, respectively. See, e.g., FIGs. 37-39. Of note, *MSH3* exon 4 skipping will not lead to NMD but will cause a shorter MSH3 product due to exon 4 containing 213 nucleotides, which will result in an in-frame exon 4-skipped mRNA product. Sequences of exemplary MSH3 splice modulators that mediate exon 4 skipping are shown in SEQ ID NOs: 291-293, which comprise MSH3 splice modulator elements SEQ ID NOs: 281-283, respectively. The present inventors observed truncated protein products for exon 4-targeted splice modulators on Western Blots probed with antibodies to visualize MSH3, providing evidence that the splice modulators were skipping the target exons as intended.

Subsequent rounds of designs tested MSH3 splice modulators targeting the skipping of exon 3 (sequences of exemplary MSH3 splice modulators are shown in SEQ ID NOs: 324-326, which comprise MSH3 splice modulator elements SEQ ID NO: 299, SEQ ID NOs: 300 + 298 + 301, and SEQ ID NOs: 302 + 298 + 303, respectively), exon 6 (sequences of exemplary MSH3 splice modulators are shown in SEQ ID NOs: 327-329, which comprise MSH3 splice modulator elements SEQ ID NO: 304, SEQ ID NOs: 305 + 298 + 306, and SEQ ID NOs: 307 + 298 + 308, respectively), exon 7 (sequences of exemplary MSH3 splice modulators are shown in SEQ ID NOs: 330-332, which comprise MSH3 splice modulator elements SEQ ID NO: 309, SEQ ID NOs: 310 + 298 + 311, and SEQ ID NOs: 312 + 298 + 313, respectively), exon 8 (sequences of exemplary MSH3 splice modulators are shown in SEQ ID NOs: 333-335, which comprise MSH3 splice modulator elements SEQ ID NO: 314, SEQ ID NOs: 315 + 298 + 316, and SEQ ID NOs: 317 + 298 + 318, respectively), and exon 15 (sequences of exemplary MSH3 splice modulators are shown in SEQ ID NOs: 336-338, which comprise MSH3 splice modulator elements SEQ ID NO: 319, SEQ ID NOs: 320 + 298 + 321, and SEQ ID NOs: 322 + 298 + 323, respectively). These exons were selected for their suitability for splice modulation, taking into consideration factors such as length of neoepitope if a protein is produced rather than transcripts undergoing NMD, splicing characteristics such as major vs minor spliceosome-mediated splicing, alternative splicing etc.

The present inventors designed and tested MSH3 splice modulators targeting the skipping of *MSH3* exon 3, 6, 7, 8, and 15. For each set, 1) a U7SmOPT molecule containing a sequence antisense to the entire length of the target exon and 12-13 nucleotides into flanking introns, 2) a U7SmOPT molecule containing a sequence antisense to the downstream exon/intron boundary and the upstream intron/exon boundary, separated by an unstructured linker, and 3) a U2

snRNA molecule containing a sequence antisense to the downstream exon/intron boundary and the upstream intron/exon boundary, separated by an unstructured linker, were tested.

Sequences of the tested *MSH3* splice modulators are set forth in SEQ ID NOs: 290, and 324-326 (FIG. 40); SEQ ID NOs: 327-329 (FIG. 41); SEQ ID NOs: 330-332 (FIG. 42); SEQ ID NOs:

5 333-335 (FIG. 43); and SEQ ID NOs: 336-338 (FIG. 44). SEQ ID NOs: indicating the corresponding asRNA region SEQ ID NO: and target region are indicated in each of respective FIGs. 40-44. The present inventors identified effective splice modulators for all target exons tested. In general, the U7SmOPT molecule containing a sequence antisense to the downstream exon/intron boundary and the upstream intron/exon boundary, separated by an unstructured
10 linker (e.g., SEQ ID NOs: 328, 331, 334, and 337), showed a particularly high degree of reduction in MSH3 expression levels, with some exception. See, e.g., FIGs. 40-44. Western blots probed with antibodies to visualize MSH3 indicated that MSH3 protein knockdown levels correspond with the RNA exon skipping efficiency profiles. Sequences of the tested *MSH3* splice modulators are set forth in SEQ ID NOs: 330-332. See, e.g., FIG. 45.

15 In some embodiments, MSH3 splice modulators disclosed herein may be used in conjunction with RNA exon editors that target HTT, as described above, or that target other genes associated with trinucleotide repeat expansion diseases. In some embodiments, MSH3 splice modulators disclosed herein may be used independently of HTT-targeting therapeutic
20 approaches. In some embodiments, MSH3 splice modulators disclosed herein may be administered to treat or prevent a trinucleotide repeat expansion disease.

Knockdown of HTT toxic species

To enhance efficacy of HTT Exon Editors described herein, the present inventors envisioned
25 combinatorial approaches that further address toxic species such as HTT1a. To this end, vectorized hybrid molecules that comprise the HTT Exon Editor, as well as a miRNA that targets exon 1 of unedited *HTT*, which elements were intended to be packaged in a single AAV, were designed and tested. See, e.g., FIG. 46. The sequences of the miRNA (HTT miRNA-1 and HTT miRNA-2) targeting unedited HTT are set forth in SEQ ID NOs: 339 and 342. The CDS of the
30 HTT Exon Editor was codon-modified to prevent interaction between the miRNA and the Exon Editor, avoiding the degradation of the Exon Editor itself as well as corrected HTT products. See, e.g., FIG. 47.

The present inventors first evaluated the efficacy of Exon Editor alone, HTT miRNA alone, and
35 Exon Editor + HTT miRNA molecules when the two modalities are combined. An HTT intron 2-targeting Exon Editor (SEQ ID NO: 204) and HTT miRNA-1 (active sequence SEQ ID NO: 339).

were tested in this experiment. SEQ ID NO: 341 corresponds to the HTT miRNA-1 (SEQ ID: 339) encoding pri-miRNA sequence in mir-33 scaffold (SEQ ID: 259, 260, 261). As shown in, e.g., FIG. 48, the proportion of trans-spliced (edited) *HTT* transcripts among all *HTT* transcripts was increased in the presence of HTT miRNA due to the reduction of total *HTT* transcript copy number. The copy number of trans-spliced edited *HTT* transcripts was, however, unchanged with or without HTT miRNA-1, indicating that there was minimal to no interaction between HTT miRNA-1 and the Exon Editor.

An additional comparison of two independent miRNAs targeting HTT exon 1, HTT miRNA-1 (active sequence SEQ ID NO: 339; encoding pri-miRNA sequence SEQ ID NO: 341) and HTT miRNA-2 (active sequence SEQ ID NO: 342), was performed. SEQ ID NO: 344 corresponds to the HTT miRNA-2 (SEQ ID: 342) encoding pri-miRNA sequence in mir-33 scaffold (SEQ ID: 259, 260, 261). The coding sequence of the HTT intron 2-targeting Exon Editor was further codon-optimized to confer and ensure resistance to both miRNAs (exemplary variant coding domain sequences are set forth in SEQ ID NOs: 59 and 349-351). No negative impact of HTT miRNA-1 or HTT miRNA-2 on trans-splicing was observed, as evidenced by unchanged (or potentially increased) trans-spliced *HTT* copy numbers and higher % trans-spliced *HTT* mRNA by the Exon Editor in the presence of miRNA. See, e.g., FIG. 49. The present inventors demonstrated similar efficiencies of HTT knockdown with HTT miRNA-1 and HTT miRNA-2, with neither miRNA negatively affecting Exon Editor performance as evidenced by unchanged copy numbers of the trans-spliced *HTT* transcripts with or without each miRNA. See, e.g., FIG. 49.

The present inventors next evaluated the HTT Exon Editor + miRNA dual molecules at the protein level. Testing of HTT intron 2-targeting Exon Editor + HTT miRNA dual hybrid molecule indicated that in the presence of the HTT miRNA (HTT miRNA-1; SEQ ID NO: 339), ~50% knockdown of unedited *HTT* transcripts was achieved (FIG. 50, left panel). Analysis of Exon Editor trans-splicing activity showed that % edited *HTT* RNA was increased in the presence of HTT miRNA due to reduction of unedited and, therefore, total *HTT* copy numbers by the miRNA (FIG. 50, right panel).

Construction of hybrid molecules

As described herein, the present inventors explored a hybrid approach wherein multiple mechanisms and/or pathogenic species were targeted. As detailed herein, these approaches can include replacement of mutant HTT via an Exon Editor, knockdown of mutant HTT and related transcripts (e.g., HTT1a) via RNAi, and reduction of MSH3 via: trans-splicing or vectorized splice modulation. These approaches, when combined, represent a multi-modal

mechanism of action that can nonetheless be delivered in a single AAV. A subset of the designs aimed at knocking down MSH3 are by design a hybrid molecule (e.g. tandem binding domain Exon Editor). Similarly, the present inventors designed the HTT-targeting miRNA with the idea that it would be combined with the HTT Exon Editor; hence, the additional consideration of codon-modifying the miRNA target site in the Exon Editor was deployed to minimize interaction between the HTT-targeting miRNA and the Exon Editor. Such combinatorial approaches may be used to achieve a high degree of therapeutic efficacy in some, if not all, HD patients. Such approaches may, indeed, confer dramatic disease-modifying effects.

10 **HTT Exon Editor + MSH3 Splice Modulator**

Although inhibition of somatic expansion alone may have a positive therapeutic effect on HD, a hybrid molecule that inhibits somatic CAG expansion and corrects *HTT* exon 1 RNA could potentially have significant therapeutic promise due to the ability of this approach to combat the disease simultaneously on multiple fronts. Any of the HTT-targeting RNA exon editors described herein may be used in combination with any one or more of the MSH3 splice modulators described herein. Any of the HTT-targeting RNA exon editors described herein may be used without also using any of the MSH3 modulators described herein, and any of the MSH3 splice modulators described herein may also be used without also using any of the HTT-targeting RNA exon editors.

20

To achieve this combinatorial effect, hybrid vectors encoding the HTT Exon Editor + MSH3 Splice modulator were designed. See, e.g., FIG. 51. Initial examinations of these dual hybrid molecules used *HTT* intron 2-targeting (HTT_intron2_12061_150) Exon Editor (SEQ ID NO: 204) and an exemplary *MSH3* exon 7 skipping Splice Modulator (SEQ ID NO: 331). However, the combinatorial effect would be similarly applicable to any combination of HTT Exon Editor and MSH3 Splice Modulator (sequences of exemplary MSH3 splice modulators are shown, e.g., in SEQ ID NOs: 330-332, which comprise MSH3 splice modulator elements SEQ ID NO: 309, SEQ ID NOs: 310 + 298 + 311, and SEQ ID NOs: 312 + 298 + 313, respectively and SEQ ID NOs: 284-293, 324-329, and 333-338. It is further understood that the combinatorial effect observed with the exemplary HTT intron 2-targeting (HTT_intron2_12061_150) Exon Editor (SEQ ID NO: 204) described here would be similarly applicable to other HTT intron 2-targeting (HTT_intron2_12061_150) Exon Editors, including those comprising any one of SEQ ID NOs: 59, 350, or 351 corresponding to the modified coding domain sequence. Accordingly, a HTT intron 2-targeting Exon Editor could comprise any one of variant coding domain sequences as set forth in any one of SEQ ID NOs: 59 and 349-351). It is further understood that the combinatorial effect observed with the exemplary HTT intron 2-targeting (HTT_intron2_12061_150) Exon

35

Editor (SEQ ID NO: 204) described here would be similarly applicable to either HTT intron 3-targeting (comprising, e.g., HTT_intron3_4223_150 or HTT_intron3_4233_150) Exon Editors or HTT intron 1-targeting (comprising, e.g., HTT_intron1_11704_100 or HTT_intron1_11724_100) Exon Editors. As shown herein, the present inventors examined Exon Editor alone, MSH3 Splice Modulator alone, and the Exon Editor + MSH3 Splice Modulator dual hybrid molecule for their trans-splicing profiles as well as MSH3 knockdown efficiencies. RT-qPCR analysis showed no change in Exon Editor performance with or without the MSH3 Splice Modulator, as indicated by similar levels of % HTT replacement. See, e.g., FIG. 52A. Similarly, the performance of the MSH3 Splice Modulator is unchanged with or without the Exon Editor. See, e.g., FIG. 52B. This experiment demonstrated that there is no interaction between the Exon Editor and the MSH3 Splice Modulator when co-delivered in the same DNA fragment.

Efficient trans-splicing requires sufficient levels of AAV delivery to the target tissue and high levels of Exon Editor expression. To achieve high levels of vector expression, self-complementary AAV (scAAV) was tested. To compare the performance of conventional single-stranded AAV (ssAAV) to scAAV, an exemplary *HTT* intron 2-targeting Exon Editor + *MSH3* exon 7 skipping Splice Modulator dual hybrid molecule was inserted into a ssAAV-compatible plasmid (containing wild-type ITRs) and a scAAV-compatible plasmid (ITR/ITR- Δ trs), packaged into AAV2 and used for transduction of HEK293 cells. Characterization of MSH3 expression, both at the RNA and protein levels, suggested that dual hybrid molecules had better performance for both modalities when packaged in scAAV (see, e.g., FIG. 53, left panel). Trans-splicing characterization also supported the higher performance of molecules packaged in scAAV compared to ssAAV, as evidenced by higher % replacement of the *HTT* transcript and corresponding elevated full-length trans-spliced HTT protein detected by Western Blot (see, e.g., FIG. 53, right panel). The present inventors also assayed the expression level of the Exon Editors and observed higher transgene expression in cells transduced with scAAV. A consideration when using scAAV vectors is the limited cargo space (about half the size compared to ssAAV), which limits its use when driving Exon Editor expression from a larger promoter (e.g., CAGGS promoter).

30

HTT Exon Editor + HTT miRNA + MSH3 Splice Modulator

The present inventors also designed, constructed, and tested triple hybrid molecules that contain multiple modalities that target multiple mechanisms and/or pathogenic species. These approaches include replacement of mutant HTT via an Exon Editor, knockdown of mutant HTT and related transcripts (e.g., HTT1a) via RNAi, and reduction of MSH3 via splice modulation. These approaches, when combined, represent a multi-modal mechanism of action that can

35

nonetheless be delivered in a single AAV. Such combinatorial approaches may be used to achieve a high degree of therapeutic efficacy in some, if not all, HD patients. Such approaches may, indeed, confer dramatic disease-modifying effects. As described above, each of these modalities are designed to minimize interaction with each other (e.g., codon modification of the Exon Editor that has resistance to the HTT miRNA).

To explore this approach, an analysis of triple hybrid molecules that contained the CAGGS promoter (SEQ ID NO: 196)-driven *HTT* intron 2-targeting Exon Editor (SEQ ID NO: 204), an exemplary *MSH3* exon 7-skipping Splice Modulator (SEQ ID NO: 331), and SEQ ID NO: 341 [HTT miRNA-1 (SEQ ID: 339) encoding pri-miRNA sequence in mir-33 scaffold (SEQ ID: 259, 260, 261)] or SEQ ID NO: 344 [HTT miRNA-2 (SEQ ID: 342) encoding pri-miRNA sequence in mir-33 scaffold (SEQ ID: 259, 260, 261)] was performed. The activity of the triple hybrid was compared to a control hybrid molecule that contained a splice mutant Exon Editor, a Splice Modulator that contained a scrambled control sequence, and a miRNA that contained a scrambled control sequence. The combinatorial effect would be similarly applicable to any combination of HTT Exon Editor, MSH3 Splice Modulator (sequences of exemplary MSH3 splice modulators are shown, e.g., in SEQ ID NOs: 330-332, which comprise MSH3 splice modulator elements SEQ ID NO: 309, SEQ ID NOs: 310 + 298 + 311, and SEQ ID NOs: 312 + 298 + 313, respectively and SEQ ID NOs: 284-293, 324-329, and 333-338), and HTT miRNA-1 or HTT miRNA-2. It is further understood that the combinatorial effect observed with the exemplary HTT intron 2-targeting (HTT_intron2_12061_150) Exon Editor (SEQ ID NO: 204) described here would be similarly applicable to other HTT intron 2-targeting (HTT_intron2_12061_150) Exon Editors, including those comprising any one of SEQ ID NOs: 59, 350, or 351 corresponding to the modified coding domain sequence. Accordingly, a HTT intron 2-targeting Exon Editor could comprise any one of variant coding domain sequences as set forth in any one of SEQ ID NOs: 59 and 349-351). It is further understood that the combinatorial effect observed with the exemplary HTT intron 2-targeting (HTT_intron2_12061_150) Exon Editor (SEQ ID NO: 204) described here would be similarly applicable to either HTT intron 3-targeting (comprising, e.g., HTT_intron3_4223_150 or HTT_intron3_4233_150) Exon Editors or HTT intron 1-targeting (comprising, e.g., HTT_intron1_11704_100 or HTT_intron1_11724_100) Exon Editors.

Trans-splicing efficiency for the HTT Exon Editor was similar with or without the MSH3 Splice Modulator (see, e.g., FIG. 54; Top panel), as previously seen in FIG. 52. In the presence of HTT-miRNA 1 or HTT-miRNA 2, the % edited *HTT* transcripts increased due to the reduction of unedited HTT transcripts (and thus total *HTT* transcripts) by the HTT miRNA (see, e.g., FIG. 54; Top panel). The performance of HTT miRNAs was assessed in FIG. 54; Middle panel, and similar

efficiencies were observed with HTT miRNA-1 and HTT miRNA-2, whether in a dual hybrid (Exon Editor + HTT miRNA) or triple hybrid (Exon Editor + HTT miRNA + MSH3 Splice Modulator) context. Evaluation of MSH3 knockdown efficiencies supported unchanged knockdown efficiency by the Splice Modulator whether in the presence of the Exon Editor or
5 HTT miRNA. See, e.g., FIG. 54; Bottom panel. These results demonstrated that there was no interaction between the Exon Editor, the HTT miRNA, and the MSH3 Splice Modulator when co-delivered in the same DNA fragment. Of note, the combinatorial effect would be similarly applicable to any combination of HTT Exon Editor, HTT miRNA and MSH3 Splice Modulator.

10 **Proof-of-mechanism in iCell GlutaNeurons**

To confirm and corroborate the efficacy of aforementioned molecules in additional cell systems in vitro, the present inventors utilized iCell GlutaNeurons (FujiFilm), human glutamatergic-enriched cortical neurons derived from induced pluripotent stem (iPS) cells. Exemplary molecules were packaged in AAV2.7m8 and used for transduction to express molecules of
15 interest in the iCell GlutaNeurons.

The present inventors tested two different AAV2.7m8 for a proof-of-mechanism experiment for three separate modalities: 1) HTT Exon Replacement by trans-splicing, 2) HTT knockdown by a miRNA targeting *HTT* exon 1, and 3) MSH3 reduction by splice modulation of *MSH3* exon 7. Of
20 note, one AAV tested (SEQ ID NO: 354) packaged a dual hybrid molecule that expressed both the *HTT* intron 2-targeting Exon Editor (SEQ ID NO: 204) and the HTT miRNA-1 (active sequence SEQ ID NO: 339; encoding pri-miRNA sequence SEQ ID NO: 341), while the other AAV tested (SEQ ID NO: 357) packaged a dual hybrid molecule that expressed the HTT intron 2-targeting Exon Editor (SEQ ID NO: 204) and the *MSH3* exon 7 skipping Splice Modulator
25 (SEQ ID NO: 331), wherein each was packaged in AAV2.7m8.

RT-ddPCR analysis measuring the levels of % edited (trans-spliced) *HTT* RNA demonstrated that without HTT miRNA, the dual hybrid molecule HTT Exon Editor + MSH3 Splice Modulator resulted in ~37% replacement by trans-splicing in iCell GlutaNeurons. In the dual hybrid
30 molecule Exon Editor + HTT miRNA, the present inventors observed % edited RNA of ~50% due in part to the reduced native *HTT* copy number as a result of the HTT miRNA. Western blot analysis of the N-terminal FLAG tag confirmed the presence of trans-spliced full-length HTT protein. See, e.g., FIG. 55.

Further analysis of unedited *HTT* transcript levels suggested that without HTT miRNA, the dual hybrid molecule HTT Exon Editor + MSH3 Splice Modulator demonstrated a 47% reduction of

unedited *HTT* RNA as a consequence of trans-splicing (replacement of native with corrected RNA) in iCell GlutaNeurons. In the presence of the *HTT* miRNA modality, the dual hybrid molecule Exon Editor + *HTT* miRNA knocked down 50% of unedited *HTT* transcripts by the miRNA, resulting in an even more favorable edited:unedited *HTT* transcript ratio. Western blot analysis to visualize total *HTT* (trans-spliced and unedited *HTT* proteins) confirmed the knockdown of unedited *HTT* with the additional *HTT* miRNA modality. See, e.g., FIG. 56.

The present inventors also tested the MSH3 splice modulation modality in iCell GlutaNeurons. Western blot analysis of MSH3 protein showed that the dual hybrid molecule Exon Editor + MSH3 exon 7 Splice Modulator (SEQ ID NO: 331) resulted in >40% knockdown of MSH3 protein in these cells. See, e.g., FIG. 57. Overall, the present inventors confirmed similar activity for each of the three HD modalities (*HTT* Exon Editor, *HTT* miRNA and MSH3 Splice Modulator) in HEK293 cells and iCell GlutaNeurons.

Additional Combinatorial Embodiments and Details Pertaining to Same:

Improving exon replacement by *HTT* pre-mRNA trans-splicing

Results presented herein (e.g., FIG. 17) show that blocking *HTT* cis-splicing events can lead to enhancement of trans-splicing efficiencies and *HTT* exon 1 replacement. Accordingly, embodiments of the present disclosure include strategies to block these cis-splicing events by anti-sense RNA (asRNA). In order to express the cis-splicing blocker from the same plasmid as an *HTT*-targeting RNA Exon Editor, modified snRNAs, such as the modified U7 snRNA (U7 Sm OPT), previously shown to act as a splice modulator, may be used. As detailed herein, snRNAs such as the U7 Sm OPT are designed by 1) changing the targeting sequence (e.g., the histone binding sequence at the 5' region of U7 snRNA) to the complementary sequence of the gene to be modified, and 2) changing the binding site (AAUUUGUCUAG; SEQ ID NO: 367; U7 Sm WT) for U7 snRNP specific proteins (Lsm10 and Lsm11) to the consensus binding sequence derived from major spliceosomal U snRNPs (AAUUUUUGGAG; SEQ ID NO: 368; U7 Sm OPT), leading to the binding of Sm proteins found in spliceosomal snRNPs to the U7 Sm OPT. In some embodiments, asRNA molecules may comprise the U1 promoter, snRNA (with asRNA sequence targeting *HTT* intron-exon boundaries and consensus Sm binding site), and a U1 terminator. Exemplary asRNA sequences include (SEQ ID NOs: 127-135). Of these, a combination of *HTT* intron 2-targeting Exon Editor (*HTT*_intron2_12061_150) with ASO6 (SEQ ID NO: 131), which blocks cis-splicing of the upstream intron, led to a particularly significant improvement in trans-splicing efficiencies *in vitro*. See, e.g., FIGs. 16 and 17.

Embodiments directed to inhibition of somatic CAG repeat expansion

Without wanting to be bound by theory, for HD and many other repeat expansion disorders, the underlying mechanism of pathogenesis is thought to be somatic expansion of the repeat tract. As detailed herein, human genetic evidence strongly supports the idea that HD clinical outcome can be explained, at least in part, by somatic instability due to somatic expansion of the CAG repeat tracts. Hence, inhibiting the underlying mechanism of pathogenesis that arises from somatic instability should alter the disease course and could be a potential therapeutic for HD and scores of other repeat disorders.

Pursuant to the objective of targeting this aspect of HD pathology, the present inventors have designed and tested strategies to treat HD at an earlier stage of pathogenesis by inhibiting somatic CAG expansion. Embodiments described herein, such as those described for inhibiting somatic CAG expansion, may be used alone or in conjunction with RNA exon editor-mediated trans-splicing to correct pathogenic mutations. Although inhibition of somatic expansion alone may have a positive therapeutic effect on HD (See Combinatorial Embodiments section above), the present inventors have tested hybrid therapeutic approaches and dual action/hybrid molecules designed to inhibit somatic CAG expansion and correct *HTT* exon 1 RNA. See, e.g., FIGs. 19 and 20. The present inventors have also shown that this hybrid therapeutic approach is effective using dual action/hybrid molecules designed to inhibit somatic CAG expansion and correct *HTT* exon 2 RNA. See, e.g., FIGs. 30-32, 52, and 53. Accordingly, embodiments described herein, such as those designed to inhibit somatic CAG expansion, may be used alone or in conjunction with RNA exon editor-mediated trans-splicing to correct pathogenic mutations.

Mechanisms controlling this process involve cellular machinery, for example, DNA mismatch repair enzymes, including *MSH2* and *MSH3*. *MSH3* is speculated to be a good target to inhibit somatic CAG repeat expansion. Here, the present inventors outline several ongoing approaches designed to reduce *MSH3* levels and inhibit somatic CAG repeat expansion.

Knockdown of *MSH3* and other genes by trans-splicing: As described herein, trans-splicing can be used to reduce expression of *MSH3* or other genes, either alone or in combination with other applications of trans-splicing.

Embodiments A1.1. Trans-splicing for knockdown of *MSH3* and other genes: As described herein, trans-splicing can be used for mRNA editing, but it is also implemented herein to reduce expression levels of a target gene. As described herein, reducing expression levels of a target gene can be achieved by trans-splicing into the target pre-mRNA or mature mRNA to produce an RNA that is quickly degraded or that cannot produce a functional protein. See, e.g., FIGs. 20-

22 and 30-32. Both 5' and 3' exon editors can achieve this outcome. A 5' exon editor, e.g., can replace one or more of the 5' exons of the target mRNA to remove the start codon and some or all of the coding sequence, replacing it with an alternative coding sequence or a non-coding sequence. A 3' exon editor can, e.g., replace one or more of the 3' exons of the target, replacing
5 them with an alternative coding sequence or a non-coding sequence. In the resulting chimeric mRNA, there may be no translation of any amino acids from the target mRNA. 5' exon editors can also, e.g., insert 5' untranslated regions (UTRs) that prevent translation. 3' exon editors can also, e.g., insert 3' UTRs that de-stabilize the transcript or prevent it from being exported from the nucleus or translated. The portion of the target mRNA remaining after the trans-splicing
10 reaction is presumed to be quickly degraded, since it will lack either a 5' cap (5' exon editor) or a 3' polyA tail (3' exon editor).

Embodiments A1.2. Tandem binding domain Exon Editors: Exon editors that edit a target, such as *HTT*, can at the same time knockdown a second gene, such as *MSH3*, through the
15 inclusion of two different binding domains, one targeting *HTT* and the other *MSH3*. Examples of how this dual effect can be achieved are presented in FIGs. 19-22 and 30-32. Briefly, a 5' *HTT* exon editor containing a 5' portion of *HTT* can trans-splice to *HTT* to reconstitute a functional wild-type *HTT* mRNA, while also trans-splicing to *MSH3* to produce a non-functional *MSH3* mRNA. In this embodiment, the chimeric non-functional *HTT-MSH3* mRNA would encode an
20 *HTT* polypeptide that terminates at the first in-frame stop codon in the *MSH3* portion of the mRNA. As a consequence, *MSH3* pre-mRNA undergoing such a trans-splicing process would be altered to become an mRNA that encodes few to no amino acids from the *MSH3* coding sequence and would likely undergo nonsense-mediated decay. This approach is applicable to binding domains targeting any *MSH3* intron. Although examples presented herein (FIGs. 19-22
25 and 30-32) use 5' exon editors, a similar approach could be taken with a 3' exon editor.

Embodiments A.1.3. Tandem RNA Exon Editors: Tandem exon editors are two different exon editors inserted into the same plasmid, AAV, or other delivery vehicle. They can be arranged in head-to-head or head-to-tail configurations. Each exon editor may have its own regulatory
30 sequences, including a promoter, or may be generated through a cleavage event from a single bi-cistronic transcript. One exon editor can target one gene or intron and the second exon editor can target a different gene or intron. In some embodiments, one exon editor can replace mutant *HTT* with wild-type *HTT* and a second exon editor can reduce *MSH3* expression. In some embodiments, one exon editor can target one intron of *HTT* and the second editor can target a
35 second intron of *HTT*, increasing the overall efficiency of *HTT* exon editing.

Embodiments A.2. Reduction of *MSH3* or other genes by RNAi in combination with *HTT*

trans-splicing: A short-hairpin RNA (shRNA) or microRNA (miRNA) designed to reduce gene expression can be added to an exon editor within the same cistron (e.g., in an intron of the exon editor) or as a separate cistron with its own regulatory sequences. The RNAi can reduce expression of the unedited target (e.g., mutant *HTT*) or another gene (e.g., *MSH3*). For the purposes of selectively reducing mutant *HTT* with RNAi, the exon editor would replace a portion of the mutant *HTT* transcript with a sequence resistant to the shRNA or miRNA. See, e.g., FIGs. 33-35, 46-50.

Embodiments A.3. *MSH3* reduction by vectorized splice modulation or translational

blockade: In some embodiments, *MSH3* is inactivated by blocking the inclusion of one or more *MSH3* exons whose absence prevents a functional protein from being translated by leading to premature translation termination. In some embodiments, *MSH3* is inactivated by incorporating sequences that are complementary to *MSH3* splice junctions into small nuclear RNA (snRNA) sequences, such as, e.g., the U7 snRNA, preventing exon inclusion. Here, the splice modulation may be vectorized, i.e., delivered by an AAV that may also include an exon editor. See, e.g., FIGs. 36-45, 51-54.

Embodiments A.3.1. Small nuclear RNA-based (e.g., U7SmOPT) asRNA against *MSH3*:

In some embodiments, *MSH3* is inactivated by incorporating sequences that are complementary to *MSH3* splice junctions into snRNA sequences, such as, e.g., the U7 snRNA. See, e.g., FIGs. 36-45.

Embodiments A.3.2. Combination *MSH3* reduction with *HTT* trans-splicing:

In some embodiments, strategies A.3.1 and others described herein are combined into the same DNA fragment, or the same fragment as an exon editor, or be co-packaged into a single AAV for co-delivery into patient tissues. See, e.g., FIGs. 33-54.

Reduction of possible *HTT* toxic species: A short-hairpin RNA (shRNA) or microRNA (miRNA) designed to reduce gene expression can be added to an exon editor as a separate cistron with its own regulatory sequences. The RNAi can reduce expression of the unedited target (e.g., mutant *HTT*). For the purposes of selectively reducing mutant *HTT* with RNAi, the exon editor would replace a portion of the mutant *HTT* transcript with a sequence resistant to the shRNA or miRNA. See FIGs. 46-50 and 54-56.

In order to address the potentiality that treatment of HD may be enhanced by using a combinatorial therapeutic approach, the present inventors have designed a hybrid, multi-pronged approach wherein multiple mechanisms and/or pathogenic species were targeted. In some embodiments, these approaches included replacement of mutant *HTT* via an exon editor, 5 knockdown of mutant *HTT* and related transcripts (e.g., *HTT1a*) via RNAi, and reduction of *MSH3* via: e.g., trans-splicing, anti-sense RNA, RNAi, vectorized splice modulation, or translation blockade, or any combination thereof. When combined into the same DNA fragment, or the same fragment as an exon editor, these approaches represent a multi-modal mechanism of action that can nonetheless be delivered in a single AAV for co-delivery into patient tissues. 10 Embodiments utilizing such a combinatorial approach may be used to achieve improved therapeutic efficacy in some, if not all, HD patients. Such an approach may confer even more disease-modifying effects.

Animal Models

15 **BACHD Mouse Model:** This animal model is a bacterial artificial chromosome (BAC)-mediated transgenic mouse model (BACHD), wherein full-length human mutant huntingtin (fl-mhtt) is expressed. It was developed expressing fl-mhtt comprising 97 glutamine repeats under the control of endogenous htt regulatory machinery on the BAC. The glutamine repeats are encoded by 97 mixed CAG-CAA repeats. BACHD mice recapitulate HD disease and disease 20 progression in humans in that they exhibit progressive motor deficits, neuronal synaptic dysfunction, and late-onset selective neuropathology, including significant cortical and striatal atrophy and striatal dark neuron degeneration. Based on the well-documented robustness of the behavioral and neuropathological phenotypes, BACHD mice are recognized as a suitable fl-mhtt mouse model for preclinical studies. BACHD mice are described in detail in Gray et al. (2008, J 25 Neuroscience 28:6182; the content of which is incorporated herein in its entirety) and commercially available.

In some embodiments, therapeutic efficacy of nucleic acid trans-splicing molecules/RNA exon editors described herein in the context of the BACHD mouse model may be measured by at 30 least one of an increase in percent replacement of pathogenic *HTT* RNA, a reduction in pathogenic *HTT*, a reduction in pathogenic *HTT* aggregates, an improvement in motor coordination and balance as measured by, e.g., a rotarod test, a decrease in forebrain atrophy, or any combination thereof.

35 **BAC-CAG Mouse Model:** This animal model is a human genomic BAC transgenic mouse model of HD that expresses human mutant huntingtin (mHTT) comprising long uninterrupted and

somatically unstable CAG repeats (120-130 pure CAG repeats) and exhibits progressive disease-related phenotypes. Unlike other mHTT transgenic models having stable, CAA-interrupted, polyglutamine-encoding repeats, BAC-CAG mice present with robust striatum-selective nuclear inclusions and transcriptional dysregulation also observed in HD patients and huntingtin knockin models. BAC-CAG are described in detail in Gu et al. (2022, Neuron 110:1173; the content of which is incorporated herein in its entirety) and commercially available. As described therein, the striatal transcriptionopathy in HD models is correlated with their uninterrupted CAG repeat length rather than the polyglutamine length. As also described therein, somatic CAG repeat instability and nuclear mHTT aggregation are best correlated with early-onset striatum-selective molecular pathogenesis and locomotor and sleep deficits, whereas repeat RNA-associated pathologies and repeat-associated non-AUG (RAN) translation may impact less selective or late pathogenic roles, respectively.

In some embodiments, therapeutic efficacy of nucleic acid trans-splicing molecules/RNA exon editors described herein in the context of the BAC-CAG mouse model may be measured by at least one of an increase in percent replacement of pathogenic *HTT* RNA, a reduction in pathogenic HTT, a reduction in pathogenic HTT aggregates, an improvement in motor coordination and balance as measured by, e.g., a rotarod test, a decrease in forebrain atrophy, or a decrease in striatum-specific transcriptionopathy, or any combination thereof.

To show in vivo proof-of-mechanism in the brain, the present inventors utilized the BAC-CAG mouse model (Gu et al. 2022). As described above, this transgenic mouse strain was engineered to contain the human mutant HTT genomic locus (>120 uninterrupted CAG repeats in HTT exon 1) and display abnormal features that resemble the disease phenotype, including nuclear HTT aggregates and transcriptional dysregulation at ≥ 12 months of age. The present inventors performed proof-of-mechanism studies for HTT trans-splicing in the BAC-CAG mouse brain with exemplary molecules packaged in AAV9. Three different AAV test articles were tested in this study: CMV promoter-driven HTT intron 2-targeting Exon Editor (SEQ ID NO: 204) packaged in self-complementary AAV (scAAV) (SEQ ID NO: 363, which comprises SEQ ID NO: 204), CMV promoter-driven HTT intron 2-targeting Exon Editor (SEQ ID NO: 204) packaged in single-stranded AAV (ssAAV) (SEQ ID NO: 363, which comprises SEQ ID NO: 204), and CAGGS promoter-driven HTT intron 2-targeting Exon Editor (SEQ ID NO: 204) packaged in ssAAV (SEQ ID NO: 364, which comprises SEQ ID NO: 204). All three test articles also contained a mouse-targeting surrogate MSH3 exon 7 Splice Modulator (SEQ ID NO: 362), co-packaged as a hybrid molecule. The entire sequence of test articles is as follows: CMV promoter-driven HTT intron 2-targeting Exon Editor packaged in scAAV (SEQ ID NO: 369), CMV promoter-driven HTT intron

2-targeting Exon Editor packaged in ssAAV (SEQ ID NO: 370), and CAGGS promoter-driven HTT intron 2-targeting Exon Editor packaged in ssAAV (SEQ ID NO: 371). To test dose response of the Exon Editor, 2 different doses were selected for injections – 1E+11 vector genomes (vgs) and 3E+11 vgs per animal, injected bilaterally. Two independent routes of administration were selected: intracerebroventricular (ICV) injection at neonatal P0 (Table 2) and intrastriatal injection at 8 weeks of age. Mouse cortex and striatum were harvested 4 weeks post-injection and the efficiencies of HTT Exon Replacement by trans-splicing and MSH3 knockdown by splice modulation were profiled by RT-ddPCR and Western Blotting.

- 10 Table 2. Study design for in vivo proof-of-mechanism study in the BAC-CAG mouse brain by ICV injection. For each treatment, the age at terminal collection was 4 weeks, and the brain region from which collection was made was the cortex striatum.

Genotype	Sex	Treatment	ssAAV vs scAAV	Serotype	Age of treatment	ROA	Dose
HEMI	matched	CMVp::HTT/MSH3 (tool) hybrid	scAAV	AAV9	P0	ICV	1.00E+11
		CMVp::HTT/MSH3 (tool) hybrid					3.00E+11
		CMVp::HTT/MSH3 (tool) hybrid	ssAAV				1.00E+11
		CMVp::HTT/MSH3 (tool) hybrid					3.00E+11
		CAGGSp::HTT/MS H3 (tool) hybrid					3.00E+11
		Vehicle					

- 15 As shown in FIG. 58, analysis of HTT trans-splicing profiles demonstrated that upwards of 30% HTT replacement was achieved in the mouse brain by neonatal ICV injection. The CAGGS promoter-driven Exon Editor outperformed its CMV promoter equivalent, as indicated by higher % HTT replacement in both the cortex and the striatum. Among the CMV promoter-driven Exon Editors, a clear dose response was observed, whereby animals that received 3E+11 vg had
- 20 higher % HTT replacement than those that received 1E+11 vg. Significantly, trans-spliced full-

length HTT protein was detected by Western blotting against the N-terminal FLAG tag. This experiment also indicated that with increasing levels of Exon Editor transcripts, higher level of trans-splicing was achieved (FIG. 59), suggesting that identifying a stronger promoter or administering a higher dose might result in even higher Exon Editor activity.

5

I. Definitions

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 coding domain 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, e.g., FIG. 3.

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 coding domain 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” refer 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 *HTT* exon) for replacement of a mutated *HTT* exon.

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).

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, the expression level of a trans-spliced RNA product is detected from RNA that is isolated from cells or tissues using RNA-seq.

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:
5 % on-target (ONT) TS = $100 \times (\text{ONT copy number} / (\text{ONT copy number} + \text{Native copy number}))$.

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 or non-targeting binding domain (e.g., a binding domain comprising, or consisting of,
10 SEQ ID NO: 7). 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.

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 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
20 separated by intervening sequences and the binding domain is still considered to be “operably linked” to the coding sequence).

As used herein, the term “coding domain” 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 coding domain may include one or more functional exons (e.g., a sequence of functional exons). In some instances, one or more functional exons of a coding domain 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 coding domain can include one or more introns (e.g., native introns) or untranslated regions
30 (UTRs, e.g., native UTRs) between or otherwise adjacent to (e.g., upstream or downstream of) exons.

As used herein, a “native 5’ *HTT* untranslated region” or “native 5’ *HTT* UTR” refers to a sequence greater than 20 nucleotides in length that has at least 90% sequence identity with a region of a native *HTT* gene (e.g., a human *HTT* gene) that is 5’ to the ATG start codon. An example of a native 5’ *HTT* untranslated region is given by the DNA sequence of SEQ ID NO: 136. An example of a modified version of a 5’ *HTT* untranslated region is given by the DNA sequence of SEQ ID NO: 192.

As used herein, a “functional sequence of 5’ *HTT* exons” refers to a nucleic acid sequence comprising one or more of *HTT* exons 1-3 (e.g., exon 1; or exon 1 and exon 2; or exons and exon 2 and exon 3) that encode a functional (biologically active) portion of *HTT* protein. In some embodiments, a “functional sequence of 5’ *HTT* exons” refers to a nucleic acid sequence comprising exon 1 of *HTT* or exons 1-2 of *HTT* or exons 1-3 of *HTT* that encodes a functional (biologically active) portion of *HTT* protein. When trans-spliced to an endogenous *HTT* exon 3’ to the binding site, a functional sequence of 5’ *HTT* exons provides expression of functional *HTT* protein (e.g., non-mutated *HTT* protein). In some instances, the functional sequence of 5’ *HTT* 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 *HTT* intron 2 and trans-splices with endogenous *HTT* exon 3 can include a functional sequence of 5’ *HTT* exons that includes exons 1 and 2 or a trans-splicing molecule that binds *HTT* intron 3 and trans-splices with endogenous *HTT* exon 4 can include a functional sequence of 5’ *HTT* exons that includes exons 1-3).

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 *HTT*, the term “functional” may be used to refer to the amount of *HTT* protein activity that is necessary to restore *HTT* activity levels to support normal cellular functions within the context of, for example, pyramidal neurons in the cortex, medium spiny neurons in the striatum, and/or hypothalamic neurons. Moreover, reducing the amount of mutant *HTT* and/or skewing the mutant:wild-type ratio will significantly contribute to reduction of neurodegeneration of pyramidal neurons in the cortex, medium spiny neurons in the striatum, and/or hypothalamic neurons. In the context of treating a condition associated with pathogenic *HTT* activity (e.g., HD) or use of a therapeutic agent comprising a nucleic acid trans-splicing molecule described herein, “functional” refers to reducing the amount of, e.g., defective (non-functional) *HTT* protein comprising polyglutamine tracts in excess of 35 or 40 glutamine repeats to eliminate one or more symptoms of a condition associated with pathogenic *HTT* activity (e.g., HD). In some

embodiments, such methods or uses lead to a decrease in pathogenic HTT activity and an increase in wild-type HTT protein activity. In some embodiments, such a decrease in pathogenic HTT protein activity decreases pathogenic HTT activity levels by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% of HTT activity (e.g., 96%, 97%, 98%, 99%, or 100%) relative to that of a control (untreated) cell in which mutated pathogenic HTT is expressed. In some embodiments, such an increase in HTT protein activity restores HTT activity levels to at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% of HTT activity (e.g., 96%, 97%, 98%, 99%, or 100%) relative to that of a cell in which HTT is present at normal, wildtype levels such as those present in cells in which non-mutated HTT is expressed.

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 cryptic splice site resistant nucleic acid sequence within a nucleic acid trans-splicing molecule is in a coding domain sequence (CDS). In some embodiments, the cryptic splice site resistant *HTT* CDS comprises, consists essentially of, or consists of exon 1 or exons 1 and 2 or exons 1, 2, and 3 of the *HTT* 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.

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. 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.

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).

5 As used herein, the term “binding site” refers to an endogenous polynucleotide sequence in the target pre-mRNA (e.g., a pre-mRNA of an endogenous gene, e.g., *HTT*) 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
10 other embodiments, the binding site is within 1-10 nucleotides longer or shorter than the binding domain (i.e., some of the nucleotides 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.

15

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.

20 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
25 a G). For instance, a binding domain that is in an anti-sense orientation to a binding site is complementary to the binding site. RNA pairing includes G pairing with U; 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
30 replace any one or more of the A's with G's or C's with T's without substantially affecting binding.

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
35 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.

With respect to sequences presented herein and in the accompanying Sequence Listing, it is understood that RNA transcripts encoded by DNA sequences comprise a uridine (U) at positions corresponding to thymidine (T) as listed in the corresponding DNA sequence. In some instances, sequences of RNA exon editor components are disclosed herein as DNA sequences. For any sequence disclosed herein as a DNA sequence, an RNA sequence with U substituted for each T in the sequence is also contemplated. Thus, if a given SEQ ID NO is identified as having a sequence that may be included in an RNA exon editor, a version of the SEQ ID NO with U substituted for each T is also contemplated.

“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 anti-sense orientation from one another (e.g., hybridized to one another).

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

As used herein, the term “mutation” may be used to refer to any aberrant nucleic acid sequence that encodes a defective RNA or protein product (e.g., a non-functional protein product, a non-biologically active protein, a protein product having reduced function, a protein product having pathogenic or 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), duplications, 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 *HTT* exon) refers to an exon containing a mutation or an exon sequence that reflects a mutation
5 in a different region, such as a cryptic exon resulting from a mutation in an intron.

The term “*HTT*” (*Huntingtin*) refers to any native *HTT* 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
10 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 *HTT* signaling. *HTT* encompasses full-length, unprocessed *HTT*, as well as any form of *HTT* that results from native processing in the cell. An exemplary human *HTT* sequence is provided as National Center for
15 Biotechnology Information (NCBI) Reference Sequence: NG_009378. In some instances, an *HTT* fragment is encoded by a therapeutic agent comprising a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 3, 59, 157, 349, 350, 351, 352, or 353 (e.g., at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to any one of SEQ
ID NOs: 3, 59, 157, 349, 350, 351, 352, or 353 a functional portion thereof, and/or a codon-
20 modified variant thereof.

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
25 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
30 polynucleotide as defined herein, wherein the homolog, isoform or transcript variant is characterized by a degree of identity or homology, respectively, as defined herein.

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
35 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, or aromatic groups in the side chains, the side chains of which can form hydrogen bridges, e.g., side chains which have a hydroxyl function. By conservative substitution, 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)).

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).

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.

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. AAV9, AAV-retro,

AAV1, AAV4, AAV8, AAV5, AAV-PHP.eB, for example, are among the AAV serotypes that are neurotropic in nature.

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).

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, and/or for accurate delivery to the nucleus, etc.

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, etc. The individual may be male or female. In one embodiment, the individual has a disease or disorder caused by a mutation in the *HTT* gene (e.g., HD). In another embodiment, the individual is at risk for developing a disease or disorder caused by a mutation in the *HTT* gene. In another embodiment, the individual has shown clinical signs of a disease or disorder caused by a mutation in the *HTT* gene, such as HD. 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, 30-40 years of age, 30-50 years of age, 40-50 years of age, 50-60 years of age, 60-70 years of age, or more than 70 years of age.

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.

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.

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 *HTT* 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.

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.

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 *HTT*. For example, in one embodiment in which striatal cells (e.g., medium spiny neurons) or cells in the cortex (e.g.,

pyramidal neurons) 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 intracerebral (IC) delivery (e.g., slow delivery injection or convection-enhanced diffusion injection), intracerebroventricular (ICV) delivery, or intrathecal delivery. In
5 some embodiments, IC injections involve stereotaxic implantation of microinjection guide sleeves to improve delivery to a specific locus in the brain. 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.

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

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
20 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 performance 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, or 7,888,112, each of which is incorporated herein by reference in its entirety. The sequence surrounding the translational start site can be converted
25 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. 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
30 other generally recognized pharmacopeia for use in animals, and more particularly in humans. 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
35 edition.

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.

- 5 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

- Provided herein are nucleic acid trans-splicing molecules useful for correcting mutations in *HTT* by replacing at least one mutated *HTT* exon with a functional *HTT* exon [e.g., a *HTT* exon 5' to the binding site, e.g., exon 1 (SEQ ID NO: 348; HTT 5' UTR + exon 1, which includes 21 CAG repeats (within normal range of repeat number), exon 2 (SEQ ID NO: 365), exon 3 (SEQ ID NO: 366 of *HTT*). Note that while the sequence of *HTT* exon 1 set forth in SEQ ID NO: 348 includes 21 CAG repeats, the number of CAG repeats in individual genomic sequences may be different (i.e., higher or lower). Corresponding nucleotide numbering positions in sequences having different numbers of CAG repeats can readily be determined by taking into account any variation in the number of CAG repeats. In some embodiments, the nucleic acid trans-splicing molecule is a 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.
- 10
15
20

- Trans-splicing molecules comprise a binding domain, a splicing domain, and a coding domain. In some embodiments, the nucleic acid trans-splicing molecule has a 5' regulatory domain having a native 5' *HTT* untranslated region (e.g., a sequence having at least 80% sequence identity with either SEQ ID NO: 136 or 192). In some embodiments, the nucleic acid trans-splicing molecule has a splice site of 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 any one of SEQ ID NOs: 37-46 and 106-112 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 any one of SEQ ID NOs: 37-46 and 106-112.
- 25
30

35

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 *HTT* exons, e.g., a functional sequence of *HTT* 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.

In some embodiments, nucleic acid trans-splicing molecules described herein are configured to correct at least one mutation in an allele of the *HTT* gene in a subject located in a 5' region of the *HTT* gene (e.g., a region 5' to intron 1, 2, or 3) by binding to target intron 1, 2, or 3 and mediating trans-splicing of a coding domain having a functional sequence of 5' *HTT* exons to an endogenous *HTT* exon 3' to the target intron. Such trans-splicing thereby repairs the defective *HTT* 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 wild-type *HTT* mRNA capable of transcribing a functional HTT protein in the cell.

15

HTT

An *HTT* gene targeted by a trans-splicing molecule described herein can contain one or multiple mutations that are associated with HD. An exemplary human *HTT* sequence is provided as National Center for Biotechnology Information (NCBI) Reference Sequence: NG_009378. 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.

25

It is anticipated that the *HTT* 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 with respect to the encoded amino acid. In other embodiments, nucleic acid base modifications which change the amino acids, e.g., to improve expression of the resulting peptide/protein are envisioned. In some embodiments, modification of allelic variations, caused by the natural degeneracy of the genetic code are envisioned.

30

Also included as modifications of *HTT* genes are analogs or modified versions of the encoded amino acid sequences. Typically, such analogs differ from the specifically identified proteins by

35

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.

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

As described herein above, HD is caused by expansion of CAG trinucleotide repeats in exon 1
10 in excess of 35 CAG repeats (incomplete penetrance) or in excess of 40 CAG repeats (juvenile or adult onset) in the *HTT* gene, which are inherited in an autosomal dominant manner. See FIG. 1. Compositions comprising trans-splicing molecules described herein can correct the expanded CAG repeats in exon 1, irrespective of how many repeats are present, because the trans-splicing molecules replace the entirety of the pathogenic exon 1 of the *HTT* gene.

Coding Domains

In some embodiments, the coding domain of a 5' trans-splicing molecule includes all *HTT* exons (e.g., functional *HTT* exons) that are 5' to the target *HTT* intron (e.g., *HTT* intron 1 [(SEQ ID NO: 348; *HTT* 5' UTR + exon 1, which includes 21 CAG repeats (within normal range of repeat
20 number) or may comprise 35-39 CAG repeats (incomplete penetrance) or 40+ CAG repeats (adult-onset or juvenile-onset HD)], intron 2 (SEQ ID NO: 365), and/or intron 3 (SEQ ID NO: 366)). For example, in embodiments in which a 5' trans-splicing molecule targets *HTT* intron 2, the coding domain may include functional *HTT* exons 1-2. In some embodiments, functional *HTT* exons 1-2 are encoded by a sequence comprising any one of SEQ ID NOs: 59, 349, 350, or 351.
25 In some embodiments, functional *HTT* exons 1-2 are encoded by a sequence comprising any one of SEQ ID NOs: 59, 349, 350, or 351, which further comprises an ATG start codon at the 5' terminal end. In some embodiments, the binding domain binds to intron 2, and the coding domain includes functional *HTT* exons 1-2.

30 In some embodiments in which a 5' trans-splicing molecule targets *HTT* intron 3, the coding domain may include functional *HTT* exons 1-3. In some embodiments, functional *HTT* exons 1-3 are encoded by a sequence comprising any one of SEQ ID NOs: 157, 352, or 353. In some embodiments, functional *HTT* exons 1-3 are encoded by a sequence comprising SEQ ID NOs: 157, 352, or 353, which further comprises an ATG start codon at the 5' terminal end. In some
35 embodiments, the binding domain binds to intron 3, and the coding domain includes functional *HTT* exons 1-3.

In some embodiments, wherein a 5' trans-splicing molecule targets *HTT* intron 1, the coding domain may include functional *HTT* exon 1. In some embodiments, functional *HTT* exon 1 is encoded by a sequence comprising SEQ ID NO: 3. In some embodiments, functional *HTT* exon 1 is encoded by a sequence comprising SEQ ID NO: 3, which further comprises an ATG start codon at the 5' terminal end. In some embodiments, the binding domain binds to intron 1, and the coding domain includes functional *HTT* exon 1.

In some embodiments, a coding domain-encoding sequence (e.g., of a transgene encoding an RTM) includes cDNA of *HTT* exons (e.g., *HTT* exons) for replacement of mutated *HTT* exon/s. For example, one or more functional *HTT* exons within the coding domain can be a cDNA sequence. In some embodiments, the entire coding domain is a cDNA sequence. Additionally, or alternatively, all or a portion of the coding domain, or one or more functional *HTT* 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 *HTT* exon).

In some embodiments, all or a portion of the coding domain or coding domain-encoding sequence, or one or more functional *HTT* 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, or 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 coding domain can be a nucleic acid sequence of up to 4,000 nucleotide bases in length.

In some embodiments, a nucleic acid trans-splicing 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 at least one nucleotide mutation relative to the endogenous RNA molecule sequence, wherein the at least one 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.

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.

Also encompassed herein is a method of increasing trans-splicing efficiency or therapeutic performance 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. In some embodiments, the cryptic splice site is identified experimentally. In some embodiments, the cryptic splice site is predicted based on in silico analysis.

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 may comprise 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
5 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
10 context of an RNA exon editor.

Further to the above, the experimental results and sequence information are analyzed generally as follows. Changes to remove cryptic splice sites identified experimentally are made by searching for and replacing certain elements of splice acceptor sites. AG sites (and more
15 strongly CAG sites) at the end of a splice acceptor site are prioritized for introduction of nucleotide changes. If an AG site is not found or could not be changed without introducing a non-synonymous mutation, then the sequence 42 to 4 base pairs upstream of the splice site is scanned for branch points (sequences matching YNAH). Any such branch point sequences identified are then analyzed and considered for introduction of one or more nucleotide
20 mutations to reduce cryptic splice site usage at the experimentally identified cryptic splice site. In addition, the sequences are 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 are then analyzed and considered for introduction of a nucleotide
25 mutation/s to reduce cryptic splice site usage at the experimentally identified cryptic splice site.

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 Example 6 of WO
30 2023/220742, which is hereby incorporated by reference in its entirety. 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.

In some embodiments, the cryptic splice site, or off-target splice site, that is changed to mitigate
35 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.

- 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.
- To address the potential for cryptic self-splicing (either by cis-splicing of AAV concatemers or inter-molecular trans-splicing), the present inventors used *in silico* prediction of self-splicing sites to mitigate such unwanted events. See, for example, Table 1.

Table 1 . Mitigation of cryptic self-splicing sites identified within the original Exon Editor.

Region	Splice Acceptor or Donor	Self-splicing mitigation change	SEQ ID NO of original sequence	SEQ ID NO of self-splicing mitigated sequence
5' UTR	Acceptor	G>T at base 20, GA>TT at base 60-61	136	192
3X FLAG	Acceptor	G>A at base 51	4	193
Upstream of MALAT1 terminator	Acceptor	AG>CT at base 2-3	195	194

20 Binding Domains

- HTT* trans-splicing molecules described herein feature a binding domain (BD) configured to bind/anneal a target *HTT* intron and/or exon. In some instances, the target *HTT* intron is *HTT* intron 2. 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 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 *HTT* intron pre-mRNA (e.g., a target *HTT* intron), which may suppress endogenous target cis-splicing while enhancing trans-splicing between the trans-splicing molecule and the target *HTT* pre-mRNA (e.g., by creating a chimeric molecule having a portion of endogenous *HTT* mRNA and a coding domain having one or more functional *HTT* exons that encode wildtype *HTT* 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 *HTT* intron pre-mRNA.

In some instances, the present invention provides trans-splicing molecules (or vectors thereof) that bind *HTT* at intron 2, e.g., wherein the nucleic acid trans-splicing molecule is configured to trans-splice a coding domain to endogenous *HTT* exon 3. 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 200, 1,500 to 2,500, or 10,500 to 12,251 of SEQ ID NO: 57.

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2 (e.g., any eight or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any ten or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 12 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 20 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 30 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 40 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 50 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of

HTT intron 2, any 100 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 150 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2, any 200 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of
5 *HTT* intron 2, or any 250 or more consecutive nucleic acids within nucleotides 1 to 200, 1,500 to 2,500, or 10,500 to 12,251 of *HTT* intron 2).

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 10,000 to 12,251 of *HTT* intron 2 (e.g., any eight or more consecutive nucleic acids
10 within nucleotides 10,000 to 12,251 of *HTT* intron 2, any ten or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 12 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 20 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 30 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 40 or more consecutive nucleic acids
15 within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 50 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 100 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 150 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, any 200 or more consecutive nucleic acids within nucleotides 10,000 to 12,251 of *HTT* intron 2, or any 250 or more consecutive nucleic
20 acids within nucleotides 10,000 to 12,251 of *HTT* intron 2).

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 11,000 to 12,251 of *HTT* intron 2 (e.g., any eight or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any ten or more consecutive nucleic acids
25 within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 12 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 20 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 30 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 40 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 50 or more consecutive nucleic acids
30 within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 100 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 150 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, any 200 or more consecutive nucleic acids within nucleotides 11,000 to 12,251 of *HTT* intron 2, or any 250 or more consecutive nucleic
35 acids within nucleotides 11,000 to 12,251 of *HTT* intron 2).

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 11,950 to 12,251 of *HTT* intron 2 (e.g., any eight or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any ten or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 12 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 20 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 30 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 40 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 50 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 100 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 150 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, any 200 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2, or any 250 or more consecutive nucleic acids within nucleotides 11,950 to 12,251 of *HTT* intron 2).

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

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 any one of SEQ ID NOs: 60-81.

25 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 any one of SEQ ID NOs: 60, 62, or 67-81.

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 any one of SEQ ID NOs: 60 or 67-81.

In some embodiments, the binding domain is a DNA sequence having at least 80% identity (e.g.,
5 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 any one of SEQ ID NOs: 67 or
70-81.

10 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.,
15 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to any one of SEQ ID NOs:
71-74 or 77-79.

In some instances, the target *HTT* intron is *HTT* intron 3. In one embodiment, the binding
domain is a nucleic acid sequence that is at least 80% complementary to (e.g., at least 85%
20 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%
25 complementary to) a sequence of the target *HTT* intron pre-mRNA (e.g., a target *HTT* intron),
which may suppress endogenous target cis-splicing while enhancing trans-splicing between the
trans-splicing molecule and the target *HTT* pre-mRNA (e.g., by creating a chimeric molecule
having a portion of endogenous *HTT* mRNA and a coding domain having one or more functional
HTT exons that encode wildtype *HTT* amino acid sequences). In one embodiment involving
30 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
35 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 *HTT* intron pre-mRNA.

In some instances, the present invention provides trans-splicing molecules (or vectors thereof) that bind *HTT* at intron 3, e.g., wherein the nucleic acid trans-splicing molecule is configured to trans-splice a coding domain to endogenous *HTT* exon 4. 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 3,100 to 4,429 of SEQ ID NO: 155.

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 3,100 to 4,429 of *HTT* intron 3 (e.g., any eight or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any ten or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 12 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 20 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 30 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 40 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 50 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 100 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 150 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, any 200 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3, or any 250 or more consecutive nucleic acids within nucleotides 3,100 to 4,429 of *HTT* intron 3).

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 4,100 to 4,429 of *HTT* intron 3 (e.g., any eight or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any ten or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 12 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 20 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 30 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 40 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 50 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 100 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 150 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3, any 200 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3).

within nucleotides 4,100 to 4,429 of *HTT* intron 3, or any 250 or more consecutive nucleic acids within nucleotides 4,100 to 4,429 of *HTT* intron 3).

In some instances, the binding site includes any six or more consecutive nucleotides within
5 nucleotides 4,100 to 4,388 of *HTT* intron 3 (e.g., any eight or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any ten or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 12 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 20 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 30 or more consecutive nucleic acids
10 within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 40 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 50 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 100 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 150 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3, any 200 or more consecutive nucleic acids
15 within nucleotides 4,100 to 4,388 of *HTT* intron 3, or any 250 or more consecutive nucleic acids
within nucleotides 4,100 to 4,388 of *HTT* intron 3).

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

20

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%,
25 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to any one of SEQ ID NOs: 158-
174.

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
30 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 any one of SEQ ID
NOs: 164-174.

35 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 any one of SEQ ID NOs: 170-174.

5

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 any one of SEQ ID NOs: 171-172.

10

15

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).

20

25

30

35

In some instances, the target *HTT* intron is *HTT* intron 1. 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 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 *HTT* intron pre-mRNA (e.g., a target *HTT* intron), which may suppress endogenous target cis-splicing while enhancing trans-splicing between the trans-splicing molecule and the target *HTT* pre-mRNA (e.g., by creating a chimeric molecule having a portion of endogenous *HTT* mRNA and a coding domain having one or more functional *HTT* exons that encode wildtype *HTT* 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 *HTT* intron pre-mRNA.

In some instances, the present invention provides trans-splicing molecules (or vectors thereof) that bind *HTT* at intron 1, e.g., wherein the nucleic acid trans-splicing molecule is configured to trans-splice a coding domain to endogenous *HTT* exon 2. 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 1000 or 11,500 to 11,850 of SEQ ID NO: 1.

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1 (e.g., any eight or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any ten or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 12 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 20 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 30 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 40 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 50 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 100 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 150 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, any 200 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1, or any 250 or more consecutive nucleic acids within nucleotides 1 to 1000 or 11,500 to 11,850 of *HTT* intron 1).

In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 11,500 to 11,850 of *HTT* intron 1 (e.g., any eight or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any ten or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 12 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 20 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 30 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 40 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 50 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 100 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 150 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1).

within nucleotides 11,500 to 11,850 of *HTT* intron 1, any 200 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1, or any 250 or more consecutive nucleic acids within nucleotides 11,500 to 11,850 of *HTT* intron 1).

5 In some instances, the binding site includes any six or more consecutive nucleotides within nucleotides 11,650 to 11,850 of *HTT* intron 1 (e.g., any eight or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any ten or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 12 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 20 or more consecutive nucleic acids
10 within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 30 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 40 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 50 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 100 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 150 or more consecutive nucleic acids
15 within nucleotides 11,650 to 11,850 of *HTT* intron 1, any 200 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1, or any 250 or more consecutive nucleic acids within nucleotides 11,650 to 11,850 of *HTT* intron 1).

In some instances, a binding domain has at least two non-overlapping sequences with at least
20 80% complementarity to the binding site.

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
25 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 any one of SEQ ID NOs: 8 or 14-21.

In some instances, the binding domain includes a nucleic acid sequence having at least 80%
30 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 any one of SEQ ID NOs: 16-21.

35

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%,
5 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity) to any one of SEQ ID NOs: 17, 18, 20, or 21.

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
10 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 any one of SEQ ID NOs: 17 or 18.

15 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).

As detailed herein, the first stage of *HTT* Exon Editor design began with the screening and
20 selection of a highly efficient BD sequence, which is complementary to the targeted pre-mRNA intron. As shown herein, BDs that target intron 2 of *HTT* and intron 3 of *HTT* exhibited the highest levels of trans-splicing and therefore, have been identified as effective elements of exemplary *HTT*-targeting Exon Editors for treatment of the HD patient population. See FIGs. 5, 10, 12, 24, and 25.

25

Splicing Domains

For 5' exon editors, the splicing domain can include a splice donor site (5' splice site) to mediate trans-splicing.

30 For 3' exon editors, the splicing domain can include a splice site, a branch point, and/or a polypyrimidine tract (PPT) 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.

35 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.

5 In one embodiment, a suitable 5' splice site comprises GTAAGT or GUAAGT.

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 binding domain and/or a
10 3' downstream sequence).

5' Untranslated Region

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
15 a native 5' *HTT* 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
20 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: 136.

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

In some embodiments, the nucleic acid trans-splicing molecule does not include a 5' untranslated region or does not include a native 5' *HTT* untranslated region.
30

5' Regulatory Domains

In some instances, the nucleic acid trans-splicing molecule is operatively linked to a 5' regulatory domain operatively linked 5' to the coding domain (e.g., directly linked to the coding domain, or linked through an intermediate domain, e.g., an untranslated region). A 5' regulatory
35 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/CMV promoter) operatively linked to a native 5' *HTT* untranslated region. In some embodiments, the 5' regulatory domain operatively linked to a native 5' *HTT* 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 SEQ ID NO: 137.

In some embodiments, the 5' regulatory domain operatively linked to a native 5' *HTT* 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 SEQ ID NO: 196.

In some embodiments, the CMV promoter is replaced by a CAGGS promoter, wherein the CAGGS promoter is used to drive expression of an RNA exon editor described herein. See, e.g., FIG. 27. As shown therein, when comparing CAGGS 5' UTR with or without HTT 5'UTR, protein translation appears to be regulated through the HTT 5' UTR when the CAGGS 5' UTR + HTT 5' UTR are operably linked. When the HTT 5 UTR is removed, leaving only the CAGGS 5' UTR, the present inventors observed stronger protein expression, which activity may be due to some element that upregulates translation in the CAGGS 5' UTR. Transduction experiments presented herein, wherein RNA exon editors were introduced into cells via an AAV, also demonstrated that the CAGGs promoter drives significant expression of RNA exons editors, MSH3 splice modulators, and miRNA encoding constructs. See, e.g., FIGs. 55, 56, 58 and 59.

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

In some embodiments, other types of promoters may be used. In some embodiments, the nucleic acid trans-splicing molecule is not operatively linked to a CMV promoter or a CAGGS promoter.

5 *Linker Domains*

As discussed herein, increasing trans-splicing efficiency remains an important goal for implementation of nucleic acid trans-splicing molecules as therapeutic agents. 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 a plurality of
10 linker sequences between the splice domain (SD) and binding domain (BD) and identified a series of exemplary linkers that improved performance relative to that of a 40mer linker to a statistically significant degree. See, for example, FIG. 8.

Further to the above, nucleic acid trans-splicing molecules may include a linker domain at one
15 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 20 nucleotides in length (e.g., between 20 and 100 nucleotides in length or between 20 and 85 nucleotides in length). In some instances, the linker domain comprises,
20 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,
25 at least 97% sequence identity, at least 98% sequence identity, at least 99% sequence identity, or 100% sequence identity) with any one of SEQ ID NOs: 37-46 or 106-112.

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
30 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. In some embodiments, the nucleic acid trans-splicing molecule does not include any of SEQ ID NOs: 37-46 or 106-112, or a linker sequence may be absent.

35 Linker domains described herein may also be included in embodiments of RNA exon editors that do not target *HTT* or *MSH3*. Embodiments disclosed herein include RNA exon editors that

include a binding domain that targets an endogenous pre-mRNA, a coding domain sequence that encodes a functional amino acid sequence, and a linker domain between the coding domain sequence and the binding domain, wherein the linker domain comprises one or more of SEQ ID NOs: 37-46 or 106-112.

5

3' Transcription Terminator Domains

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 *HTT* 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: 5. 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: 6.

15
20

In some embodiments, the nucleic acid trans-splicing molecule does not include a 3' transcription terminator domain or does not include a MALAT1-derived transcription terminator.

Splicing Domains

In some embodiments, exemplary RNA exon editors described herein include those that comprise binding domains that bind to intron 1 of *HTT*, wherein such exemplary intron 1-binding RNA exon editors may comprise any one of SEQ ID NOs: 23-36 and 47-56, which sequences comprise the coding sequence (e.g., SEQ ID NO: 3), splice domain, linker domain, binding domain, and 3' transcription terminator). In some embodiments, the coding sequence further comprises an ATG start codon at the 5' terminus.

25
30

In some embodiments, exemplary RNA exon editors described herein include those that comprise binding domains that bind to intron 2 of *HTT*, wherein such exemplary intron 2-binding RNA exon editors may comprise any one of SEQ ID NOs: 83-105 and 113-125, which sequences comprise the coding sequence (e.g., SEQ ID NOs: 59, 349, 350, or 351), splice

35

domain, linker domain, binding domain, and 3' transcription terminator). In some embodiments, the coding sequence further comprises an ATG start codon at the 5' terminus.

In some embodiments, exemplary RNA exon editors described herein include those that
 5 comprise binding domains that bind to intron 3 of *HTT*, wherein such exemplary intron 3-binding RNA exon editors may comprise any one of SEQ ID NOs: 175-191, which sequences comprise the coding sequence (e.g., any one of SEQ ID NOs: 157, 352, or 353), splice domain, linker domain, binding domain, and 3' transcription terminator). In some embodiments, the coding sequence further comprises an ATG start codon at the 5' terminus.

10

In some embodiments, exemplary RNA exon editors described herein include those that comprise binding domains that bind to *HTT* intron 1 and binding domains that bind an *MSH3* intron, wherein such exemplary hybrid/dual *HTT/MSH3* RNA exon editors may comprise any one of SEQ ID NOs: 149-154, which sequences comprise a coding sequence (e.g., SEQ ID NO:
 15 3), splice domain, linker domain, binding domain, and 3' transcription terminator. In some embodiments, the coding sequence further comprises an ATG start codon at the 5' terminus. In some embodiments, exemplary hybrid/dual *HTT/MSH3* RNA exon editors comprise a binding domain capable of binding to both an *HTT* intron and an *MSH3* intron.

20

In some embodiments, exemplary RNA exon editors described herein include those that comprise binding domains that bind to *HTT* intron 2 and binding domains that bind an *MSH3* intron, wherein such exemplary hybrid/dual *HTT/MSH3* RNA exon editors may comprise any one of SEQ ID NOs: 212-223, which sequences comprise the coding sequence (e.g., SEQ ID NO: 59), splice domain, linker domain, binding domain, and 3' transcription terminator). In some
 25 embodiments, the coding sequence further comprises an ATG start codon at the 5' terminus. In some embodiments, exemplary hybrid/dual *HTT/MSH3* RNA exon editors comprise a binding domain capable of binding to both an *HTT* intron and an *MSH3* intron.

30

In some embodiments, an exemplary RNA exon editor includes a binding domain that binds to an *HTT* intron (e.g., intron 2) and a binding domain that binds to an *MSH3* intron (e.g., intron 5 or intron 15). In some embodiments, the *MSH3* binding domain comprises intron5_213_100 (SEQ ID NO: 140), intron5_188_150 (SEQ ID NO: 209), intron15_6523_120 (SEQ ID NO: 144), or intron15_6498_150 (SEQ ID NO: 210). In some embodiments, the *HTT* binding domain comprises *HTT_intron2_12061_150* (SEQ ID NO: 95). In some embodiments, the *MSH3* binding
 35 domain is 5' to the *HTT* binding domain. In some embodiments, the *MSH3* binding domain is 3' to the *HTT* binding domain. In some embodiments, a MALAT1 terminator is between the *MSH3*

binding domain and the *HTT* binding domain. In some embodiments, a MALAT1 terminator is not between the *MSH3* binding domain and the *HTT* binding domain. In some embodiments, a MALAT1 terminator is 3' to both the *MSH3* binding domain and the *HTT* binding domain.

5 In some embodiments, an exemplary RNA exon editor includes a binding domain that binds to an *MSH3* intron (e.g., intron 5 or intron 15). In some embodiments, the *MSH3* binding domain comprises intron5_213_100 (SEQ ID NO: 140), intron5_188_150 (SEQ ID NO: 209), intron15_6523_120 (SEQ ID NO: 144), or intron15_6498_150 (SEQ ID NO: 210). In some
10 embodiments, an *MSH3*-targeting RNA exon editor further comprises any splice domain disclosed herein, any 3X UBS sequence described herein, any AU-rich element described herein, any linker domain described herein, and/or any terminator sequence disclosed herein. In some embodiments, an *MSH3*-targeting RNA exon editor is administered in conjunction with an *HTT*-targeting exon editor. In some embodiments, an *MSH3*-targeting exon editor is not used in conjunction with an *HTT*-targeting exon editor.

15 In some embodiments, *MSH3* expression is reduced by a miRNA targeting *MSH3* mRNA. In some embodiments, the pri-miRNA comprises one or more of mir-30a [scaffold 5' (SEQ ID NO: 227); scaffold 3' (SEQ ID NO: 228); loop (SEQ ID NO: 229)], mir155 [5' scaffold (SEQ ID NO: 230); 3' scaffold (SEQ ID NO: 231); loop (SEQ ID NO: 232)], or mir-33 [scaffold 5' (SEQ ID NO: 259), scaffold 3' (SEQ ID NO: 260); loop (SEQ ID NO: 261)]. In some embodiments, the pri-miRNA comprises one or more of SEQ ID NOs: 234, 235, 238-241, or 262-269. In some
20 embodiments, the miRNA active sequence comprises, consists essentially of, or consists of any one of SEQ ID NOs: 224, 244, 246, 248, 250, 252, 254, 256, or 257 or a sequence at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical
25 to any one of SEQ ID NOs: 224, 244, 246, 248, 250, 252, 254, 256, or 257. The miRNA targeting *MSH3* mRNA may be used in combination with any of the RNA exon editors disclosed herein including, for example, RNA exon editors that target *HTT*, *MSH3*, or both. The miRNA targeting *MSH3* mRNA may be used without any RNA exon editor disclosed herein, but may instead be used independently to reduce *MSH3* expression. The miRNA targeting *MSH3* mRNA may be
30 used in a method of treating or preventing a trinucleotide repeat expansion disorder.

In some embodiments, *MSH3* expression is reduced by snRNA-based antisense RNA, which can induce exon skipping during pre-mRNA processing. In some embodiments, the snRNA construct comprises one or more of the following: SEQ ID NO: 274, which targets the junction between
35 *MSH3* intron 1 and exon 2; SEQ ID NO: 275, which targets the junction between *MSH3* exon 2 and intron 2; SEQ ID NOs: 278, 301, or 303, which target the junction between *MSH3* intron 2

and exon 3; SEQ ID NO: 279, 300, or 302, which target the junction between *MSH3* exon 3 and intron 3; SEQ ID NO: 281, which targets the junction between *MSH3* intron 3 and exon 4; SEQ ID NO: 282, which targets the junction between *MSH3* exon 4 and intron 4; SEQ ID NOs: 306 or 308, which target the junction between *MSH3* intron 5 and exon 6; SEQ ID NOs: 305 or 307, which targets the junction between *MSH3* exon 6 and intron 6; SEQ ID NOs: 311 or 313, which target the junction between *MSH3* intron 6 and exon 7; SEQ ID NOs: 310 or 312, which target the junction between *MSH3* exon 7 and intron 7; SEQ ID NOs: 316 or 318, which target the junction between *MSH3* intron 7 and exon 8; SEQ ID NOs: 315 or 317, which target the junction between *MSH3* exon 8 and intron 8; SEQ ID NOs: 321 or 323, which target the junction between *MSH3* intron 14 and exon 15; or SEQ ID NOs: 320 or 322, which target the junction between *MSH3* exon 15 and intron 15.

In some embodiments, one, two, three, four, five, or more of the above asRNA constructs are used (e.g., administered to a patient) in conjunction. For example, in some embodiments, a construct targeting the intron 1 – exon 2 junction and a construct targeting the exon 2 – intron 2 junction are used in conjunction; a construct targeting the intron 2 – exon 3 junction and a construct targeting the exon 3 – intron 3 junction are used in conjunction; a construct targeting the intron 3 – exon 4 junction and a construct targeting the exon 4 – intron 4 junction are used in conjunction; a construct targeting the intron 5 – exon 6 junction and a construct targeting the exon 6 – intron 6 junction are used in conjunction; a construct targeting the intron 6 – exon 7 junction and a construct targeting the exon 7 – intron 7 junction are used in conjunction; a construct targeting the intron 7 – exon 8 junction and a construct targeting the exon 8 – intron 8 junction are used in conjunction; or a construct targeting the intron 14 – exon 15 junction and/or a construct targeting the exon 15 – intron 15 junction are used in conjunction.

In some embodiments, a single asRNA construct that targets two intron – exon junctions is used. In some embodiments, the asRNA construct includes a sequence that is at least partially complementary to an entire exon sequence plus a portion of intronic sequence on either side of the exon, such as, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides of intronic sequence on either side of the exon sequence. Such constructs may target, for example, the intron – exon junctions on either side of any one of *MSH3* exons 2, 3, 4, 5, 6, 7, 8, or 15. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 324 (In3/Ex3/In2 sequence; comprises SEQ ID NO: 299), which targets the entire length of exon 3 and the flanking junctions on either side. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 327 (In6/Ex6/In5 sequence; comprises SEQ ID NO: 304), which targets the entire length of exon 6 and the flanking junctions

- on either side. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 330 In7/Ex7/In6 sequence; comprises SEQ ID NO: 309), which targets the entire length of exon 7 and the flanking junctions on either side. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 333 (In8/Ex8/In7 sequence; comprises SEQ ID NO: 314), which targets the entire length of exon 8 and the flanking junctions on either side. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 336 (In15/Ex15/In14 sequence; comprises SEQ ID NO: 319), which targets the entire length of exon 15 and the flanking junctions on either side.
- 10 In some embodiments, a single asRNA construct that targets two intron – exon junctions is used. In some embodiments, the asRNA construct includes a sequence that is at least partially complementary to a 5' intron – exon junction sequence, a sequence that is at least partially complementary to a 3' intron – exon junction sequence, and an unstructured linker joining these two sequences. Such constructs may target, for example, the intron – exon junctions on either side of *MSH3* exon 2, 3, 4, 5, 6, 7, 8, or 15. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 325 [U7SmOPT In3/Ex3 (SEQ ID NO: 300) + linker + Ex3/In2 (SEQ ID NO: 301)], which targets the junctions on either side of exon 3. In some embodiments, the asRNA construct is a U2 snRNA construct comprising SEQ ID NO: 326 [U2 In3/Ex3 (SEQ ID NO: 302) + linker + Ex3/In2 (SEQ ID NO: 303)], which targets the junctions on either side of exon 3. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 328 [U7SmOPT In6/Ex6 (SEQ ID NO: 305) + linker + Ex6/In5 (SEQ ID NO: 306)], which targets the junctions on either side of exon 6. In some embodiments, the asRNA construct is a U2 snRNA construct comprising SEQ ID NO: 329 [U2 In6/Ex6 (SEQ ID NO: 307) + linker + Ex6/In5 (SEQ ID NO: 308)], which targets the junctions on either side of exon 6.
- 25 In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 331 [U7SmOPT In7/Ex7 (SEQ ID NO: 310) + linker + Ex7/In6 (SEQ ID NO: 311)], which targets the junctions on either side of exon 7. In some embodiments, the asRNA construct is a U2 snRNA construct comprising SEQ ID NO: 332 [U2 In7/Ex7 (SEQ ID NO: 312) + linker + Ex7/In6 (SEQ ID NO: 313)], which targets the junctions on either side of exon 7. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 334 [U7SmOPT In8/Ex8 (SEQ ID NO: 315) + linker + Ex8/In7 (SEQ ID NO: 316)], which targets the junctions on either side of exon 8. In some embodiments, the asRNA construct is a U2 snRNA construct comprising SEQ ID NO: 335 [U2 In8/Ex8 (SEQ ID NO: 317) + linker + Ex8/In7 (SEQ ID NO: 318)], which targets the junctions on either side of exon 8. In some embodiments, the asRNA construct is a U7SmOPT construct comprising SEQ ID NO: 337 [U7SmOPT In15/Ex15 (SEQ ID NO: 320) + linker + Ex15/In14 (SEQ ID NO: 321)], which targets the junctions on either side of exon 15. In

some embodiments, the asRNA construct is a U2 snRNA construct comprising SEQ ID NO: 338 [U2 In15/Ex15 (SEQ ID NO: 322) + linker + Ex15/In14 (SEQ ID NO: 323)], which targets the junctions on either side of exon 15.

5 Some embodiments of *MSH3* splice modulator constructs include, operatively linked, a sequence encoding a small nuclear RNA (snRNA) sequence (e.g., a U7 Sm OPT sequence or a U2 snRNA sequence) and a sequence encoding an antisense RNA that promotes exon skipping of a target exon of *MSH3* pre-mRNA. The exon skipping may introduce frameshifts and/or premature stop codons, which may induce nonsense mediated decay or otherwise impair

10 production of functional MSH3. The antisense RNA that promotes exon skipping of a target exon may target one or both of the 5' exon-intron junction and a 3' exon-intron junction of the target exon of the *MSH3* pre-mRNA. As used herein, an antisense RNA is said to "target" a particular exon-intron junction if it has sufficient complementarity to a sequence surrounding the exon-intron junction that it is capable of promoting skipping of the target exon during pre-mRNA

15 processing. In some embodiments, an antisense RNA sequence includes a contiguous stretch of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, or 50 nucleotides that is 100% complementary to a contiguous stretch of nucleotides of the same length on the pre-mRNA that includes an exon-intron junction. In some embodiments, an antisense RNA sequence includes a contiguous stretch of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,

20 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, or 50 nucleotides that is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% complementary to a contiguous stretch of nucleotides of the same length on the pre-mRNA that includes an exon-intron junction. In some embodiments, an antisense RNA sequence includes a contiguous stretch of 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,

25 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, or 50 nucleotides that is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% complementary to a contiguous stretch of nucleotides of the same length on the pre-mRNA that includes an exon-intron junction or is within 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides of an exon-intron junction. A skilled person will understand that when less than full

30 complementarity is present, exon skipping functionality may be enhanced by increasing the length of the contiguous stretch of partially complementary nucleotides. A skilled person will also understand that GC content of the of the sequence surrounding an exon-intron junction may influence the ability of an antisense RNA sequence to effectively target the exon-intron junction and induce skipping. A higher GC content increases the strength of annealing, which

35 may lead to a smaller required stretch of complementarity.

In some embodiments, an MSH3 splice modulator comprises an antisense RNA sequence that targets two exon-intron junctions. In some embodiments, an antisense RNA sequence that targets two exon-intron junctions comprises a sequence that anneals to most or all of an entire exon sequence of a pre-mRNA. In some embodiments, the antisense RNA is 100% complementary to the entire pre-mRNA exon sequence. In some embodiments, the antisense RNA is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% complementary to the pre-mRNA exon sequence. In some embodiments, the antisense RNA further comprises nucleotide sequences that are at least partially complementary to intron sequences upstream and/or downstream of the targeted pre-mRNA exon sequence. In some embodiments, such intron sequences are at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides in length or more. In some embodiments, the antisense RNA sequence comprises a sequence that is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% complementary to such intron sequences.

In some embodiments, an antisense RNA sequence that targets two exon-intron junctions of a target exon comprises (a) two non-adjacent sequences that each anneal to a sequence surrounding an exon-intron junction of the target exon and (b) a linker sequence between the two non-adjacent sequences that has a low degree of complementarity (e.g., less than 50%) to the target exon. In some embodiments, the linker sequence comprises at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, or 50 nucleotides or from 10-30, 10-50, 15-25, or 18-22 nucleotides. In some embodiments, the linker has less than 70, 60, 50, 40, or 30% complementarity to all contiguous stretches of the target exon sequence that are the same length as the linker.

In some embodiments, any of the MSH3 splice modulators disclosed herein may be used in combination with any HTT RNA exon editors disclosed herein. In some embodiments, the MSH3 splice modulators disclosed herein are not used in combination with any RNA exon editors disclosed herein. Any of the MSH3 splice modulators described herein may be used independently in methods of treating trinucleotide repeat expansion disorders.

In some embodiments, multiple constructs are encoded on a single vector, such as, for example, an AAV vector. Any of the exon editors, miRNA, and asRNA sequences described herein can be combined on a single vector. In some embodiments, two or more of an exon editor, an miRNA, and an asRNA construct are encoded on a single vector. In some embodiments, the exon editor, miRNA, and/or asRNA, as applicable, target the same gene. For example, a single vector may

encode an *HTT* exon editor for generating corrected *HTT* mRNA and an *HTT*-targeting miRNA and/or an *HTT*-targeting asRNA to reduce the amount of defective HTT. In some embodiments, a single vector encodes an *HTT* exon editor (e.g., an exon editor comprising any one of SEQ ID NOs: 83-104, SEQ ID NOs: 113-125, SEQ ID NOs: 175-191, SEQ ID NOs: 199-206, or SEQ ID NOs: 23-36) and an *HTT*-targeting miRNA (e.g., SEQ ID NO: 339 and/or 342). In some embodiments, the vector comprises the sequence set forth in any one of SEQ ID NOs: 354 or 355 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to such sequences. In some embodiments, the exon editor, miRNA, and/or asRNA, as applicable, target different genes. In some embodiments, for example, a single vector encodes an *HTT* exon editor for generating corrected *HTT* mRNA and an *MSH3* exon editor, *MSH3* miRNA, and/or an *MSH3* asRNA (e.g., an snRNA construct) for knocking down *MSH3* expression. In some embodiments, a single vector encodes an *HTT* exon editor (e.g., an exon editor comprising any one of SEQ ID NOs: 83-104, SEQ ID NOs: 113-125, SEQ ID NOs: 175-191, SEQ ID NOs: 199-206, or SEQ ID NOs: 23-36) and an *MSH3*-targeting asRNA (e.g., a construct comprising any one or more of SEQ ID NOs: 284-293 or 324-338). In some embodiments, the vector comprises the sequence set forth in SEQ ID NO: 356 or 357 or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to such sequences. In some embodiments, the vector encodes an *HTT* exon editor (e.g., an exon editor comprising any one of SEQ ID NOs: 83-104, SEQ ID NOs: 113-125, SEQ ID NOs: 175-191, SEQ ID NOs: 199-206, or SEQ ID NOs: 23-36), an *MSH3*-targeting asRNA (e.g., a construct comprising any one or more of SEQ ID NOs: 284-293 or 324-338, and an *HTT*-targeting miRNA (e.g., SEQ ID NO: 339 and/or 342). In some embodiments, the vector comprises the sequence set forth in SEQ ID NO: 358 or 359, or a sequence having at least 80, 85, 90, 95, 96, 97, 98, or 99% identity to such sequences.

Different therapeutic modalities disclosed herein may be combined in various combinations, whether they are encoded on the same vector or on different vectors. For example, any of the HTT exon editors disclosed herein with a binding domain that targets intron 1 can be combined with any of the *MSH3*-targeting asRNA constructs disclosed herein and/or any of the *HTT*-targeting miRNA constructs disclosed herein. In some embodiments, the *HTT* exon editor with a binding domain that targets intron 1, the *MSH3*-targeting asRNA, and the *HTT*-targeting miRNA are included on the same vector. As another example, any of the HTT exon editors disclosed herein with a binding domain that targets intron 2 can be combined with any of the *MSH3*-targeting asRNA constructs disclosed herein and/or any of the *HTT*-targeting miRNA constructs disclosed herein. In some embodiments, the *HTT* exon editor with a binding domain that targets intron 2, the *MSH3*-targeting asRNA, and the *HTT*-targeting miRNA are included on the same vector. As another example, any of the HTT exon editors disclosed herein with a binding domain

that targets intron 3 can be combined with any of the *MSH3*-targeting asRNA constructs disclosed herein and/or any of the *HTT*-targeting miRNA constructs disclosed herein. In some embodiments, the *HTT* exon editor with a binding domain that targets intron 3, the *MSH3*-targeting asRNA, and the *HTT*-targeting miRNA are included on the same vector.

5

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) or any combination thereof. In one embodiment, the nucleic acid trans-splicing molecule includes
10 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
15 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
20 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.

25 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.

30 *Cell Line Assays*

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 *HTT* pre-mRNA at a sufficient level.

35 *RNA Exon Editor Screening Platform*

As described herein, 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 and/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. Both approaches are described below.

RNA exon editor screening in individual format:

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.

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 (*HTT*) 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.

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

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

At the protein level, TS activity is measured via Western blot analysis applied to protein extracted from cell or tissue samples. Beta-actin, which is a cytoskeletal protein, or tubulin, may

be used as a loading control. For those constructs that include a tag, e.g., a FLAG tag at their N-terminus, an antibody (Ab) specific for the tag (e.g., a FLAG-specific Ab) can be used to probe Western blots to assess ONT protein levels.

5 III. Vectors

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
10 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.
15 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.

In certain embodiments described herein, the trans-splicing molecule is delivered to the
20 selected cells, e.g., neuronal 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 neuronal 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
25 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
30 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.

35

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.

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 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
5 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
10 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
15 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
20 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
25 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
30 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
35 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 is 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.

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.

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.

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.

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.

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.

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 some embodiments, the promoter is specific for expression of the transgene in neuronal cells. In some embodiments, the promoter is specific for expression in cortical neurons (e.g., pyramidal neurons of the cortex). In some embodiments, the promoter is specific for expression of the transgene in striatal neurons (medium spiny neurons of the striatum). In some embodiments, the promoter is specific for expression of the transgene in hypothalamic neurons. In some embodiments, the transgene is expressed in at least one of the cell types or cells.

In another embodiment, the promoter is the native promoter for the target gene to be expressed. Useful promoters include, without limitation, the promoter CAGGS and neuronal specific promoters, including, without limitation, a human synapsin 1 gene promoter, a neuron-specific enolase (NSE) promoter, human synapsin 1 promoter, a CaMK kinase promoter, or an MeCP2 promoter. Other suitable promoters comprise inducible promoters, wherein such promoters initiate transcription only when the host cell is exposed to a stimulus which acts as a trigger for activating the promoter.

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 select among these, and other, expression control sequences without departing from the scope described herein.

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.

In some embodiments, 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 neuron-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 the 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.

In some embodiments, 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 or (B) 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
5 to increase backbone length and prevent reverse packaging of non-functional AAV genomes.

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
10 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.

In some embodiments, the CMV promoter is replaced by a CAGGS promoter, wherein the
15 CAGGS promoter is used to drive expression of an RNA exon editor described herein. See, e.g., FIG. 27. As shown therein, when comparing CAGGS 5' UTR with or without HTT 5'UTR, protein translation appears to be regulated through the HTT 5' UTR when the CAGGS 5' UTR + HTT 5' UTR are operably linked. When the HTT 5 UTR is removed, leaving only the CAGGS 5' UTR, the present inventors observed stronger protein expression, which activity may be due to some
20 element that upregulates translation in the CAGGS 5' UTR. Transduction experiments presented herein, wherein RNA exon editors were introduced into cells via an AAV, also demonstrated that the CAGGS promoter drives significant expression of RNA exons editors. See, e.g., FIGs. 55, 56, 58 and 59.

25 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
30 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
35 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.

5

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),

10 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

15 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

20 Provided herein are pharmaceutical compositions including a nucleic acid trans-splicing molecule, a proviral plasmid, or a rAAV comprising any of the *HTT* nucleic acid trans-splicing molecules described herein. In some embodiments, the pharmaceutical composition includes any of the 5' trans-splicing molecules described herein.

25 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.

30 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., a neuron). 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

35 agents, buffers, carriers, adjuvants, diluents, etc.

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
5 embodiment, the carrier is a 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.

In other embodiments, compositions containing trans-splicing molecules described herein
10 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 HD described herein comprises a recombinant adeno-associated vector carrying a nucleic acid sequence encoding a 5' trans-splicing molecule as described herein,
15 under the control of regulatory sequences which express the trans-splicing molecule in a neuronal cell of a mammalian subject, 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.

In yet another exemplary embodiment, the composition comprises a rAAV virus comprising any
20 of the *HTT* trans-splicing molecules described herein for *HTT* gene correction, the nucleic acid sequence under the control of a promoter which directs expression of the trans-splicing molecule in neuronal cells of the brain (e.g., neurons in the cortex, striatum, and/or hypothalamus), wherein the composition is formulated with a carrier and additional components
25 suitable for intracerebral delivery (e.g., via slow delivery or convection-enhanced infusion) or intracerebroventricular delivery. 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
30 instructions for administering the composition to an individual, e.g., as a treatment for HD.

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
35 pharmaceutical composition prior to administration.

V. Methods and Uses

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 *HTT*, and/or modulating expression of *HTT*, in a target cell (e.g., a neuron, e.g., pyramidal neurons of the cortex, medium spiny neurons in the striatum, and/or hypothalamic neurons) of an individual in, e.g., methods for treating diseases or disorders associated with mutations in the *HTT* gene, such as HD, including delaying or ameliorating symptoms associated with HD.

- 10 In some embodiments, symptoms of HD include, without limitation, involuntary jerking or writhing movements (chorea); muscle problems, such as rigidity or muscle contracture (dystonia); slow or unusual eye movements; impaired gait, posture and balance; difficulty with speech or swallowing. In some embodiments, symptoms of HD include, the following categories: muscular (e.g., abnormality walking, increased muscle activity, involuntary movements, problems with coordination, loss of muscle, and/or muscle spasms); cognitive (e.g., amnesia, delusion, lack of concentration, mental confusion, slowness in activity, and/or difficulty thinking and understanding); behavioral (e.g., compulsive behavior, fidgeting, irritability, or lack of restraint); psychological (e.g., delirium, depression, hallucination, and/or paranoia); and mood (e.g., anxiety, apathy, and/or mood swings). Additional symptoms commonly observed in HD patients include memory loss, tremor, and/or weight loss.

- Some embodiments of treatment methods described herein, or of constructs and/or molecules for use in methods of treatment or in the preparation of medicaments for treatment, involve targeting MSH3 without also targeting *HTT*. Some embodiments involve methods of treating or preventing trinucleotide repeat expansion disorders by administering to a subject a therapeutic agent that reduces MSH3 expression. In some embodiments, the trinucleotide repeat expansion disorder is a polyglutamine disease. In some embodiments, the polyglutamine disease is dentatorubropallidoluysian atrophy, Huntington's disease, spinal and bulbar muscular atrophy, spinocerebellar ataxia type 1, spinocerebellar ataxia type 2, spinocerebellar ataxia type 3, spinocerebellar ataxia type 6, spinocerebellar ataxia type 7, spinocerebellar ataxia type 17, or Huntington's disease-like 2. In some embodiments, the trinucleotide repeat expansion disorder is Huntington's disease. In some embodiments, the trinucleotide repeat expansion disorder is a non-polyglutamine disease. In some embodiments, the non-polyglutamine disease is fragile X syndrome, fragile X-associated tremor/ataxia syndrome, fragile XE mental retardation, Friedreich's ataxia, myotonic dystrophy type 1, spinocerebellar ataxia type 8, spinocerebellar

ataxia type 12, oculopharyngeal muscular dystrophy, Fragile X-associated premature ovarian failure, FRA2A syndrome, FRA7A syndrome, or early infantile epileptic encephalopathy.

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
5 furthermore useful for expressing functional *HTT*, and/or modulating expression of *HTT*, in a target cell (e.g., a neuron, e.g., pyramidal neurons of the cortex, medium spiny neurons in the striatum, and/or hypothalamic neurons) of an individual as applied, e.g., to their use for treating diseases or disorders associated with mutations in the *HTT* gene, such as HD, including
10 delaying or ameliorating symptoms associated with HD as described herein or to their use in the preparation of a medicament for the treatment of diseases or disorders associated with mutations in the *HTT* gene, such as HD, including delaying or ameliorating symptoms associated with HD as described herein. Such methods and uses involve contacting a target *HTT* gene (e.g., *HTT* pre-mRNA) with a trans-splicing molecule as described herein [e.g., a 5' trans-splicing
15 molecule, or a mixture of 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 coding domain of the trans-splicing molecule is spliced to the target *HTT* pre-mRNA to replace a part of the targeted pre-mRNA carrying one or more defects or mutations, with a biologically functional (i.e., healthy), or normal or wildtype or corrected mRNA
20 of the targeted gene, in order to correct expression of *HTT* in the target cell. Thus, the methods and compositions are used to treat the HD pathologies associated with the specific mutations such as, e.g., CAG repeats in excess of 40 repeats on the genomic level and polyglutamine stretches in excess of 40 glutamines in proteins transcribed and translated from genomic CAG repeats in excess of 40 repeats (disease causing expansion of genomic CAG repeats).

25 In some embodiments, provided herein are methods of expressing functional *HTT* 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 a cultured cell (e.g., a neuronal cell or precursor thereof) and the treated cultured cell reimplanted in the individual. In another embodiment, the method involves administering an rAAV carrying any of the 5' *HTT* trans-splicing molecules (including dual 5' *HTT/MSH3* trans-splicing molecules – tandem binding domains and tandemly arranged 5' *HTT* and 5' or 3' *MSH3*
30 trans-splicing molecules); 5' or 3' *MSH3* trans-splicing molecules that induce NMD; *MSH3* splice modulators; pri-miRNA comprising *MSH3* miRNA; or pri-miRNA comprising *HTT* miRNA
35

described herein or any combination thereof. In still another embodiment, the method involves administering a mixture of an rAAV carrying a 5' *HTT* trans-splicing molecule and an rAAV 5' *MSH3* trans-splicing molecule. Each of these embodiments may be combined with ASOs (e.g., SEQ ID NO: 131) that reduce cis-splicing for either of *HTT* or *MSH3*, thereby promoting trans-splicing to respective trans-splicing molecules. In some embodiments, ASOs that reduce cis-splicing for either one of or each of *HTT* and *MSH3* are used in combination with 5' *HTT* trans-splicing molecules (including dual 5' *HTT/MSH3* trans-splicing molecules) or combinations of 5' *HTT* trans-splicing molecules and 5' *MSH3* trans-splicing molecules. In some embodiments, anti-sense RNA ASOs that reduce cis-splicing for either one of or each of *HTT* and *MSH3* are used in combination with 5' *HTT* trans-splicing molecules (including dual 5' *HTT/MSH3* trans-splicing molecules) or combinations of 5' *HTT* trans-splicing molecules and 5' *MSH3* trans-splicing molecules. These methods comprise administering to an individual in need thereof an effective concentration of a composition of any of those described herein.

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

Methods of the invention include selecting a single trans-splicing molecule based on the location of a single mutation in *HTT* (e.g., a mutation of one allele of the individual). As described herein, the causative mutations associated with HD comprise the pathological expansion (greater than 35 or greater than 40 CAG repeats) of the CAG repeats in exon 1 of the *HTT* gene. Thus, in some embodiments, methods of the invention include administering a single trans-splicing molecule to correct the pathological expansion (greater than 35 or greater than 40 CAG

repeats) of the CAG repeats in exon 1 of the *HTT* gene, e.g., without regard to the location of any other mutations that may exist in the other allele.

Additionally presented are methods involving selecting a single 5' trans-splicing molecule to correct the pathological expansion (greater than 35 or greater than 40 CAG repeats) of the CAG repeats in exon 1 of the *HTT* pre-mRNA and reduce levels of *MSH3*, such that a single trans-splicing molecule capable of being packaged in an AAV vector is capable of correcting the pathological expansion of CAG repeats in *HTT* at the level of RNA and reducing levels of cellular *MSH3* to slow or inhibit expansion of the CAG repeats in *HTT* gene in the genome.

In other embodiments, provided herein are methods for correcting mutations within a *HTT* gene using two trans-splicing molecules—a 5' *HTT* trans-splicing molecule and a 5' *MSH3* trans-splicing molecule.

Nucleic acid trans-splicing molecules described herein and vectors, proviral plasmids, and AAV comprising same, as well as compositions comprising such nucleic acid trans-splicing molecules and vectors, proviral plasmids, and AAV comprising same are for use in medical treatment, in particular for use in the treatment of HD. In some embodiments, when using, e.g., an AAV vector (or other gene therapy vector) the AAV vector may be administered via direct infusion into the brain. In some embodiments, direct infusion comprises an intrathecal infusion of the AAV vector into the cerebrospinal fluid. Intrathecal infusion offers an efficient delivery mode into the CNS, wherein neurons can be targeted. In some embodiments, striatal and cortical structures may be targeted via intrastriatal convection enhanced diffusion (CED) delivery via injections into the striatum. In some embodiments, injections may be directed to the striatum and the thalamus to provide greater coverage of the structures of the brain implicated in HD. In some embodiments, AAV vectors may be delivered intrastrially or intrastrially and intrathalamically via CED injections into the striatum or the striatum and the thalamus. Such injections may be performed using magnetic resonance imaging-guided injections. Such methods for treatment are particularly useful for human subjects having HD. Such treatment involves human subjects having HD, including those having a genetic predisposition for developing HD that do not exhibit symptoms of HD. Accordingly, in some embodiments, treatment of human subjects with HD may include the treatment of any human subject carrying a Huntingtin allele with more than 35 CAG repeats.

In some embodiments, an effective concentration of a recombinant adeno-associated virus carrying a trans-splicing molecule as described herein ranges between about 10^8 and

- 10¹³ 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 10⁹ and 10¹³ vg/mL. In another embodiment, the effective concentration is about 1.5 x 10¹¹ vg/mL. In another embodiment, the effective concentration is about 5 x 10¹¹ vg/mL. In one
5 embodiment, the effective concentration is about 1.5 x 10¹⁰ vg/mL. In another embodiment, the effective concentration is about 2.8 x 10¹¹ vg/mL. In yet another embodiment, the effective concentration is about 1.5 x 10¹² vg/mL. In another embodiment, the effective concentration is about 1.5 x 10¹³ vg/mL.
- 10 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 brain. An effective dosage of a recombinant adeno-associated virus carrying a trans-splicing molecule as described herein ranges between about 10⁸ and 10¹³ vector genomes (vg) per dose (i.e., per injection). In one embodiment, the dosage ranges
15 between 10⁹ and 10¹³ vg. In another embodiment, the effective dosage is about 1.5 x 10¹¹ vg. In another embodiment, the effective dosage is about 5 x 10¹¹ vg. In one embodiment, the effective dosage is about 1.5 x 10¹⁰ vg. In another embodiment, the effective dosage is about 2.8 x 10¹¹ vg. In yet another embodiment, the effective dosage is about 1.5 x 10¹² vg. In another embodiment, the effective concentration is about 1.5 x 10¹³ vg. Still other dosages in these
20 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.
- 25 In some embodiments, the composition may be delivered in a volume of from about 50 µL 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, the age of the recipient, and the desired effect of the method. In one embodiment, the volume is about 50 µL. In another embodiment, the volume is about 70 µL. In another embodiment, the volume is about 100 µL. In another
30 embodiment, the volume is about 125 µL. In another embodiment, the volume is about 150 µL. In another embodiment, the volume is about 175 µL. In yet another embodiment, the volume is about 200 µL. In another embodiment, the volume is about 250 µL. In another embodiment, the volume is about 300 µL. In another embodiment, the volume is about 350 µL. In another embodiment, the volume is about 400 µL. In another embodiment, the volume is about 450 µL.
35 In another embodiment, the volume is about 500 µL. 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.

In some instances, treatments and uses described herein replace 10% or more of the target
5 *HTT* mRNA in the target cell (e.g., 11% or more of the target *HTT* mRNA in the target cell(s),
12% or more of the target *HTT* mRNA in the target cell(s), 13% or more of the target *HTT* mRNA
in the target cell(s), 14% or more of the target *HTT* mRNA in the target cell(s), 15% or more of
the target *HTT* mRNA in the target cell(s), 16% or more of the target *HTT* mRNA in the target
cell(s), 17% or more of the target *HTT* mRNA in the target cell(s), 18% or more of the target *HTT*
10 mRNA in the target cell(s), 19% or more of the target *HTT* mRNA in the target cell(s). In some
instances, treatments and uses described herein replace 20% or more of the target *HTT* mRNA
in the target cell (e.g., 21% or more of the target *HTT* mRNA in the target cell(s), 22% or more of
the target *HTT* mRNA in the target cell(s), 23% or more of the target *HTT* mRNA in the target
cell(s), 24% or more of the target *HTT* mRNA in the target cell(s), 25% or more of the target *HTT*
15 mRNA in the target cell(s), 26% or more of the target *HTT* mRNA in the target cell(s), 27% or
more of the target *HTT* mRNA in the target cell(s), 28% or more of the target *HTT* mRNA in the
target cell(s), 29% or more of the target *HTT* mRNA in the target cell(s), 30% or more of the
target *HTT* mRNA in the target cell(s), 31% or more of the target *HTT* mRNA in the target cell(s),
32% or more of the target *HTT* mRNA in the target cell(s), 33% or more of the target *HTT* mRNA
20 in the target cell(s), 34% or more of the target *HTT* mRNA in the target cell(s), 35% or more of
the target *HTT* mRNA in the target cell(s), 36% or more of the target *HTT* mRNA in the target
cell(s), 37% or more of the target *HTT* mRNA in the target cell(s), 38% or more of the target *HTT*
mRNA in the target cell(s), 39% or more of the target *HTT* mRNA in the target cell(s), 40% or
more of the target *HTT* mRNA in the target cell(s), 41% or more of the target *HTT* mRNA in the
25 target cell(s), 42% or more of the target *HTT* mRNA in the target cell(s), 43% or more of the
target *HTT* mRNA in the target cell(s), 44% or more of the *HTT* mRNA in the target cell(s), 45%
or more of the target *HTT* mRNA in the target cell(s), 46% or more of the target *HTT* mRNA in
the target cell(s), 47% or more of the target *HTT* mRNA in the target cell(s), 48% or more of the
target *HTT* mRNA in the target cell(s), 49% or more of the target *HTT* mRNA in the target cell(s),
30 or 50% or more of the target *HTT* mRNA in the target cell(s)). In some instances, treatments and
uses described herein replace 50% or more of the target *HTT* mRNA in the target cell (e.g., 51%
or more of the target *HTT* mRNA in the target cell(s), 52% or more of the target *HTT* mRNA in
the target cell(s), 53% or more of the target *HTT* mRNA in the target cell(s), 54% or more of the
target *HTT* mRNA in the target cell(s), 55% or more of the target *HTT* mRNA in the target cell(s),
35 56% or more of the target *HTT* mRNA in the target cell(s), 57% or more of the target *HTT* mRNA
in the target cell(s), 58% or more of the target *HTT* mRNA in the target cell(s), 59% or more of

the target *HTT* mRNA in the target cell(s), 60% or more of the target *HTT* mRNA in the target cell(s), 61% or more of the target *HTT* mRNA in the target cell(s), 62% or more of the target *HTT* mRNA in the target cell(s), 63% or more of the target *HTT* mRNA in the target cell(s), 64% or more of the target *HTT* mRNA in the target cell(s), 65% or more of the target *HTT* mRNA in the target cell(s), 66% or more of the target *HTT* mRNA in the target cell(s), 67% or more of the target *HTT* mRNA in the target cell(s), 68% or more of the target *HTT* mRNA in the target cell(s), 69% or more of the target *HTT* mRNA in the target cell(s), 70% or more of the target *HTT* mRNA in the target cell(s), 71% or more of the target *HTT* mRNA in the target cell(s), 72% or more of the target *HTT* mRNA in the target cell(s), 73% or more of the target *HTT* mRNA in the target cell(s), 74% or more of the *HTT* mRNA in the target cell(s), 75% or more of the target *HTT* mRNA in the target cell(s), 76% or more of the target *HTT* mRNA in the target cell(s), 77% or more of the target *HTT* mRNA in the target cell(s), 78% or more of the target *HTT* mRNA in the target cell(s), 79% or more of the target *HTT* mRNA in the target cell(s), or 80% or more of the target *HTT* mRNA in the target cell(s) (e.g., 81% or more of the target *HTT* mRNA in the target cell(s), 82% or more of the target *HTT* mRNA in the target cell(s), 83% or more of the target *HTT* mRNA in the target cell(s), 84% or more of the target *HTT* mRNA in the target cell(s), 85% or more of the target *HTT* mRNA in the target cell(s), 86% or more of the target *HTT* mRNA in the target cell(s), 87% or more of the target *HTT* mRNA in the target cell(s), 88% or more of the target *HTT* mRNA in the target cell(s), 89% or more of the target *HTT* mRNA in the target cell(s), 90% or more of the target *HTT* mRNA in the target cell(s), 91% or more of the target *HTT* mRNA in the target cell(s), 92% or more of the target *HTT* mRNA in the target cell(s), 93% or more of the target *HTT* mRNA in the target cell(s), 94% or more of the target *HTT* mRNA in the target cell(s), 95% or more of the target *HTT* mRNA in the target cell(s), 96% or more of the target *HTT* mRNA in the target cell(s), 97% or more of the target *HTT* mRNA in the target cell(s), 98% or more of the target *HTT* mRNA in the target cell(s), 99% or more of the target *HTT* mRNA in the target cell(s), or 100% or more of the target *HTT* mRNA in the target cell(s).

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.

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
5 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
10 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
15 composition is administered when more than 95% of the target cells are functioning or remaining.

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 yet
20 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 neuronal cells, such as administration of neurotrophic factors, anti-oxidants, and/or anti-
25 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.

For use in these methods, the volume and viral titer of each injection is determined individually
30 and may be the same or different from other injections performed in, e.g., the brain. The dosages, administrations, and regimens may be determined by the attending physician given the teachings of this disclosure.

VI. Examples

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.

5

Example 1: Trans-Splicing Efficiencies of HTT targeting RNA Exon Editors

Intron 1-targeting, intron 2-targeting, or intron 3-targeting Exon Editors were designed and tested for efficiency of replacement of the mutant *HTT* exon 1 (FIG. 3). In brief, HEK293 cells were transfected with *HTT* intron 1-targeting, intron 2-targeting or intron 3-targeting RNA Exon Editors that target various regions of the respective intron. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR. Exon editors produced by the construct depicted in FIG. 4 comprised a 5' UTR, exon 1 coding sequence, splice donor site, a linker, a binding domain, and a terminator sequence. The binding domain was varied to target different positions along intron 1 of *HTT*. *HTT* intron 1-targeting RNA Exon Editors which exhibit varying levels of trans-splicing efficiencies (% replacement) depending on where the binding domain targets in the intron were tested. FIGs. 5 and 6 shows the % replacement for the tested binding domains. Results shown relate to exemplary *HTT* intron 1-targeting RNA Exon Editors comprising the indicated binding domain targets, wherein the 5' UTR comprises the *HTT* 5' UTR (SEQ ID NO: 136) and the linker comprises the 40mer linker (SEQ ID NO: 37). NBD_150 is a control editor in which the binding domain targeting *HTT* is replaced with a binding domain that does not target *HTT*. The results demonstrate that targeting the 3' end of intron 1 near the branchpoint correlates with increased trans-splicing efficiency into the *HTT* pre-mRNA.

Exon editors are shown in FIG. 7 wherein expression was driven by a CMV promoter. The exon editor depicted comprised the *HTT* 5' UTR, N-terminal 3X FLAG tag, exon 1 coding sequence, a splice donor site, a linker, a binding domain (*HTT_intron1_11704_100*), and a terminator sequence. The activity of the exemplary *HTT* intron 1-targeting RNA Exon Editors that included various linkers was measured (FIG. 8). As shown therein, some linkers increased trans-splicing efficiency relative to the 40mer linker in *HTT* intron 1 (*HTT_intron1_11704_100*) Exon Editors. Exon editors depicted in FIG. 9 comprised a 5' UTR, exons 1-2 coding sequences, a splice donor site, a linker, a binding domain, and a terminator sequence. The binding domain was varied to target different positions along intron 2 of *HTT*. The activity of various exemplary *HTT* intron 2-targeting RNA Exon Editors, which exhibit varying levels of trans-splicing efficiencies (% replacement) depending on where the binding domain targets in the intron was measured is shown in FIGs. 10-11. Results shown relate to exemplary *HTT* intron 2-targeting RNA Exon Editors comprising the indicated binding domain targets, wherein the 5' UTR comprises the *HTT*

5' UTR and the linker comprises the 40mer linker. NBD, a control editor in which the binding domain targeting HTT was replaced with a binding domain that does not target HTT. A splice mutant, a control editor lacking a functional splice donor site was also tested.

5 HTT intron 2-targeting RNA Exon Editors targeting the region upstream of the branchpoint were designed to target upstream of the intron 2 branchpoint and vary in binding domain length (FIG. 12A). Exon Editor expression was driven by a CMV promoter. The Exon Editors comprised the HTT 5' UTR, N-terminal 3X FLAG tag for on-target protein detection, exons 1-2 coding sequence, a splice donor site, the 41mer_2 linker, the indicated binding domain, and a
10 terminator sequence. The binding domain was varied to target different lengths upstream of the branchpoint in intron 2. The length of the binding domain influenced targeting efficiencies, based on % replacement calculations (FIG. 12B). Analyses examining the association of binding domain length in the region of the intron upstream of the branchpoint with functionality identified binding domain lengths ranging from 125-200 nt as having the highest relative trans-splicing
15 efficiencies.

Constructs encoding HTT intron 2-targeting RNA Exon Editors with varying linkers were tested (FIG. 13). The Exon Editor expression was driven by a CMV promoter. Exemplary Exon Editors comprised the HTT 5' UTR, N-terminal 3X FLAG tag, exon 1-2 coding sequences, a splice donor
20 site, a linker, a binding domain, (HTT_intron2_12061_150), and a terminator sequence. Exon Editors comprising the indicated linkers did not exhibit significantly different trans-splicing efficiencies relative to the 40mer linker in HTT intron 2 (HTT_intron2_12061_150) (FIG. 14). Constructs encoding HTT Intron 3 targeting RNA Exon Editors were tested (FIG. 23). The activity of various exemplary HTT intron 3-targeting RNA Exon Editors, which show varying levels of
25 trans-splicing efficiencies (% replacement) depending on the binding site within the intron targeted by the binding domain was measured (FIG. 24). A direct comparison of HTT intron 2-targeting Exon Editors with HTT intron 3-targeting Exon Editors was performed (FIG. 25).

Self-splicing mitigation does not affect trans-splicing efficiency of HTT Exon Editors. Table 1
30 presents cryptic splice sites identified in the original exon editor and the sequence changes made in self-splicing mitigated exon editors (FIG. 26). In brief, HEK293 cells were transfected with HTT intron 2-targeting (HTT_intron2_12061_150) RNA Exon Editors with or without self-splicing mitigation. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR (FIG. 26).

35

RT-qPCR and Western blot image of lysates from HEK293 cells transfected with HTT RNA Exon Editors testing 2 promoters and different 5' UTR combinations were produced (FIG. 27). In short, HEK293 cells were transfected with N-terminally FLAG-tagged HTT Exon Editors driven by either the CMV or CAGGS promoter, with or without the HTT 5'UTR, and testing wild-type
5 (GTAAGT) splice site targeting intron 2 (HTT_intron2_12061_150), splice mutant targeting intron 2 (HTT_intron2_12061_150), or wild-type splice site with a non-targeting binding domain (NBD). RNA from these cells were subject to RT-qPCR (FIG. 27 upper panel). α -FLAG antibody was used to detect protein generated following successful trans-splicing (which comprises an N-terminal FLAG epitope) in whole cell lysates (FIG. 27 lower panel). ONT refers to successfully
10 trans-spliced on-target HTT protein and NSP refers to non-spliced protein.

Non-spliced protein (NSP) was reduced in a combinatorial manner by the inclusion of three tandem repeats of U1 snRNA binding site (3X UBS; SEQ ID NO: 345) and an AU-rich element (ARE; SEQ ID NO: 346) in an exemplary 5' HTT intron 1-targeting Exon Editor (FIG. 28). In short,
15 HEK293 cells were transfected with HTT intron 1-targeting (HTT_intron1_11704_100) RNA Exon Editors that were varied at their linker region to include the indicated NSP reduction elements. Cells were harvested 48 hours post-transfection and subjected to Western Blot analysis. ONT refers to successfully trans-spliced on-target HTT protein and NSP refers to non-spliced protein. In HTT intron 1-targeting Exon Editors, it was demonstrated that the inclusion of 3X UBS in the
20 linker domain reduced NSP levels by ~75% relative to that of the baseline 40mer-only linker control. Inclusion of ARE in the linker reduced the levels of NSP by ~40% relative to that of the baseline 40mer-only linker control. The combination of 3X UBS and ARE in the Exon Editor reduced the NSP level by ~88% relative to that of the baseline 40mer-only linker control.

Non-spliced protein (NSP) was reduced in a combinatorial manner by the inclusion of three tandem repeats of U1 snRNA binding site (3X UBS; SEQ ID NO: 345) and an AU-rich element (ARE; SEQ ID NO: 346) in an exemplary 5' HTT intron 2-targeting Exon Editor (FIG. 29). In short, HEK293 cells were transfected with HTT intron 2-targeting (HTT_intron2_12061_150) RNA Exon Editors that were varied at their linker region to include the indicated NSP reduction elements.
30 Cells were harvested 48 hours post-transfection, assayed for trans-splicing efficiencies by RT-qPCR (FIG. 29 upper panel) or subjected to Western Blot analysis (FIG. 29 lower panel). In HTT intron 2-targeting Exon Editors, the present inventors demonstrated that the inclusion of 3X UBS in the linker domain reduced NSP levels by ~ 38% relative to that of the baseline 40mer-only linker control. Inclusion of ARE in the linker reduced the levels of NSP by ~33% relative to that of
35 the baseline 40mer-only linker control. The combination of 3X UBS and ARE in the Exon Editor

worked combinatorially in reducing the NSP level by approximately 66% as compared to the baseline 40mer-only linker control.

Example 2: RNA Exon Editors with antisense oligonucleotides (ASOs)

- 5 A trans-splicing reaction and competition thereof with respect to cis-splicing was tested (FIG. 15). Anti-sense oligonucleotides (ASOs) were designed to block competing cis-splicing sites (ASO8-10), as well as cis-splicing sites for the upstream exon (ASO2-7) (FIG. 16). Each of these ASOs was co-transfected with the indicated HTT intron 2-targeting Exon Editor (HTT_intron2_12061_150) and assayed for trans-splicing efficiency in vitro in HEK293 cells. In
- 10 brief, HEK293 cells were co-transfected with an exemplary HTT intron 2-targeting RNA Exon Editor Construct (REEC) and ASOs designed to block the competing cis-splicing site or ASOs designed to block the splicing of the upstream intron. Cells were harvested 48 hours post-transfection and assayed for trans-splicing efficiencies by RT-qPCR.
- 15 Percent replacement activity of an exemplary HTT intron 2-targeting Exon Editor (HTT_intron2_12061_150) in combination with the indicated ASOs was measured (FIG. 17). ASO6, which was designed to block the cis-splicing of the upstream intron, led to an improvement in trans-splicing efficiencies in vitro. RT-qPCR and Western blot images (probed for the N-terminal FLAG epitope) of whole cell lysates from HEK293 cells transfected with HTT
- 20 RNA Exon Editors comprising the indicated elements were produced (FIG. 18 upper panel). It is noteworthy that on-target protein detection levels correlate with trans-splicing efficiency as represented by % Replacement of HTT RNA. HEK293 cells were transfected with N-terminally FLAG-tagged HTT Exon Editors that have a range of activity based on qPCR assays (FIG. 18 upper panel). Anti-FLAG antibody was used to detect protein generated following successful
- 25 trans-splicing in whole cell lysates (FIG. 18 lower panel). The intensity of the FLAG on-target (ONT) band scales with the relative performance of the Exon Editor based on qPCR.

- A combination strategy to correct mutant HTT by trans-splicing and knockdown unedited HTT species (including HTT1a) with a microRNA (miRNA) was tested (FIG. 46). A vectorized hybrid
- 30 molecule that combines a HTT Exon Editor with a miRNA targeting unedited HTT mRNA was used (FIG. 47). A short-hairpin RNA (shRNA) or microRNA (miRNA) designed to reduce HTT gene expression was added to an Exon Editor within the same cistron (e.g., in an intron of the Exon Editor) or as a separate cistron with its own regulatory sequences. The RNAi was designed to reduce expression of the unedited target (e.g., mutant HTT). For the purposes of selectively
- 35 reducing unedited HTT with RNAi, the Exon Editor comprised a HTT coding domain sequence

comprising a sequence-altered portion that renders edited HTT resistant to the shRNA or miRNA.

- Trans-splicing and HTT knockdown profiles of HTT Exon Editor, HTT miRNA-1, and a dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 are presented in FIG. 48. As shown therein, FIG. 48A) presents % trans-spliced (edited) HTT transcripts in all HTT transcripts and FIG. 48B) presents copy numbers of unedited and trans-spliced (edited) HTT transcripts for each of the conditions tested. In short, HEK293 cells were transfected with HTT intron 2-targeting RNA Exon Editor, HTT miRNA-1, and a dual hybrid molecule of HTT intron 2-targeting Exon Editor + HTT miRNA-1 as indicated in FIG. 48. Cells were harvested 48 hours post-transfection and RNA was subjected to RT-qPCR. The dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), comprised SEQ ID NO: 341 and SEQ ID NO: 204.
- Trans-splicing and HTT knockdown profiles of molecules containing HTT miRNA-1 and HTT miRNA-2 in vitro were produced (FIG. 49A-C). In short, HEK293 cells were transfected with HTT intron 2-targeting RNA Exon Editor, and dual hybrid molecules of HTT intron 2-targeting Exon Editor + HTT miRNA-1 or HTT miRNA-2. Cells were harvested 48 hours post-transfection and RNA was subjected to RT-qPCR analysis for a trans-splicing profile (FIG. 49A), HTT knockdown profile (FIG. 49B), and HTT copy number analysis (FIG 49C). The dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), comprised SEQ ID NO: 341 and SEQ ID NO: 204. The dual hybrid molecule of HTT Exon Editor and HTT miRNA-2 (SEQ ID NO: 355), comprised SEQ ID NO: 344 and SEQ ID NO: 204.
- MiRNA-1 successfully knocked down unedited *HTT* transcripts with minimal interaction with the Exon Editor and edited HTT transcript (FIG. 50). In short, HEK293 cells were transfected with a HTT Exon Editor +/- HTT miRNA-1 dual hybrid molecules. Cells were harvested 48 hours post-transfection, assayed for *HTT* knockdown and trans-splicing efficiencies by RT-qPCR (upper panels) or subjected to Western Blot analysis (lower panels). The dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), comprised SEQ ID NO: 341 and SEQ ID NO: 204.

Example 3: RNA Exon Editors with somatic CAG expansion inhibitors

- A potential mechanism of a hybrid therapeutic approach was designed and tested to treat HD. The hybrid therapeutic approach combined agents that inhibit somatic CAG expansion (e.g., by MSH3 reduction) with RNA Exon Editors targeting *HTT* pre-mRNA (FIG. 19). RNA Exon Editors

targeting *HTT* pre-mRNA served to replace any mutant *HTT* RNA that might have been produced from DNA that escaped inhibition of the somatic expansion process.

A potential mechanism of action of a tandem binding domain RNA Exon Editor targeting *HTT* and *MSH3* pre-mRNA was tested (FIG. 20). Expression of an Exon Editor was driven by a CMV promoter. Such exemplary Exon Editors comprised the *HTT* 5' UTR, N-terminal 3X FLAG tag, *HTT* exon 1 coding sequence, splice donor site, a linker, an *MSH3* binding domain (targeting, e.g., intron 5 or intron 15 of *MSH3* pre-mRNA), an *HTT* binding domain (e.g., *HTT_intron1_11704_100*), and a terminator sequence. The *HTT* binding domain targeted the Exon Editor to produce the corrected *HTT* RNA after successful trans-splicing, while the *MSH3* binding domain targeted the Exon Editor to produce a chimeric *HTT* exon 1-*MSH3* RNA molecule with a premature stop codon which was subject to nonsense-mediated decay (NMD) and led to the subsequent reduction of *MSH3* expression.

Tandem binding domain RNA Exon Editors targeting *HTT* and *MSH3* exhibited successful trans-splicing to both pre-mRNAs (FIG. 21A-21C). RT-qPCR of *HTT* on-target (ONT) trans-splicing efficiency (via *HTT* binding domain) (FIG. 21A), *HTT*-*MSH3* chimeric trans-splicing efficiency (via *MSH3* binding domain)(FIG. 21B), and *MSH3* RNA transcript expression level (FIG. 21C), on HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 1 and *MSH3* intron 5 was performed.

Tandem binding domain RNA Exon Editors targeting *HTT* and *MSH3* exhibited successful trans-splicing to both pre-mRNAs (FIG 22A-22C). RT-qPCR of *HTT* on-target (ONT) trans-splicing efficiency (via *HTT* binding domain)(FIG. 22A), *HTT*-*MSH3* chimeric trans-splicing efficiency (via *MSH3* binding domain)(FIG. 22B), and *MSH3* RNA transcript expression level (FIG. 22C), on HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 1 and *MSH3* intron 15 was performed.

A mechanism of action of a tandem binding domain RNA Exon Editor targeting *HTT* intron 2 and *MSH3* pre-mRNA was tested (FIG. 30). The Exon Editor expression was driven by a CMV promoter and contained the *HTT* 5' UTR, N-terminal 3X FLAG tag, *HTT* exon 1 and exon 2 coding sequence, splice donor site, the linker, *MSH3* binding domain (targeting intron 5 or intron 15 of *MSH3* pre-mRNA), *HTT* binding domain (*HTT_intron2_12061_150*), and a terminator sequence. The *HTT* binding domain targeted the Exon Editor to produce the corrected *HTT* RNA after successful trans-splicing, while the *MSH3* binding domain targeted the Exon Editor to produce a chimeric *HTT* exon 1+2 -*MSH3* RNA molecule with a premature stop codon which

was subject to nonsense-mediated decay (NMD) and led to the subsequent reduction of MSH3 expression.

RT-qPCR profiles of HTT trans-splicing and HTT-MSH3 chimera production (via MSH3 trans-splicing) in tandem binding domain Exon Editors were produced (FIG. 31A-31B). RT-qPCR of
5 HTT on-target (ONT) trans-splicing efficiency (via *HTT* intron 2-targeting binding domain) (FIG. 31A) and HTT-MSH3 chimeric trans-splicing efficiency (via *MSH3* intron 5-targeting binding domain) (FIG. 31B), in HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 2 and *MSH3* intron 5 was performed. The binding domain targeting *HTT*
10 intron 2 was HTT_intron2_12061_150 for all Exon Editors tested here, while
MSH3_intron5_213_100 and MSH3_intron5_188_150 were tested for the MSH3 intron 5-targeting binding domains. Binding domains were positioned in tandem, and the order of the binding domains was as indicated. An Exon Editor was also tested for each MSH3 binding domain with the MALAT1 triple helix terminator placed between the two tandem binding
15 domains.

RT-qPCR profiles of HTT trans-splicing and HTT-MSH3 chimera production (via MSH3 trans-splicing) in tandem binding domain Exon Editors were produced (FIG. 32A-32B). RT-qPCR of
20 HTT on-target (ONT) trans-splicing efficiency (via *HTT* intron 2-targeting binding domain) (FIG. 32A), and HTT-MSH3 chimeric trans-splicing efficiency (via *MSH3* intron 15-targeting binding domain) (FIG. 32B), in HEK293 cells transfected with tandem binding domain RNA Exon Editor targeting *HTT* intron 2 and *MSH3* intron 15 was performed. The binding domain targeting *HTT* intron 2 was HTT_intron2_12061_150 for all Exon Editors tested here, while
MSH3_intron15_6523_120 and MSH3_intron15_6498_150 were tested for the *MSH3* intron 15-targeting binding domains. Binding domains were positioned in tandem, and the order of the binding domains was as indicated. An Exon Editor was also tested for each MSH3 binding domain with the MALAT1 triple helix terminator placed between the two tandem binding
25 domains.

30 MSH3 can be knocked down by miRNAs that target the *MSH3* mRNA and degrade the transcript (FIG. 33). Western blot analysis of *MSH3* exon 23-targeting RNAi constructs was performed (FIG. 34). Constructs containing *MSH3* exon 23-targeting miRNA active sequence TTAATCCATAACTCCTTGC (SEQ ID NO: 224) were analyzed. ImageJ analysis was performed on the Western blots to analyze MSH3 protein knockdown (upper panel). U6 promoter-driven
35 shRNAs and CMV promoter-driven pri-miRNA mimics were designed and tested. Variations include the strand placement (5' arm or 3' arm) of the guide strand, including a bulge in the

stem structure, and varying the miRNA scaffold. Negative controls include constructs that contain a non-targeting sequence or a no hairpin loop control.

RT-qPCR and Western blot analysis of *MSH3*-targeting RNAi constructs were performed (FIG.

5 35). Constructs containing miRNAs targeting different regions of the *MSH3* transcript were analyzed. ImageJ analysis was performed on the Western blots to analyze MSH3 protein knockdown. CMV promoter-driven pri-miRNA mimics targeting different exonic sequences of *MSH3* were designed and tested.

10 MSH3 knockdown by small nuclear RNA (snRNA)-based antisense RNA (asRNA) was tested (FIG. 36). MSH3 can be inactivated by antisense RNAs encoded in a snRNA scaffold that anneal to *MSH3* splice junctions, preventing exon inclusion. This leads to exon skipping and the generation of a premature stop codon, ultimately causing NMD of the *MSH3* transcript. Relative *MSH3* RNA expression levels of exon 1 – exon 2 and exon 2 – exon 3 junctions in

15 MSH3 splice modulators targeting exon 2 skipping were measured (FIG. 37). Relative *MSH3* RNA expression levels of exon 2 – exon 3 and exon 3 – exon 4 junctions in MSH3 splice modulators targeting exon 3 skipping were measured (FIG 38). Relative *MSH3* RNA expression levels of exon 3 – exon 4 and exon 4 – exon 5 junctions in MSH3 splice modulators targeting

20 exon 4 skipping were measured (FIG. 39). Relative *MSH3* RNA expression levels of exon 2 – exon 3 and exon 3 – exon 4 junctions in MSH3 splice modulators targeting exon 3 skipping were measured (FIG. 40). Relative *MSH3* RNA expression levels of exon 5 – exon 6 and exon 6 – exon 7 junctions in MSH3 splice modulators targeting exon 6 skipping were measured (FIG. 41). Relative *MSH3* RNA expression levels of exon 6 – exon 7 and exon 7 – exon 8 junctions in

25 MSH3 splice modulators targeting exon 7 skipping were measured (FIG. 42). The splice modulator transcript was SEQ ID NO: 331. The anti-sense RNA (asRNA) comprised therein were In7/Ex7-1 (asRNA region SEQ ID NO: 310) + linker + Ex7/In6-1 (asRNA region SEQ ID NO: 311). Relative *MSH3* RNA expression levels of exon 7 – exon 8 and exon 8 – exon 9 junctions in

30 MSH3 splice modulators targeting exon 8 skipping were measured (FIG. 43). Relative *MSH3* RNA expression levels of exon 14 – exon 15 and exon 15 – exon 16 junctions in MSH3 splice modulators targeting exon 15 skipping were measured (FIG. 44).

MSH3 exon 7 splice modulators show reduction of *MSH3* RNA and protein levels (FIG. 45A-C).

In short, HEK293 cells were transfected with snRNA-based splice modulators designed to skip *MSH3* exon 7. Cells were harvested 48 hours post-transfection, assayed for MSH3 knockdown

35 by RT-qPCR (FIG. 45B) or subjected to Western Blot analysis (FIG. 45C).

Reduction of MSH3 by splice modulation in combination with HTT trans-splicing tested (FIG. 51). Results showed the performance of HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecules (FIG. 52A-B). In short, HEK293 cells were transfected with a HTT Exon Editor +/- MSH3 Splice Modulator. Cells were harvested 48 hours post-transfection and assayed for trans-splicing profiles (FIG. 52A) and MSH3 knockdown profiles (FIG. 52B) by RT-qPCR. The HTT Exon Editor and MSH3 Splice Modulator dual hybrid molecules were MSH3 Splice Modulator + HTT Exon Editor (SEQ ID NO: 356), which comprised SEQ ID NO: 331 and SEQ ID NO: 204.

A comparison of self-complementary AAV (scAAV) to single-stranded AAV (ssAAV) using a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule was performed (FIG. 53). In short, HEK293 cells were transduced with scAAV or ssAAV expressing a HTT Exon Editor + *MSH3* exon 7-skipping Splice Modulator dual hybrid molecule. AAV2 serotype was used. Cells were harvested 48 hours after transduction and subjected to RT-qPCR and Western Blot analysis. The HTT Exon Editor and MSH3 Splice Modulator dual hybrid molecule (SEQ ID NO: 357), comprised SEQ ID NO: 331 and SEQ ID NO: 204, in a head-to-head orientation.

The performance of HTT Exon Editor + HTT miRNA + MSH3 Splice Modulator triple hybrid molecules compared to its controls was tested (FIG. 54; Top, Middle, and Bottom panels). In short, HEK293 cells were transfected with HTT Exon Editor +/- HTT miRNA-1 or HTT miRNA-2 +/- MSH3 Splice Modulator. Cells were harvested 48 hours post-transfection and assayed for % trans-spliced HTT transcripts (FIG. 54; Top panel), HTT knockdown profiles (FIG. 54; Middle panel), and MSH3 knockdown profiles by RT-qPCR (FIG. 54; Bottom panel). Control hybrid molecules comprising an Exon Editor with a splice donor mutation, a Splice Modulator that contains a scrambled asRNA sequence, or a miRNA that contains a scrambled asRNA, were also tested. "1" for HTT miRNA indicates HTT miRNA-1, "2" for HTT miRNA indicates HTT miRNA-2 was used. SM referred to splice mutant. Scr referred to scrambled control. miR-33 was used for the miRNA scaffold. The Dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), comprised SEQ ID NO: 341 and SEQ ID NO: 204. The Dual hybrid molecule of HTT Exon Editor and HTT miRNA-2 (SEQ ID NO: 355), comprised SEQ ID NO: 344 and SEQ ID NO: 204. The HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecules comprised MSH3 Splice Modulator + HTT Exon Editor (SEQ ID NO: 356), which comprised SEQ ID NO: 331 and SEQ ID NO: 204. The HTT Exon Editor + HTT miRNA + MSH3 Splice Modulator triple hybrid molecule (SEQ ID NO: 358), comprised SEQ ID NO: 331, SEQ ID NO: 341, and SEQ ID NO: 204. The HTT Exon Editor + HTT miRNA + MSH3 Splice Modulator triple hybrid molecule (SEQ ID NO: 359), comprised SEQ ID NO: 331, SEQ ID NO: 344, and SEQ ID NO: 204.

35

HTT miRNA knockdown profiles in iCell GlutaNeurons were measured by RT-ddPCR and Western blotting (FIG. 56). In short, iCell GlutaNeurons were transduced with a HTT Exon Editor + HTT miRNA dual hybrid molecule or a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule, both packaged in AAV2.7m8. Cells were harvested for RT-ddPCR and Western Blots
5 after 18 days. The dual hybrid molecule of HTT Exon Editor and HTT miRNA-1 (SEQ ID NO: 354), comprised SEQ ID NO: 341 and SEQ ID NO: 204. The HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule (SEQ ID NO: 357), comprised SEQ ID NO: 331 and SEQ ID NO: 204, in a head-to-head orientation.

10 A MSH3 knockdown profile of the MSH3 Splice Modulator in iCell GlutaNeurons as measured by RT-ddPCR and Western blotting was produced (FIG. 57). In short, iCell GlutaNeurons were transduced with a HTT Exon Editor + HTT miRNA-1 dual hybrid molecule or a HTT Exon Editor + MSH3 Splice Modulator dual hybrid molecule, both packaged in AAV2.7m8. Cells were harvested for RT-ddPCR and Western Blots after 18 days.

15 HTT trans-splicing profiles in the BAC-CAG mouse brain were produced (FIG. 58). In short, neonatal ICV injections (at 1E+11 or 3E+11 vg/animal) in BAC-CAG mice were performed with the indicated HD molecules packaged in AAV9. Mouse cortex and striatum were harvested 4 weeks post-injection and the efficiencies of HTT Exon Replacement by trans-splicing profiled by
20 RT-ddPCR and Western Blotting. As shown in FIG. 58, analysis of HTT trans-splicing profiles demonstrated that upwards of 30% HTT replacement was achieved in the mouse brain by neonatal ICV injection. The CAGGS promoter-driven Exon Editor outperformed its CMV promoter equivalent, as indicated by higher % HTT replacement in both the cortex and the striatum. Among the CMV promoter-driven Exon Editors, a clear dose response was observed,
25 whereby animals that received 3E+11 vg had higher % HTT replacement than those that received 1E+11 vg. Significantly, trans-spliced full-length HTT protein was detected by Western blotting against the N-terminal FLAG tag. This experiment also indicated that with increasing levels of Exon Editor transcripts, higher level of trans-splicing was achieved (FIG. 59), suggesting that identifying a stronger promoter or administering a higher dose might result in
30 even higher Exon Editor activity. In addition, a relationship between Exon Editor RNA copy number and trans-splicing efficiency (% HTT replacement) in the BAC-CAG mouse brain was observed (FIG 59).

While preferred embodiments of the present disclosure have been shown and described herein,
35 it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in

the art without departing from the disclosure. It should be understood that various alternatives to the embodiments of the present disclosure may be employed in practicing the present disclosure. It is intended that the following embodiments define the scope of the present disclosure and that methods and structures within the scope of these embodiments and their equivalents be covered thereby.

VII. ENUMERATED EMBODIMENTS

1. A nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain comprising HTT exon 1 and HTT exon 2;
 - (b) a splicing domain; and
 - (c) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 2.
2. The nucleic acid trans-splicing molecule of embodiment 1, wherein the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 60-81.
3. A nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain comprising HTT exon 1;
 - (b) a splicing domain; and
 - (c) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 1, and wherein the binding domain comprises any one of SEQ ID NOs: 8-19.
3. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the coding domain comprises, consists essentially of, or consists of HTT exon 1 or HTT exon 1 and HTT exon 2.
4. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the coding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 3 or 59.
5. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, wherein the coding domain, the splicing domain, and the binding domain are operatively linked in a 5'-to-3' direction.

6. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a linker, wherein the coding domain, splicing domain, linker, and binding domain are operatively linked in a 5'-to-3' direction.
7. A nucleic acid trans-splicing molecule of embodiment 6, wherein the linker comprises, consists essentially of, or consists of:
- a sequence ranging from 20 to 50 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine;
 - a sequence ranging from 20 to 45 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine; or
 - a sequence ranging from 22 to 42 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.
8. The nucleic acid trans-splicing molecule of any one of embodiments 6 or 7, wherein the linker comprises, consists essentially of, or consists of any one of:
- SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38;
 - SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39;
 - SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; or
 - SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41.
9. The nucleic acid trans-splicing molecule of embodiment 6, wherein the linker comprises, consists essentially of, or consists of any one of:
- SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42;
 - SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43;
 - SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44;
 - SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45;
 - SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46;
 - SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106;
 - SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107;
 - SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108;
 - SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109;
 - SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110;
 - SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111; or

SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112.

10. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a triple helix terminator, wherein the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator are operatively linked in a 5'-to-3' direction.
11. The nucleic acid trans-splicing molecule of embodiment 10, wherein the triple helix terminator comprises, consists essentially of, or consists of SEQ ID NO: 5 or a sequence having at least 90% identity to SEQ ID NO: 5.
12. The nucleic acid trans-splicing molecule of embodiment 11, wherein the triple helix terminator comprises, consists essentially of, or consists of SEQ ID NO: 6.
13. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a 5' untranslated region (5' UTR), wherein the 5' UTR, the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.
14. The nucleic acid trans-splicing molecule of embodiment 13, wherein the 5' UTR is an *HTT* 5' UTR.
15. The nucleic acid trans-splicing molecule of embodiment 14, wherein the *HTT* 5' UTR comprises, consists essentially of, or consists of SEQ ID NO: 136 or a sequence having at least 90% identity to SEQ ID NO: 136.
16. The nucleic acid trans-splicing molecule of any one of the preceding embodiments, further comprising a sequence encoding an epitope tag, wherein the 5' UTR, when present, the epitope tag, the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.
17. The nucleic acid trans-splicing molecule of embodiment 16, wherein the sequence encoding the epitope tag comprises, consists essentially of, or consists of a SEQ ID NO: 4.
18. A nucleic acid trans-splicing molecule comprising:

- (a) a coding domain comprising HTT exon 1 and HTT exon 2;
(b) a splicing domain;
(c) a linker; and
(d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target
5 intron comprises intron 2.
19. The nucleic acid trans-splicing molecule of embodiment 18, wherein the binding domain
comprises, consists essentially of, or consists of any one of SEQ ID NOs: 60-81.
- 10 20. A nucleic acid trans-splicing molecule comprising:
(a) a coding domain comprising HTT exon 1;
(b) a splicing domain;
(c) a linker; and
(d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target
15 intron comprises intron 1.
21. The nucleic acid trans-splicing molecule of embodiment 20, wherein the binding domain
comprises, consists essentially of, or consists of any one of SEQ ID NOs: 8-21.
- 20 22. The nucleic acid trans-splicing molecule of any one of embodiments 18-21, wherein the
linker comprises, consists essentially of, or consists of any one of SEQ ID NOs:
SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38;
SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39;
SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40;
25 SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41;
SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42;
SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43;
SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44;
SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45;
30 SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46;
SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106;
SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107;
SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108;
SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109;
35 SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110;
SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111; or

SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112.

23. The nucleic acid trans-splicing molecule of embodiment 22, further comprising a triple helix terminator, wherein the coding domain, the splicing domain, the linker, the binding domain, and the triple helix terminator are operatively linked in a 5'-to-3' direction; and optionally, further comprising a 5' UTR, wherein the 5' UTR, when present, the coding domain, the splicing domain, the linker, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.
24. A nucleic acid trans-splicing molecule comprising a linker, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 20 to 50 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.
25. The nucleic acid trans-splicing molecule of embodiment 24, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 20 to 45 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.
26. The nucleic acid trans-splicing molecule of embodiment 24 or embodiment 25, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 22 to 42 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.
27. The nucleic acid trans-splicing molecule of any one of embodiments 24-26, wherein the linker comprises, consists essentially of, or consists of:
- SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38;
SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39;
SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; or
SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41.
28. A nucleic acid trans-splicing molecule comprising a linker, wherein the linker comprises, consists essentially of, or consists of:

- SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42;
SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43;
SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44;
SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45;
5 SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46;
SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106;
SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107;
SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108;
SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109;
10 SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110;
SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111; or
SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112.
29. The nucleic acid trans-splicing molecule of any one of embodiments 1-28, further
15 comprising a binding domain that binds a target intron of an *MSH3* pre-mRNA or a second
nucleic acid trans-splicing molecule comprising a binding domain that binds a target intron
of an *MSH3* pre-mRNA.
30. The nucleic acid trans-splicing molecule of embodiment 29, wherein the *MSH3* target intron
20 comprises any one of intron 5 or intron 15 of *MSH3*.
31. The nucleic acid trans-splicing molecule or the second nucleic acid trans-splicing molecule
of any one of embodiments 29 or 30, wherein the binding domain that binds a target intron
of an *MSH3* pre-mRNA comprises, consists essentially of, or consists of any one of SEQ ID
25 NOs: 140, 142, 143, 144, 146, or 147.
32. The nucleic acid trans-splicing molecule or a second nucleic acid trans-splicing molecule of
any one of embodiments 29-31, wherein the nucleic acid trans-splicing molecule comprises
any one of SEQ ID NOs: 149 - 154.
30
33. A ribonucleic acid trans-splicing molecule comprising any one of SEQ ID NOs: 23-36, 47-
56, 83-105, or 113-125.
34. A ribonucleic acid trans-splicing molecule transcribed from the nucleic acid trans-splicing
35 molecule of any one of embodiments 1-33.

35. The nucleic acid trans-splicing molecule of any one of embodiments 1-34, wherein the *HTT* pre-mRNA comprises at least one mutation associated with Huntington's Disease (HD).
36. The nucleic acid trans-splicing molecule of any one of embodiments 1-35, wherein the at
5 least one mutation associated with HD comprises an expansion of CAG repeats in an *HTT* gene allele.
37. The nucleic acid trans-splicing molecule of embodiment 36, wherein the expansion of CAG repeats in an *HTT* gene allele comprises greater than 35 CAG repeats.
- 10 38. The nucleic acid trans-splicing molecule of any one of embodiments 35-37, wherein the at least one mutation associated with HD is autosomal dominant.
39. The nucleic acid trans-splicing molecule of any one of embodiments 36-38, wherein the at
15 least one mutation associated with HD is expressed in at least one of cortical pyramidal neurons, striatal medium spiny neurons, or hypothalamic neurons.
40. A vector comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-39.
- 20 41. The vector of embodiment 40, wherein the vector comprises a 5' regulatory domain operatively linked 5' to the coding domain.
42. The vector of embodiment 41, wherein the 5' regulatory domain is operatively linked to a 5'
25 untranslated region.
43. The vector of any one of embodiments 41-42, wherein the 5' regulatory domain comprises a constitutive promoter or a tissue specific promoter.
- 30 44. The vector of embodiment 43, wherein the constitutive promoter is a CMV promoter.
45. A proviral plasmid comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-39.

46. An adeno-associated virus (AAV) comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-39, wherein the AAV optionally comprises a 5' regulatory domain operatively linked 5' to the nucleic acid trans-splicing molecule.
- 5 47. The AAV of embodiment 46, wherein the AAV comprises a 5' regulatory domain operatively linked 5' to the coding domain.
48. The AAV of any one of embodiments 46 or 47, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.
- 10 49. The AAV of any one of embodiments 46-48, wherein the 5' regulatory domain comprises a constitutive promoter.
50. The AAV of embodiment 49, wherein the constitutive promoter is a CMV promoter.
- 15 51. The AAV of any one of embodiments 46-50, wherein the AAV exhibits neuronal tropism.
52. The AAV of any one of embodiments 46-50, wherein the AAV is AAV9, AAV8, AAV5, or AAV2.
- 20 53. A composition comprising the nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, or the AAV of any one of embodiments 46-52.
- 25 54. The composition of embodiment 53, comprising a pharmaceutically acceptable excipient.
55. The composition of any one of embodiments 53 or 54, further comprising at least one anti-sense oligonucleotide or a construct that encodes at least one anti-sense RNA that inhibits cis-splicing of the HTT pre-mRNA.
- 30 56. The composition of embodiment 55, wherein the at least one anti-sense oligonucleotide comprises any one of SEQ ID NOs: 126-135 or the construct that encodes the at least one anti-sense RNA binds to a target sequence bound by any one of SEQ ID NOs: 126-135.

57. The composition of any one of embodiments 55-56, wherein the at least one anti-sense oligonucleotide comprises SEQ ID NO: 131 or the construct that encodes the at least one anti-sense RNA binds to a target sequence bound by SEQ ID NO: 131.
- 5 58. A method of expressing biologically active HTT in a target cell to restore functional levels of HTT protein in the target cell, the method comprising transducing the target cell with the nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57.
- 10 59. The method of embodiment 58, wherein at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, or at least 45% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced.
- 15 60. The method of embodiment 59, wherein at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, or at least 85% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced.
- 20 61. The method of embodiment 60, wherein at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced.
- 25 62. The method of any one of embodiments 58-61, wherein functional levels of HTT are restored in the target cell by expressing biologically functional HTT protein.
- 30 63. A method of reducing expression of HTT comprising a polyglutamine repeat exceeding 35 consecutive glutamine residues in a subject, the method comprising transfecting or transducing a target cell, more particularly a neuron, in the subject with the nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57.
- 35 64. A method of correcting at least one mutation in an HTT exon sequence in an HTT pre-mRNA 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 1-39, the vector of any one

of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57.

- 5 65. A method of treating HD in a subject in need thereof, the method comprising administering to the subject the nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57 in a therapeutically effective amount.
- 10 66. The method of any one of embodiments 58-65, the method comprising administration of the nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57 to the subject's brain.
- 15 67. The method of any one of embodiments 58-66, wherein the subject is a mammal, preferentially a rodent, non-human primate, or a human.
- 20 68. The method of any one of embodiments 63-67, wherein the subject is genetically predisposed to have HD or has been diagnosed with HD.
- 25 69. The nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57 for use in preventing or treating HD in a subject in need thereof.
- 30 70. The nucleic acid trans-splicing molecule of any one of embodiments 1-39, the vector of any one of embodiments 40-44, the proviral plasmid of embodiment 45, the AAV of any one of embodiments 46-52, or the composition of any one of embodiments 53-57 for use in the preparation of a medicament for the treatment or prevention of HD in a subject in need thereof.
- 35 71. A method comprising introducing into a cell a nucleic acid trans-splicing molecule configured to splice to both a first target pre-mRNA and a second target pre-mRNA, wherein the splicing to the first target pre-mRNA corrects a defect in the first target pre-mRNA, and

wherein the splicing to the second target pre-mRNA introduces a defect in the second target pre-mRNA.

72. The method of embodiment 71, wherein the nucleic acid trans-splicing molecule comprises
5 a first binding domain configured to target an intron of the first target pre-mRNA and a
second binding domain configured to target an intron of the second target pre-mRNA.
73. The method of embodiment 71 or 72, wherein the nucleic acid trans-splicing molecule
further comprises a coding domain sequence comprising a functional sequence of one or
10 more exons of the first target pre-mRNA that corrects the defect in the first target pre-
mRNA.
74. The method of any one of embodiments 71 to 73, wherein the defect in the second target
pre-mRNA comprises a frameshift in the coding sequence of the second target pre-mRNA.
15
75. The method of embodiment 74, wherein the frameshift creates a premature termination
codon in the second target pre-mRNA.
76. The method of any one of embodiments 71 to 75, wherein the defect comprises the
20 endogenous start codon of the second target pre-mRNA being eliminated.
77. The method of any one of embodiments 71 to 76, wherein the defect comprises an inserted
5' UTR that prevents translation of the protein encoded by the second target pre-mRNA.
- 25 78. The method of any one of embodiments 71 to 77, wherein the defect comprises an inserted
3' UTR that destabilizes the pre-mRNA or prevents export of the second target pre-mRNA
from the nucleus.
79. The method of any one of embodiments 71 to 78, wherein the defect comprises elimination
30 of a 5' cap or a 3' polyA tail from the second target pre-mRNA.
80. The method of any one of embodiments 71 to 79, wherein the defect causes nonsense-
mediated decay of the second target pre-mRNA.

81. The method of any one of embodiments 71 to 80, wherein the introducing causes the abundance of gene product of the second target pre-mRNA in the cell to be reduced compared to the abundance of the gene product before the introducing.
- 5 82. A method of reducing the abundance of a protein in a cell, the method comprising introducing into the cell a nucleic acid trans-splicing molecule that introduces a defect into a pre-mRNA that encodes the protein.
83. The method of embodiment 82, wherein the defect comprises one or more of the following:
- 10 (a) a frameshift introduced into the coding sequence of the pre-mRNA;
(b) elimination of the endogenous start codon of the pre-mRNA;
(c) introduction of a premature stop codon into the coding sequence of the pre-mRNA;
(d) replacement of the endogenous coding sequence of the pre-mRNA with an alternative coding sequence;
- 15 (e) insertion of a 5' UTR that prevents translation of the endogenous coding sequence of the pre-mRNA;
(f) insertion of a 3' UTR that destabilizes the pre-mRNA;
(g) insertion of 3' UTR that prevents the pre-mRNA from being exported from the nucleus;
- 20 (h) elimination of a 5' cap from the pre-mRNA; or
(i) elimination of a 3' polyA tail from the pre-mRNA.
84. The method of embodiment 82 or 83, wherein the protein is MSH3.
- 25 85. The method of embodiment 84, wherein the nucleic acid trans-splicing molecule comprises a binding domain that binds to an intron of the pre-mRNA.
86. The method of embodiment 85, wherein the nucleic acid trans-splicing molecule comprises a heterologous coding domain sequence.
- 30

CLAIMS

1. An HTT nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain comprising HTT exon 1 and HTT exon 2;
 - (b) a splicing domain; and
 - 5 (c) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 2.
2. The HTT nucleic acid trans-splicing molecule of claim 1, wherein the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 60-81, or a
10 sequence having at least 90% identity to any one of SEQ ID NOs: 60-81.
3. An HTT nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain comprising HTT exons 1-3;
 - (b) a splicing domain; and
 - 15 (c) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 3.
4. The HTT nucleic acid trans-splicing molecule of claim 3, wherein the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 158-174, or a
20 sequence having at least 90% identity to any one of SEQ ID NOs: 158-174.
5. An HTT nucleic acid trans-splicing molecule comprising:
 - (a) a coding domain comprising HTT exon 1;
 - (b) a splicing domain; and
 - 25 (c) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 1, and wherein the binding domain comprises any one of SEQ ID NOs: 8-21.
6. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, wherein
30 the coding domain comprises, consists essentially of, or consists of HTT exon 1; HTT exon 1 and HTT exon 2; or HTT exons 1-3.
7. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, wherein
35 the coding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 3, 59, 157, or 349-353 or a sequence having at least 90% identity to SEQ ID NOs: 3, 59, 157, or 349-353.

8. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, wherein the coding domain, the splicing domain, and the binding domain are operatively linked in a 5'-to-3' direction.
- 5
9. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a linker, wherein the coding domain, splicing domain, linker, and binding domain are operatively linked in a 5'-to-3' direction.
- 10
10. A HTT nucleic acid trans-splicing molecule of claim 9, wherein the linker comprises, consists essentially of, or consists of:
- a sequence ranging from 20 to 50 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine;
- 15
- a sequence ranging from 20 to 45 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine; or
- a sequence ranging from 22 to 42 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.
- 20
11. The HTT nucleic acid trans-splicing molecule of any one of claims 9 or 10, wherein the linker comprises, consists essentially of, or consists of any one of:
- SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38;
- 25
- SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39;
- SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; or
- SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41.
12. The HTT nucleic acid trans-splicing molecule of claim 9, wherein the linker comprises,
- 30
- consists essentially of, or consists of any one of:
- SEQ ID NO: 37 or a sequence having at least 90% identity to SEQ ID NO: 37;
- SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42;
- SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43;
- SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44;
- 35
- SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45;
- SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46;

SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106;
SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107;
SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108;
SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109;
5 SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110;
SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111;
SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112;
SEQ ID NO: 197 or a sequence having at least 90% identity to SEQ ID NO: 197; or
SEQ ID NO: 198 or a sequence having at least 90% identity to SEQ ID NO: 198.

10

13. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a triple helix terminator, wherein the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator are operatively linked in a 5'-to-3' direction.

15

14. The HTT nucleic acid trans-splicing molecule of claim 13, wherein the triple helix terminator comprises, consists essentially of, or consists of SEQ ID NO: 5 or a sequence having at least 90% identity to SEQ ID NO: 5.

20

15. The HTT nucleic acid trans-splicing molecule of claim 13, wherein the triple helix terminator comprises, consists essentially of, or consists of SEQ ID NO: 6.

16. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a 5' untranslated region (5' UTR), wherein the 5' UTR, the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.

25

17. The HTT nucleic acid trans-splicing molecule of claim 16, wherein the 5' UTR is an *HTT* 5' UTR.

30

18. The HTT nucleic acid trans-splicing molecule of claim 17, wherein the *HTT* 5' UTR comprises, consists essentially of, or consists of any one of SEQ ID NO: 136 or 192 or a sequence having at least 90% identity to any one of SEQ ID NO: 136 or 192.

35

19. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a sequence encoding an epitope tag, wherein the 5' UTR, when present, the

epitope tag, the coding domain, the splicing domain, the linker, when present, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.

5 20. The HTT nucleic acid trans-splicing molecule of claim 19, wherein the sequence encoding the epitope tag comprises, consists essentially of, or consists of a SEQ ID NO: 4.

21. An HTT nucleic acid trans-splicing molecule comprising:

- 10 (a) a coding domain comprising HTT exon 1 and HTT exon 2;
(b) a splicing domain;
(c) a linker; and
(d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 2.

15 22. The HTT nucleic acid trans-splicing molecule of claim 21, wherein the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 60-81, or a sequence having at least 90% identity to any one of SEQ ID NOs: 60-81.

23. An HTT nucleic acid trans-splicing molecule comprising:

- 20 (a) a coding domain comprising HTT exons 1-3;
(b) a splicing domain;
(c) a linker; and
(d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target intron comprises intron 3.

25

24. The HTT nucleic acid trans-splicing molecule of claim 23, wherein the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 158-174, or a sequence having at least 90% identity to any one of SEQ ID NOs: 158-174.

30 25. An HTT nucleic acid trans-splicing molecule comprising:

- (a) a coding domain comprising HTT exon 1;
(b) a splicing domain;
(c) a linker; and
(d) a binding domain that binds a target intron of an HTT pre-mRNA, wherein the target
35 intron comprises intron 1.

26. The HTT nucleic acid trans-splicing molecule of claim 25, wherein the binding domain comprises, consists essentially of, or consists of any one of SEQ ID NOs: 8-21, or a sequence having at least 90% identity to any one of SEQ ID NOs: 8-21.
- 5 27. The HTT nucleic acid trans-splicing molecule of any one of claims 21-26, wherein the linker comprises, consists essentially of, or consists of any one of SEQ ID NOs:
SEQ ID NO: 37 or a sequence having at least 90% identity to SEQ ID NO: 37;
SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38;
SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39;
10 SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40;
SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41;
SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42;
SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43;
SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44;
15 SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45;
SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46;
SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106;
SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107;
SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108;
20 SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109;
SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110;
SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111;
SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112;
SEQ ID NO: 197 or a sequence having at least 90% identity to SEQ ID NO: 197; or
25 SEQ ID NO: 198 or a sequence having at least 90% identity to SEQ ID NO: 198.
28. The HTT nucleic acid trans-splicing molecule of claim 27, further comprising a triple helix terminator, wherein the coding domain, the splicing domain, the linker, the binding domain,
30 and the triple helix terminator are operatively linked in a 5'-to-3' direction; and optionally, further comprising a 5' UTR, wherein the 5' UTR, when present, the coding domain, the splicing domain, the linker, the binding domain, and the triple helix terminator, when present, are operatively linked in a 5'-to-3' direction.
- 35 29. A nucleic acid trans-splicing molecule comprising a linker, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 20 to 50 nucleotides long,

wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.

- 5 30. The nucleic acid trans-splicing molecule of claim 29, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 20 to 45 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74% guanines interspersed with thymidine/uridine.

10

31. The nucleic acid trans-splicing molecule of claim 29 or claim 30, wherein the linker comprises, consists essentially of, or consists of a sequence ranging from 22 to 42 nucleotides long, wherein the linker comprises 60-80% guanines interspersed with thymidine/uridine; 65-75% guanines interspersed with thymidine/uridine; or 66-74%
15 guanines interspersed with thymidine/uridine.

15

32. The nucleic acid trans-splicing molecule of any one of claims 29-31, wherein the linker comprises, consists essentially of, or consists of:
SEQ ID NO: 38 or a sequence having at least 90% identity to SEQ ID NO: 38;
20 SEQ ID NO: 39 or a sequence having at least 90% identity to SEQ ID NO: 39;
SEQ ID NO: 40 or a sequence having at least 90% identity to SEQ ID NO: 40; or
SEQ ID NO: 41 or a sequence having at least 90% identity to SEQ ID NO: 41.

20

33. A nucleic acid trans-splicing molecule comprising a linker, wherein the linker comprises,
25 consists essentially of, or consists of:
SEQ ID NO: 37 or a sequence having at least 90% identity to SEQ ID NO: 37;
SEQ ID NO: 42 or a sequence having at least 90% identity to SEQ ID NO: 42;
SEQ ID NO: 43 or a sequence having at least 90% identity to SEQ ID NO: 43;
SEQ ID NO: 44 or a sequence having at least 90% identity to SEQ ID NO: 44;
30 SEQ ID NO: 45 or a sequence having at least 90% identity to SEQ ID NO: 45;
SEQ ID NO: 46 or a sequence having at least 90% identity to SEQ ID NO: 46;
SEQ ID NO: 106 or a sequence having at least 90% identity to SEQ ID NO: 106;
SEQ ID NO: 107 or a sequence having at least 90% identity to SEQ ID NO: 107;
SEQ ID NO: 108 or a sequence having at least 90% identity to SEQ ID NO: 108;
35 SEQ ID NO: 109 or a sequence having at least 90% identity to SEQ ID NO: 109;
SEQ ID NO: 110 or a sequence having at least 90% identity to SEQ ID NO: 110;

25

30

35

SEQ ID NO: 111 or a sequence having at least 90% identity to SEQ ID NO: 111;
SEQ ID NO: 112 or a sequence having at least 90% identity to SEQ ID NO: 112;
SEQ ID NO: 197 or a sequence having at least 90% identity to SEQ ID NO: 197; or
SEQ ID NO: 198 or a sequence having at least 90% identity to SEQ ID NO: 198.

5

34. The HTT nucleic acid trans-splicing molecule of any one of claims 1-28, further comprising a binding domain that binds a target intron of an *MSH3* pre-mRNA.
35. The HTT nucleic acid trans-splicing molecule of claim 34, wherein the *MSH3* target intron
10 comprises any one of intron 5 or intron 15 of *MSH3*.
36. The HTT nucleic acid trans-splicing molecule of any one of claims 34 or 35, wherein the binding domain that binds a target intron of an *MSH3* pre-mRNA comprises, consists essentially of, or consists of any one of SEQ ID NOs: 140, 142, 144, 146, 209, or 210 or a
15 sequence having at least 90% identity to any one of SEQ ID NOs: 140, 142, 144, 146, 209, or 210.
37. The HTT nucleic acid trans-splicing molecule of any one of claims 34-36, wherein the nucleic acid trans-splicing molecule comprises any one of SEQ ID NOs: 149 – 154 or SEQ ID
20 NOs: 212-223, or a sequence having at least 90% identity to any one of SEQ ID NOs: 149 – 154 or SEQ ID NOs: 212-223.
38. The HTT nucleic acid trans-splicing molecule of any one of the preceding claims, further comprising a nucleic acid sequence encoding a pri-miRNA that comprises a microRNA
25 (miRNA) sequence specific for exon 1 of endogenous HTT mRNA, wherein exon 1 of the nucleic acid trans-splicing molecule comprises a change in nucleotide sequence that impairs binding of the miRNA to mRNA encoded at least in part by the nucleic acid trans-splicing molecule.
- 30 39. The HTT nucleic acid trans-splicing molecule of claim 38, wherein the miRNA sequence comprises any one of SEQ ID NOs: 339 or 342 or a nucleic acid sequence having at least 90% identity to any one of SEQ ID NOs: 339 or 342.
- 35 40. The HTT nucleic acid trans-splicing molecule of claim 39, wherein the nucleic acid sequence encoding the pri-miRNA comprises any one of SEQ ID NOs: 341 or 344.

41. The HTT nucleic acid trans-splicing molecule of any one of claims 38 to 40, wherein the pri-miRNA comprises a mir-33 scaffold sequence.
42. The HTT nucleic acid trans-splicing molecule of any one of claims 38 to 40, wherein the pri-miRNA comprises a mir-30a scaffold sequence, a mir-30a loop sequence, a mir-155 scaffold sequence, a mir-155 loop sequence, a mir-33 scaffold sequence, or a mir-33 loop sequence.
43. The HTT nucleic acid trans-splicing molecule of claim 42, wherein the mir-30a scaffold sequence comprises a 5' scaffold sequence set forth in SEQ ID NO: 227 or a 3' scaffold sequence set forth in SEQ ID NO: 228; wherein the mir-30a loop sequence comprises SEQ ID NO: 229; wherein the mir-155 scaffold sequence comprises a 5' scaffold sequence set forth in SEQ ID NO: 230 or a 3' scaffold sequence set forth in SEQ ID NO: 231; wherein the mir-155 loop sequence comprises SEQ ID NO: 232; wherein the mir-33 scaffold sequence comprises a 5' scaffold sequence set forth in SEQ ID NO: 259 or a 3' scaffold sequence set forth in SEQ ID NO: 260; or wherein the mir-33 loop sequence comprises SEQ ID NO: 261.
44. An MSH3 exon skipping nucleic acid construct comprising, operatively linked:
- (a) a sequence encoding an antisense RNA that promotes exon skipping of a target exon of MSH3 pre-mRNA, wherein the target exon is any one of MSH3 exons 2-4, 6-8, or 15, wherein the target exon comprises a 5' exon-intron junction and a 3' exon-intron junction sequence; and
 - (b) a sequence encoding a small nuclear RNA (snRNA) sequence.
45. The MSH3 exon skipping nucleic acid construct of claim 44, further comprising a U1 promoter and a U1 terminator operatively linked to (a) and (b).
46. The MSH3 exon skipping nucleic acid construct of claim 44 or 45, wherein the snRNA is a modified snRNA.
47. The MSH3 exon skipping nucleic acid construct of claim 46, wherein the modified snRNA comprises a U7 Sm OPT sequence or a U2 snRNA sequence.
48. The MSH3 exon skipping nucleic acid construct of any one of claims 44-47, wherein the antisense RNA targets either the 5' exon-intron junction or the 3' exon-intron junction of the target exon.

49. The MSH3 exon skipping nucleic acid construct of claim 44, wherein the antisense RNA comprises, consists essentially of, or consists of any one of SEQ ID NOs: 274, 275, 276, 277, 278, 279, 280, 300, 302, 301, 303, 281, 282, 306, 308, 305, 307, 311, 313, 310, 312, 316, 318, 315, 317, 321, 323, 320, or 322 or a sequence having at least 90% identity to any one of SEQ ID NOs: 274, 275, 276, 277, 278, 279, 280, 300, 302, 301, 303, 281, 282, 306, 308, 305, 307, 311, 313, 310, 312, 316, 318, 315, 317, 321, 323, 320, or 322.
50. The MSH3 exon skipping nucleic acid construct of claim 44, comprising any one of SEQ ID NOs: 284, 285, 286, 287, 288, 289, 290, 325, 326, 291, 292, 328, 329, 331, 332, 334, 335, 337, and 338.
51. The MSH3 exon skipping nucleic acid construct of any one of claims 44-47, wherein the antisense RNA targets both the 5' exon-intron junction and the 3' exon-intron junction.
52. The MSH3 exon skipping nucleic acid construct of claim 51, wherein the antisense RNA comprises a sequence that is at least 80% complementary to the entire sequence of the target exon.
53. The MSH3 exon skipping nucleic acid construct of claim 51, wherein the antisense RNA further comprises:
- (a) a sequence that is at least 80% complementary to a 5-nucleotide sequence upstream of the 5' exon-intron junction; and
 - (b) a sequence that is at least 80% complementary to a 5-nucleotide sequence downstream of the 3' exon-intron junction.
54. The MSH3 exon skipping nucleic acid construct of claim 52 or 53, wherein the antisense RNA comprises any one of SEQ ID NOs: 299, 304, 309, 314, or 319, or a sequence having at least 90% identity to any one of SEQ ID NOs: 299, 304, 309, 314, or 319.
55. The MSH3 exon skipping nucleic acid construct of any one of claims 51-53, comprising any one of SEQ ID NOs: 324, 327, 330, 333, or 336.
56. The MSH3 exon skipping nucleic acid construct of claim 51, wherein the antisense RNA comprises, operatively linked in a 5' to 3' direction:
- (a) a sequence that targets the 3' exon-intron junction;
 - (b) a linker sequence of at least 15 nucleotides that does not anneal to the target exon; and

(c) a sequence that targets the 5' exon-intron junction.

57. The MSH3 exon skipping nucleic acid construct of claim 56, wherein the linker sequence is less than 50% complementary to all sequences of the target exon of the same length as the linker.

58. The MSH3 exon skipping nucleic acid construct of claim 56 or 57, wherein the antisense RNA comprises any one of SEQ ID NOs: 300, 301, 302, 303, 305, 306, 307, 308, 310, 311, 312, 313, 315, 316, 317, 318, 320, 321, 322, or 323, or any combination thereof.

59. The MSH3 exon skipping nucleic acid construct of any one of claims 56-58, comprising any one of SEQ ID NOs: 325, 326, 328, 329, 331, 332, 334, 335, 337, or 338.

60. The MSH3 exon skipping nucleic acid construct of any one of claims 44-59, wherein the antisense RNA targets MSH3 exon 7.

61. The MSH3 exon skipping nucleic acid construct of claim 60, comprising SEQ ID NO: 309.

62. The MSH3 exon skipping nucleic acid construct of claim 60, comprising from 5' to 3':
(a) SEQ ID NO: 310 (In7/Ex7 asRNA), SEQ ID NO: 298 (linker), and SEQ ID NO: 311 (In7/Ex7 asRNA); or
(b) SEQ ID NO: 312 (In7/Ex7 asRNA), SEQ ID NO: 298 (linker), and SEQ ID NO: 313 (In7/Ex7 asRNA).

63. The MSH3 exon skipping nucleic acid construct of claim 60, wherein the antisense RNA comprises any one of SEQ ID NOs: 309, 310, 311, 312, or 313, or any combination thereof or a sequence having at least 90% identity to any one of SEQ ID NOs: 309, 310, 311, 312, or 313.

64. The MSH3 exon skipping nucleic acid construct of claim 60, comprising at least one of SEQ ID NOs: 330-332, or any combination thereof.

65. An MSH3 miRNA nucleic acid construct comprising a sequence encoding a pri-miRNA that comprises a scaffold sequence, a loop sequence, and a miRNA sequence that targets endogenous MSH3 mRNA, wherein:
(a) the scaffold sequence is derived from mir-30a, mir-33, or mir-155;
(b) the loop sequence is derived from mir-22, mir-30a, mir-33, or mir-155; and

(c) the miRNA sequence comprises any one of SEQ ID NOs: 224, 244, 246, 248, 250, 252, 254, 256, or 257 or a sequence having at least 90% identity to any one of SEQ ID NOs: 224, 244, 246, 248, 250, 252, 254, 256, or 257.

5 66. The MSH3 miRNA nucleic acid construct of claim 65, wherein the scaffold sequence comprises any one of SEQ ID NOs: 227, 228, 230, 231, 259 or 260.

67. The MSH3 miRNA nucleic acid construct of claim 65 or 66, wherein the loop sequence comprises any one of SEQ ID NOs: 229, 232, or 261.

10

68. The MSH3 miRNA nucleic acid construct of any one of claims 65-67, wherein the pri-miRNA sequence comprises any one of SEQ ID NOs: 234, 235, 238-241, or 262-269.

15

69. The MSH3 miRNA nucleic acid construct of any one of claims 65-68, wherein the sequence encoding the pri-miRNA is operatively linked to a U6 promoter or a CMV promoter.

70. An MSH3 nucleic acid trans-splicing molecule comprising:

(a) coding domain sequence;

(b) a splicing domain; and

20

(c) a binding domain that binds a target intron of an MSH3 pre-mRNA;

wherein the coding domain sequence is not an MSH3 coding domain sequence.

71. The nucleic acid trans-splicing molecule of claim 70, wherein the coding domain sequence comprises a sequence that results in a frameshift in a mature MSH3 mRNA when trans-spliced into the MSH3 pre-mRNA.

25

72. The nucleic acid trans-splicing molecule of claim 70 or 71, wherein the coding domain sequence comprises one or more of exons 1, 2, and 3 of HTT.

30 73. The nucleic acid trans-splicing molecule of any one of claims 70-72, wherein the target intron of the MSH3 pre-mRNA is intron 5 or intron 15.

74. The nucleic acid trans-splicing molecule of any one of claims 70 to 73, wherein the binding domain comprises any one of SEQ ID NOs: 140, 142, 144, 146, 209, or 210, or a sequence having at least 90% identity to any one of SEQ ID NOs: 140, 142, 144, 146, 209, or 210.

35

75. An HTT trans-splicing and MSH3 exon skipping nucleic acid construct comprising:
- (a) the HTT nucleic acid trans-splicing molecule of any one of claims 1-28; and
 - (b) the MSH3 exon skipping nucleic acid construct of any one of claims 44-64.
- 5 76. The HTT trans-splicing and MSH3 exon skipping nucleic acid construct of claim 75, wherein
- (a) and (b) are comprised on a single vector.
77. The HTT trans-splicing and MSH3 exon skipping nucleic acid construct of claim 76, wherein
- 10 the single vector is an AAV vector.
78. The HTT trans-splicing and MSH3 exon skipping nucleic acid construct of claim 77,
- comprising any one of SEQ ID NOs: 356, 357, 363, or 364.
79. The HTT trans-splicing and MSH3 exon skipping nucleic acid construct of claim 77, wherein
- 15 the AAV vector is a scAAV or ssAAV vector.
80. The HTT trans-splicing and MSH3 exon skipping nucleic acid construct of claim 79,
- comprising any one of SEQ ID NOs: 369, 370, and 371.
- 20 81. An HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct
- comprising:
- (a) the HTT nucleic acid trans-splicing molecule of any one of claims 38-43; and
 - (b) the MSH3 exon skipping nucleic acid construct of any one of claims 44-64.
- 25 82. The HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of
- claim 81, wherein (a) and (b) are comprised on a single vector.
83. The HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of
- claim 82, comprising any one of SEQ ID NOs: 358 or 359.
- 30 84. The HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of
- claim 82 or 83, wherein the single vector is an AAV vector.
85. An HTT trans-splicing and MSH3 miRNA nucleic acid construct comprising:
- 35 (a) the HTT nucleic acid trans-splicing molecule of any one of claims 1-28; and
- (b) the MSH3 miRNA nucleic acid construct of any one of claims 65-69.

86. The HTT trans-splicing and MSH3 miRNA nucleic acid construct of claim 85, wherein (a) and (b) are comprised on a single vector.
87. The HTT trans-splicing and MSH3 miRNA nucleic acid construct of claim 86, comprising any one of SEQ ID NOs: 354 or 355.
88. The HTT trans-splicing and MSH3 miRNA nucleic acid construct of claim 86 or 87, wherein the vector is an AAV vector.
89. An AAV vector comprising the HTT nucleic acid trans-splicing molecule of any one of claims 38-43.
90. The AAV vector of claim 89, comprising any one of SEQ ID NOs: 356, 357, 363, or 364.
91. An AAV vector comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 and 34-43 or the nucleic acid trans-splicing molecule of any one of claims 29-33.
92. A ribonucleic acid trans-splicing molecule comprising any one of SEQ ID NOs: 23-36, 47-56, 83-105, 113-125, 175-191, or 199-206.
93. A ribonucleic acid trans-splicing molecule transcribed from the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 39-43, or the nucleic acid trans-splicing molecule of any one of claims 29-33.
94. The HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43, wherein the HTT pre-mRNA comprises at least one mutation associated with Huntington's Disease (HD).
95. The HTT nucleic acid trans-splicing molecule of claim 94 wherein the at least one mutation associated with HD comprises an expansion of CAG repeats in an *HTT* gene allele.
96. The HTT nucleic acid trans-splicing molecule of claim 95, wherein the expansion of CAG repeats in an *HTT* gene allele comprises greater than 35 CAG repeats.
97. The HTT nucleic acid trans-splicing molecule of any one of claims 94-96, wherein the at least one mutation associated with HD is autosomal dominant.

- 5 98. The HTT nucleic acid trans-splicing molecule of any one of claims 94-97, wherein the at least one mutation associated with HD is expressed in at least one of cortical pyramidal neurons, striatal medium spiny neurons, or hypothalamic neurons.
- 10 99. A vector comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84, or the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88.
- 15 100. A vector comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43.
- 20 101. The vector of claim 100, wherein the vector comprises a 5' regulatory domain operatively linked 5' to the coding domain.
102. The vector of claim 101, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.
- 25 103. The vector of claim 101 or 102, wherein the 5' regulatory domain comprises a constitutive promoter or a tissue specific promoter.
104. The vector of claim 103, wherein the constitutive promoter is a CMV promoter or a CAGGS promoter.
- 30 105. A proviral plasmid comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and
- 35

MSH3 exon skipping nucleic acid construct of any one of claims 81-84, or the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88.

- 5 106. An adeno-associated virus (AAV) comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT
10 miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84, or the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88.
- 15 107. An adeno-associated virus (AAV) comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43, wherein the AAV optionally comprises a 5' regulatory domain operatively linked 5' to the nucleic acid trans-splicing molecule.
- 20 108. The AAV of claim 107, wherein the AAV comprises a 5' regulatory domain operatively linked 5' to the coding domain.
- 25 109. The AAV of any one of claims 107 or 108, wherein the 5' regulatory domain is operatively linked to a 5' untranslated region.
- 30 110. The AAV of any one of claims 107-109, wherein the 5' regulatory domain comprises a constitutive promoter.
- 35 111. The AAV of claim 110, wherein the constitutive promoter is a CMV promoter or a CAGGS promoter.
112. The AAV of any one of claims 107-111, wherein the AAV exhibits neuronal tropism.
113. The AAV of any one of claims 107-111, wherein the AAV is AAV9, AAV8, AAV5, AAV2, AAV7, or AAV2.7m8, AAV-retro, AAV1, AAV4, or AAV-PHP.eB.
114. A composition comprising the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA

nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; or the AAV of any one of claims 89-91 or 106-113.

115. The composition of claim 114, comprising a pharmaceutically acceptable excipient.

116. The composition of any one of claims 114 or 115, further comprising at least one anti-sense oligonucleotide or a construct that encodes at least one anti-sense RNA that inhibits cis-splicing of the HTT pre-mRNA.

117. The composition of claim 116, wherein the at least one anti-sense oligonucleotide comprises any one of SEQ ID NOs: 126-135 or the construct that encodes the at least one anti-sense RNA binds to a target sequence bound by any one of SEQ ID NOs: 126-135.

118. The composition of claim 116 or 117, wherein the at least one anti-sense oligonucleotide comprises SEQ ID NO: 131 or the construct that encodes the at least one anti-sense RNA binds to a target sequence bound by SEQ ID NO: 131.

119. A method of expressing biologically active HTT in a target cell to restore functional levels of HTT protein in the target cell, the method comprising transducing the target cell with the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104, the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the composition of any one of claims 114-118.

120. The method of claim 119, wherein at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, or at least 45% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced.
- 5 121. The method of claim 120, wherein at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, or at least 85% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced.
- 10 122. The method of claim 121, wherein at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% of the HTT pre-mRNA comprising at least one mutation associated with HD in the target cell is replaced.
- 15 123. The method of any one of claims 119-122, wherein functional levels of HTT are restored in the target cell by expressing biologically functional HTT protein and/or mutant *HTT* RNA and related transcripts (e.g., *HTT1a*) are reduced.
- 20 124. A method of reducing expression of HTT comprising a polyglutamine repeat exceeding 35 consecutive glutamine residues in a subject, the method comprising transfecting or transducing a target cell, more particularly a neuron, in the subject with the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the composition of any one of claims 114-118.
- 30 125. A method of correcting at least one mutation in an HTT exon sequence in an HTT pre-mRNA in a target cell of a subject, the method comprising administering to the subject the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the
- 35

HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the composition of any one of claims 114-118.

126. A method of treating Huntington's disease in a subject in need thereof, the method comprising administering to the subject the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the composition of any one of claims 114-118 in a therapeutically effective amount.

127. The method of any one of claims 119-126, the method comprising administration of the HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the composition of any one of claims 114-118 to the subject's brain.

128. The method of any one of claims 119-127, wherein the subject is a mammal, preferentially a rodent, non-human primate, or a human.

129. The method of any one of claims 124-128, wherein the subject is genetically predisposed to have HD or has been diagnosed with HD.
130. The HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the
5 nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic
10 acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the composition of any one of claims 114-118 for use in preventing or treating HD in a subject in need thereof.
- 15
131. The HTT nucleic acid trans-splicing molecule of any one of claims 1-28 or 34-43; the nucleic acid trans-splicing molecule of any one of claims 29-33; the MSH3 exon skipping nucleic acid construct of any one of claims 44-64; the MSH3 miRNA nucleic acid construct of any one of claims 65-69; the MSH3 nucleic acid trans-splicing molecule of any one of
20 claims 70-74; the HTT trans-splicing and MSH3 exon skipping nucleic acid construct of any one of claims 75-80; the HTT trans-splicing, HTT miRNA, and MSH3 exon skipping nucleic acid construct of any one of claims 81-84; the HTT trans-splicing and MSH3 miRNA nucleic acid construct of any one of claims 85-88; the vector of any one of claims 99-104; the proviral plasmid of claim 105; the AAV of any one of claims 89-91 or 106-113; or the
25 composition of any one of claims 114-118 for use in the preparation of a medicament for the treatment or prevention of HD in a subject in need thereof.
132. A method comprising introducing into a cell a nucleic acid trans-splicing molecule configured to splice to both a first target pre-mRNA and a second target pre-mRNA, wherein
30 the splicing to the first target pre-mRNA corrects a defect in the first target pre-mRNA, and wherein the splicing to the second target pre-mRNA introduces a defect in the second target pre-mRNA.
133. The method of claim 132, wherein the nucleic acid trans-splicing molecule comprises a
35 first binding domain configured to target an intron of the first target pre-mRNA and a second binding domain configured to target an intron of the second target pre-mRNA.

134. The method of claim 132 or 133, wherein the nucleic acid trans-splicing molecule further comprises a coding domain sequence comprising a functional sequence of one or more exons of the first target pre-mRNA that corrects the defect in the first target pre-mRNA.
- 5
135. The method of any one of claims 132-134, wherein the defect in the second target pre-mRNA comprises a frameshift in the coding sequence of the second target pre-mRNA.
136. The method of claim 135, wherein the frameshift creates a premature termination codon in the second target pre-mRNA.
- 10
137. The method of any one of claims 132-136, wherein the defect comprises the endogenous start codon of the second target pre-mRNA being eliminated.
138. The method of any one of claims 132-137, wherein the defect comprises an inserted 5' UTR that prevents translation of the protein encoded by the second target pre-mRNA.
- 15
139. The method of any one of claims 132-138, wherein the defect comprises an inserted 3' UTR that destabilizes the pre-mRNA or prevents export of the second target pre-mRNA from the nucleus.
- 20
140. The method of any one of claims 132-139, wherein the defect comprises elimination of a 5' cap or a 3' polyA tail from the second target pre-mRNA.
141. The method of any one of claims 132-140, wherein the defect causes nonsense-mediated decay of the second target pre-mRNA.
- 25
142. The method of any one of claims 132-141, wherein the introducing causes the abundance of gene product of the second target pre-mRNA in the cell to be reduced compared to the abundance of the gene product before the introducing.
- 30
143. A method of reducing the abundance of a protein in a cell, the method comprising introducing into the cell a nucleic acid trans-splicing molecule that introduces a defect into a pre-mRNA that encodes the protein.
- 35
144. The method of claim 143, wherein the defect comprises one or more of the following:

- 5 (a) a frameshift introduced into the coding sequence of the pre-mRNA;
(b) elimination of the endogenous start codon of the pre-mRNA;
(c) introduction of a premature stop codon into the coding sequence of the pre-mRNA;
(d) replacement of the endogenous coding sequence of the pre-mRNA with an
alternative coding sequence;
(e) insertion of a 5' UTR that prevents translation of the endogenous coding sequence of
the pre-mRNA;
(f) insertion of a 3' UTR that destabilizes the pre-mRNA;
10 (g) insertion of 3' UTR that prevents the pre-mRNA from being exported from the
nucleus;
(h) elimination of a 5' cap from the pre-mRNA; or
(i) elimination of a 3' polyA tail from the pre-mRNA.

15 145. The method of claim 143 or 144, wherein the protein is MSH3.

146. The method of claim 145, wherein the nucleic acid trans-splicing molecule comprises a
binding domain that binds to an intron of the pre-mRNA.

20 147. The method of claim 146, wherein the nucleic acid trans-splicing molecule comprises a
heterologous coding domain sequence.